Economics of Innovation and New Technology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gein20

# Firm-level contributions to the R&D intensity distribution: evidence and policy implications

**Sebastiano Cattaruzzo, Agustí Segarra-Blasco & Mercedes Teruel**

**To cite this article:** Sebastiano Cattaruzzo, Agustí Segarra-Blasco & Mercedes Teruel (2024) Firm-level contributions to the R&D intensity distribution: evidence and policy implications, Economics of Innovation and New Technology, 33:1, 45-65, DOI: 10.1080/10438599.2022.2140658

**To link to this article:** https://doi.org/10.1080/10438599.2022.2140658

Published online: 03 Nov 2022.

ECONOMICS OF INNOVATION AND NEW TECHNOLOGY
2024, VOL. 33, NO. 1, 45–65
https://doi.org/10.1080/10438599.2022.2140658

RESEARCH ARTICLE

# Firm-level contributions to the R&D intensity distribution: evidence and policy implications

Sebastiano Cattaruzzo, Agustí Segarra-Blasco and Mercedes Teruel

Department of Economics, Universitat Rovira i Virgili, GRIT & ECO-SOS, Reus, Spain

**ABSTRACT**

This paper decomposes the Spanish aggregate R&D distribution to disentangle the contributions of R&D public financing, gazelle firms, and financial constraints. Applying the Chernozhukov, Fernández-Val and Melly (2013) distribution regression approach, we estimate the contributions of these components at each point of the distribution. The analysis is carried out for two periods, pre-crisis 2004–2008 and post-crisis 2009–2014. We thereby introduce a comparative perspective that allows us to consider possible business cycle effects. Our findings show that the main explanatory factors of the significant post-crisis drop in Spanish aggregate R&D are changes in the public financing scheme and the decreased contribution of gazelles. Our results provide a rigorous analysis of Spanish R&D, hint at a possible transmission channel for reduced business dynamism, and offer interesting insights for policymaking.

**ARTICLE HISTORY**
Received 19 November 2021
Accepted 19 October 2022

**KEYWORDS**
R&D; decomposition; distribution; intensity; gazelles; financial; constraints

**JEL CODES**
O30; O33; L20

## 1. Introduction

Many studies have analysed the determinants of firms' internal R&D expenditure – see, for example, Montresor and Vezzani (2015) or Coad (2019). Generally, they contributed to establish stylized facts such as the heterogeneity of firm-level R&D investment (Coad et al. 2021), the incidence of non-observable characteristics on the R&D effort (Cohen and Klepper 1992), and the non-homogenous nature of R&D activity (Czarnitzki and Hottenrott 2011; Barge-Gil and López 2014). The topic is important since R&D investment expands a country's technological frontier and firm-level productivities. Thus, understanding how the distribution of R&D investments is influenced by different firms' characteristics and by the economic cycle is fundamental. Despite the many analyses on the topic, there is a gap in the literature on how common firm-level factors such as receiving public subsidies, being a gazella firm, or being financially constrained, affect a country's aggregate R&D distribution in each point of the distribution. Some recent multi-level studies have stressed the importance of considering aggregate impacts using micro-data (Di Giovanni et al. 2014).[1] Here, we propose a decomposition method based on distribution regression to explore Spanish firms undertaking R&D (henceforth, R&D firms).

Exploiting PITEC[2] data and applying the decomposition method developed by Chernozhukov et al. (2013), we investigate the determinants of change in the distribution of R&D intensities. The focus on intensities, rather than pure quantities, is both to correct for obvious scale effects, but also to follow the model developed by Cohen and Klepper (1992). Specifically, we quantify to what extent the variations are attributable to changes of observable characteristics in relation to the total observed change. Instead of standard estimation methods centred around the mean, a

**CONTACT** Sebastiano Cattaruzzo seba.cattaruzzo@unive.it Department of Economics, Universitat Rovira i Virgili, GRIT & ECO-SOS, Av. Universitat, 1, 43204, Reus, Spain
Supplemental data for this article can be accessed online at https://doi.org/10.1080/10438599.2022.2140658

precise reckoning for the whole distribution is preferable. Our approach entails estimation and inference procedures to compute appropriately chosen counterfactuals that allow the decomposition of differences (change over time in our case) in a given distribution.

This technique allows us to quantify both the levels and the changes in the contribution of given factors to the composition and evolution of Spanish internal R&D. In particular, we isolate three firm-level determinants: the level of R&D public financing, the share of 'gazelles',[3] and the extent of financial constraints (FCs). Also, given the 2003–2014 data coverage, we split the sample in two phases, one where aggregate R&D intensity is increasing and *vice versa* for the other. Thus, the analysis sheds light on the role that the above factors play in medium-term R&D aggregate fluctuations in relation to business cycle movements. Our results suggest that public financing and the contribution of gazelles are the two main determinants of the distribution's shape. The contribution from gazelles decreases considerably over time. Curiously, FCs impact negatively in the period of expansion, while they are unimportant in the period of contraction.

For many reasons, Spain is a perfect candidate for this analysis (COTEC 2018). Until 2008, its R&D expenditure was converging to EU levels, but it then started to diverge again (see Figure A-1 in Supplementary material for a graphical representation of the trend). Between 2008 and 2016, the Spanish economy experienced a loss of 30% in its public R&D budget and a reduction of 43% in the number of enterprises performing R&D activities. Also, private R&D expenditure experienced the effects of the crisis, making Spain the only one of the four major EU economies (Italy, Spain, France and Germany) whose R&D investment decreased continuously between 2009 and 2014 (Xifré 2018). This had a dramatic impact on the path to convergence.

Public financing, which has been sustaining private expenditures, was a major contributor. Up to 2007, more than 90% of the allocated budget was invested, but since then the trend has been decreasing, reaching the historical minimum of only 46.6% in 2017 (COTEC 2018). Furthermore, the EU recommends that private sector expenditure should be approximately two-thirds of the total, while it is currently only 47% in Spain.

This paper makes several contributions. Methodologically, we employ a compelling approach for distribution decomposition. Although widely applied in labour economics, as far as we know this is its first application in innovation studies and it constitutes an improvement on the previous quantile regression approaches. At the theoretical level, we provide an analysis framework whose aim is to isolate the most consistent determinants of R&D investment. Empirically, we show how several factors explain the levels and evolution of aggregate R&D intensity in different phases of the business cycle. Finally, from a policy perspective, the results are relevant in shaping innovation policy toward the long-pursued 3% R&D target proposed in the 2000 Lisbon Agenda, then incorporated in the policies of the Horizon 2020 program (Hervás Soriano and Mulatero 2010; Veugelers and Cincera 2015). Although the subsequent inclusion also implied a revision of the related implementation guidelines, European countries are still far away from the objective and the reasons remain mostly structural (Demircioglu et al. 2019; Moncada-Paternò-Castello and Grassano 2022). Indeed, as noted by some authors, setting R&D intensity targets is largely debatable and its effectiveness and achievability are still left to economic investigation (Carvalho 2018). Nevertheless, as this practice still exists and is present across several economies, we propose the insights presented in this analysis as key tools to properly set an intensity-based policy. Concluding, these results suggest solutions for a reduction of the long-standing EU-US gap in R&D intensities. Our analysis captures the role of public subsidies, high-growth firms (henceforth HGFs), and young innovative companies (henceforth YICs) in diminishing such gaps.

The rest of the work proceeds as follows. Section 2 summarizes the most relevant literature on R&D heterogeneity and on its determinants with the aim of contextualizing our research questions. Section 3 contains a description of the data and relevant statistics, while Section 4 explains the econometric methodology. Section 5 shows our main empirical results and Section 6 concludes and discusses the policy implications.

## 2. Conceptual framework

Internal R&D intensity depends on characteristics that have been largely analysed in the literature (Hall and Hayashi 1989; Hall 1993; Aghion et al. 2005; Breschi and Malerba 1997; Griffiths and Webster 2010). Although firm-level structural variables (i.e. industry concentration, market share, ownership structure, or lagged performance) explain well the levels of R&D intensity (Crépon et al. 1998; Jefferson et al. 2006), this study opts for a selection of policy-relevant variables.

Since Cohen and Levin (1989), it has been clear that robust conclusions on firm-specific determinants are elusive. Notably, the most regular findings are the procyclical nature of R&D (Barlevy 2007) and the fact that the R&D distribution directly affects countries' performance (Falk 2007). This section reports the evidence on R&D heterogenous firm-level behaviours, presents the variables that may most affect the distribution, and connects these concepts to recent literature, particularly the one related to reduced business dynamism and productivity slowdown.

### *2.1. Heterogeneity of the R&D intensity distribution*

The literature on R&D intensity is extremely rich in both empirical and theoretical contributions. Although Cohen and Klepper (1992) proposed a reliable probabilistic approach to model its distribution, most studies still rely on the use of production functions augmented to accommodate knowledge capital as input. Montresor and Vezzani (2015) demonstrate clearly how this approach is fragile and how its estimations vary significantly for each point of the distribution. Indeed, one of the main conclusions reached by Cohen and Klepper (1992) was that the heterogeneity in firm-level R&D intensities could not be explained by observable firm characteristics, but rather by unobserved ones, which the authors interpreted as R&D-related expertise.

Coad (2019) shows evidently how heterogeneity in firm-level R&D intensities is an empirical fact, and that this property is resistant to sectoral disaggregation. The results imply that even in high-tech sectors there are firms whose intensity can be increased (and vice versa for low-tech sectors). Also, Evangelista (2006) presented evidence on the widespread heterogeneity in both services and manufacturing sectors, with the former surprisingly dominating the latter. This has important implications for R&D policies, as the objective should be to correctly target the policy effort and to increase overall R&D intensities, rather than trying to pursue deeper and more complicated transformations which favour only the high-tech sectors, leaving low- and medium-tech ones behind.

Factors like the supply of different product lines, the presence of varying innovation opportunities, the diversity in knowledge bases due to cumulativeness, and varying exogenous rates of technological progress, have a direct impact on firm-level R&D intensities, even within the same sector. Only by assuming that all of these have similar patterns across rival firms, would it be possible to imagine convergence in R&D intensities, but this is not the case in the real world and heterogeneity is a widespread phenomenon.

### *2.2. Firm-level determinants: public subsidies, gazelles, and financial constraints*

Although empirical studies largely recognize the role of unobservables in the explanation of firm-level R&D investment, some factors can be identified. For instance, it is a stylized fact that investment in R&D is conducted at a sub-optimal level, thus requiring public intervention to provide support to firms. As reported by the OECD (2016), subsidies are the main tool of public R&D policy for SMEs. Good design and implementation of the public financing scheme are essential for making it an effective tool (Appelt et al. 2016; Soete et al. 2022). Furthermore, there have been extensive debates among scholars to establish whether public support may lead to crowding-out effects (Zúñiga-Vicente et al. (2014) and Aristei et al. (2017) for a survey on the evidence).[4] Despite the large debate on the effects of R&D subsidies on R&D investment, generally R&D subsidies exert a positive effect on R&D investments (Huergo and Moreno 2017). Interestingly, there is a lack of consensus on which firms benefit more from this kind of support. On the one hand, some papers find that for firms with low R&D intensity, R&D subsidies have a high net benefit, whereas the net benefit of R&D subsidies is small in companies with high R&D intensity (Busom et al. 2014). On the other hand, some papers (Akcigit and Kerr 2018; Acemoglou et al. 2018) show the importance of the heterogeneous R&D activity and the positive incidence on the allocation of subsidies. Acemoglu et al. (2018) studied innovation policies in models with incumbent firm-type heterogeneity with respect to innovation ability. Both papers identify the existence of high-ability firms that deserve receiving higher stimulus to expand. Finally, also because of the cumulative nature of knowledge, it is likely that the effects are greater for firms with a high R&D intensity, which have already internal routines in place to exploit them. Firms positioned in the right-tail of the R&D distribution are implicitly highly engaged in innovative search, which is by nature risky and uncertain. Thus, these firms are more needed to secure funding that goes beyond their internal and external capital, and that allows them to keep their cutting-edge positions. Despite the evidence derived from the literature does not point at a clear consensus on which R&D firms benefit more from the subsidies, a considerable majority of the authors suggest that the effect might be asymmetric and concentrated in more R&D-intensive firms. This brings us to the first hypothesis.

**Hypothesis 1:** R&D subsidies positively affect the R&D intensity distribution, but their effect is concentrated on the right-tail.

Another relevant factor is the presence of gazelles, firms characterized by their dynamism influencing countries' innovation (and employment) patterns (Brown et al. 2017). When considering gazelles, various definitions and groupings can be found. Here, we consider possibly the broadest possible definition that includes HGFs (both in terms of employment and sales) and young innovative companies. Despite their very broad definition (Delmar et al. 2003), the main characteristics of HGFs are that (i) they are younger, (ii) they have a quasi-homogenous presence across sectors, (iii) they have a tendency to be more innovative and (iv) they are involved in international markets (Moreno and Coad 2015; Teruel et al. 2021). YICs are younger than 6 years, with fewer than 250 employees, and operate at least at 15% of R&D intensity. Almost by definition,[5] these are the companies that should foster aggregate productivity growth thanks to their large innovation focus and disruptive approaches (Schneider and Veugelers 2010; Czarnitzki and Delanote 2013). Due to their organizational profiles, HGFs and YICs are ideally located to be major influencers in the distribution of R&D intensities in a given country.

**Hypothesis 2:** Given their small dimension and innovation-focused strategies, gazelle firms' influence on the R&D intensity distribution is positive and concentrated on the right-tail.

Thirdly, another factor consists of financial constraints (henceforth, FCs). As argued by Dosi (1990), finance is directly connected to the possibilities and ways of conducting innovative activities. FCs influence all firms to some extent, but especially innovative ones due to both uncertainty and information asymmetries between borrowers and lenders. This and the presence of high sunk costs, high risks and, appropriability issues, and possible negative externalities may hurt and induce lower investments among highly innovative firms (Arrow 1962; Mina et al. 2013).

More recently, Howell (2016) finds that FCs hinder R&D-intensive firms from commercializing their research activities. According to the author, the main reason is that these firms face (unanticipated) costs that prevent them from appropriating necessary complementary assets. Furthermore, the presence of higher financial barriers for firms in high-tech industries is much more likely as their innovation projects are hardly evaluated by outside observers, since experience or observed past realizations can offer little guidance in assessing the prospects of truly new innovative projects. In this line, García-Quevedo et al. (2018) find that innovation projects during the concept stage are more hindered by FCs. The main reason is the high uncertainty of these projects. This factor tends to become increasingly influential during economic downturns and for SMEs, especially R&D-intensive ones (North et al. 2013; Brown and Petersen 2015; Lee et al. 2015).

Despite the potential asymmetrical incidence of financial constraints according to each firm's R&D intensity, there are no empirical analyses exploring the unequal relationship between FCs across the R&D distribution. Nevertheless, looking at the actual use of financial resources and their connection with growth indicators, Cattaruzzo and Teruel (2022) found that firms in the right and in the left tails of the distribution have dramatically different impacts of leverage on their growth performance. Following this intuition and given the previous evidence connecting (lagged) R&D intensity to firm growth (Falk 2012), we expect that highly R&D intensive firms may be more harmed than those with lower intensity, which have more easiness in obtaining financial resources to solve more ordinary and tangible issues.

**Hypothesis 3:** Financial constraints are more impactful on the right-tail of R&D intensity distribution, and these impacts become more severe in the contractionary phase.

## 3. Data

### *3.1. Database and statistics*

Our database is PITEC, *Panel de Innovación Tecnológica*, a yearly project conducted by the Spanish Statistical Office and the Spanish Foundation for Science and Technology. Based on the Community Innovation Survey framework, thanks to its representativeness and structure it is one of the most analysed data sources in innovation studies (De Marchi 2012; Barge-Gil and López 2014; Segarra and Teruel 2014; Audretsch et al. 2014; Costa-Campi et al. 2014; Cainelli et al. 2015; Del Río et al. 2016; Kunapatarawong and Martinez-Ros 2016; Marzucchi and Montresor 2017; Coad et al. 2021).[6] The database contains both SMEs and large firms. Its sampling frame includes all innovative firms with more than 200 employees, while for firms with fewer than 200 employees, it is based on random selection among all firms in this size category.

We focus on all firms performing R&D activity, regardless of their sector. Following Cohen and Klepper (1992), the R&D intensity distribution shows extremely regular patterns across sectors, thus suggesting the existence of a common stochastic process within industries. Further, as highlighted by Tether (2005), also the service sector is a locus of innovation. We only remove firms reporting abnormal R&D intensities.[7] Our sample comprises 4366 observations in the years 2004 and 2008, and 3308 observations in the years 2009 and 2014 – in total 15,348 year-firm pairs.

To clarify the definition of firm-level variables and also ease their interpretation in the subsequent context of distributional decomposition Table 1 reports each firm-level determinant under consideration and its corresponding concept at the distribution level. Finally Table 2 reports the list of the explanatory variables and their definitions.

Table A-2 (Supplementary materials) shows how the main variables of interest fluctuated over the period of observation. The sum of governmental R&D financing schemes roughly followed the Spanish business cycle trend. This resulted in a considerable increase during the period 2004–2008 and a contraction in the period 2009–2014. Despite the internal R&D investment following roughly the same trend (see Table A-1, Supplementary materials), the same does not hold for R&D intensities, which diminished substantially between 2004 and 2009, while recovering only marginally in 2014. There was, however, little change in the share of gazelles, which diminished in percentage by little. Finally, FCs seem to have slightly increased during the recovery of Spanish economy.

**Table 1.** Correspondences between firm-level and distributional variables.

| Firm-level variables | Distribution-wise correspondent |
|---|---|
| Binary gazelle status | Share and location of gazelle firms |
| Binary financial constraint status | Share and location of financially constrained firms |
| Public subsidy amount | Quantity and recipients of subsidies |

Note: we use the term 'location' to refer to the firms' location along the R&D intensity distribution.

**Table 2.** Variable definitions.

| Factor name | Definition |
|---|---|
| Internal R&D intensity | Ratio of internal R&D expenditure to sales level (in logs) |
| R&D public financing | Internal R&D expenditure financed with public funding (in €) |
| Gazelles | Either one of the three conditions below holds: |
| – HGF sales | – Quantile-based definition including the top decile of the unconditional sales growth distribution. |
| – HGF employment | – Quantile-based definition including the top decile of the unconditional employment growth distribution. |
| – YIC | – Firms with fewer than 250 employees, less than six years old and at least 15% of R&D intensity. |
| Financial constraints | Firms declaring lack of either internal or external funding as highly relevant. |
| Firm-level characteristics | |
| – sales level | Volume of turnover (in €). |
| – size | Number of employees. |
| – sector-normalized sales growth | Year sales growth minus 2-digit sector average growth.[8] |

As a measure of R&D, we focus on intensities in terms of sales, rather than pure investment. Doing so, we correct for likely scale effects by following the extensive literature on the topic (Leonard 1971; Grabowski and Baxter 1973; Dosi 1988). Figures 1 and 2 show the empirical differences along the distributions of interest. Further, given that testing differences on the mean (or other points of the distribution) using traditional approaches (e.g. the t-test or the Kolmogorov–Smirnov test) would generate inevitable biases due to the different sample sizes across the years, and due to the cross-time dependence of incumbent firms' observations, we opt for a different approach. We set up a simple, quantile panel regression model where the dependent variable is *R&D intensity*, while as explanatory variables we include only the categorical variable *year* (see Table 3). Through this, we can explore the difference between years across the whole distribution, while considering serial correlations and distributional specificities.

First, looking at the kernel densities, it is quite clear that from 2004 there has been a small location shift toward the right, such that the aggregate distributions relative to 2009 and 2014 exhibit a

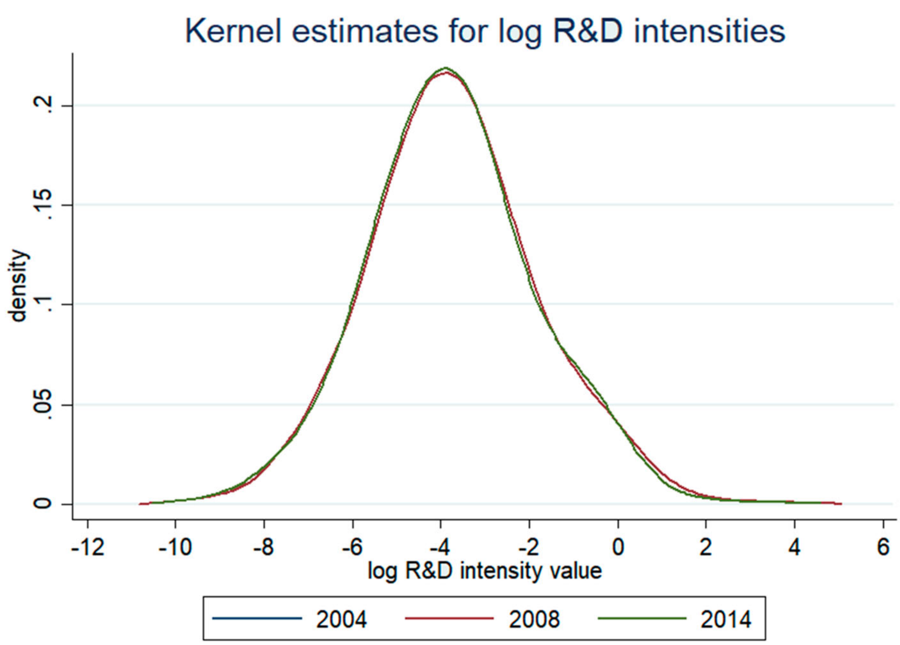


**Figure 1.** Kernel density of the distribution of interest (2004, 2008, 2014).

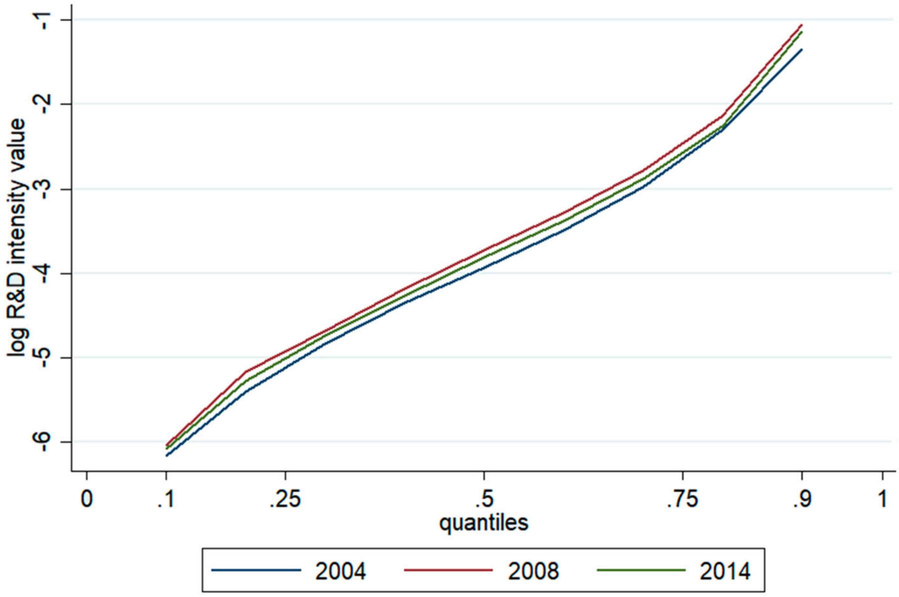


**Figure 2.** Observed quantile functions.

higher mode value. In addition, the unconditional analysis shows that the changes that took place in the expansionary period are quite concentrated around average and slightly above-average R&D-performing firms. On the contrary, looking at the contractionary period, the movements also involve the right-tail which is populated by highly technological enterprises. The quantiles with low R&D-intensive firms seem to be largely unaffected by time variations, suggesting a smaller role for them in the distribution.

## 4. Econometric approach

### *4.1. Methodology and types of counterfactuals*

Methods such as the Kitagawa–Blinder–Oaxaca decomposition have a long tradition in economics.[9] The seminal works by Oaxaca (1973) and Blinder (1973) paved the way for the development of methods aimed at going beyond the simple distributional mean by decomposing a distribution according to appropriately chosen components.[10] Recently, Chernozhukov et al. (2013) overcame the apparent inapplicability of quantile regressions by developing distribution regressions (Fortin et al. 2011).

Different from the standard quantile regression approaches, the approach developed by Chernozhukov et al. (2013) has substantial perks. On the one hand, econometrically, it contains the limit laws and a complete inference theory for the estimators, which were lacking in Machado and Mata (2005). On the other hand, in terms of practical application, the estimators are also easier to use in a variety of different contexts, as they well adapt to different types of data, like duration regression for

**Table 3.** Statistical differences across the distributions.

| | Location | Scale | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2008 | Y | Y | N | N | N | N | Y | Y | Y | N | N |
| 2014 | Y | Y | N | N | N | Y | Y | Y | Y | Y | Y |

Note: the refence category is the R&D intensity distribution in 2004, estimations are conducted using xtqreg package, which follows the method of Machado & Santos-Silva (2019). Differences are reported as such if statistically significant with *p*-value less than 5%.

instance. Finally, the authors also showed how in the presence of strong conditional heteroskedasticity of the dependent variable and of sizeable mass points, the quantile regression model is outperformed by their approach. Particularly, computational efficiency, less bias in presence of mass points, and no requirement of smoothness of the conditional density function, as the approximation is conducted pointwise, are the main reasons for the choice.

As reported by Chernozhukov et al. (2013), it is possible to identify three cases of counterfactual effects (henceforth CE). First, when the covariates do not vary, but the conditional distribution does (Type 1, or coefficient effect). Second, when the covariate distribution varies but the conditional distribution does not (Type 2, or characteristic effect). And finally, when both the conditional and the covariate distribution vary (Type 3). In this work, we first focus on Type 2, looking at how different values of covariates would have affected the distribution, and then, we appraise the reliability of the empirical model by introducing Type 3 effects.[11]

## *4.2. Identification strategy and our decomposition*

Following Chernozhukov et al. (2013) and Fortin et al. (2011), we develop an identification strategy to decompose the changes in the R&D distribution for Spain. We conduct the analysis for two opposed phases of the economic cycle.

In line with the reviewed literature on economics of innovation, the factors suspected to influence the evolution of R&D dynamics are governmental financing, the role of gazelles, financially constrained firms and, finally, certain firms' characteristics. We include firm size in terms of employees and their level of sales. Additionally, we take into account their sales growth rates normalized by sector. In doing so, we control for demand sectoral variations, but we also remove biasing factors such as inflation, while controlling for individual firms' growth (Coad and Grassano 2019; Bianchini et al. 2017). This argument holds even more for young and small companies (García-Quevedo et al. 2014). Finally, this variable also controls for firms' relative performances, as well as for differences in opportunities between sectors (Gkotsis and Vezzani 2022).

Starting from the estimation of Type 2 CE, we develop the decomposition relative to the years 2004–2008, the expansionary period.[12] This implies the estimation of how the 2008 internal R&D distribution would look, assuming that the covariates of interest take the 2004 values.

Suppose that $F_{RD\_int,\ (a,b,c,d)}$ corresponds to the counterfactual distribution of log R&D intensities, when: (a) the public financing scheme is as in year *a*, (b) the gazelles are those in year *b*, (c) financial constraints apply as in year *c*, and, (d) firms' characteristics are as in year *d*. Thanks to the law of iterated probabilities, it is possible to decompose the observed change in the distribution of R&D intensity between two years (2004, year 0, and 2008, year 1) into the sum of the above four effects, as follows:[13]

$$
\begin{aligned}
(1)\quad & F_{RD_{int\_1}|(1,1,1,1)} - F_{RD_{int\_1|(0,0,0,0)}} = [F_{RD_{int\_1}|(1,1,1,1)} - F_{RD_{int\_1}|(0,1,1,1)}] + [F_{RD_{int\_1}|(0,1,1,1)} - F_{RD_{int\_1}|(0,0,1,1)}] \\
& +[F_{RD_{int\_1}|(0,0,1,1)} - F_{RD_{int\_1}|(0,0,0,1)}] + [F_{RD_{int\_1}|(0,0,0,1)} - F_{RD_{int\_1}|(0,0,0,0)}]
\end{aligned}
$$

In more detail, this requires the identification and estimation of the following counterfactuals:

– $F_{RD_{int\_1}|(0,1,1,1)}(y) = \int F_{RD_{int\_1}|(1,1,1,1)}(y) \cdot dF_{RD_{int|(0,1,1,1)}}(x)$, corresponding to the distribution of R&D intensities that would prevail in 2008 if firms were subject to the public financing scheme of 2004;[14]
– $F_{RD_{int\_1}|(0,0,1,1)}(y) = \int F_{RD_{int\_1}|(1,1,1,1)}(y) \cdot dF_{RD_{int|(0,0,1,1)}}(x)$, corresponding to the distribution of R&D intensities that would prevail for firms in 2008 if firms were subject to the public financing scheme of 2004 and the gazelles were those of 2004.
– $F_{RD_{int\_1}|(0,0,0,1)}(y) = \int F_{RD_{int\_1}|(1,1,1,1)}(y) \cdot dF_{RD_{int|(0,0,0,1)}}(x)$, corresponding to the distribution of R&D intensities that would prevail in 2008 if gazelles were the ones of 2004 and firms were subject to the public financing scheme and financial constraints of 2004.

Following these estimations, our aim is to appraise how much of the observed differences between the two conditional distributions are explained by actual variations of the covariates, and how much is left to unobservability. This falls under the Type 3 CE:

$$(2)\quad F_{RD_{int\_1}|(1,1,1,1)} - F_{RD_{int\_0}|(0,0,0,0)} = [F_{RD_{int\_1}|(1,1,1,1)} - F_{RD_{int\_1}|(0,0,0,0)}] + [F_{RD_{int\_1}|(0,0,0,0)} - F_{RD_{int\_0}|(0,0,0,0)}]$$

From Equation (2), Type 3 CE consists of the sum of Type 1 and Type 2 effects, where the latter is the subject of the first part of the analysis. Indeed, it corresponds to the total differences emerging through the proposed conditional model and consists of: (i) a 'characteristics effect' due to changes of the selected determinants (the firm-level features), and (ii) a 'coefficient effect', which is imputable to changes in the estimated parameters relative to the dependent variable distribution. While the coefficient effect can be interpreted quite straightforwardly, the same does not hold for the characteristic effect, which may hide simple intercept shifts or an altered coefficient in any of our covariates. Finally, despite tracking some time dynamics, the model falls into the static category. In this, potential endogeneity problems may arise due to the omitted variable bias, some measurement error, or simultaneity. Excluding the chance of diffused measurement errors, as firms are supposed to report their official accounts, we are left with the possibility of omitted variable bias and of simultaneity. Regarding the former, we include the control vector of firms' characteristics with the explicit aim of accounting for both observables and non-observables correlations. Concerning possible simultaneity biases, we argue that it is hard to categorically exclude them. Nevertheless, we stress the associative nature of our model and how the standard errors and the estimated effects are robust across estimations involving different years and specifications, and how intrinsically the plug-in principle on which the method is based, reduces some of the potential sources of endogeneity by using lagged variables.

## 5. Results

The empirical results first address how the distribution would have changed if its covariates had been those of the starting year (type 2 CE). This is replicated for both the pre- and post-crisis, respectively in subsections 5.1 and 5.2. Then, in subsection 5.3 we quantify how much the observed change was explained by variation in the chosen covariates (type 3 CE). Finally, subsection 5.4 summarizes our results under a comparative perspective.

### *5.1. Pre-crisis (2004–2008): the expansionary phase*

Starting our decomposition analysis Table 4 and Figure A-3 report the results. The first terms correspond to the effects that governmental financing had on shaping the R&D distribution. An equivalent question would be: how would the distribution of R&D in 2008 look if public financing were of the same level and given to the same firms as they were in 2004? As expected, the answer is lower. There are three clear patterns. First, there is a positive association between the new financing scheme and firms' technological intensity. Second, the progressiveness with which these subsidies influenced the distribution is also clear. Finally, most of these effects are concentrated among top R&D firms.

Similarly, our results regarding the influence of gazelles hint at a progressive lessening of these 'hyper-firms' contribution to R&D. It appears that past gazelles were able to contribute more to the aggregate technological effort of the country. Curiously, these effects are concentrated not only among the more R&D intensive, but also in the less intensive, quantiles. The magnitude of the effects differs significantly, the 90/50 ratio being roughly 7-fold, but it is interesting to see how gazelles contribute quite extensively to R&D intensity.

Looking at FCs, we detect a prominent negative impact of 2004 constraints on 2008 firms, implying that in 2008, they would perform significantly worse than the previous FCs. Also, although the

**Table 4.** Results relative to the impact that each factor of choice in 2004 would have had on the 2008.

| Quantile | R&D Public financing | Gazelle firms | Financial constraints | Firms' characteristics |
|---|---|---|---|---|
| 0.1 | −0.023*** | −0.009 | −0.019*** | 0.072*** |
| | (0.007) | (0.008) | (0.005) | (0.01) |
| 0.2 | −0.027*** | −0.004 | −0.031*** | 0.056*** |
| | (0.007) | (0.007) | (0.005) | (0.007) |
| 0.3 | −0.031*** | 0.007 | −0.036*** | 0.042*** |
| | (0.007) | (0.008) | (0.006) | (0.007) |
| 0.4 | −0.035*** | 0.019** | −0.038*** | 0.031*** |
| | (0.007) | (0.008) | (0.006) | (0.005) |
| 0.5 | −0.040*** | 0.029*** | −0.040*** | 0.024*** |
| | (0.007) | (0.009) | (0.006) | (0.005) |
| 0.6 | −0.043*** | 0.041*** | −0.040*** | 0.017*** |
| | (0.007) | (0.01) | (0.006) | (0.006) |
| 0.7 | −0.051*** | 0.067*** | −0.039*** | 0.007 |
| | (0.008) | (0.014) | (0.007) | (0.008) |
| 0.8 | −0.066*** | 0.106*** | −0.041*** | −0.01 |
| | (0.01) | (0.017) | (0.008) | (0.011) |
| 0.9 | −0.095*** | 0.190*** | −0.038*** | −0.044** |
| | (0.015) | (0.024) | (0.01) | (0.017) |

Note: Number of observations: 4366. Significance levels corresponding to *$p < 0.1$, **$p < 0.05$, and ***$p < 0.01$. Standard errors are computed on 100 bootstrap repetitions.

estimated coefficients are roughly uniform across the distribution, a notable exception is for the least R&D-intensive firms, whose investment (in relation with their assets) have been less affected by FCs. This represents only a partial confirmation of Hypothesis 3, as we find that for this period all firms are affected uniformly, with the only difference that those occupying the first decile that is less affected. Nevertheless, the line of interpretation of the result is valid, and as theory and previous evidence suggests, low R&D-intensive firms have less difficulties in obtaining funding than those pursuing more intensive (and uncertain) objectives.

Finally, firms' characteristics control for possible standard cofounding factors. Particularly, firms' main attributes (such as size, sales, and sector-normalized growth) played a significant role in shaping the aggregate distribution. Interestingly, the effects show a strong asymmetry. The upper part of the distribution would show lower performance if firms' characteristics were those of 2004, while the lower quantiles would be greater. All of this suggests that Spanish firms became more R&D-intensive during the expansionary phase, increasing their presence in the top-performing quantiles.

### *5.2. After-crisis (2009–2014): the contractionary phase*

This section presents the changes in the effect of our key factors during a contraction phase (Table 5). The global economic downturn hit Spain with great strength and led to a considerable contraction in both R&D investment and subsidies (Cruz-Castro and Sanz-Menéndez 2016). Hence, it is particularly interesting to assess what might have happened if the Spanish government had been able to maintain the pre-crisis public financing scheme intact. It appears that it would have generated a higher performance among firms' R&D intensity. This is also in line with the general countercyclical effects found in the literature (Aristei et al. 2017). Again, the progressive effect of the scheme that emerged in the previous estimation is confirmed, the top-performing quantiles are those that would have benefited the most.

Subsection 5.1. showed that the gazelles' contribution slowed down considerably in 2008 as compared to 2004. The same is true when comparing 2009 gazelles with those of 2014. This evidences a general decrease in the capacity of these firms to contribute to aggregate R&D, which is now likely to be sustained by other types of firms. Interestingly, focusing on the 90/50 ratio as a proxy for (positive) outliers' dispersion, this is about twice as big as the same ratio measured during the expansionary phase (from 7 to 14). This hits at a major concentration of the effects on the right part of the distribution. Besides the observed decline in gazelles' relevance for R&D dynamics, it is also clear that

**Table 5.** Results relative to the impact that each factor of choice in 2009 would have had on the 2014 low.

| Quantile | R&D public financing | Gazelle firms | Financial constraints | Firms' characteristics |
|---|---|---|---|---|
| 0.1 | 0.026*** | −0.002 | −0.001 | −0.009 |
| | (0.008) | (0.008) | (0.006) | (0.012) |
| 0.2 | 0.023*** | 0.000 | 0.002 | −0.006 |
| | (0.008) | (0.007) | (0.005) | (0.009) |
| 0.3 | 0.024*** | −0.001 | −0.001 | −0.002 |
| | (0.008) | (0.007) | (0.004) | (0.007) |
| 0.4 | 0.032*** | 0.004 | 0.000 | −0.003 |
| | (0.009) | (0.008) | (0.005) | (0.006) |
| 0.5 | 0.036*** | 0.007 | 0.000 | 0.001 |
| | (0.011) | (0.008) | (0.005) | (0.006) |
| 0.6 | 0.042*** | 0.015 | −0.001 | −0.004 |
| | (0.012) | (0.01) | (0.006) | (0.008) |
| 0.7 | 0.057*** | 0.029** | −0.002 | −0.012 |
| | (0.015) | (0.012) | (0.008) | (0.01) |
| 0.8 | 0.087*** | 0.044*** | −0.002 | −0.021* |
| | (0.021) | (0.015) | (0.009) | (0.013) |
| 0.9 | 0.122*** | 0.097*** | −0.003 | −0.029* |
| | (0.03) | (0.022) | (0.009) | (0.017) |

Note: Number of observations: 3308. Significance levels corresponding to *$p < 0.1$, **$p < 0.05$, and ***$p < 0.01$. Standard errors are computed on 100 bootstrap repetitions.

gazelles show low resiliency during the crisis, struggling hard to maintain their usual contributions to aggregate R&D distribution.

FCs had a distinct trend during the expansionary phase, but the corresponding results for the contractionary period differ considerably. Unexpectedly, financial constraints have almost no role in affecting R&D intensities of Spanish firms in the period 2009–2014. Indeed, as for the previous period, some negative effects are detected, but the absence of statistical significance under several different statistical tests suggests large levels of dispersion in how FCs impacted firm-level contributions (and decisions). Thus, we find no confirmation of Hypothesis 3 second part. The fact that imposing 2009-FCs on 2014 firms shows no relation with the R&D intensity variations along the distribution could look puzzling. Nevertheless, a reasonable explanation is that in 2009, Spanish public financing reached its peak (see Table A-2), and that this intervention strongly alleviated firms' financial struggles.

Finally, during the contractionary phase, the asymmetric effect that emerged through the expansion vanishes. Instead, we find a negative impact, but more strongly concentrated around two quantiles (0.8 and 0.9). This is interpreted as a deterioration of firms' characteristics, becoming less technologically prone and showing a lower propensity to invest intensively in R&D, over the period from 2009 to 2014.

### 5.3. Aggregate model decomposition (type 3 CE) and the role of unobservables

This subsection explores the decomposition from a more aggregate approach and quantifies how much of the observed changes are explained by our conditional model and how much of it is due to unobserved variation.

Recalling that in 2008–2009 the distribution was at its peak, the left side of Figure 3 presents the Type 3 CE that emerges changing both the covariates and the conditional distribution of 2008 from the ones of 2004. Trivially, the effect would be negative, and the distribution would be lower. On the right side of the figure, the same difference emerges regarding 2014 by considering the covariates and conditional distribution of 2009. The impact of the decreased contribution of gazelles to the right-tail of the distribution is already clear.

Nevertheless, it is possible to go into considerable detail, decomposing these overall differences into Type 1 and Type 2 CE, recalling that Type 2 corresponds to what is outlined in the previous two subsections (characteristic effect), while Type 1 (coefficient effect) reports on the effect attributable to unobserved characteristics. Figure 4 and Table 6 present these estimations.

Furthermore, we observe differences in the relative magnitudes of these effects between the two phases under analysis. The effect of our selected determinants in the contractionary phase dominates the same effect estimated for the expansionary phase. This is also clear from Table 4, where the relative magnitudes are roughly two-fold.

Generally, the fact that the coefficient effect is so strong in both periods suggests the influence that unobservable firm-level characteristics have on R&D intensities. This portion corresponds to what Cohen and Klepper (1992) defined as R&D-related expertise. Further, these characteristics weigh more in the expansionary period, where an increase in intensity was expected, rather than in the contractionary one. Finally, in both specifications, this effect decreases across quantiles, making the top R&D-intensive quantiles less subject to these economic fluctuations. Economically, this can be explained by the fact that these quantiles are mostly composed of science-based firms, or R&D specialists (Cattaruzzo 2020), whose main (and sometimes only) business focus is introducing novelties to the market.

### 5.4. Comparison and interpretation

Table 7 summarizes the empirical findings regarding the effects of the selected determinants (Type 2 CE). In addition to the estimated coefficients, we also report the associated percentages of the total estimated change.

The most interesting asymmetries relate to the role that public financing and gazelles had in shaping the distribution of Spanish internal R&D. Indeed, in the expansionary period, public R&D financing from the years 2008 and 2009 was strongly positive and progressive in incentivizing firms' technological intensity. This and the higher coefficients estimated for the higher intensity deciles confirm hypothesis 1 regarding the asymmetric effect of public subsidies. Additional evidence for the finding is that in 2007 the usage of public R&D financing reached its peak, and then started a monotonic decline that led to only half of the allocated national budget for R&D financing being taken up by enterprises. This is in line with previous findings and likely due to changes in the application and format of the public financing, which led to much greater inefficiency

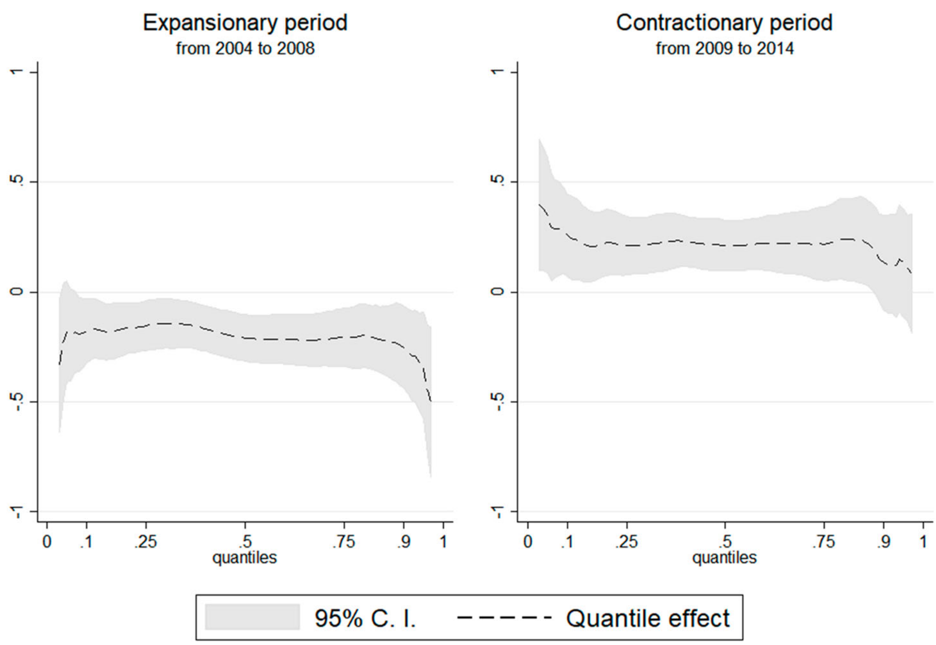


**Figure 3.** Overall differences for the conditional model.

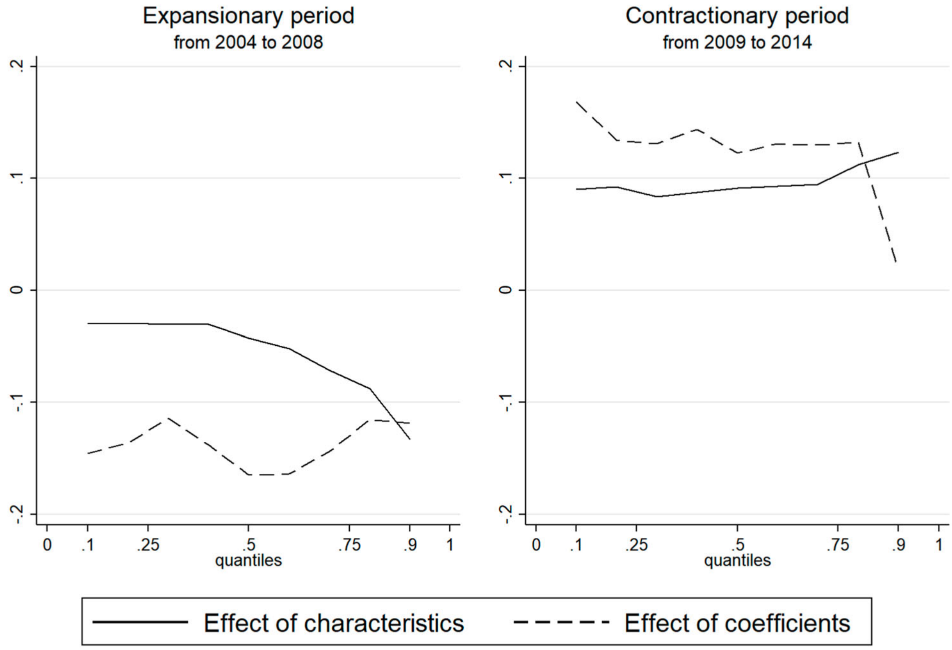


**Figure 4.** Model decomposition.

(Cruz-Castro and Sanz-Menéndez 2016). Similarly, and especially in the expansionary period, gazelles played a leading role but their overall contribution to R&D declined from 2004 to 2014. This evidence confirms Hypothesis 2, which postulated a key role for gazelles in shaping the distribution dynamics of the right-tail. Further, this can be interpreted very fruitfully recalling the vein of study on 'reduced business dynamism'.

It is well known that firms located in the right-tail are largely responsible for aggregate dynamics, thus HGFs or gazelles are of utmost importance (Bijnens and Konings 2020). Recent economic evidence points at a widespread reduced business dynamism in advanced economies. The decline has been identified for countries such as the USA (Decker et al. 2014, 2016, 2017 and 2020; Guzman and Stern 2020), Australia (Bakhtiari 2017), Belgium (Bijnens and Konings 2020), Turkey (Akcigit et al. 2020), Canada (Macdonald 2014), and Portugal (Sarmento and Nunes 2010). More recently, a multi-country analysis for 18 countries and 22 industries with data covering the last two decades demonstrated how the phenomenon is both common and globally diffuse (Calvino et al. 2020). Still lacking a stylized explanation, researchers have proposed various theories.

**Table 6.** Aggregate decomposition results.

| | Characteristics effect – explained by the determinants | | Coefficient effect – explained by unobservables | |
|---|---|---|---|---|
| Quantiles | "2004–2008" | "2009–2014" | "2004–2008" | "2009–2014" |
| 0.1 | 17.1% | 34.9% | 82.9% | 65.1% |
| 0.2 | 17.7% | 40.8% | 82.3% | 59.2% |
| 0.3 | 21.0% | 39.0% | 79.0% | 61.0% |
| 0.4 | 18.1% | 37.9% | 81.9% | 62.1% |
| 0.5 | 20.5% | 42.7% | 79.5% | 57.3% |
| 0.6 | 24.0% | 41.5% | 76.0% | 58.5% |
| 0.7 | 33.1% | 42.2% | 66.9% | 57.8% |
| 0.8 | 43.2% | 45.8% | 56.8% | 54.2% |
| 0.9 | 52.9% | 86.6% | 47.1% | 13.4% |

**Table 7.** Results summary table.

| | Expansionary period | | | | | Contractionary period | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Effect of | | | | | Effect of | | | |
| Quantile | Total estimated change | Public financing | Gazelle firms | Financial constraint | Firms' chars. | Total estimated change | Public financing | Gazelle firms | Financial constraint | Firms' char. |
| **0.1** | 0.021 | −0.023 | −0.010 | −0.019 | 0.072 | 0.013 | 0.026 | −0.002 | −0.001 | −0.009 |
| | | −106.9% | −45.5% | −87.3% | 339.7% | | 201.8% | −19.1% | −9.4% | −73.2% |
| **0.3** | −0.017 | −0.031 | 0.007 | −0.036 | 0.042 | 0.021 | 0.024 | −0.001 | −0.001 | −0.002 |
| | | 177.6% | −40.9% | 208.0% | −244.6% | | 118.2% | −6.9% | −2.5% | −8.8% |
| **0.5** | −0.028 | −0.040 | 0.029 | −0.040 | 0.024 | 0.044 | 0.036 | 0.007 | 0.000 | 0.001 |
| | | 145.9% | −104.4% | 145.6% | −87.1% | | 82.7% | 15.7% | 0.4% | 1.2% |
| **0.7** | −0.016 | −0.051 | 0.067 | −0.039 | 0.007 | 0.072 | 0.057 | 0.029 | −0.002 | −0.012 |
| | | 310.9% | −408.6% | 239.3% | −41.6% | | 80.0% | 39.8% | −2.7% | −17.1% |
| **0.9** | 0.014 | −0.095 | 0.190 | −0.038 | −0.044 | 0.188 | 0.122 | 0.097 | −0.003 | −0.029 |
| | | −694.1% | 1394.1% | −278.2% | −321.8% | | 64.8% | 51.9% | −1.4% | −15.4% |

Note: The coefficient estimates are reported from the above estimations, while percentage contributions to the total variation is computed and reported in the second line of each cell. Lastly, the total estimated change has a rough correspondence to the estimated 'characteristics effect', as different decomposition orderings would return different pointwise estimates for the non-statistically significant effects.

Among these, the reduced presence of gazelles is one effective candidate. When looking at the causes, it is virtually impossible to isolate a specific one. Nevertheless, technological effort, as proxied by R&D intensity, is a key component of market dynamics, since it allows starting and follower firms to challenge the industry leaders, thus fostering competition and faster growth. Complementing the theoretical insights proposed in the model by De Ridder (2019), we find confirmation of the fact that ' … the relationship between aggregate R&D and aggregate growth depends on how R&D is distributed across firms'. Indeed, this puts considerable importance on the way R&D subsidies schemes are designed in terms of firm-level heterogeneity and it suggests that they should strive for creating dynamic markets where not only incumbent and wealthy firms perform R&D investments, but also smaller and well-selected firms are receivers of support to compete. Thus, tools aimed at providing support such as targeted R&D subsidies can be very valuable (Akcigit et al. 2020). Both empirical and theoretical evidence point at a reduction in knowledge diffusion, especially between frontier and laggard firms, as a coherent explanation for the reduction in business dynamism (Akcigit and Ates 2021). In this context, the role of right-tail firms (i.e. gazelles) and of R&D subsidies are primary factors that can stimulate more dynamism.

Finally, if FCs were indeed impacting firms at the beginning of the expansionary period, what emerged is that the same did not happen during the strong contraction to the Spanish industrial system. The pieces of evidence confirm partially Hypothesis 3. On the one hand, we do confirm that there exist asymmetries in the relation between finance (or lack thereof) and innovative effort across the distribution of R&D intensity. On the other hand, we reject the strengthening of this phenomenon when firms face a contractionary period.

Recalling that these estimations derive from a counterfactual framework, a likely explanation is that the public financing scheme employed in the expansionary period alleviated the financial needs of most firms. Thus, the constraints in place in 2009 are not relevant in explaining variations of R&D intensities. This would be in line with Bartz and Winkler (2016), who find that younger and smaller firms have the capacity to procure additional funding to pursue projects granting them advantages during economic downturns. The authors also highlight the role of the lending sector, banks particularly, and of governmental support of bank-mediated liquidity.

## 6. Conclusion

Cohen and Klepper (1992) offered a seminal way to look at the R&D intensity distribution, where much of its anatomy is explained by non-observables, managerial skills, a phenomenon that is common across industries. In a complementary fashion, our work identifies and quantifies specific firm-level contributions, also looking at how these dynamically change over time and according to distinct phases of the business cycle. Given the wide coverage and timing of PITEC, the Spanish R&D aggregate distribution before and after the 2008 global economic downturn constitutes an insightful research setting. There, the detailed decomposition shows robust emerging patterns in terms of the chosen firm-level determinants and modelling approach.

We apply the technique developed by Chernozhukov et al. (2013), which offers comprehensive explanations of the individual factors' contribution to each part of the R&D distribution. This allows us to consider its skewed nature and largely heterogeneous composition. We analyse how four determinants (public financing, gazelles, FCs, and firms' characteristics) contribute to shaping the overall distribution of R&D. In line with Cohen and Klepper (1992), our results show a dominance of unobservables for all quantiles other than the top R&D intensive firms. This implies that, for non-R&D specialists, the determinants under study are not the main drivers of their investment decision, rather they are driven by other economic fluctuations. Looking at the individual determinants, we identify public financing and the gazelles' contribution as the main determinants in shaping the distribution of interest. In particular, the contribution from gazelles decreases considerably over the period from 2004 to 2014.

There are several potential explanations for this decreased contribution that can be mostly identified in structural factors affecting economies. First, the rising importance of intangible and digital assets can lead to increases in market shares and power of the best-performing firms, posing additional barriers to growth (De Ridder 2019). Additionally, the effect of globalization and the predominance of global value chains cannot be neglected since they deeply affect firm dynamics and entry mechanisms. This not only extends the existing empirical findings on reduced business dynamism across developed economies (Decker et al. 2016 2017; Akcigit and Ates 2021), but also constitutes a possible transmission channel (R&D → innovation → productivity). Our evidence is in line with the steady declines in business dynamism showed by Calvino et al. (2020) over the last two decades, even after accounting for the role of the business cycle.

The present study is not without limitations, as some methodological and modelling choice have been made to keep the results interpretable and robust. With this aim, the number of variables of interest has been limited and further inclusion of other relevant variables could be a matter of investigation. Additionally, although the coverage in terms of investment is very high, having information on all R&D investors in a country could allow the application of even more refined and precise techniques. Also, the variables under study, despite they have been already analysed jointly under similar perspectives, are not exempt from possible endogeneity concerns. To reduce the influence of this phenomenon, we performed several robustness checks that show the stability of the estimates. Although the non-causal scope of the paper makes the issue slightly less relevant, it is important to keep the concern in mind. Finally, our analysis is restricted to the Spanish context. As explained above, Spain is a moderate and slow-growing innovator. Future studies should aim at also re-investigating the issue of innovation persistence in other low- and high-innovative countries to have a more complete picture.

### 6.1. Policy implications

Our results suggest the development of at least two different policy actions. First, gazelles are considered as important players in innovative and dynamic markets but due to their quasi-random growth process, supporting them has always been difficult. Policies aimed at promoting, not only their presence, but also their contribution in terms of innovative performance should be pursued. Second, Spain is the EU country that has cut public R&D spending the most after the 2007 crisis. Since 2017, public investment in R&D has been recovering, but its level has not reached the levels before the 2008 crisis. Given the distributional nature of our study, both actions could help in the pursuit for convergence of R&D intensities across EU countries, in which Spain is lagging behind. According to the 2021 European Innovation Scoreboard, the innovative capacity of the EU-27 economies grew at an annual rate of 8.9% between 2012 and 2019.

Despite this, the erratic evolution of innovation policies in countries such as Spain has limited the ability of the EU to reverse the decline in R&D spending and innovation in the Euro area. More generally, it appears evident that simply setting R&D targets accompanied by general scope support policies is not the way to go. Contrarily, we stress how the modern econometric tools and the great availability of firm-level data would allow for the development of bottom-up, micro-derived policies and targets that together can effectively increase aggregate intensity.

Also, the political importance that knowledge policies bring is increasing and there is shared awareness that recessions are historically triggered for successful and disruptive innovations focusing on grand societal challenges, such as climate change and the 'green' revolution. Considering all of these, the correct design of innovation policy is of key and utmost importance to compete on a global scale. To do so and create a policy that is fitter to face sustainability and digitalization challenges, two elements are necessary: (1) better coordination at different levels (i.e. regional-national-supranational), and (2) better tailoring of these schemes.

Particularly on this second point, the intrinsic heterogeneity of firms in each country makes it fundamental to control for distributional dynamics as excessive concentration in the distribution of R&D

intensities is inevitably harmful to entry-exit firm dynamics, hindering especially the smallest, youngest, and possibly more innovative ones. Particularly, recession times should be accompanied by generous and targeted subsidy schemes given the reluctance and lower convenience in financing innovation investment with external funding. In doing so, a proper and theoretically optimal allocation of R&D subsidies should be pursued also following the evidence presented here. This should be considered as an additional and complementary piece of information for a bottom-up tailoring of public support policy design.

In conclusion, R&D spending is concentrated in a limited number of companies, industries, and countries. Our analysis shows that gazelle-like companies are undervalued, and their centrality suggests a re-evaluation (Kuhlmann and Rip 2018). This study offers evidence for the size of their impact on the whole innovative capacity. Although these young and dynamic companies are not the ones that invest the most in R&D in absolute terms, they are often the ones that generate the most disruptive innovations, thereby affecting the rest of the productive network. These actors require better access to financial resources and their needs should be considered by policymakers. Their actions not only generate large knowledge and technology spill-overs, but also exhibit greater transformation potential than non-gazelle companies.

## Notes

1. Over the last decade a wide number of investigations has highlighted the importance of multi-level studies and proposed estimations based on micro (or sectoral) data, but which can be quantified in relation to the most common aggregate measures (Gabaix 2011; Acemoglu et al. 2012; Di Giovanni and Levchenko 2010; Di Giovanni et al. 2014; Carvalho and Gabaix 2013; Foerster et al. 2011). These studies tend to rely on either input-output matrices or quite long-term micro data.
2. The access to data from the *Panel de Innovación Tecnológica* (PITEC) promoted a proliferation of works on various aspects related to the drivers and effects of R&D among Spanish companies. Particularly, they covered issues such as the drivers of innovation, firm growth and high-growth firms, R&D and cooperation strategies, barriers to innovation, and the influence of the business cycle (López-García et al. 2013; Costa-Campi et al. 2014; García-Quevedo et al. 2018), among other topics.
3. This group of firms will be formally defined in Section 2, but we anticipate that its definition derives from Birch (1987) and all the subsequent works on high-growth firms and young innovative companies.
4. This question, however, is outside the scope of this paper. Our aim is to disentangle the effect that the public financing scheme has on each part of the R&D distribution.
5. Indeed, one could argue that the inclusion of these companies is almost tautological, given their definition. Nevertheless, we argue that YICs, together with HGFs, are extremely important actors in the economy and offering results that hold for a larger group of firms is also a way to offer more powerful and easier to implement policy suggestions. Given the erratic nature of firms, and of dynamic ones especially, this argument is even more prominent. For the sake of robustness, we repeated the estimations of the paper also without YICs, and the statistically significant results are extremely correlated. Results are available upon request to the authors.
6. See Table A-1 for a description of the representativeness of the total national investment in internal R&D.
7. Particularly, we remove firms declaring 1€ of sales while investing large amounts on R&D and firms declaring R&D intensities higher than 100%.
8. Sectors follow the CNAE-2009 classification, which can be directly linked to the more detailed NACE classification.
9. Typical applications of the original method in the industrial literature regard the decomposition of productivity (i.e. Fariñas and Ruano 2004).
10. Kitagawa (1955) was the first to propose this type of decomposition.
11. Despite its advantages, the approach has one main limitation: counterfactuals may not be enough for causal insights. Particularly, the present analysis cannot be considered out of the influence of possible confounding and/or selection on variables, which prevent us from interpreting the results as causal. Thus, the results are commented in an associative manner.
12. It is possible to obtain the same decomposition for the contractionary period by substituting the appropriate time indexes.
13. The results of these decompositions can be order independent. We run the estimations for each possible order of the decomposition. Table A-4 shows the results remain unchanged.
14. Firm-level public financing data are available only for 2003 and 2005. For the baseline version, we use 2003 public financing data but, in the robustness checks, we explore the sensitivity of the findings to alternative

choices. All relevant results show correlations higher than 90%, thus we can safely conclude that this does not affect our findings.

## Acknowledgements

We thank Alex Coad, José García-Quevedo, Clemens Domnick, Francesco Rentocchini, Alexander Tübke, James Gavigan, Elisabeth Nindl, Ramón Compaño, Lorenzo Napolitano, and other participants of the 26th SPRU PhD forum (Brighton, 14–15 May 2020), of the Workshop on Industrial and Public Economics (Reus, 4–5 February 2021), of the XXIII Applied Economics Meeting (online, 3–4 June 2021), EU-SPRI conference (online, 9–11 June 2021) and of the EC-JRC Seminar (Seville, 20 June 2022) for useful feedback.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was supported by Horizon 2020 Framework Programme [grant number 713679]; Xarxa de Referència en Economia Aplicada (XREAP); Consolidated Group of Research [grant number 2017-SGR-00493]; Universitat Rovira i Virgili [grant number 2019PFR-URV-B2-80].

## ORCID

*Sebastiano Cattaruzzo* http://orcid.org/0000-0001-8857-8314
*Agustí Segarra-Blasco* http://orcid.org/0000-0001-6316-3171
*Mercedes Teruel* http://orcid.org/0000-0002-4104-7679

## References

Acemoglu, D., U. Akcigit, N. Bloom, and W. Kerr. 2018. "Innovation, Reallocation and Growth." *American Economic Review* 108 (11): 3450–3491.

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi. 2012. "The Network Origins of Aggregate Fluctuations." *Econometrica* 80 (5): 1977–2016.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt. 2005. "Competition and Innovation: An Inverted-U Relationship." *The Quarterly Journal of Economics* 120 (2): 701–728.

Akcigit, U., Y. E. Akgunduz, S. M. Cilasun, E. Ozcan-Tok, and F. Yilmaz. 2020. "Facts on Business Dynamism in Turkey." *European Economic Review* 128: 103490.

Akcigit, U., and S. T. Ates. 2021. "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory." *American Economic Journal: Macroeconomics* 13 (1): 257–298.

Akcigit, U., and W. Kerr. 2018. "Growth Through Heterogeneous Innovations." *Journal of Political Economy* 126 (4): 1374–1443.

Appelt, S., M. Bajgar, C. Criscuolo, and F. Galindo-Rueda. 2016. "R&D Tax Incentives: Evidence on Design, Incidence and Impacts." *OECD Science, Technology and Industry Policy Papers* 32: 1–43.

Aristei, D., A. Sterlacchini, and F. Venturini. 2017. "Effectiveness of R&D Subsidies During the Crisis: Firm-Level Evidence Across EU Countries." *Economics of Innovation and New Technology* 26 (6): 554–573.

Arrow, K. J. 1962. "The Economic Implications of Learning by Doing." *Review of Economic Studies* 29: 155–173.

Audretsch, D. B., A. Segarra, and M. Teruel. 2014. "Why Don't all Young Firms Invest in R&D?" *Small Business Economics* 43 (4): 751–766.

Bakhtiari, S. 2017. "Entrepreneurship Dynamics in Australia: Lessons from Micro-Data", *Research Paper No. 5/*2017, Commonwealth of Australia.

Barge-Gil, A., and A. López. 2014. "R&D Determinants: Accounting for the Differences Between Research and Development." *Research Policy* 43 (9): 1634–1648.

Barlevy, G. 2007. "On the Cyclicality of Research and Development." *American Economic Review* 97 (4): 1131–1164.

Bartz, W., and A. Winkler. 2016. "Flexible or Fragile? The Growth Performance of Small and Young Businesses During the Global Financial Crisis – Evidence from Germany." *Journal of Business Venturing* 31 (2): 196–215.

Bianchini, S., G. Bottazzi, and F. Tamagni. 2017. "What Does (not) Characterize Persistent Corporate High-Growth?" *Small Business Economics* 48 (3): 633–656.

Bijnens, G., and J. Konings. 2020. "Declining Business Dynamism in Belgium." *Small Business Economics* 54 (4): 1201–1239.

Birch, D. G. W. 1987. "Job Creation in America: How Our Smallest Companies Put the Most People to Work". University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship.

Blinder, A. S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *Journal of Human Resources* 8 (4): 436–455.

Breschi, S., and F. Malerba. 1997. "Sectoral Innovation Systems: Technological Regimes, Schumpeterian Dynamics, and Spatial Boundaries." *Systems of Innovation: Technologies, Institutions and Organizations* 1: 130–156.

Brown, R., S. Mawson, and C. Mason. 2017. "Myth-busting and Entrepreneurship Policy: The Case of High Growth Firms." *Entrepreneurship & Regional Development* 29 (5–6): 414–443.

Brown, J. R., and B. C. Petersen. 2015. "Which Investments Do Firms Protect? Liquidity Management and Real Adjustments When Access to Finance Falls Sharply." *Journal of Financial Intermediation* 24 (4): 441–465.

Busom, I., B. Corchuelo, and E. Martínez-Ros. 2014. "Tax Incentives … or Subsidies for Business R&D?" *Small Business Economics* 43 (3): 571–596.

Cainelli, G., V. De Marchi, and R. Grandinetti. 2015. "Does the Development of Environmental Innovation Require Different Resources? Evidence from Spanish Manufacturing Firms." *Journal of Cleaner Production* 94: 211–220.

Calvino, F., C. Criscuolo, and R. Verlhac. 2020. "Declining Business Dynamism: Structural and Policy Determinants." *OECD Science, Technology and Industry Policy Papers* 94: 1–76.

Carvalho, A. 2018. "Wishful Thinking About R&D Policy Targets: What Governments Promise and What They Actually Deliver." *Science and Public Policy* 45 (3): 373–391.

Carvalho, V., and X. Gabaix. 2013. "The Great Diversification and its Undoing." *American Economic Review* 103 (5): 1697–1727.

Cattaruzzo, S. 2020. "On R&D Sectoral Intensities and Convergence Clubs", JRC Working Papers on Corporate R&D and Innovation No 01/2020, Joint Research Centre.

Cattaruzzo, S., and M. Teruel. 2022. "On the Heterogeneity of the Long-Term Leverage-Growth Relationship: A Cross-Country Analysis of Manufacturing Firms." *Structural Change and Economic Dynamics* 62: 552–565.

Chernozhukov, I. Fernández-Val and B. Melly. 2013. "Inference on Counterfactual Distributions." *Econometrica* 81 (6): 2205–2268.

Coad, A. 2019. "Persistent Heterogeneity of R&D Intensities Within Sectors: Evidence and Policy Implications." *Research Policy* 48 (1): 37–50.

Coad, A., and N. Grassano. 2019. "Firm Growth and R&D Investment: SVAR Evidence from the World's Top R&D Investors." *Industry and Innovation* 26 (5): 508–533.

Coad, A., A. Segarra, and M. Teruel. 2021. "A Bit of Basic, a Bit of Applied? R&D Strategies and Firm Performance." *The Journal of Technology Transfer* 46: 1758–1783.

Cohen, W. M., and S. Klepper. 1992. "The Anatomy of Industry R&D Intensity Distributions." *The American Economic Review* 82 (4): 773–799.

Cohen, W. M., and R. C. Levin. 1989. "Empirical Studies of Innovation and Market Structure"." In *Handbook of Industrial Organization – Vol.2*, edited by R. Schmalensee and R. Willig, 1059–1107. Amsterdam: North Holland.

Costa-Campi, M. T., N. Duch-Brown, and J. Garcia-Quevedo. 2014. "R&D Drivers and Obstacles to Innovation in the Energy Industry." *Energy Economics* 46: 20–30.

COTEC. 2018. "Tecnología e innovación en España: informe COTEC 2018", Madrid: Fundación COTEC para la innovación.

Crépon, B., E. Duguet, and J. Mairessec. 1998. "Research, Innovation and Productivity: An Econometric Analysis at the Firm Level." *Economics of Innovation and New Technology* 7 (2): 115–158.

Cruz-Castro, L., and L. Sanz-Menéndez. 2016. "The Effects of the Economic Crisis on Public Research: Spanish Budgetary Policies and Research Organizations." *Technological Forecasting and Social Change* 113: 157–167.

Czarnitzki, D., and J. Delanote. 2013. "Young Innovative Companies: The New High-Growth Firms?" *Industrial and Corporate Change* 22 (5): 1315–1340.

Czarnitzki, D., and H. Hottenrott. 2011. "Financial Constraints: Routine Versus Cutting Edge R&D Investment." *Journal of Economics & Management Strategy* 20 (1): 121–157.

Decker, R. A., J. Haltiwanger, R. Jarmin, and J. Miranda. 2014. "The Role of Entrepreneurship in US job Creation and Economic Dynamism." *Journal of Economic Perspectives* 28 (3): 3–24.

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda. 2016. "Where Has All the Skewness Gone? The Decline in High-Growth (Young) Firms in the US." *European Economic Review* 86: 4–23.

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda. 2017. "Declining Dynamism, Allocative Efficiency, and the Productivity Slowdown." *American Economic Review* 107 (5): 322–326.

Decker, Ryan A, John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2020. "Changing Business Dynamism and Productivity: Shocks versus Responsiveness." *American Economic Review* 110 (12): 3952–3990.

Delmar, F., P. Davidsson, and W. B. Gartner. 2003. "Arriving at the High-Growth Firm." *Journal of Business Venturing* 18 (2): 189–216.

Del Río, P., C. Peñasco, and D. Romero-Jordán. 2016. "What Drives Eco-innovators? A Critical Review of the Empirical Literature Based on Econometric Methods." *Journal of Cleaner Production* 112: 2158–2170.
De Marchi, V. 2012. "Environmental Innovation and R&D Cooperation: Empirical Evidence from Spanish Manufacturing Firms." *Research Policy* 41 (3): 614–623.
Demircioglu, M. A., D. B. Audretsch, and T. F. Slaper. 2019. "Sources of Innovation and Innovation Type: Firm-level Evidence from the United States." *Industrial and Corporate Change* 28 (6): 1365–1379.
De Ridder, M. 2019. "Market Power and Innovation in the Intangible Economy", Discussion Papers 1907, Centre for Macroeconomics (CFM).
Di Giovanni, J., and A. A. Levchenko. 2010. "Putting the Parts Together: Trade, Vertical Linkages, and Business Cycle Comovement." *American Economic Journal: Macroeconomics* 2 (2): 95–124.
Di Giovanni, J., A. A. Levchenko, and I. Mejean. 2014. "Firms, Destinations, and Aggregate Fluctuations." *Econometrica* 82 (4): 1303–1340.
Dosi, G. 1988. "Sources, Procedures, and Microeconomic Effects of Innovation." *Journal of Economic Literature* 26 (3): 1120–1171.
Dosi, G. 1990. "Finance, Innovation and Industrial Change." *Journal of Economic Behavior & Organization* 13 (3): 299–319.
Evangelista, R. 2006. "Innovation in the European Service Industries." *Science and Public Policy* 33 (9): 653–668.
Falk, M. 2007. "R&D Spending in the High-Tech Sector and Economic Growth." *Research in Economics* 61 (3): 140–147.
Falk, M. 2012. "Quantile Estimates of the Impact of R&D Intensity on Firm Performance." *Small Business Economics* 39 (1): 19–37.
Fariñas, J. C., and S. Ruano. 2004. "The Dynamics of Productivity: A Decomposition Approach Using Distribution Functions." *Small Business Economics* 22 (3): 237–251.
Foerster, A. T., P.-D. G. Sarte, and M. W. Watson. 2011. "Sectoral vs. Aggregate Shocks: A Structural Factor Analysis of Industrial Production." *Journal of Political Economy* 119 (1): 1–38.
Fortin, N., T. Lemieux, and S. Firpo. 2011. "Decomposition Methods in Economics." In *Handbook of Labor Economics*, Vol. 4, edited by Orley Ashenfelter and David Card, 1–102. Elsevier.
Gabaix, X. 2011. "The Granular Origins of Aggregate Fluctuations." *Econometrica* 79 (3): 733–772.
García-Quevedo, J., G. Pellegrino, and M. Vivarelli. 2014. "R&D Drivers and Age: Are Young Firms Different?" *Research Policy* 43 (9): 1544–1556.
García-Quevedo, J., A. Segarra, and M. Teruel. 2018. "Financial Constraints and the Failure of Innovation Projects." *Technological Forecasting & Social Change* 127: 127–140.
Gkotsis, P., and A. Vezzani. 2022. "The Price Tag of Technologies and the 'Unobserved' R&D Capabilities of Firms." *Economics of Innovation and New Technology* 31 (5): 339–361.
Grabowski, H. G., and N. D. Baxter. 1973. "Rivalry in Industrial Research and Development: An Empirical Study." *Journal of Industrial Economics* 21 (3): 209–235.
Griffiths, W., and E. Webster. 2010. "What Governs Firm-Level R&D: Internal or External Factors?" *Technovation* 30 (7–8): 471–481.
Guzman, J., and S. Stern. 2020. "The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 32 US States, 1988–2014." *American Economic Journal: Economic Policy* 12 (4): 212–243.
Hall, B. 1993. "The Stock Market's Valuation of R&D Investment During the 1980s." *American Economic Review* 83 (2): 259–264.
Hall, B., and F. Hayashi. 1989. *Research and Development as an Investment* (No. w2973), National Bureau of Economic Research.
Hervás Soriano, F., and F. Mulatero. 2010. "Knowledge Policy in the EU: From the Lisbon Strategy to Europe 2020." *Journal of the Knowledge Economy* 1 (4): 289–302.
Howell, A. 2016. "Firm R&D, Innovation and Easing Financial Constraints in China: Does Corporate tax Reform Matter?" *Research Policy* 45 (10): 1996–2007.
Huergo, E., and L. Moreno. 2017. "Subsidies or Loans? Evaluating the Impact of R&D Support Programmes." *Research Policy* 46 (7): 1198–1214.
Jefferson, G. H., B. Huamao, G. Xiaojing, and Y. Xiaoyun. 2006. "R&D Performance in Chinese Industry." *Economics of Innovation and New Technology* 15 (4–5): 345–366.
Kitagawa, E. M. 1955. "Components of a Difference Between Two Rates." *Journal of the American Statistical Association* 50 (272): 1168–1194.
Kuhlmann, S., and A. Rip. 2018. "Next-generation Innovation Policy and Grand Challenges." *Science and Public Policy* 45 (4): 448–454.
Kunapatarawong, R., and E. Martínez-Ros. 2016. "Towards Green Growth: How Does Green Innovation Affect Employment?" *Research Policy* 45 (6): 1218–1232.
Lee, N., H. Sameen, and M. Cowling. 2015. "Access to Finance for Innovative SMEs Since the Financial Crisis." *Research Policy* 44 (2): 370–380.
Leonard, W. N. 1971. "Research and Development in Industrial Growth." *Journal of Political Economy* 79 (2): 232–256.

López-García, P., J. M. Montero, and E. Moral-Benito. 2013. "Business Cycles and Investment in Productivity-Enhancing Activities: Evidence from Spanish Firms." *Industry and Innovation* 20 (7): 611–636.
Macdonald, R. 2014. "Business Entry and Exit Rates in Canada: A 30-year Perspective", *Economic Insights* No. 38, Statistics Canada.
Machado, José A. F, and José Mata. 2005. "Counterfactual decomposition of changes in wage distributions using quantile regression." *Journal of Applied Econometrics* 20 (4): 445–465.
Machado, J. A., and J. S. Silva. 2019. "Quantiles via Moments." *Journal of Econometrics* 213 (1): 145–173.
Marzucchi, A., and S. Montresor. 2017. "Forms of Knowledge and Eco-innovation Modes: Evidence from Spanish Manufacturing Firms." *Ecological Economics* 131: 208–221.
Mina, A., H. Lahr, and A. Hughes. 2013. "The Demand and Supply of External Finance for Innovative Firms." *Industrial and Corporate Change* 22 (4): 869–901.
Moncada-Paternò-Castello, P., and N. Grassano. 2022. "The EU vs US Corporate R&D Intensity Gap: Investigating Key Sectors and Firms." *Industrial and Corporate Change* 31 (1): 19–38.
Montresor, S., and A. Vezzani. 2015. "The Production Function of Top R&D Investors: Accounting for Size and Sector Heterogeneity with Quantile Estimations." *Research Policy* 44 (2): 381–393.
Moreno, F., and A. Coad. 2015. "High-growth Firms: Stylized Facts and Conflicting Results." *Entrepreneurial Growth: Individual, Firm, and Region* 17: 187–230.
North, D., R. Baldock, and F. Ullah. 2013. "Funding the Growth of UK Technology-based Small Firms Since the Financial Crash: Are There Breakages in the Finance Escalator?" *Venture Capital* 15: 237–260.
Oaxaca, R. 1973. "Male-female Wage Differentials in Urban Labor Markets." *International Economic Review* 14: 693–709.
OECD. 2016. "OECD Science, Technology and Innovation Outlook", 2016.
Sarmento, E., and A. Nunes. 2010. "Entrepreneurship Performance Indicators for Employer Enterprises in Portugal", *Temas Económicos*.
Schneider, C., and R. Veugelers. 2010. "On Young Highly Innovative Companies: Why They Matter and How (Not) to Policy Support Them." *Industrial and Corporate Change* 19 (4): 969–1007.
Segarra, A., and M. Teruel. 2014. "High-growth Firms and Innovation: An Empirical Analysis for Spanish Firms." *Small Business Economics* 43 (4): 805–821.
Soete, L., B. Verspagen, and T. H. Ziesemer. 2022. "Economic Impact of Public R&D: An International Perspective." *Industrial and Corporate Change* 31 (1): 1–18.
Teruel, M., A. Coad, C. Domnick, F. Flachenecker, P. Harasztosi, M. L. Janiri, and R. Pal. 2021. "The Birth of New HGEs: Internationalization Through New Digital Technologies." *Journal of Technology Transfer*, doi:10.1007/s10961-021-09861-6.
Tether, B. S. 2005. "Do Services Innovate (Differently)? Insights from the European Innobarometer Survey." *Industry & Innovation* 12 (2): 153–184.
Veugelers, R., and M. Cincera. 2015. "The Impact of Horizon 2020 on Innovation in Europe." *Intereconomics - Forum* 50: 2–9.
Xifré, R. 2018. "Spanish Investment in R&D + I in the Wake of the Crisis: Public Versus Private Sector." *Spanish Economic and Financial Outlook (SEFO)* 7 (4): 67–79.
Zúñiga-Vicente, JÁ, C. Alonso-Borrego, F. J. Forcadell, and J. I. Galán. 2014. "Assessing the Effect of Public Subsidies on Firm R&D Investment: A Survey." *Journal of Economic Surveys* 28 (1): 36–67.