ICAE

Instituto Complutense de Análisis Económico

When More Is Not Better: Heterogeneous Dose–Response Effects of R&D Subsidies by Firm Size

Diego Sancho-Bosch
Department of Economic Analysis
Universidad Complutense de Madrid
Spain

Elena Huergo
ICAE – Department of Economic Analysis
Universidad Complutense de Madrid
Spain

Abstract

This paper examines how the level of public R&D subsidies and firm size jointly influence firms' net R&D investment. Using data on Spanish manufacturing firms from 2008 to 2018, we estimate parametric and non-parametric dose–response functions after applying entropy weighting to balance covariate distributions across treatment levels. The results reveal an inverted U-shaped relationship between subsidy intensity and net R&D expenditure for small, medium-sized, and large firms, but not for very large firms, which display a negative linear pattern. We also find substantial heterogeneity in subsidy effects within both the SME and large-firm categories, and show that the public funding share of R&D expenditure at which the positive impact of subsidies peaks declines markedly with firm size. These findings suggest that support schemes should implement progressively lower maximum subsidy rates, rather than relying on only two distinct caps for SMEs and larger firms. Overall, the results underscore firm size as a critical determinant of innovation policy effectiveness and provide practical guidance for optimizing subsidy design.

Keywords: R&D support, policy evaluation, dose-response, entropy balancing.

JEL Classification: L24; L25; O32; R11

ICAE Working Paper # 0925

December, 2025

ISSN: **2341-2356**
WEB DE LA COLECCIÓN:
http://www.ucm.es/fundamentos-analisis-economico2/documentos-de-trabajo-del-icae

When More Is Not Better: Heterogeneous Dose–Response Effects of R&D Subsidies by Firm Size

by

Diego Sancho-Bosch and Elena Huergo*

Abstract

This paper examines how the level of public R&D subsidies and firm size jointly influence firms' net R&D investment. Using data on Spanish manufacturing firms from 2008 to 2018, we estimate parametric and non-parametric dose–response functions after applying entropy weighting to balance covariate distributions across treatment levels. The results reveal an inverted U-shaped relationship between subsidy intensity and net R&D expenditure for small, medium-sized, and large firms, but not for very large firms, which display a negative linear pattern. We also find substantial heterogeneity in subsidy effects within both the SME and large-firm categories, and show that the public funding share of R&D expenditure at which the positive impact of subsidies peaks declines markedly with firm size. These findings suggest that support schemes should implement progressively lower maximum subsidy rates, rather than relying on only two distinct caps for SMEs and larger firms. Overall, the results underscore firm size as a critical determinant of innovation policy effectiveness and provide practical guidance for optimising subsidy design.

Keywords: R&D support, policy evaluation, dose-response, entropy balancing

JEL classification: L24; L25; O32; R11.

---

*Sancho: Department of Economic Analysis, Universidad Complutense de Madrid, 28223, Spain (e-mail: diesan07@ucm.es); Huergo: Department of Economic Analysis, GRIPICO and ICAE. Universidad Complutense de Madrid, 28223, Spain. (e-mail: ehuergoo@ucm.es)

This research has been partially funded by the Spanish Ministry of Economy and Competitiveness (project PID2020-112984GB-C21). Diego Sancho also received financial support from Universidad Complutense de Madrid and Banco Santander (CT58/21-CT59/21).

1. Introduction

A large and continually expanding body of literature has examined public subsidy programs designed to foster business research, development, and innovation (R&D)[1] from multiple perspectives, exploring their effects across a wide range of countries and employing diverse methodological approaches. Empirical studies have traditionally assessed the effectiveness of these instruments by estimating their impact on firms' R&D investments. In general, such analyses rest on the assumption that public subsidies are effective when they induce additional private R&D expenditure (see reviews by Zuñiga-Vicente et al., 2014; Becker, 2015; and Dimos and Pugh, 2016).

From a methodological point of view, most of this literature has estimated the impact of public subsidies in terms of a single average treatment effect based on a binary distinction between firms receiving and not receiving support (Bellucci et al., 2019; Acebo et al., 2022; Heijs et al., 2022). However, the effect of public subsidies may vary depending on the amount of subsidy received, i.e., the 'dose' of the treatment.

The existing studies that apply a framework to account for the full distribution of subsidy levels are still scarce and find heterogeneous results, reflecting differences in national contexts and the specific features of the subsidy program (Dai and Cheng, 2015; Marino et al., 2016; Cerulli and Potì, 2016; Nilsen et al., 2020; Cerulli et al., 2022; Kiman and Jongmin; 2022). Nonetheless, they consistently confirm the non-linearity of the effect of financial aid intensity on firms' R&D efforts or outcomes. In many cases, the shape of the relationship between the subsidy amount and R&D investment is inverted U-shaped, which points to the existence of a threshold at which the direction of the effect shifts. This evidence raises the question of whether there exists an optimal subsidy level that maximizes its impact on net R&D investment, defined as business R&D spending net of public subsidy funding. Furthermore, it is worth considering whether such an optimal level of subsidy, if any, varies according to firm size.

In the case of selective public subsidies, the amount of the R&D grant is typically directly related to the scale of the funded project, which in turn tends to be associated with firm size.[2] In this context, an important question is whether the subsidy rate should be uniform across project sizes or instead higher for smaller projects or firms, since firms of different scales face distinct strengths and weaknesses that shape both barriers to R&D investment and the effectiveness of innovation policy interventions (Hewitt-Dundas, 2006; Akcigit, 2009; Ortega-Argilés et al., 2009).

In this regard, previous evidence mostly suggests that smaller firms respond more positively than larger firms to support policies in terms of net R&D expenditure (Becker, 2019; Bloom et al., 2019). However, studies that explicitly place firm size at the centre of

[1] Throughout the paper, we use the term R&D to encompass firms' research, development, and innovation activities.

[2] Following Colombo, Grilli, and Murtinu (2011), we classify a subsidy as "selective" if it is a public subsidy allocated through a competitive mechanism that entails an ex ante assessment of the firm's R&D project by the public agency.

the analysis of heterogeneity—as opposed to treating the number of employees merely as a control variable—present more nuanced findings.[3] These studies mainly test whether the average treatment effect on the treated differs between SMEs and large firms. For instance, González and Pazó (2008), Huergo and Moreno (2017) and Criscuolo et al. (2019) find that the positive effects of subsidies are more pronounced among SMEs than larger firms, whereas Afcha and Quevedo (2016) find stronger effects on inputs among larger firms. Herrera and Sánchez-González (2013), who distinguish among small, medium, and large firms, conclude that public policies are effective for the first two groups but not for the largest firms. None of these studies accounts for the amount of subsidy awarded to the firms when analysing the impact of this instrument of public support.

In practice, following the General Block Exemption Regulation (GBER), most European countries include in their national programs maximum subsidy rates that differ depending on company size. Some examples are the Innovative UK Smart Grants programme, the BPiFrance programme, the Enterprise Ireland R&D fund or Business Finland funding programmes, which establish lower maximum rates for large firms. In other countries, the firm-size dimension is addressed through R&D grant schemes specifically targeted at SMEs, such as the Small Business Innovation Research (SBIR) and Small Business Technology Transfer (SBTR) programmes in the United States, the Central Innovation Programme for SMEs (ZIM) in Germany, or the National Research Council of Industrial Research Assistance Program (NRC IRAP) in Canada.[4] This underscores the interest in examining not only whether public subsidies generate additional private investment, but also whether the amount or intensity of such aid affects net R&D investment differently depending on firm size.

The aim of this study is to contribute to the existing literature by integrating these dimensions. First, we estimate the causal response of firms' net R&D expenditure to the amount of subsidy received and examine whether there is a threshold beyond which the subsidy no longer generates positive effects. Second, we perform this analysis across four firm-size groups to assess whether the optimal level of subsidy—if it exists—differs among small, medium-sized, large, and very large firms. Specifically, we classify firms as 'small' if they employ 49 or fewer workers, 'medium-sized' if they employ between 50 and 249, 'large' if they employ between 250 and 499, and 'very large' if they employ more than 500. This classification allows for a systematic comparison that extends beyond the size groups used in previous studies. To the best of our knowledge, no prior study has simultaneously examined these dimensions—continuous treatment, non-linear effects, and size heterogeneity.

[3] Among these studies, some choose to focus exclusively on SMEs, which precludes comparisons with other segments (Bellucci et al., 2019, Tingvall and Videnord, 2020).
[4] For further details, see European Commission (2014), UK Research and Innovation (n.d.), BPiFrance (n.d.), Enterprise Ireland (n.d.), Business Finland (n.d.), National Research Council (2008), BMWi – Federal Ministry for Economic Affairs and Energy (n.d.), and National Research Council Canada (n.d.).

This analysis represents a step forward in the evaluation of innovation policy as it offers a more precise understanding of the effects of public financial support on recipient firms. Moreover, it provides useful insights for policymakers, enabling them to more accurately determine the appropriate level of funding to allocate depending on the size of the firm applying for support.

To conduct this research, we use data on Spanish manufacturing firms from the Survey on Business Strategies (Encuesta Sobre Estrategias Empresariales, ESEE), which covers the entire Spanish territory from 2008 to 2018. Our objective is to assess the impact of public subsidies on beneficiary firms, drawing on a dataset that comprises 1,996 observations of firms that received subsidies for R&D activities.

Methodologically, to address the potential selection problem associated with estimating the causal response to the treatment dose, we adopt a two-step strategy. First, we use entropy weights as a pre-processing method for the sample of treated firms (i.e., those that receive subsidies intended to promote R&D). Entropy weighting involves a reweighting scheme that assigns a scalar weight to each treated unit in the sample so that the covariates in the model are balanced, meaning that their distribution is similar across all levels of the treatment. Second, we use these weights to estimate both parametric and non-parametric dose-response functions (DRF) in order to enhance the robustness of the results obtained.

Our results confirm the existence of a non-linear relationship between the level of subsidies and net R&D expenditure, which holds for small, medium-sized, and large firms, but not for very large ones, which display a negative and linear pattern. For the former groups, this relationship takes a concave (inverted U-shaped) form. We also find that the public funding share of R&D expenditure at which the positive effect of subsidies peaks—and beyond which it begins to diminish—declines markedly with firm size. This suggests that support schemes should apply progressively lower maximum subsidy rates, rather than simply differentiating between SMEs and larger firms. Furthermore, the effects of subsidies also vary within both the SME and large-firm categories, indicating that a more finely tuned calibration of funding intensity is required.

Our paper contributes to the recent literature applying DRF approaches to examine the effects of R&D subsidies as a function of the amount of public support received by firms (Nilsen et al., 2020; Cerulli et al., 2022; Kiman and Jongmin, 2022). However, previous studies differ from ours in several important respects. First, we model the dose as a genuinely continuous variable, whereas Nilsen et al. (2020) and Cerulli et al. (2022) discretize subsidy intensity into quartiles and deciles, respectively. Although Kiman and Jongmin (2022) also adopt a continuous specification, they balance covariate distributions across treatment levels using a Generalized Propensity Score, whereas our entropy-weighted preprocessing approach proves to be more effective (Vegetabile et al., 2021). Second, we focus on firm size as the key source of heterogeneity in the policy's effects, whereas Kiman and Jongmin (2022) restrict their analysis to SMEs. In the case of Nilsen et al. (2020), the emphasis is on firms' prior innovation trajectories, while

Cerulli et al. (2022), although considering some contextual factors, do not introduce explicit firm-level differentiators. Moreover, these two studies focus on the impact of public subsidies in terms of output additionality, whereas our analysis concentrates on the input side of R&D activities.

A further novelty of our analysis is that we provide evidence across the whole firm-size spectrum—rather than focusing solely on SMEs or applying two broad sample cuts—thereby uncovering systematic size-related heterogeneity. Finally, we identify group-specific interior optima, approximating the threshold subsidy level at which the effect peaks for each size group, which provides actionable guidance for calibrating subsidy intensity by firm size.

The paper is structured as follows. The next section reviews the literature on the advantages and disadvantages that firms of different sizes have when undertaking innovation activities and how these characteristics may influence the impact of subsidies on net R&D investment. The subsequent sections describe the methodology employed, the institutional framework of the subsidy schemes and the source of information. Finally, Sections 5 and 6 present the results and outline the main conclusions.

2. Firm Size: Innovation, Financial Constraints and the Amount of Public Support

One of the main justifications for public intervention to promote R&D is the correction of market failures that constrain firms' ability to initiate or maintain innovative projects, thereby creating a negative gap between the socially optimal and the private level of investment (Nelson 1959; Arrow 1962; Bloom, 2019).[5] A firm's size plays a crucial role in understanding this issue as different barriers to closing this gap may arise depending on firm size. These include the availability of sufficient financial resources and technological capabilities, absorptive and organisational capacity and the presence of economies of scale and scope (Hewitt-Dundas, 2006; Akcigit, 2009; Ortega-Argilés et al., 2009).

This debate has a long intellectual lineage in the Schumpeterian tradition. Schumpeter's early work emphasised the entrepreneur and new entrants as primary drivers of innovation, whereas his later perspective highlighted the role of large, established firms leveraging scale and accumulated knowledge to sustain innovation-led competition (Schumpeter, 1934; 1942). Building on this legacy, evolutionary and sectoral analyses distinguish between Schumpeter Mark I, where innovation is decentralised, driven by entrants and smaller firms in turbulent environments with low entry barriers, and Schumpeter Mark II, where innovation is concentrated within large incumbents that benefit from economies of scale, accumulated capabilities and stronger appropriability

[5] Some authors extend this view by emphasizing that public financing of innovation is not only intended to compensate for private underinvestment, but also to create and shape markets along the innovation chain (Mazzucato and Semieniuk, 2017).

regimes (Pavitt, 1984; Breschi et al., 2000). This distinction underscores that the relationship between firm size and innovation is not universal but depends on the technological regime, the structure of opportunities and the extent of knowledge cumulativeness in each sector.

Consistent with these Schumpeterian patterns, organisational dynamics also change with firm size, shaping how innovative processes are structured and executed within firms. In organisational terms, as firms grow, they tend to lose managerial and structural flexibility. At the same time, innovation processes within the firm become increasingly formalised, making knowledge transfer more difficult and delaying project execution (Vossen, 1998; Keum and See, 2017). In smaller businesses, moreover, ownership and management are often intertwined (Rothwell, 1989). Combined with a less bureaucratic innovation environment (Ortega-Argilés et al., 2009) and behavioural advantages among employees (Patterson and Kerrin, 2014). This suggests that smaller firms may be better positioned to adapt to the innovation demands of the market and identify niche opportunities that allow them to pursue innovation more efficiently given their available resources (Bellucci et al., 2019).

These facts would imply that smaller firms tend to exhibit higher labour productivity in innovation-related activities than their larger counterparts (Akcigit, 2009; Hall et al., 2009), generating, ceteris paribus, increasing returns for each additional monetary unit invested (Cohen and Klepper, 1992), especially in high-tech sectors (Kim, 2018). Along these lines, authors such as Pavitt et al. (1987) and Hall et al. (2009) emphasise the R&D capabilities of small firms, providing evidence of higher relative R&D expenditure. At the same time, smaller firms tend to exploit their technological knowledge more efficiently within the market for technology. This organisational orientation enhances the economic use of patents and contributes to a more dynamic and accessible innovation environment, in contrast to the more rigid and internally focused strategies typically adopted by larger corporations (Gambardella et al., 2007).

Nevertheless, larger firms also benefit from several advantages, which are directly or indirectly related to their access to financial resources for undertaking research and development projects (Hall and Lerner, 2010; Meuleman and De Maeseneire, 2012). As is well known, the presence of sunk and adjustment costs, together with the very nature of this type of investment, typically requires substantial investment in fixed assets (Athreye et al., 2021). In addition to financing the innovation project itself, firms may also need resources to ensure the appropriability of the benefits generated by the innovation (Hall and Lerner, 2010). First, they must cover the monetary costs associated with patenting, such as administrative expenses for creating and maintaining a patent portfolio, or the cost of hiring specialised personnel to manage this portfolio and assess its market viability. Second, it is essential for firms to have access to physical and logistical assets, including distribution and supplier networks, manufacturing and storage infrastructure and marketing capabilities (Teece, 1986). Third, once the innovation is commercialised, substantial capital is often required for complementary strategies, such as brand development, trade secrecy, and advertising campaigns that target consumers

(Cassiman and Veugelers, 2002). Meeting all these costs may be overwhelming for firms with limited resources or no access to external finance (Hewitt-Dundas, 2006).

Taken together, these factors, along with the fact that innovation occurs in a dynamic and uncertain environment, means that firms usually finance most of their R&D expenditure with internal funds (Kamien and Schwartz, 1978). If external finance is needed, these firms are generally able to turn to capital markets and make use of financial leverage mechanisms in line with their resource capacity.

In this context, large firms not only have greater internal capacity to fund these investments but also possess sufficient tangible assets to serve as collateral when seeking loans in financial markets.[6] By contrast, smaller firms often face financial constraints due to more limited internal resources and greater difficulty in accessing conventional external finance.[7] In fact, recent evidence supports that SMEs are more reluctant to patent their innovations compared to larger firms, due to cost barriers rather than issues related to innovation quality or patent infringement (Athreye et al., 2021).

To summarise, with regard to R&D investment, smaller firms tend to benefit from organisational and behavioural advantages, whereas firms with larger workforces typically enjoy resource-related advantages (Rothwell, 1985). Furthermore, following Hewitt-Dundas (2006) and Meuleman and De Maeseneire (2012), the financial and logistical constraints faced by small firms are more difficult to overcome than the limitations faced by large firms, such as a lack of internal staff experience or resistance among employees to learning and adapting in order to implement new innovation projects. As firms grow, they often experience a transitional phase in which increasing organisational complexity and formalisation can reduce flexibility and slow decision-making before scale-related innovation advantages fully materialise. Evidence shows that expanding firms tend to introduce more structured routines and coordination layers, which can support innovation implementation yet constrain early-stage exploration and market adaptation (Damanpour and Aravind, 2012).

In this context, it is reasonable to assume that the impact of public financial support for R&D also varies according to the size of the recipient firm as R&D subsidies may help to alleviate firms' financial constraints. This support can operate through both direct and indirect channels.

Receiving public funding directly can ease the substantial financial burden required to undertake innovation investment internally, reduce managers' risk aversion and enhance competitiveness in the firm's market. As a result, such support may help increase the number of SMEs willing to engage in innovation, enabling them to initiate their projects.

---

[6] Usually, the most valuable assets of small innovative companies tend to be intangible assets (patents, ideas, skilled personnel, etc.) that are difficult to use as collateral because of the lack of knowledge of quantification by potential creditors (Meuleman and De Maeseneire, 2012). This fact can lead to a disparity between the value of these assets for the companies and the loan granted (Fischer and Ringler, 2014).

[7] However, work such as that of Bellucci et al. (2023) shows that the patent ownership of companies can attract other types of financing for innovation projects, such as venture capital financing.

Evidence suggests that a significant proportion of firms would not carry out any form of innovation investment without receiving a subsidy (Hall et al., 2009; Becker, 2015).

Public subsidies can also ease the financial burden of R&D projects through their signalling effect, partly mitigating the problem of asymmetric information between firms and private investors or credit institutions. This, in turn, facilitates access to external finance for supported firms or significantly reduces its cost (Meuleman and De Maeseneire, 2012). Furthermore, in periods of economic recession, public subsidies aimed at promoting R&D in SMEs can be particularly effective, given that during such periods, external financing tends to be redirected towards larger, more secure companies (Grandi et al., 2024).

However, large established innovators typically plan recurrent and long-term R&D activities as part of their strategic planning cycles. These firms often maintain sustained and structured innovation efforts and respond to competitive pressures by investing in R&D to consolidate technological leadership and strengthen their absorptive capacity for innovation (Klette and Griliches, 2000; Griffith et al., 2004). As a result, public subsidies may not necessarily trigger additional research expenditure, but instead risk replacing private funds already committed to innovation (Heijs et al., 2020).

To date, the existing empirical evidence does not provide a definitive answer regarding the relationship between firm size and the effects of R&D subsidies. While some studies report stronger effects for smaller firms relative to larger ones (González and Pazó, 2008; Huergo and Moreno, 2017; Criscuolo et al., 2019), others find the opposite pattern (Afcha and Quevedo, 2016). This apparent inconsistency may partly stem from treating firms dichotomously as either SMEs or large firms. As argued throughout this section, although these groups share certain similarities, there are reasons to expect distinct innovation patterns across the firm-size distribution and, therefore, also to observe different uses and impacts of public support.

This study aims to shed light on this issue by distinguishing four firm-size categories—small, medium-sized, large, and very large firms—rather than adopting the conventional dichotomy between SMEs and non-SMEs. No specific hypotheses are proposed regarding the impact of R&D subsidies within each firm-size group; instead, the analysis is intended to be exploratory, allowing the data to speak for themselves so that the findings may serve as a foundation for future research.

3. Methodology

As previously discussed, a continuous treatment refers to a situation in which all units receive the same type of treatment, but at a different level (intensity or amount). At present, there are several methodological procedures that address the estimation of a dose–response function, that is, a function that characterizes the relationship between the treatment dose and the outcome variable of interest. Among these, notable approaches include the Generalized Propensity Score (Hirano and Imbens, 2004), the Generalized

Boosted Modeling approach (Zhu et al., 2015), the doubly robust estimation for continuous treatments (Kennedy et al., 2017), and the Entropy Balancing for Continuous Treatments method (Tübbicke, 2020, 2022).

The primary aim of these methodologies is not to estimate the 'level treatment effect', understood as the difference between the potential outcome for a unit treated at a specific dose and the outcome it would have had in the absence of treatment (Callaway et al., 2024). Instead, the focus lies on estimating the causal response, that is, the change in the potential outcome that results from a variation in the treatment dose.

However, in estimating this causal response, a potential selection problem may arise, as receiving a specific level of support likely depends on other covariates—in our framework, firm characteristics—that also influence the outcome variable. Since the treatment variable is continuous, it may span a wide range of values, which can introduce inconsistencies when comparing treated firms that receive different doses. This dimensionality problem of confounding factors becomes more pronounced as the number of treatment levels increases, leading to a progressive loss of precision in estimating the relationship with the outcome variable, a well-documented trade-off (Lechner, 2001; Crump et al., 2009).

To address this challenge, the aforementioned methodologies aim to optimise covariate balance—that is, to reweight the sample observations so that the conditional means of the covariate, namely the mean of each covariate conditional on the treatment level, are balanced across different levels of the treatment. By adjusting the observation weights, the covariates become effectively independent of the treatment in the weighted sample, thereby approximating the conditions of a randomized assignment.

In most cases, these methods involve reprocessing or weighting sample observations in such a way that the correlation between the treatment dose and each covariate is minimised while preserving the original marginal distributions required for valid inference. For this study, we employ Entropy Balancing for Continuous Treatments (EBCT), in which weighting is performed using entropy-based weights.

Accordingly, our estimation procedure involves two main steps. First, entropy weights are calculated and used to pre-process the treated sample based on the covariates, which capture the characteristics of the firms under study. Second, these entropy weights are applied to the estimation of the dose–response function. These two steps are explained in detail below.

3.1 Entropy Weights in a Continuous Treatment Framework

Entropy balancing for continuous treatments is a pre-processing procedure that involves selecting balancing weights, w_i, for each observation i so that they deviate as little as possible from a set of base weights, q_i, according to an entropy-based distance metric. At the same time, they achieve zero correlation between the treatment variable and the covariates in the reweighted sample.

Formally, following Tübbicke (2022), this requires solving the following optimisation problem:[8]

$$\min_{w} \; H(w) = \sum_{i|T_i>0} w_i \ln\left(\frac{w_i}{q_i}\right)$$

$$s.t.: \;\; \sum_{i|T_i>0} w_i g(p, X_i, T_i) = 0 \qquad [1]$$

$$\sum_{i|T_i>0} w_i = 1 \; w_i > 0 \;\; \forall\, i \mid T_i > 0$$

where $H(w)$ is the loss function to be minimised, T_i is the treatment variable and X_i is the vector of covariates that may affect both the treatment and the outcome variables. In this case, the loss function, which measures the distance between the distribution of the estimated weights W = [w_i, . . ., w_{n0}] and the base weights given by the vector Q = [q_i, . . ., q_{n0}], is the entropy divergence function (Kullback, 1959).

In our study, the base weights are uniform across all observations, i.e., $q_i = \frac{1}{N}$ where N is the number of treated observations. Based on this criterion, the balancing weights are chosen so as to deviate as little as possible from the reference weights while achieving zero correlation in the reweighted sample. In other words, the weights that satisfy the constraints of this optimisation problem simultaneously preserve the unconditional means of the treatment and covariates and also eliminate the correlation between the treatment variable and the covariates.

In the constraint of the optimisation problem, $g(p, X_i, T_i) = \left[X_i', T_i, \dots, T_i^p, X_i' \cdot T_i, \dots, X_i' \cdot T_i^p\right]'$ is a vector that may include higher-order terms p of the treatment variable and its interactions with the covariates. In our estimations, we use a second-order specification. By incorporating second-order terms, we approximate the independence between T_i and X_i more closely, thereby reducing potential sources of bias.

Correct identification of the model using this methodology relies on the fulfilment of the following assumptions. First, the Conditional Independence Assumption (CIA)—also known as the assumption of selection on observables—must hold. That is, given a set of covariates, the potential outcomes must be independent of the treatment dose. In this way, if the covariates are balanced across different levels of the treatment, any observed differences in outcomes can be attributed solely to the treatment.

To assess this assumption, we examine the R-squared values of the model's independent variables before and after entropy balancing (Tübbicke, 2023). A low R-squared suggests that the proportion of variance in the covariates is not systematically related to the

[8] To simplify notation, the expressions assume that the treatment amount and covariates are standardised to zero mean.

treatment, thereby supporting the plausibility of conditional independence (Austin, 2019; Tübbicke, 2023).

Secondly, the Stable Unit Treatment Value Assumption (SUTVA) must also be satisfied. This assumption states that the outcome for each unit depends solely on its own treatment level.

According to the literature, the use of entropy weighting offers several advantages. First, it allows for the extension of the lack of correlation between covariates and the treatment variable to higher-order sample moments and can yield notable improvements in terms of reducing both mean squared error and bias when used in conjunction with dose–response functions (Tübbicke, 2022). Second, entropy weights can eliminate many forms of bias in observational studies and are particularly effective compared to other statistical inference techniques, such as the Generalized Boosted Modelling approach, the Generalized Propensity Score (GPS), and Ordinary Least Squares (OLS) regression implemented through an inverse probability weighting scheme (Vegetabile et al., 2021). Third, entropy weighting can substantially reduce coverage error in survey-based datasets[9] (as in our case), thereby enhancing the reliability of subsequent statistical inference (Watson and Elliot, 2016).

3.2 Dose–Response Function

Once the entropy weights have been obtained, we proceed to estimate the dose-response function using the sample reweighted with using these balancing weights. Note that, in this second step, the covariates are not included as additional explanatory variables in the specification of the dose–response function; this is because they have already been accounted for in the computation of the entropy weights applied to obtain the reweighted sample.

The dose–response function can be estimated using both parametric and non-parametric methods. To assess the robustness of our results, we employ three different approaches: OLS regression, Tobit censored regression and local polynomial regression. In the first two procedures, which are parametric, following the results in previous literature (Dai and Cheng, 2015; Nilsen et al., 2020; Cerulli et al. 2022; Kiman and Jongmin; 2022), we include a quadratic term for the treatment variable to test for U-shaped or inverted U-shaped relationships between treatment and outcome and to identify the treatment dose at which the direction of the effect changes, if any. Specifically, the model estimated by OLS is:

$$E[Y_i|T_i] = \alpha_0 + \alpha_1 T_i + \alpha_2 T_i^2 + \varepsilon_i \qquad [3]$$

[9] Coverage error is the error that occurs when certain elements of the population under study are not included in the sample selection process.

where $E[Y_i|T_i]$ represents the expected value of the dependent variable Y_i given the value of the treatment variable T_i, and ε_i represents the error term.

The choice of the second methodology—the censored Tobit regression—is driven by the definition of the outcome variable, which in our analysis is the logarithm of net R&D expenditure. This variable presents in our sample a small portion of zero observations, which are associated with theoretically implausible negative values for net R&D expenditure, that are replaced with zero when applying the logarithmic transformation to avoid attrition. As explained in the following section, this issue arises primarily because of a timing mismatch between the awarding of the subsidy and its accounting by the firm. For this reason, we choose to censor the lower bound of the sample, restricting it to strictly positive values of the outcome variable, such that:

$$Y_i = \begin{cases} Y_i^* \; if \; Y_i^* > 0 \\ 0 \; if \; Y_i^* \leq 0 \end{cases} \quad [4]$$

Finally, we use a second-degree local polynomial regression, which is a non-parametric approach that allows for a more flexible estimation of the functional relationship between the treatment variable and the observed outcome. This method is based on a locally weighted linear regression, in which a polynomial of order 2 is fitted around each point of the treatment variable using nearby observations. To implement this local fitting, a kernel weighting function is applied — in our case, the Epanechnikov kernel — which determines how weights are distributed within a certain neighbourhood around the treatment value being estimated. The size of this neighbourhood is controlled by a bandwidth parameter, which is selected through an asymptotically optimal data-driven method.

All these methods are applied to the full sample as well as separately to groups defined by firm size, in order to examine whether the relationship between the treatment dose and the outcome variable varies across firm sizes.

It is worth noting that, although the data used in the analysis cover the period 2008–2018, the equations presented above do not include a time subscript as the data are treated as a pooled sample. This implies that the impact of the R&D subsidies is measured in the same year in which the firm is treated. As we explain in the following section, this decision is based on the fact that treated observations are unevenly distributed across years and firm size groups, making the use of panel data techniques unfeasible.

For the same reason, we also exclude the use of a difference-in-differences (DiD) methodology from this study. In our case, dividing the data into different periods and groups to perform a DiD analysis may lead to issues related to insufficient observations, resulting in biased estimates, limited predictive power, or both (Brewer et al., 2018). Second, there is no specific period within the sample that would allow us to study the before-and-after effect of a particular public subsidy as the subsidy system was in place throughout the entire study period. Lastly, to the best of our knowledge, there is no statistical method that allows for the integration of entropy weights with DiD estimations in the context of continuous treatments.

4. Institutional Framework and Data Sources

4.1 Spanish Institutional Context

As in many other European countries, Spanish firms have had access to both direct and indirect instruments of public financial support for business R&D since the early 1980s. However, the design and implementation of these policy instruments have evolved significantly over time.

Regarding indirect support, the Spanish R&D tax system has followed a mixed design that combines deductions based on both total (volume-based) and incremental R&D expenditure (OECD, 2023; Appelt et al., 2020). The scheme applies to most costs associated with R&D investment, including personnel, raw materials, externally contracted R&D, and asset depreciation. In 2007, tax credit rates were reduced relative to previous years, although new mechanisms were introduced to further stimulate innovation. Since 2013, firms have been allowed to monetize or request refunds for unused tax credits, provided they meet certain conditions such as maintaining employment levels and reinvesting in R&D activities. Additionally, in 2014 a 40% reduction in social security contributions was introduced for research staff, reinforcing non-fiscal incentives for R&D employment (Government of Spain, 2014). Firms receiving R&D subsidies may still claim tax credits for eligible expenditures. Collectively, these measures aim to sustain fiscal support for research while ensuring consistency with broader national budgetary and EU policy frameworks.

In terms of direct support, between 2008 and 2018, three national R&D plans were implemented, covering the intervals 2008–2011, 2013–2016, and 2017–2020, respectively (Ministry of Economy and Competitiveness, 2013). Across these plans, the financing of business R&D projects by the Spanish central government was subject to an ex-ante evaluation process for proposal selection and was channelled primarily through the grant and partially reimbursable loan schemes managed by the Centre for the Development of Industrial Technology (CDTI). In 2008, the direct funding provided by the CDTI to companies for the development of their R&D projects reached 917.29 million euros, including R&D projects, aerospace projects, and CENIT initiatives (CDTI, 2008).

In parallel, another significant portion of public support was channelled through regional calls for proposals framed within the European Regional Development Fund (ERDF) Operational Programmes (2007–2013 and 2014–2020). During the second programming period, the regions aligned their aid schemes with the Smart Specialisation Strategy (RIS3), which in practice entailed EU co-financing rates of 50–80% and the adoption of common requirements regarding traceability and outcome evaluation (European Commission, 2012).

This process of harmonizing eligibility and evaluation criteria across calls issued by different public administrations had already taken an initial step in 2008, following the entry into force of Commission Regulation (EC) No 800/2008 of 6 August 2008, declaring certain categories of aid compatible with the common market under Articles 87 and 88 of the Treaty (known as the General Block Exemption Regulation, or GBER 2008). This regulation introduced a harmonized classification of eligible research and development activities, subdividing R&D aid into industrial research, experimental development, and feasibility studies.

The category of technological development previously used in many national and regional programs (including those in Spain) was aligned with "experimental development", while the distinction between technological development and technological innovation ceased to have a separate legal basis. Consequently, Spanish administrations—both national and regional—adapted their calls for proposals, merging the two categories into a single one referred to as "technological or experimental development projects", in accordance with the Commission's definitions.

In particular, in 2008 the CDTI allocated approximately two-thirds of the total funding to this new combined category, with grant intensities reaching up to 25% of project costs. CDTI also began issuing official certifications classifying projects as either R&D or Technological Innovation, thereby allowing firms to claim corresponding tax credits directly—certifications that are binding for the tax authorities.

4.2 Database

The database used in this study is drawn from the Survey on Business Strategies (Encuesta Sobre Estrategias Empresariales, ESEE), which provides annual firm-level information on strategic decisions and capabilities, enabling longitudinal analysis of Spanish manufacturing firms' participation in public support programmes and their effects. The survey is highly representative, as participation in the survey is mandatory for all firms with more than 200 employees, while smaller firms are selected through stratified sampling and represent approximately 5% of all national firms that have between 10 and 200 employees.

Although the ESEE contains information dating back to 1990, we restrict the analysis to the period 2008–2018. As described previously, eligibility and evaluation criteria of Spanish subsidy programs have changed over time, which suggests exercising caution when jointly considering observations of firms that have received subsidies at very different points in time. In this regard, the greater harmonization of criteria brought about by the implementation of the GBER 2008 suggests using year 2008 as the temporal starting point for the analysis.

To analyse the effects by firm size, we group companies in this database into four categories: 'small' (49 or fewer employees), 'medium-sized' (between 50 and 249 employees), 'large' (between 250 and 499 employees), and 'very large' (500 or more employees). This classification is consistent with the EU definition of SMEs (2020), which builds on the European Commission's 2003 recommendations on business size categorisation. It also allows for easier comparisons with studies using other international classification. For example, the U.S. Small Business Administration (SBA) defines firms with fewer than 500 employees as small, and those above this threshold as large.

For our analysis, we consider all firms that received a subsidy in any of the years within the period under study, regardless of whether they obtained the subsidy as a stand-alone instrument or as part of a policy mix that also included other instruments of financial support.[10] The final sample consists of 1,996 annual firm-level observations, each representing a treated firm in a given year. Accordingly, a firm may appear multiple times in the sample, depending on the number of years in which it was treated.

Note that, as shown in Table 1, the distribution of firms by the number of consecutive years receiving R&D subsidies and by size category is highly uneven. Firms participating for at least two consecutive years represent only 46% of small firms, whereas the number is around 75% for very large firms. For this reason, we treat the entire sample as a pooled dataset. Introducing firm fixed effects in a panel data specification would imply losing the information from firms that received support in only one year within the period under study, which account for 41% of the sample (54% in the case of small firms). In terms of the number of observations, the sample distribution is as follows: 14.2% correspond to small firms, 44.2% to medium-sized firms, 18.4% to large firms, and 23.1% to very large firms. This composition by firm size deviates considerably from that of the original database, thereby precluding the estimation of a single dose–response function for the full sample.

4.3 Model Variables

The treatment variable in our analysis is the annual amount of subsidy received by a firm to carry out its R&D activities. For the analysis, this variable is transformed into logarithms, which ensures a normal distribution, a necessary condition for applying the methodology described in the following section.[11] The amount of public funding granted as R&D subsidies ranges from €2,200 to €31 million over the period of study. The average subsidy is €59,000 for small firms, €125,000 for medium-sized firms, €250,000 for large firms, and €473,000 for very large firms.

[10] Although the effect of the subsidy may differ when it is the only source of public support (Marino et al., 2016), excluding firms that benefited from a policy mix could introduce attrition bias and undermine the external validity of the study. In any case, the ESEE database does not provide information on the monetary amounts received from subsidised loans, which prevents for analysing both sources of public support (subsidies and loans) separately.

[11] To avoid the influence of outliers, extreme values (the top and bottom 1%) are removed, resulting in the loss of 33 observations.

The outcome variable used to assess the impact of the treatment is the annual logarithm of net business R&D expenditure, which corresponds to the total innovation investment minus the public subsidy received through grant programmes.

It should be noted that the information on R&D expenditures collected in the ESEE refers to firms rather than projects. Naturally, in a given year, a firm may undertake more than one project, particularly in the case of large firms. Due to the anonymization of firms in the database, we do not know how many projects each firm carries out, nor whether the awarded subsidy corresponds to a single project or to multiple ones. Moreover, the database does not provide information on project duration, so it is not possible to determine whether the subsidy received in year t corresponds to a project completed within that year or to one extending over several years.

This may be the reason why a small portion of observations in the sample correspond to negative values for net R&D expenditure: some firms may not annualise the public investment received and instead report the total amount in a single year, even if the project spans multiple periods.

To avoid losing observations that may be relevant to the analysis when applying the logarithmic transformation, we recode these negative values to €1, so that the logarithmic transformation yields a value of 0. This approach has several advantages over simply removing the observations. First, retaining these atypical—but not erroneous—values may be useful for our estimation or for the weighting of entropy scores because of their characteristics in the covariates. Second, since no other treated firm that receives a subsidy reports a net R&D expenditure of €1, the recoded observations are uniquely identifiable within the estimations, allowing for the application of methods that control for this fact.[12]

Regarding the covariates used in the analysis, and following previous literature (Heijs et al., 2020), we primarily consider intrinsic firm characteristics that may influence the outcome variable in addition to the subsidy: age, sales, size, productivity, patent grants, and innovation generation. We also include year, and sectoral dummies based on Pavitt's (1984) taxonomy, reflecting the Schumpeterian insight that industry-specific technological regimes shape firms' innovative dynamics. Moreover, following the literature reviewed in Section 2, we include two financial variables: the long-term leverage ratio and the ratio of tangible fixed assets to total liabilities. The inclusion of the former is justified by the need to finance R&D-related costs internally, whereas the latter captures the use of tangible assets as collateral to secure external funding.

Finally, in the robustness analysis presented at the end of the results section, we also include the variable for personnel costs as a complement to the number of employees. This variable indirectly reflects the human capital component of R&D effort through the associated labour costs (Hall and Lerner, 2010), given the potential heterogeneity in wage structures within firms.

[12] For example, the censored Tobit regression described in Section 3.

Table 2 presents the definitions and basic descriptive statistics of all variables in our model, both for the full sample and by firm-size group.

5. Estimates and Results

5.1 Reweighting Using Entropy Weights

Before proceeding with the estimation of the dose–response function, we must assess the covariate balance and the extent of information loss in the sample due to the application of entropy weights. If balance is properly achieved, then, after weighting, the distribution of confounding factors prior to treatment exposure should resemble that of the target population for which the inference is drawn.

To carry out this assessment, we perform several statistical tests, the results of which are presented in Table 3.[13] First, we examine Pearson correlations to determine whether entropy weighting has succeeded in eliminating significant correlations between the treatment variable (logarithm of the monetary amount received from R&D subsidies) and the covariates included in our model.[14]

Second, we assess balance via the conditional mean to ensure that the marginal distributions of the treatment variable and the covariates remain unchanged in the weighted distribution. To obtain the conditional mean of each variable, we estimate an OLS regression of the treatment variable on the model covariates before and after applying the entropy weights (Vegetabile et al., 2021).

As shown in Table 3, balancing the sample using entropy weights has substantially reduced the correlations observed in the unweighted sample, bringing them down to values that are virtually zero. In the reweighted sample, all Pearson correlation coefficients are below 0.1 and are therefore acceptable as evidence of the effectiveness of this balancing procedure. This confirms that, after reweighting, there are no systematic differences between the treatment variable and the covariates. In other words, the entropy weights have successfully equalised the distribution of the covariates across the different treatment levels, ensuring that their marginal distributions remain balanced.

5.2 Effect of Subsidy Dose on Net R&D Expenditure

Having established that the entropy weights are satisfactory, we proceed to estimate the dose–response function that characterises the average effect of each subsidy dose on firms' net R&D expenditure. To test the robustness of our results, we estimate this dose–response function using three different econometric approaches, applying entropy weights in all cases to reweight the sample. The first approach is a parametric analysis

[13] These same tests to corroborate the suitability of the entropic weights in each of the subsamples by size ranges considered can be found in the supplementary tables OA1 and OA2 in the Online Annex.

[14] We recall that, in order to perform the balancing with entropy weights, it is necessary to solve the optimisation problem, which requires minimising the uniform weights of the Kullback function. Therefore, it is necessary that these weights be subject to zero correlation with respect to the treatment variable.

based on OLS regression. The second employs a Tobit censored regression model to account for potential bias that arises from the conversion of anomalous negative values to zero in the dependent variable. The third approach is a non-parametric method based on local polynomial regression.

The coefficients estimated using the parametric methods are summarised in Table 4, while Figure 1 displays the dose–response functions obtained using the three methodologies. Regarding the parametric estimations, Columns (1) and (2) of Table 4 present the results for the full sample, whereas Columns (3) to (10) refer to the subsamples by firm-size category. Regardless of the specific sample, the coefficients in the Tobit estimations—both positive and negative—are larger in magnitude than those from the OLS models, which is consistent with estimated dose–response functions in Figure 1. The function estimated using the Tobit censored regression exhibits a slightly more pronounced curvature in its lower section than those obtained through OLS and local polynomial regression. This underscores the importance of retaining the zero values used to recode the anomalous negative values for net R&D expenditure. Nevertheless, similar patterns are obtained with all methodologies.

The results in Columns (1) and (2) of Table 4 indicate that, regardless of the estimation method, failing to account for firm size introduces such heterogeneity that the subsidy effect becomes statistically insignificant. This finding is also consistent with the composition of our sample by firm size, which, as discussed in the previous section, deviates from that of the original database. While size-specific estimations preserve the representativeness of the ESEE, the estimation that pools the entire sample combines firms with unequal degrees of representativeness.

As for the size-specific samples, both estimated parameters and dose-response functions appear to confirm a quadratic relationship between the amount of the subsidy (in logs) and net R&D expenditure (also in logs) in the case of small, medium-sized and large firms, whereas for very large firms it becomes negative and linear.[15]

However, as suggested by Lind and Mehlum (2010) and Haans et al. (2016), while the parametric estimates for the first three groups of firms may indicate a concave relationship, this is a necessary but not sufficient condition for confirming the existence of an inverted U-shaped relationship in the model. The observed correlation may simply result from the inclusion of the quadratic terms, potentially generating spurious marginal effects that could lead to a misinterpretation of the model's true linearity. For a truly inverted U-shaped relationship, that is, a quadratic relationship with a concave parabola, the function must exhibit a positive slope at the left-hand side followed by a negative slope at the right-hand side across a specific interval of values, which implies testing two combined null hypotheses.

To do so, we use the methodology proposed by Lind and Mehlum (2010), who adopt the framework developed by Sasabuchi (1980) for testing a composite null hypothesis.

[15] Additional tests (not reported here) confirm the absence of higher-order curvature for very large firms, supporting the linear relationship for this group.

Accordingly, they propose a test in which the composite hypothesis posits that the outcome variable increases at low levels of the treatment and decreases at high levels, or vice versa, depending on whether the test concerns an inverted-U-shaped or U-shaped relationship. The second component of Lind and Mehlum's methodology involves calculating the turning point of the function within the observed range of the treatment variable. To identify this extremum point (the maximum in the case of inverted U-shaped relationships), they propose using 95% Fieller confidence intervals, constructed for the ratio of the relevant model parameters (Fieller, 1954).

Lind and Mehlum's U-shape tests in Table 4 confirm the existence of an inverted U-shaped relationship between public subsidies and firms' net R&D investment. In addition, the points of maximum impact would correspond to dose amounts of €7,111, €8,103, and €28,000 for small, medium-sized, and large firms, respectively. Beyond these points, further increases in R&D subsidies would no longer yield additional investment.

As shown in Table 5, the estimated optimal doses are relatively small compared with the average subsidy within each firm-size stratum. Accordingly, the estimated turning-point doses should be interpreted as lower bounds of the true optimal levels. In this regard, the shares of observations below the estimated optimal doses—8.5%, 2.2%, and 8.5% (14.3%, 5.7%, and 12.5% when considering the upper limit of the Fieller interval)—for small, medium-sized, and large firms, respectively, suggest that a non-trivial proportion of firms may have been underfunded relative to the level of public support required to achieve the maximum net R&D investment response.

To facilitate the discussion, Table 5 also reports, for each size-specific sample, the normalized position of the estimated optimal subsidy within the observed range of subsidy amounts, the average annual share of R&D expenditure financed through subsidies, and the implicit share at the estimated optimal dose. This implicit share is computed in two different ways: (i) as the average of the shares of firms that received subsidy amounts close to the estimated optimum,[16] and (ii) as the average of the shares that represent the optimal subsidy over the R&D expenditure that would result from the sum of this subsidy and the observed net R&D investment.

These numbers reveal two main patterns. On the one hand, the optimal subsidy rate is estimated to range between 36.3% and 41.0% for small firms, 31.1% and 31.8% for medium-sized firms, and 17.0% and 24.8% for large firms. This indicates that the intensity of public funding required to achieve the maximum effect in terms of net R&D expenditure decreases substantially with firm size. On the other hand, since the subsidy shares at the mean level exceed the estimated optimal intensities across small, medium-sized, and large firms, it is likely that many firms would have achieved higher net investment with lower subsidy amounts.

---

[16] Since the results obtained for the optimal point are based on estimates, there is no firm whose subsidy amount exactly matches the value at which the effect is maximised. For this reason, we use a margin of ±0.1 logarithmic units around that point.

Delving into what happened for each size category, we see that, in the case of small firms, the position of optimal subsidy within the observed range is around 19.6%, while the effect of the subsidy peaks earlier (around 12% within the range) in the case of medium-sized firms. This evidence suggests that small firms are able to maintain positive returns to public support over a wider range of subsidy intensities than medium-sized firms. This pattern is consistent with the argument that the less bureaucratic structures and closer alignment between ownership and management enable small firms to react more flexibly to market opportunities and to identify niche areas where innovation can be pursued effectively (Bellucci et al., 2019).

In addition, the group of medium-sized firms shows a lower proportion of observations below the optimal subsidy level. In this line, firms of intermediate size may often lose part of the managerial and structural flexibility that characterises small firms, while not yet benefiting from the economies of scale or possessing the level of resources available to larger organisations. As a result, medium-sized firms appear to be "caught in the middle," which may explain their relatively weaker response to public subsidies compared with the other size groups.

In turn, large firms (employing between 250 and 499 workers) exhibit a more balanced relationship between (net) private and public investment in the implementation of their innovation projects. For this group, the normalized position of the optimal subsidy amounts to 20% of the observed range of subsidy intensities within this size category. This likely reflects the fact that, compared with smaller firms, they possess a broader set of resources that enhances their credibility with both credit institutions and private investors, allowing them to finance innovation activities more easily and at a lower cost (Hall and Lerner, 2010; Meuleman and De Maeseneire, 2012). Consequently, they are better positioned to manage the appropriability of the benefits and surplus generated by their innovations. They are also better equipped to absorb sunk and adjustment costs, as well as the intrinsic characteristics of this type of investment, which typically entails substantial capital commitments in fixed and logistical assets (Athreye et al., 2021). Moreover, these companies are less affected by some of the key disadvantages they share with firms that employ more than 500 workers. They do not have such rigid managerial structures and are not so large as to experience the same degree of organisational difficulties as very large firms, which facilitates internal knowledge transfer (Vossen, 1998; Keum and See, 2017).

Finally, firms classified as very large appear to exhibit a negative linear relationship between the subsidy dose and net R&D investment. This trend suggests that, when public support is substantial enough to cover most of a project's costs, firms may reallocate or scale back their own previously planned resources, effectively substituting public funds for private ones (Heijs et al., 2020). Because of their considerable size, it is common for very large firms to allocate a portion of their private resources to R&D activities. They do this, among other reasons, to avoid falling behind and losing market share to competitors who may launch more innovative products, and to strengthen their absorptive capacity for innovation (Klette and Griliches, 2000; Griffith et al., 2004). As a result, they often have their innovation budgets defined well in advance (typically across multiple periods) regardless of whether or not they receive public support. These firms may also have access to privileged information that allows them to anticipate with greater accuracy whether they are likely to benefit from public funding instruments. In addition, they tend to have large, well-prepared and well-informed teams, enabling them to be aware of such public programmes and prepare funding applications more efficiently, which is often a resource-intensive process.

5.3 Robustness Checks

Our robustness checks consist of three procedures. First, we take into account that the ESEE—the database used in this study—requires the participation of all firms with more than 200 employees. These are classified as large firms within this statistical source, while smaller firms are selected through stratified sampling and represent approximately 5% of all national firms. This creates a discrepancy between the ESEE and the guidelines of the European Union in the definition of medium-sized firms (50 to 200 employees and 50 to 250 employees, respectively).

To assess whether the size-group definitions affect the results, we re-estimate the previous models by defining medium-sized firms as those with 50 to 199 employees and large firms as those with 200 to 499 employees.

As shown in Table 6, under these new definitions, approximately one third of the firms previously classified as medium-sized are now reclassified as large. However, the estimated coefficients confirm that there is no statistically significant difference between using one grouping or the other. Therefore, the main results are not affected by this sampling feature of the database.

The second robustness check involves testing the sensitivity of the model results to the inclusion or exclusion of the following covariates in the estimation of entropy weights: number of employees, sales and personnel costs.[17] The results, presented in Tables OA3 to OA6 in the Online Annex, indicate that the models shown in Table 4 are the most efficient in terms of R-squared, pseudo R-squared and the Akaike Information Criterion. These models include firm size (and its square) and sales (and its square) as independent variables for the estimation of entropy weights.

Moreover, for each firm size group, the estimated coefficients are similar across models that include different sets of independent variables. Thus, from a statistical standpoint, there are no significant differences that result from the inclusion or exclusion of firm size, sales or personnel costs in the entropy weighting procedure.

[17] The sector fixed effects (Pavitt), the dichotomous variables that indicate patent applications and product or process innovations, the age of the company and the variables that reflect its financial situation are maintained in all specifications.

Our third robustness check examines whether our baseline results are affected by the timing of both the treatment and outcome variables. Given the structure of our database, in the previous models the impact of the R&D subsidy is measured in the same year in which the firm receives it. However, in many cases, the funded projects last more than one year. In addition, most public subsidy programmes in Spain provide funding for a maximum of three years and allow firms to choose between two main payment modalities: upfront payment (regardless of whether the subsidised project is annual or multiannual) or instalment payments aligned with the project's duration. In this latter case, it is common for the largest disbursement to be made at the beginning of the project, with payments in subsequent years being smaller in amount.

Therefore, our initial specification may be problematic, as it would associate smaller impacts with the initial year of the project and larger impacts with the final years. This issue becomes particularly relevant for high-value R&D subsidies, which are typically associated with large-scale innovation projects that may take several periods to implement, thereby challenging the validity of the one-year investment assumption.

To shed further light on this issue, we re-estimate our models under the assumption that firms receiving subsidies in consecutive years are engaged in multi-annual projects, with the subsidy payments distributed over the duration of those projects.[18] Specifically, we identify firms whose financing lasted up to a maximum of three years and impute the averages of both the subsidy amount received and the firm's net R&D expenditure, considering them as the treatment dose and the outcome, respectively.[19] Estimated dose-response functions for this sample are broadly consistent with those obtained in the previous estimations (see Figure 2), confirming our main findings. The overall trend remains very similar across the different firm-size groups. This suggests that accounting for consecutive awarding of public funding does not substantially alter the observed patterns, reinforcing the stability of our results in terms of the shape of the relationship and the existence of points of maximum impact for small, medium-sized and large firms.

6. Conclusions

Throughout this study, we have integrated two dimensions that have received limited attention in the literature evaluating innovation-related subsidy programs. On the one hand, we examine the effect of public R&D subsidies on firms' net R&D investment, accounting for the entire range of subsidy intensities among supported firms. Our objective is not to estimate the effect of the R&D subsidy in terms of the difference between the net private innovation investment of a firm supported with a specific amount of subsidy and the investment it would have had in the absence of the grant. Instead, we

[18] In most cases, public subsidy schemes provide funding for a maximum of two to three years.

[19] Approximately 25% of firms in the sample declare having obtained public support for two or three consecutive years. Firms reporting more than four consecutive years of subsidy receipt (less than 15% in the sample) are excluded from this analysis, as they are likely chaining together funding for different projects.

estimate the causal response to the treatment as the change in net investment that results from a variation in the subsidy amount.

On the other hand, we apply this continuous approach across four firm-size strata, allowing for a non-linear relationship between the subsidy amount and net R&D expenditure within each group, and identifying, where applicable, the thresholds at which the positive effect of subsidies on net private investment is maximized. We have explored this aspect because firms tend to exhibit size-related advantages and disadvantages that may either mitigate or exacerbate their R&D activities.

The empirical analysis relies on a dataset comprising 1,996 observations of Spanish manufacturing firms that received public subsidies to support their innovation activities during the period 2008-2018. Our methodology involves estimating parametric and non-parametric dose-response functions for subsamples of firms grouped into four size categories—small, medium-sized, large, and very large firms. To address the potential problem associated with estimating the causal response to treatment intensity, we use entropy weights as a pre-processing method for the sample of treated firms, thereby balancing covariate distributions across treatment levels.

The main conclusions of our analysis can be summarised as follows. First, we provide empirical evidence that, for most firms, there is a non-linear relationship between the level of subsidy allocated and its effect on net R&D investment. While this finding is consistent with the results of previous international studies, our data specifically indicate an inverted U-shaped relationship between the subsidy level and net R&D expenditure for small, medium-sized, and large firms, but not for very large ones, which exhibit a linear pattern.

Second, the proportion of public funding relative to R&D expenditure required to generate the maximum effect on net investment appears to decline substantially with firm size. This finding supports the notion that the design of R&D subsidy programmes should incorporate maximum subsidy rates that decrease with firm size—a practice already adopted by some public agencies, although typically restricted to a distinction between SMEs and large firms—.

Third, the estimated optimal intensities for small, medium-sized, and large firms suggest that, for a considerable share of firms, higher levels of net R&D investment could have been achieved with lower subsidy amounts. However, they also indicate that, particularly in the case of small and large firms, a non-trivial proportion of companies may have been underfunded relative to the level of public support required to reach the maximum net investment.

Finally, in the case of Spain, the effect of subsidies differs between small and medium-sized firms and between large and very large firms. Specifically, the positive effects of subsidies are more persistent for small firms than for medium-sized ones, and for large firms than for very large firms, which even display a negative relationship between the subsidy amount and net R&D investment. This finding suggests that the design of subsidy schemes to support business R&D should not rely solely on the SME–large firm distinction but should instead incorporate a more nuanced calibration.

Overall, these results confirm firm size as a key source of heterogeneity in the impact of innovation subsidies on net private R&D investment. They also contribute to a more refined understanding of how subsidy intensity should vary across firm sizes to enhance policy effectiveness.

One aspect that remains to be studied is whether our conclusions could be extended to firms in sectors other than manufacturing, or to other types of direct financial support for innovation activities, such as public loans. For this latter analysis, it would be necessary to have information on not only whether the company has been supported through those instruments but also the amount of money received through each of them. In addition, the limited number of extreme values in our sample at either end of the subsidy distribution makes it difficult to accurately estimate the causal response for firms at the lower and upper tails. Finally, our database provides information at the firm level instead of at the project level, which would allow us to account for the duration of R&D projects. This data constraints imply that our estimated optimal subsidies should be interpreted with caution, as they likely represent lower bounds of the true optimal amounts. These limitations suggest a starting point for future research.

References

Acebo, E., Miguel-Dávila, J. A., & Nieto, M. (2022). Do financial constraints moderate the relationship between innovation subsidies and firms' R&D investment? *European Journal of Innovation Management*, *25*(2), 347-364. https://doi.org/10.1108/EJIM-07-2020-0286.

Afcha, S., & García-Quevedo, J. (2016). The impact of R&D subsidies on R&D employment composition. *Industrial and Corporate Change*, *25*(6), 955-975. https://doi.org/10.1093/icc/dtw008.

Akcigit, U. (2009). Firm size, innovation dynamics and growth. Paper presented at the meeting of the Society for Economic Dynamics. Retrieved from https://EconPapers.repec.org/RePEc:red:sed009:1267.

Appelt, S., Galindo-Rueda, F., González Cabral, A., & Pereira, C. (2020). *R&D tax incentives: Evidence on design, incidence and impacts* (JRC Science for Policy Report). European Commission, Joint Research Centre. https://joint-research-centre.ec.europa.eu/publications/rd-tax-incentives-evidence-design-incidence-and-impacts_en

Arrow, K. J. (1962). Economic welfare and the allocation of resources for invention. In: Rowley, C. K. (eds.): *Readings in Industrial Economics*. Palgrave, London. Chapter 13, 219-236. https://doi.org/10.1007/978-1-349-15486-9_13.

Athreye, S. S., Fassio, C., & Roper, S. (2021). Small firms and patenting revisited. *Small Business Economics*, *57*(1), 513-530. https://doi.org/10.1007/s11187-020-00323-1.

Austin, P. C. (2019). Assessing covariate balance when using the score with quantitative or continuous exposures. *Statistical Methods in Medical Research*, *28*(5), 1365-1377. https://doi.org/10.1177/0962280218756159.

Becker, B. (2015). Public R&D policies and private R&D investment: a survey of the empirical evidence. *Journal of Economic Surveys*, *29*, 917-942. https://doi.org/10.1111/joes.12074.

Becker, B. (2019). The impact of innovation policy on firm innovation and performance: A review of recent research developments. ifo DICE Report, ISSN 2511-7823, ifo Institut – Leibniz-Institut für Wirtschaftsforschung an der Universität München, München, 17(4), 10-15. https://hdl.handle.net/10419/216282.

Bellucci, A., Pennacchio, L., & Zazzaro, A. (2019). Public R&D subsidies: Collaborative versus individual place-based programs for SMEs. *Small Business Economics*, *52*, 213-240. https://doi.org/10.1007/s11187-018-0017-5.

Bellucci, A., Fatica, S., Georgakaki, A., Gucciardi, G., Letout, S., & Pasimeni, F. (2023). Venture capital financing and green patenting. *Industry and Innovation*, *30*(7), 947-983. https://doi.org/10.1080/13662716.2023.2228717.

Bloom, N., Van Reenen, J., & Williams, H. (2019). A toolkit of policies to promote innovation. *Journal of Economic Perspectives, 33*(3), 163–84. https://doi.org/10.1257/jep.33.3.163.

BMWi – Federal Ministry for Economic Affairs and Climate Action. (n.d.). *Central Innovation Programme for SMEs (ZIM)*. https://www.zim.de.

Bpifrance. (n.d.). *Innovation and R&D support programmes*. https://www.bpifrance.fr.

Breschi, S., Malerba, F., & Orsenigo, L. (2000). Technological regimes and Schumpeterian patterns of innovation. *The economic journal*, *110*(463), 388-410. https://www.jstor.org/stable/2566240.

Brewer, M., Crossley, T. F., & Joyce, R. (2018). Inference with difference-in-differences revisited. *Journal of Econometric Methods, 7*(1), 20170005. https://doi.org/10.1515/jem-2017-0005.

Business Finland. (n.d.). *Research, development and innovation funding*. https://www.businessfinland.fi.

Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. (2024). Difference-in-differences with a continuous treatment. Paper No. w32117. National Bureau of Economic Research. https://doi.org/10.3386/w32117.

Cassiman, B., & Veugelers, R. (2002). R&D cooperation and spillovers: Some empirical evidence from Belgium. *American Economic Review*, *92*(4), 1169-1184. https://doi.org/10.1257/00028280260344704.

Centre for the Development of Industrial Technology (CDTI). (2008). *Annual report 2008*. Ministry of Science and Innovation, Government of Spain.

Cerulli, G. & Potì, B. (2016). Explaining firm sensitivity to R&D subsidies within a dose-response model: The role of financial constraints, real cost of investment, and strategic value of R&D. DEM Working Papers 2016/09.

Cerulli, G., Corsino, M., Gabriele, R. & Giunta, A. (2022). A dose-response evaluation of a regional R&D subsidies policy. *Economics of Innovation and New Technology, 31*(3), 173-190. https://doi.org/10.1080/10438599.2020.1792604.

Cohen, W. M., & Klepper, S. (1992). The tradeoff between firm size and diversity in the pursuit of technological progress. *Small Business Economics*, *4*, 1-14. https://doi.org/10.1007/BF00402211.

Colombo, M. G., Grilli, L. & Murtinu, S. (2011). R&D subsidies and the performance of high-tech start-ups. *Economics Letters*, 112 (1): 97–99. https://doi.org/10.1016/j.econlet.2011.03.007.

Criscuolo, C., Martin, R., Overman, H. G., & Van Reenen, J. (2019). Some causal effects of an industrial policy. *American Economic Review*, *109*(1), 48-85. https://doi.org/10.1257/aer.20160034.

Crump, R. K., Hotz, V. J., Imbens, G. W., & Mitnik, O. A. (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, *96*(1), 187-199. https://doi.org/10.1093/biomet/asn055.

Dai, X., & Cheng, L. (2015). The effect of public subsidies on corporate R&D investment: An application of the generalized propensity score. *Technological Forecasting and Social Change*, *90*, 410–419. https://doi.org/10.1016/j.techfore.2014.04.014.

Damanpour, F., & Aravind, D. (2012). Managerial innovation: Conceptions, processes and antecedents. *Management and organization review*, *8*(2), 423-454. https://doi.org/10.1111/j.1740-8784.2011.00233.x.

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.

Enterprise Ireland. (n.d.). *Research and development funding programmes*. https://www.enterprise-ireland.com.

Enterprise Ireland. (2025). *Overview of R&D Fund eligibility and claims*. https://www.enterprise-ireland.com/…

European Commission. (2012). *Guide to research and innovation strategies for smart specialisation (RIS3)*. Publications Office of the European Union. https://data.europa.eu/doi/10.2776/65746.

European Commission. (2014). *Commission Regulation (EU) No 651/2014 of 17 June 2014 declaring certain categories of aid compatible with the internal market (General Block Exemption Regulation—GBER)*. Official Journal of the European Union. http://data.europa.eu/eli/reg/2014/651/2023-07-01.

Fieller, E. C. (1954). Some problems in interval estimation. *Journal of the Royal Statistical Society, Series B, 16*(2), 175–185. https://doi.org/10.1111/j.2517-6161.1954.tb00159.x.

Fischer, T., & Ringler, P. (2014). What patents are used as collateral? An empirical analysis of patent reassignment data. *Journal of Business Venturing*, *29*(5), 633-650. https://doi.org/10.1016/j.jbusvent.2014.04.002.

Gambardella, A., Giuri, P., & Luzzi, A. (2007). The market for patents in Europe. *Research policy, 36*(8), 1163-1183. https://doi.org/10.1016/j.respol.2007.07.006.

Government of Spain. (2014). *Royal Decree 475/2014 of 13 June on reductions in social security contributions for research staff. Official State Gazette (BOE).* https://www.boe.es/buscar/doc.php?id=BOE-A-2014-6351

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.

Grandi, P., Belin, J., Darriet, E., & Guille, M. (2024). SMEs, R &D financing, and credit shocks. *Small Business Economics, 63*(1), 325-347. https://doi.org/10.1007/s11187-023-00814-x.

Griffith, R., Redding, S., & Reenen, J. V. (2004). Mapping the two faces of R&D: Productivity growth in a panel of OECD industries. *Review of economics and statistics, 86*(4), 883-895. https://doi.org/10.1162/0034653043125194.

Hall, B. H., Lotti, F., & Mairesse, J. (2009). Innovation and productivity in SMEs: Empirical evidence for Italy. *Small Business Economics, 33*, 13-33. https://doi.org/10.1007/s11187-009-9184-8.

Hall, B. H., & Lerner, J. (2010). The financing of R&D and innovation. In *Handbook of the Economics of Innovation* (Vol. 1, pp. 609-639). North-Holland. https://doi.org/10.1016/S0169-7218(10)01014-2.

Haans, R. F., Pieters, C., & He, Z. L. (2016). Thinking about U: Theorizing and testing U-and inverted U-shaped relationships in strategy research. *Strategic management journal*, *37*(7), 1177-1195. https://doi.org/10.1002/smj.2399.

Heijs, J., Buesa, M., Vergara, D. M., Gutiérrez, C., Díaz, G. A., & Picoita, A. J. G. (2020). *Innovación, crecimiento y competitividad: El papel de la política tecnológica en España*. Funcas. ISBN: 9788417609467.

Heijs, J., Guerrero, A. J. & Huergo, E. (2022). Understanding the heterogeneous additionality of R&D subsidy programs of different government levels. *Industry and Innovation, 29*(4), 533-563. https://doi.org/10.1080/13662716.2021.1990024.

Herrera, L., & Sánchez-González, G. (2013). Firm size and innovation policy. *International Small Business Journal*, *31*(2), 137-155. https://doi.org/10.1177/0266242611405553.

Hewitt-Dundas, N. (2006). Resource and capability constraints to innovation in small and large plants. *Small Business Economics*, *26*, 257-277. https://doi.org/10.1007/s11187-005-2140-3.

Hirano, K., & Imbens, G. W. (2004). The propensity score with continuous treatments. In Gelman, A. & Meng, X.L. (Eds.): *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives: An Essential Journey with Donald Rubin's Statistical Family,* 73-84. Wiley-Blackwell. https://doi.org/10.1002/0470090456.ch7.

Huergo, E., & Moreno, L. (2017). Subsidies or loans? Evaluating the impact of R&D support programmes. *Research Policy*, *46*(7), 1198-1214. https://doi.org/10.1016/j.respol.2017.05.006.

UK Research and Innovation. (n.d.). *Innovate UK: Business innovation funding*. https://www.ukri.org/councils/innovate-uk/

Kamien, M. I., & Schwartz, N. L. (1978). Self-Financing of an R and D Project. *The American Economic Review*, *68*(3), 252-261. https://www.jstor.org/stable/1805258.

Kennedy, E. H., Ma, Z., McHugh, M. D., & Small, D. S. (2017). Non-parametric methods for doubly robust estimation of continuous treatment effects. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 79*(4), 1229-1245. https://doi.org/10.1111/rssb.12212.

Keum, D. D., & See, K. E. (2017). The influence of hierarchy on idea generation and selection in the innovation process. *Organization Science*, *28*(4), 653-669. https://doi.org/10.1287/orsc.2017.1142.

Kim, K. (2018). Diminishing returns to R&D investment on innovation in manufacturing SMEs: do the technological intensity of industry matter? *International Journal of Innovation Management*, *22*(07), 1850056. https://doi.org/10.1142/S1363919618500561.

Kiman, K., & Jongmin, Y. (2022). Linear or nonlinear? Investigation an affect of public subsidies on SMEs R&D investment. *Journal of the Knowledge Economy*, *13*(3), 2519-2546. https://doi.org/10.1007/s13132-021-00823-9.

Klette, T. J., & Griliches, Z. (2000). Empirical patterns of firm growth and R&D investment: A quality ladder model interpretation. *The Economic Journal*, *110*(463), 363-387. https://www.jstor.org/stable/2566239.

Kullback, S. (1959). *Information theory and statistics*. New York, NY: John Wiley & Sons. ISBN 9780486696843.

Lechner, M. (2001). Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In *Econometric evaluation of labour market policies* (pp. 43-58). Heidelberg: Physica-Verlag HD. https://doi.org/10.1007/978-3-642-57615-7_3.

Lind, J. T., & Mehlum, H. (2010). With or without U? The appropriate test for a U-shaped relationship. *Oxford Bulletin of Economics and Statistics*, *72*(1), 109-118. https://doi.org/10.1111/j.1468-0084.2009.00569.x.

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.

Mazzucato, M., & Semieniuk, G. (2017). Public financing of innovation: new questions. *Oxford Review of Economic Policy*, (1), 24-48. https://doi.org/10.1093/oxrep/grw036.

Ministry of Economy and Competitiveness. (2013). *Spanish strategy for science, technology and innovation 2013–2020*. Government of Spain.

National Research Council. (2008). *An Assessment of the SBIR Program*. Washington, DC: The National Academies Press. https://doi.org/10.17226/11989.

National Research Council of Canada. (n.d.). *Support for technology innovation (Industrial Research Assistance Program — IRAP)*. Government of Canada. https://nrc.canada.ca/en/support-technology-innovation.

Meuleman, M., & De Maeseneire, W. (2012). Do R&D subsidies affect SMEs' access to external financing? *Research Policy, 41*(3), 580-591. https://doi.org/10.1016/j.respol.2012.01.001.

Nelson, R. R. (1959). The simple economics of basic scientific research. *The Journal of Political Economy, 67*(3), 297–306. http://www.jstor.org/stable/1827448.

Nilsen, Ø. A., Raknerud, A. & Iancu, D. C. (2020). Public R&D support and firm performance: A multivariate dose-response analysis. *Research Policy, 49*(7), 104067. https://doi.org/10.1016/j.respol.2020.104067.

OECD. (2023). *OECD R&D tax incentives database*. Organisation for Economic Co-operation and Development. https://www.oecd.org/sti/rd-tax-stats/

Ortega-Argilés, R., Vivarelli, M., & Voigt, P. (2009). R&D in SMEs: A paradox? *Small Business Economics, 33*, 3-11. https://doi.org/10.1007/s11187-009-9187-5.

Patterson, F., & Kerrin, M. (2014). Characteristics and behaviours associated with innovative people in small-and medium-sized enterprises. In *Handbook of research on small business and entrepreneurship* (pp. 187-206). Edward Elgar Publishing. https://doi.org/10.4337/9781849809245.00020.

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.

Pavitt, K., Robson, M., & Townsend, J. (1987). The size distribution of innovating firms in the UK: 1945-1983. *The Journal of Industrial Economics*, 297-316. https://doi.org/10.2307/2098636.

Rothwell, R. (1985). Venture finance, small firms and public policy in the UK. *Research Policy*, *14*(5), 253-265. https://doi.org/10.1016/0048-7333(85)90008-3.

Rothwell, R. (1989). Small firms, innovation and industrial change. *Small Business Economics*, 51-64. https://doi.org/10.1007/BF00389916.

Sasabuchi, S. (1980). A test of a multivariate normal mean with composite hypotheses determined by linear inequalities. *Biometrika, 67*(2), 429–439. https://doi.org/10.1093/biomet/67.2.429.

Schumpeter, J. A. (1934). *The theory of economic development: An inquiry into profits, capital, credit, interest and the business cycle*. Cambridge, MA: Harvard University Press. ISBN: 9780674879904.

Schumpeter, J. A. (1942). *Capitalism, socialism and democracy*. New York, NY: Harper & Brothers. ISBN: 0061330086.

Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, *15*(6), 285-305. https://doi.org/10.1016/0048-7333(86)90027-2.

Tingvall, P. G. & Videnord, J. (2020). Regional differences in effects of publicly sponsored R&D grants on SME performance. S*mall Business Economics, 54*(4), 951-969. https://doi.org/10.1007/s11187-018-0085-6.

Tübbicke, S. (2022). Entropy balancing for continuous treatments. *Journal of Econometric Methods*, *11*(1), 71-89. https://doi.org/10.1515/jem-2021-0002.

Tübbicke, S. (2023). ebct: Using entropy balancing for continuous treatments to estimate dose–response functions and their derivatives. *The Stata Journal*, *23*(3), 709-729. https://doi.org/10.1177/1536867X231196291.

Vegetabile, B. G., Griffin, B. A., Coffman, D. L., Cefalu, M., Robbins, M. W., & McCaffrey, D. F. (2021). Nonparametric estimation of population average dose-response curves using entropy balancing weights for continuous exposures. *Health Services and Outcomes Research Methodology*, *21,* 69-110. https://doi.org/10.1007/s10742-020-00236-2.

Vossen, R. W. (1998). Relative strengths and weaknesses of small firms in innovation. *International Small Business Journal*, *16*(3), 88-94. https://doi.org/10.1177/0266242698163005.

Watson, S. K., & Elliot, M. (2016). Entropy balancing: A maximum-entropy reweighting scheme to adjust for coverage error. *Quality & Quantity*, *50*, 1781-1797. https://doi.org/10.1007/s11135-015-0235-8.

Zhu, Y., Coffman, D. L. & Ghosh, D. (2015). A boosting algorithm for estimating generalized propensity scores with continuous treatments. *Journal of Causal Inference, 3*(1), 25–40. https://doi.org/10.1515/jci-2014-0022.

Zúñiga-Vicente, J. Á., 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.

Zhu, Y., Coffman, D. L. & Ghosh, D. (2015). A boosting algorithm for estimating generalized propensity scores with continuous treatments. *Journal of Causal Inference, 3*(1), 25–40. https://doi.org/10.1515/jci-2014-0022.

Table 1. Consecutive awarding of R&D subsidies. 2008-2018

	Small	**Medium-sized**	**Large**	**Very large**	**Total**
Firms awarded:					
- only 1 year	80	138	59	28	305
- only 2 consecutive years	36	72	30	20	158
- only 3 consecutive years	16	43	18	14	91
- 4 or more consecutive years	17	79	36	52	184
Total firms	149	332	143	114	738
(Observations)	(285)	(882)	(368)	(461)	(1996)

Notes: This table reports the number of firms by number of consecutive years in which they have received public R&D subsidies during the period 2008-2018.

Table 2. Description and basic statistics of the model variables

Variable	Description	Min.	Max.	Average: All firms	Average: Small firms	Average: Medium firms	Average: Large firms	Average: Very large firms
Subsidy (log.)	Amount of R&D subsidy (€)	7.69	17.25	12.19	10.98	11.74	12.43	13.07
Net R&D expenditure (log.)	Total R&D expenditure – R&D subsidy (€)	0	19.86	10.09	7.87	9.12	11.06	12.53
Pavitt 1: 'Traditional producer of consumer goods' (0/1)	= 1 if the firm's activity corresponds to CNAE-09: 151–223, 261–268 or 361–372	0	1	0.30	0.43	0.32	0.20	0.25
Pavitt 2: 'Producer of intermediate goods' (0/1)	= 1 if the firm's activity corresponds to CNAE-09: 271–287	0	1	0.14	0.14	0.15	0.16	0.11
Pavitt 3: 'Specialised supplier' (0/1)	= 1 if the firm's activity corresponds to CNAE-09: 251–252, 291–297 or 300–335	0	1	0.18	0.22	0.20	0.12	0.16
Pavitt 4: 'Scale-intensive/assembly-based producer' (0/1)	= 1 if the firm's activity corresponds to CNAE-09: 311–343	0	1	0.25	0.14	0.21	0.33	0.35
Pavitt 5: 'Science- and R&D-intensive producer' (0/1)	= 1 if the firm's activity corresponds to CNAE-09: 241–247 or 351–355	0	1	0.13	0.07	0.12	0.18	0.13
Firm size (log.)	Number of employees	2.20	9.47	5.33	3.38	4.83	5.84	7.10
Firm age (log.)	Number of years since the company was founded	0.69	4.72	3.37	3.20	3.40	3.39	3.40
Productivity (log.)	Value added (€) / Number of employees	6.56	12.92	10.98	10.80	10.91	11.05	11.19
Patents (0/1)	= 1 if the firm registers patents in Spain or abroad	0	1	0.17	0.07	0.13	0.18	0.32
Innovations (0/1)	= 1 if the firm achieves product or process innovations	0	1	0.76	0.64	0.74	0.79	0.83
Long-term leverage (ratio)	Long-term external funds / equity	0	195.36	1.01	0.78	1.08	0.79	1.20
Tangible fixed assets (ratio)	Tangible fixed assets / total liabilities	0.001	5.26	0.66	0.58	0.66	0.71	0.66
Sales (log.)	Total sales (€)	12.86	22.71	17.43	15.13	16.85	17.97	19.54
Personnel costs (log.)	Gross wages plus severance payments and employer contributions	12.34	20.26	15.80	13.62	15.24	16.38	17.74
Observations				1,996	285	882	368	461

Notes: (0/1) indicates that the variable is binary. (log.) means that the variable is transformed by applying the natural logarithm. All monetary variables have been deflated to reflect constant 2010 prices.

Table 3. Covariate balance before and after entropy weighting

	Unbalanced sample				Balanced sample			
Variable	Pearson correlation	Conditional mean	t-Student	p-value	Pearson correlation	Conditional mean	t-Student	p-value
Pavitt 1	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Pavitt 2	-0.072	0.028	0.324	0.749	0.000	0.000	0.000	1.000
Pavitt 3	0.010	0.248	3.000	0.003	0.000	0.000	0.000	1.000
Pavitt 4	0.203	0.176	2.174	0.030	0.000	0.000	0.000	1.000
Pavitt 5	0.072	0.005	0.054	0.961	0.000	0.000	0.000	1.000
Firm size (log.)	0.456	-4.898	-1.487	0.163	0.000	0.000	0.000	1.000
Firm age	0.043	-0.027	-0.094	0.931	0.000	0.000	0.000	1.000
Productivity (log.)	0.173	0.031	0.475	0.639	0.000	0.000	0.000	1.000
Patents	0.145	0.002	0.030	0.978	0.000	0.000	0.000	1.000
Innovations	0.100	0.020	0.349	0.736	0.000	0.000	0.000	1.000
Long-term leverage	0.061	0.077	2.178	0.030	0.000	0.000	0.000	1.000
Tangible fixed assets	0.004	0.036	0.071	0.486	0.000	0.000	0.000	1.000
Sales (log.)	0.446	0.463	8.166	0.000	0.000	0.000	0.000	1.000
Personnel costs (log.)	0.304	0.268	3.274	0.000	0.000	0.000	0.000	1.000
Observations	1,996				1,996			

Notes: The coefficients of the conditional mean and its statistics (t-Student and p-value) are derived from the OLS regression between the treatment variable and the covariates of the model.

Table 4. Effect of subsidies on net R&D expenditure. Parametric methods on the reweighted sample

	All firms		**Small firms**		**Medium firms**		**Large firms**		**Very large firms**	
Estimation method	OLS	Tobit	OLS	Tobit	OLS	Tobit	OLS	Tobit	OLS	Tobit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Subsidy	-0.954	2.069	10.370***	17.445***	6.821***	11.525***	9.726***	12.862***	-2.161	-1.419
	(1.459)	(1.791)	(3.076)	(5.643)	(1.758)	(2.907)	(2.812)	(3.708)	(2.241)	(2.668)
Subsidy squared	-0.027	-0.169**	-0.584***	-0.944***	-0.378***	-0.607***	-0.474***	-0.617***	0.034	-0.002
	(0.060)	(0.075)	(0.145)	(0.263)	(0.077)	(0.126)	(0.119)	(0.152)	(0.090)	(0.102)
U-shape test			1.62		1.23		1.81			
[p-value]			[0.05]		[0.10]		[0.03]			
Optimal subsidy			8.87		9.00		10.24			
[value in €]			[7,115]		[8,103]		[28,001]			
Fieller interval			[7.65; 9.41]		[7.81; 9.61]		[9.11; 10.74]			
Observations	1996	1996	285	285	882	882	368	368	461	461
Censored observations		429		77		221		62		69
LR $X^2(2)$	92.73***	229.11***	49.55***	65.91***	115.94***	168.01***	37.26***	80.98***	25.74***	45.49***
R2	0.113		0.217		0.180		0.200		0.095	
Pseudo R2		0.018		0.040		0.033		0.036		0.015

Notes: We report estimated coefficients. ***: p-value<0.01, **: p-value<0.05. Standard errors are in parentheses. The U-shape test is the Lind and Mehlum's (2010) intersection –union test used to assess U-shaped or inverted U-shaped relationships. Optimal subsidy refers to the treatment dose at which the maximum effect on net R&D expenditure is estimated to occur. The Fieller interval corresponds to 95% confidence interval for the estimated optimal subsidy following Fieller's (1954) method. U-shape tests and optimal subsidies are not presented in columns (1) and (9) because estimated coefficients reject the existence of a quadratic relationship.

Table 5. Summary of results for discussion

	Small firms	Medium firms	Large firms	Very large firms
Optimal subsidy (log.) (% over average subsidy)	8.87 (80.8%)	9.00 (76.7%)	10.24 (82.4%)	-
Position of optimal subsidy within the observed range	19.6%	12.0%	20.0%	
% observations below optimal subsidy (below upper limit of Fieller interval)	8.5% (14.3%)	2.2% (5.7%)	8.5% (12.5%)	-
Subsidy share at sample average	63.50%	54.32%	43.75%	34.89%
Subsidy share at optimal subsidy (i)	36.26%	31.83%	16.96%	-
Subsidy share at optimal subsidy (ii)	40.96%	31.07%	24.78%	-

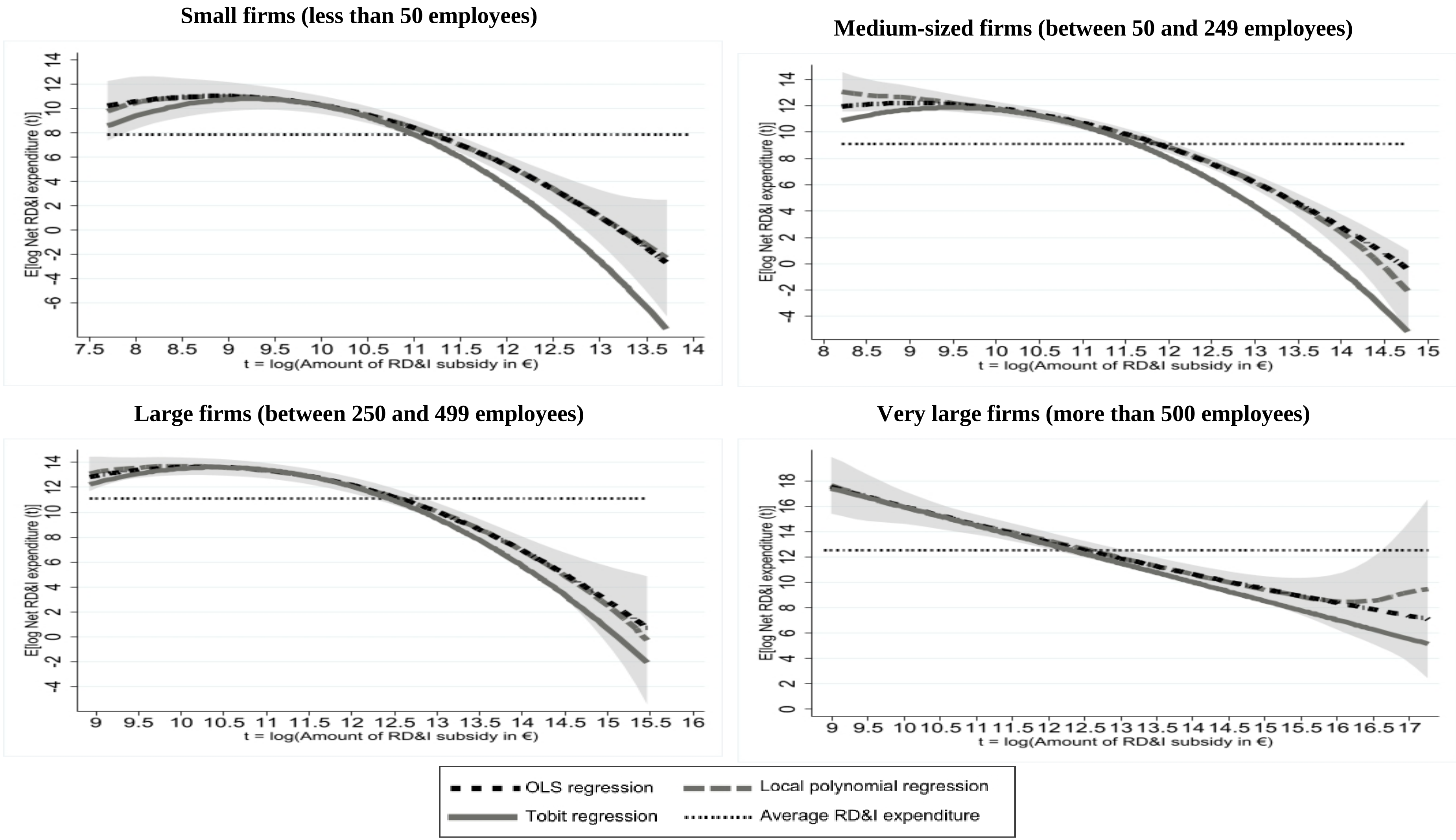
Notes: Optimal subsidy refers to the treatment dose at which the maximum effect on net R&D expenditure is estimated to occur in OLS models. The position of the optimal subsidy within the observed range is computed as follows: (optimal subsidy – minimum subsidy) / (maximum subsidy – minimum subsidy). The Fieller interval corresponds to 95% confidence interval for the estimated threshold following Fieller's (1954) method. The subsidy share at the optimal subsidy is computed as: (i) the average of the shares of firms that received subsidy amounts within a ±0.1 log range around the estimated optimum, and (ii) as the average of the shares that represent the optimal subsidy over the R&D expenditure that would result from the sum of this subsidy and the observed net R&D investment.

Table 6. Effect of subsidies on net R&D expenditure. Estimates with different size group definitions

	EU medium-sized firms		**ESEE medium-sized firms**		**Difference in coefficients(*)**	
Estimation method	OLS	Tobit	OLS	Tobit		
	(1)	(2)	(3)	(4)	(1) vs (3)	(2) vs (4)
Subsidy	6.821***	11.525***	7.365***	13.174***	-0.544	-1.649
	(1.758)	(2.907)	(1.947)	(3.385)	[-0.207]	[-0.369]
Subsidy squared	-0.378***	-0.607***	-0.404***	-0.686***	0.026	0.079
	(0.077)	(0.126)	(0.085)	(0.148)	[0.226]	[0.406]
Observations	882	882	711	711		
Censored observations		221		186		
LR X^2	115.94***	168.01***	94.54***	137.54***		
R-Squared	0.180		0.180			
Pseudo R-squared		0.033		0.033		
	EU large firms		**ESEE large firms**		**Difference in coefficients(*)**	
Estimation method	OLS	Tobit	OLS	Tobit		
	(5)	(6)	(7)	(8)	(5) vs (7)	(6) vs (8)
Subsidy	9.726***	12.862***	8.745***	12.003***	0.981	0.859
	(2.812)	(3.708)	(2.372)	(3.271)	[0.266]	[0.163]
Subsidy squared	-0.474***	-0.617***	-0.439***	-0.589***	-0.035	-0.028
	(0.119)	(0.152)	(0.102)	(0.136)	[-0.223]	[-0.155]
Observations	368	368	539	539		
Censored observations		62		97		
LR X^2	37.26***	80.98***	51.40***	103.99		
R-squared	0.200		0.177			
Pseudo R-squared		0.036		0.031		

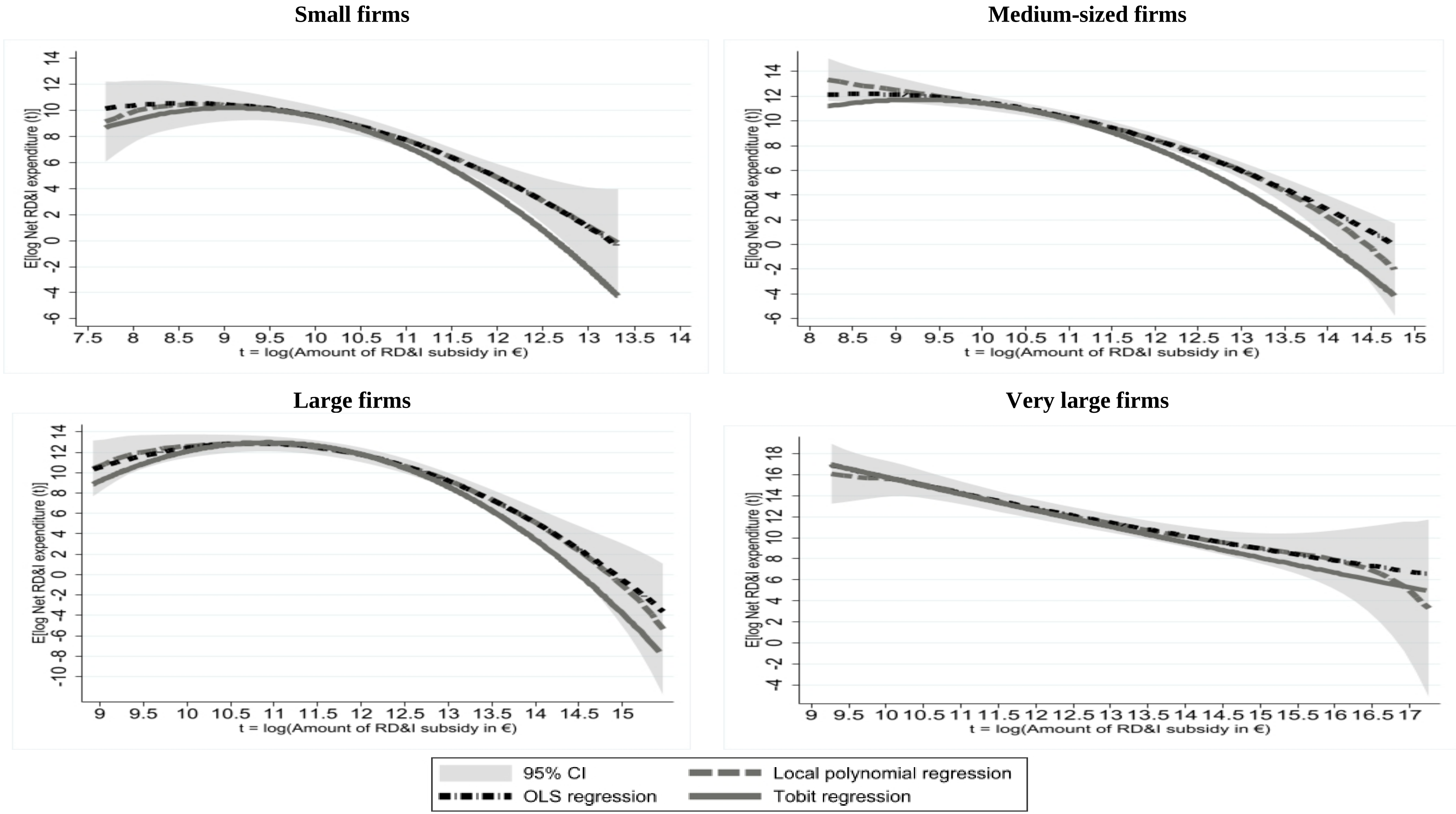
Notes: We report estimated coefficients. ***: p-value < 0.01, **: p-value < 0.05, *: p-value < 0.1. Standard deviations in parentheses. EU medium-sized firms: from 50 to 249 employees. ESEE medium-sized firms: from 50 to 199 employees. EU large firms: from 250 to 499 employees. ESEE large firms: from 200 to 499 employees. (*) Difference between the estimated coefficients for the two groups considered. The values in square brackets represent the associated t-statistics.

Figure 1. Dose-response function. Impact of subsidies on net R&D expenditure by size stratum



Notes: The confidence intervals, estimated at the 95% level from the non-parametric (local polynomial) regression, are shown as shaded areas in the graphs. The mean of net R&D expenditures in each graph corresponds to the mean of that variable in companies of that size range.

Figure 2. Dose-response function accounting for consecutive participation



Notes: The confidence intervals, estimated at the 95% level from the non-parametric (local polynomial) regression, are shown as shaded areas in the graph