

Related diversification and R&D intensity dynamics

César Alonso-Borrego^{a,1}, Francisco Javier Forcadell^{b,*} 

^a Universidad Carlos III de Madrid, Department of Economics, 28903 Getafe, Madrid, Spain

^b Universidad Rey Juan Carlos, Department of Management, Paseo de los Artilleros s/n, 28032 Madrid, Spain

A B S T R A C T

Using longitudinal data of Spanish manufacturing companies, we study the dynamic, bidirectional relationship between firm research and development (R&D) intensity and corporate diversification in an organic growth setting. Our empirical approach accounts for the different sources of endogeneity. Although we find a positive linear effect of R&D intensity on related diversification, the effect of related diversification on R&D intensity assumes the form of an inverted U. Thus, the effect of related diversification on R&D intensity is positive, but marginally decreases for moderate levels of related diversification. Such an effect can become negative, however, for high levels of related diversification. Additionally, as a consequence of dynamics, the effects after one year are substantially lower than the overall effects that occur over several years.

Keywords:

Diversification
R&D intensity
Dynamics
Organic growth
Endogeneity
Panel data

1. Introduction

The relationship between R&D intensity and corporate diversification has attracted considerable attention in empirical research on strategic management over the last couple of decades (e.g., Chatterjee and Wernerfelt, 1991; Chen, 1996; Miller, 2006; Silverman, 1999). Although there is pervasive evidence for a linear and positive effect of related diversification on R&D intensity, empirical evidence for the effect of R&D intensity on diversification is mixed. Additionally, most empirical work has concentrated on unidirectional relationships, and evidence on the potential feedback between related diversification and R&D intensity is lacking.

We attempt to reconcile contradictory findings concerning the link between corporate diversification and R&D. We posit a dynamic bi-directional hypothesis between related diversification and R&D intensity, and we evaluate this relationship at the empirical level. The dynamic nature of such a relationship is sustained by the concepts of synergies and economies of scope. This bi-directional link emphasises the endogenous character of the relationship between corporate diversification and technological resources (Baldwin and Scott, 1987; Miller, 2004).

Most empirical studies of corporate diversification have focused on the experience of U.S.-based companies (Wan and Hoskisson, 2003). However, the institutional environment in which firms operate influences their dominant growth mode. Studies, in which both organic growth and external growth are considered, do not distinguish between the effects of these alternative growth modes. Thus, isolating the effect of organic growth by analysing a sample of companies that makes use of only this growth mode is worthwhile.

We evaluated the theoretical hypothesis using information supplied by the Survey of Business Strategies, a representative sample of Spanish manufacturing companies, between 1990 and 2001. The availability of longitudinal firm-level panel data permits us to consider the dynamic features of R&D intensity and diversification decisions as well as allows for the lagged effects distributed over time. This permits us to distinguish between the direct effect after one year and the full effect, which requires several years. To operationalise the simultaneous decisions regarding R&D intensity and diversification as well as the potential feedback between them, we estimate a bivariate vector auto-regression (VAR) for R&D intensity and related diversification, augmented by additional covariates. To increase the robustness of our results, we also controlled for two potential sources of endogeneity: censoring and unobserved firm heterogeneity. The failure to account for either of these two sources of endogeneity can lead to misleading conclusions.

Our empirical results provide evidence in favour of a bi-directional relationship between corporate diversification and R&D intensity at two levels. First, an increase in R&D intensity positively affects related diversification; innovation increases firm incentives to establish businesses in related activities. Second, related

* Corresponding author. Tel.: +34 914888049; fax: +34 914887786.

E-mail addresses: cesar.alonso@uc3m.es (C. Alonso-Borrego),

franciscojavier.forcadell@urjc.es (F.J. Forcadell).

¹ Tel.: +34 916249749; fax: +34 916249875.

diversification has a positive effect on R&D intensity, at least for moderate levels of related diversification. In this situation, firms undertaking related diversification may be more likely to innovate to consolidate their position in new activities and make them more efficient in implementing R&D expenditures. However, we find that the positive effect of related diversification on R&D intensity is marginally decreasing and may be offset when related diversification reaches a certain level. Our results suggest the effect of related diversification on R&D intensity exhibits an inverted-U shaped form.

2. Theory and hypothesis

The link between diversification strategy and R&D intensity has been empirically analysed in several studies, which are summarised in Table 1. Empirical studies have adopted a unidirectional approach, and therefore, provide conflicting evidence. The common finding is that diversification has a positive effect on R&D intensity. Among the few exceptions to this finding are the results reported by Hill and Hansen (1991), who concentrated on very particular industries, and those reported by Miller (2004), who focused on an external growth framework. Evidence is contradictory with regard to the effect of R&D intensity on diversification. Some authors find a positive effect (Davis and Thomas, 1993; Hoskisson and Johnson, 1992), while other authors find a negative effect (Hitt et al., 1996; Stimpert and Duhaime, 1997).

There are several possible reasons for these striking differences. First, the measures of these two strategic variables, diversification and R&D intensity, differ across studies. In particular, many studies do not distinguish between related and unrelated diversification, whose consequences can differ remarkably. Second, most studies do not distinguish between growth modes, yet the dominant growth mode can influence corporate decisions on both diversification and R&D intensity. Third, assumptions about the time schedule of when the effects occur may affect the results. Fourth, differences in the methodological approaches can affect the conclusions. The different approaches used in earlier studies, as shown in Table 1, had a high probability of affecting the empirical findings. Nonetheless, and more importantly, none of the previous studies analysed the bidirectional relationship of these two variables. Therefore, they ignored the potential feedback between these two strategic decisions. To further the research, we propose and test a new hypothesis that posits a dynamic bi-directional relationship between related diversification and R&D.

Economies of scope and synergies play a key role in diversifying firms, in general, and also in the particular case of the mutual relationship between related diversification and R&D. The analysis of economies of scope and synergies can be considered from the viewpoint of outputs and costs (see Tanriverdi and Venkatraman, 2005). From the production perspective, input complementarities may increase the value of a multiproduct firm in comparison to a single-product firm through super-additive value synergies among the firm's different businesses (Davis and Thomas, 1993). From the cost perspective, economies of scope may reduce the unit costs of a multiproduct firm in comparison to single-product firms by sharing resources between the firm's businesses (Teece, 1982) in the joint production process. Synergies and economies of scope are suitable for a firm that diversifies among related activities but are not relevant in cases of unrelated diversification.

From a dynamic resource-based view, a firm involved in several related businesses should efficiently use its current resources, particularly its technological resources, and generate sufficient resources to make future strategies viable. This implies a long-term dynamic interaction between related diversification and techno-

logical resources.² When a firm increases its degree of related diversification by entering in a new business, the firm takes advantage of its excess resources and acquires complementary resources needed to operate (Chatterjee, 1990). Research indicates the closer the relation with the previous business of the firm, the lower the costs of entering in a new business (Yip, 1982).

If the firm grows through related diversification, the new activities can take advantage of the existing technological resources, which will be exploited to a greater extent. This effect arises as a result of synergies and economies of scope (Teece, 1982) derived from the use of the firm R&D in their different businesses. An increase in R&D investment may boost related diversification, as long as it improves the capacity to exploit the available technological resources. Further, the firm technology affects the firm's diversification strategy (Silverman, 1999) because the greater the R&D investment, the greater the related diversification (Burgelman, 1983).

In addition, the specificity of R&D can influence the firm strategy on related diversification. Such specificity, which can be due to (path-dependent) learning inside the firm and firm-specific capabilities, may generate appropriability, both through protection from imitation and through the use of complementary resources (Helfat, 1994; Teece, 1986). Therefore, highly specific R&D resources can provide the firm further incentives to related diversification, to improve appropriability.

Related diversification, as part of a long-term corporate strategy, tends to increase the expected return from R&D in two instances: the greater the diversification, the better the capacity to use research outputs (Teece, 1980). Consequently, a higher degree of related diversification can favour R&D investment and technology adoption (Hill and Snell, 1988; Chen, 1996). However, a firm entering or strengthening a given business area might undertake technology investments to improve its competitive position in that business (Itami and Numagami, 1992; Lunn and Martin, 1986; Scherer, 1984). Therefore, a firm introducing a certain degree of related diversification may have incentives to increase its R&D.

There are some forces that may induce an opposite effect of related diversification on R&D effort. These forces have to do with the efficiency gains in R&D exploitation associated with (related) diversification strategies (Baysinger & Hoskisson, 1989). From the resource-based view, the optimal amount of additional resources (particularly, technological resources) tends to decrease with the number of related businesses. Technological knowledge is not limited to a particular business but extended to related businesses (Helfat and Raubitschek, 2000). Thus, the complementarity of technological resources in related activities (Helfat, 1997) can eventually reduce the optimal R&D intensity. In other words, the same R&D investment has a greater impact the higher the diversification, due to synergies and economies of scope (Baysinger and Hoskisson, 1989; Miller, 2004), so the necessity to increase R&D is reduced.

As a consequence, related diversification may produce counteracting effects on R&D. We have argued that related diversification boosts innovation, thus increasing R&D investment. However, if related diversification increases the efficiency in the use of technological resources, the need of further R&D effort may be moderated.³ These two conflicting forces suggest that the effect of related diversification on R&D behaves as an inverted U-shaped

² One explanation of Penrose (1995) for the process of related diversification over time refers to efficiency gains from learning that result in excess resources, which can be used to diversify.

³ Although R&D expenditure keeps increasing, it could be the case that R&D intensity, defined as the ratio of firm R&D expenditure to total sales, falls. This phenomenon is due to the relative increase in sales, associated with the firm additional related businesses, which is larger than the increase in R&D expenditure because of the aforementioned economies of scope.

Table 1
The unidirectional relationship between R&D and corporate diversification. Empirical evidence.

Type of relationship	Study	Diversification type	Sign	Growth type	Sample	Comments
R&D/Diversification Association	Bettis (1981)	Related	+	No distinction	USA	Related-constrained and Related-linked categories
	Bettis and Mahajan (1985)	Related	+	No distinction	USA	Related-constrained and Related-linked categories
	Lecraw (1984)	Related	+	No distinction	Canada	Related-constrained and Related-linked categories
R&D → Diversification	Chatterjee and Singh (1999)	Related	+	Organic, External	USA	No influence of R&D on mode of growth
	Chatterjee and Wernerfelt (1991)	Related	+	No distinction	USA	
	Delios and Beamish (1999)	Related	+	No distinction	Japan	
	Hill and Hansen (1991)	Related Unrelated	–	No distinction	USA	Pharmaceutical industry. Diversification and R&D compete on the funds
	Miller (2004)	Related Unrelated	–	External, n.s. for organic	USA	Positive relation between related diversification and organic growth
	Miller (2006)	Related	+	No distinction	USA	Negative effect after controlling for firms having no R&D or in R&D intensive industries
	Montgomery and Hariharan (1991)	Related	+	No distinction	USA	Computer and PC industries Detailed level of technological resources analysis
	Schoenecker and Cooper (1998)	Unrelated	–	No distinction	USA	
Silverman (1999)	Related	+	No distinction	USA		
Diversification → R&D	Baysinger and Hoskisson (1989)	Related Unrelated	n.s. –	No distinction	USA	Pharmaceutical industry. Modified version of the concentric index to estimate synergy
	Baysinger et al. (1991)	Related Unrelated	n.s. n.s.	No distinction	USA	
	Davis and Thomas (1993)	Related	+	No distinction	USA	R&D per employee
	Hill and Snell (1989)	Unrelated	n.s.	No distinction	USA	
	Hitt et al. (1991)	Unrelated	–	External	USA	
	Hitt et al. (1996)	No distinction	–	External	USA	Diversification as control variable
	Hitt et al. (1997)	No distinction	–	No distinction	USA	
	Hoskisson and Hitt (1988)	Related Unrelated	n.s. n.s.	No distinction		Dominant business firms have a higher relative R&D
	Hoskisson and Johnson (1992)	Related Unrelated	+	External	USA	Restructuring effect on diversification strategy and R&D
	Stimpert and Duhaime (1997)	Related (4)	–	No distinction	USA	

function. Under such a scheme, the effect of related diversification on R&D could be positive, but marginally decreasing and it may eventually become negative for high levels of related diversification. Such non-linearity, which means a more general framework, may help to reconcile the empirical evidence of earlier research on this topic (see Table 1).

Two additional elements of the relationship between R&D intensity and corporate diversification must still be considered: the growth mode and the time schedule. Firms can grow through different modes: internal or organic, external, and cooperative (Tsang, 2000). Earlier empirical studies on this topic did not distinguish between external and organic growth. In particular, the link between related diversification and R&D has never been analysed in an organic growth setting. The lack of control for the growth mode may lead to ambiguous conclusions about this link. The growth mode may affect the extent of diversification and its relation with R&D intensity. In particular, whereas internal or organic growth usually encourages related diversification (Chatterjee and Singh, 1999) and vice-versa, external growth is usually associated with unrelated diversification (Chang and Singh, 1999). Furthermore, under the transaction costs economics rationale, the possession of a highly specific set of resources (e.g., technological resources) is associated with organic growth.

The dominant growth mode also depends on the transaction costs associated with different governance structures (Williamson, 1975). Governance structures are influenced by firm characteristics and, very specifically, by the institutional environment in which firms are involved (Hoskisson et al., 2004). Unlike countries subject to English common law, nations ruled by French civil law, such as Spain, suffer from relatively poor investor and creditor protection and weaker law enforcement (La Porta et al., 1997). Thus, capital markets in these latter countries are typically less developed and narrower (Canals, 1997); the median value of the annual merger and acquisitions (M&A) activity in the US during the nineties was twice that of the M&A activity in Europe during the same time period.⁴ The legal system may then determine the prevailing growth mode in each country. Mayer and Whittington (2003) advocate the consideration of alternative environments to study the topic of diversification. Clearly, the choices of the companies under study impact the empirical findings. Most empirical studies have been focused on US companies, where external growth is the dominant growth mode. Companies located in most European countries are more prone to organic growth (Capaldo et al., 2009).

⁴ Calculation based on data from the Thomson Financial Securities Data Corporation.

Our approach acknowledges the dynamic nature of the relationship between corporate R&D intensity and related diversification. These two strategic decisions have an inter-temporal nature because their consequences have a lasting influence on the organisation and performance of the firm. The inter-temporal framework, whereby the firm takes into account the effect of both R&D intensity and related diversification on current and future performance when making its strategic decisions, has been acknowledged by, among others, Helfat and Eisenhardt (2004), in the context of related diversification decisions. Such decisions are constrained by the fact that changes in strategic variables entail costs to the firm that affect the time period in which the firm will adjust to the new target levels of such variables.⁵ Learning effects and organisational issues (Grant and Jammine, 1988; Bergh, 1995) are among the primary causes of such costs. In the economics literature, such costs are labelled “adjustment costs” (Helfat and Eisenhardt, 2004).⁶ The major consequence of adjustment costs is that the behaviour of strategic variables exhibits substantial inertia, which is reflected by the fact that the current level of the strategic variable is strongly affected by its past value. In this kind of context, the effect of any variable on the level of the strategic variable is not instantaneous but is distributed over several years. Consequently, the instantaneous effect of a variable change, or even the effect after one year, may be substantially lower than the overall effect.

2.1. Hypothesis

Hypothesis 1. In a context of organic growth, R&D intensity and related diversification show a bidirectional and dynamic relationship. Therefore, the overall effect of a variable change takes several years to be completed.

Hypothesis 1a. R&D intensity has a linear and positive influence on the level of related diversification.

Hypothesis 1b. Related diversification influences R&D intensity through an inverted-U shaped function. Hence, although the effect of related diversification on R&D intensity is initially positive, it marginally decreases and may become negative when there is a sufficiently high level of related diversification.

3. Econometric issues

To test the hypothesis, we must first characterise a dynamic model for R&D and diversification decisions that acknowledges the simultaneity of both decisions. Our concern for the importance of dynamic feedback has been expressed in earlier work. The lack of accounting for such feedback effects (because of the use of a single cross-section) results in the failure to establish a clear relationship between diversification and R&D. In his review of empirical work, Bergh (1995) asserted that change over time has not been accounted for empirically. Hill and Hansen (1991) contended that the underlying causal relationship cannot be determined by cross-sectional studies.

We allow for the potential feedback between R&D intensity and diversification decisions through a dynamic model for the joint decisions of both variables, conditional on further covariates.

The availability of longitudinal data allows us to characterise the potential dynamic effects. In particular, our model consists of a bivariate-augmented VAR model (see Holtz-Eakin et al., 1988) for R&D intensity and related diversification, where each of the two equations contains both lagged measures of the two variables.

Formally, denoting related diversification and R&D effort as d and r , respectively, X as the vector of further covariates, and using i to index companies and t to denote time periods, the equation for each of the two strategic decisions can be written as follows:

$$\begin{aligned} d_{it} &= \alpha_{d1} + \rho_{d1}d_{i,t-1} + \rho_{d2}d_{i,t-2} + \delta_{d1}r_{i,t-1} + \delta_{d2}r_{i,t-2} + X'_{it}\beta_d + u_{dit} \\ r_{it} &= \alpha_{r1} + \rho_{r1}r_{i,t-1} + \rho_{r2}r_{i,t-2} + \delta_{r1}d_{i,t-1} + \delta_{r2}d_{i,t-2} + \pi_{r1}d_{i,t-1}^2 \\ &\quad + \pi_{r2}d_{i,t-2}^2 + X'_{it}\beta_r + u_{rit} \end{aligned} \quad (1)$$

We have used the sub indices d and r to distinguish the coefficients in the corresponding equations for diversification and R&D effort; u_{dit} and u_{rit} are the corresponding error terms of the diversification and R&D equations. This dynamic structure will have two major consequences. First, the interrelation between the two decision variables may follow a distributed lag structure, and therefore the whole interrelation effect may take some time. Second, the total or full effect of a change in any of the right-hand-side variables will take several years to be completed, due to the presence of lagged values of the dependent variable on the right-hand side of each equation.

We took advantage of the empirical work in this area from Merino and Rodríguez (1997), Gedajlovic and Shapiro (2002), and Miller (2006). Merino and Rodríguez (1997) studied the discrete decision of whether or not to diversify, and these authors were particularly concerned with the inconsistency bias of the estimates due to the disregard of the unobserved firm heterogeneity. They propose a proper econometric treatment that requires panel data. Gedajlovic and Shapiro (2002) studied the role of agency problems in profitability, taking into account dynamic effects, and Miller (2006) reported a positive relationship between technological diversification and performance based on a random effects model.

The variables that we are considering are strategic; managers base their decisions on expectations about future firm performance (Hamilton and Nickerson, 2003) so that managerial decisions are endogenous. There are two potential sources of endogeneity, unobserved heterogeneity and censoring, which must be considered to obtain appropriate estimates that guarantee robust results.

The problem of unobserved heterogeneity arises from the existence of relevant unobserved firm characteristics that are potentially correlated with the existing covariates. Under the assumption that firm-specific unobserved characteristics are invariant over time, the availability of longitudinal or panel data may yield consistent estimates by means of a first-differences transformation that removes the unobserved heterogeneity term, whereas the parameters of interest remain unchanged after such a transformation. The fact that there are endogenous variables among the covariates (the lagged dependent variable, among others) requires an instrumental variable procedure. The obvious instruments are the lagged values of the covariates, which are uncorrelated with the error term.⁷ We can therefore apply a

⁵ In particular, if the optimal targets among the strategic variables imply large changes, it may be more profitable to distribute such changes over a longer time period until reaching the target level, instead of making the full change instantaneously, which would involve higher costs to the firm.

⁶ Empirically, this is reflected by the fact that lagged values of the strategic variables will appear among the covariates of the equations characterizing such a strategic decision. As a consequence, the immediate effect of a change in any covariate will differ from the full effect (the total accumulated effect).

⁷ Valid instruments should be uncorrelated with the error term of the first-differences transformation, and correlated with the explanatory endogenous variables. In a framework of dynamic decisions of firms, any firm-level variable associated with firm decisions or performance is hardly uncorrelated with the error term (which reflects, among other things, unobserved firm-specific shocks affecting the corresponding decision variables, either related diversification or R&D effort). This usually invalidates contemporaneous values of firm-level as instruments. Anderson and Hsiao (1982) and Arellano and Bond (1991), among others, prove that with uncorrelated errors or moving-average errors, lagged values of the explanatory

generalised method of moments (GMM) estimator (see Hansen, 1982; Arellano and Bond, 1991). In particular, our estimation approach consists of the system-GMM estimator proposed by Arellano and Bover (1995) and developed by Blundell and Bond (1998).

The second source of endogeneity arises because each of the dependent variables (R&D intensity and diversification) is censored. If we estimate each equation, restricting to the subsample of observations those in which the dependent variable is strictly positive at two consecutive periods, t and $t-1$, a sample selection bias may be introduced in our estimates. The proper treatment requires characterisation of the firm's discrete decision on whether or not to innovate (diversify) as well as the firm's decision regarding the amount of R&D intensity (diversification).

The censoring problem arises because the decision variables can be positive, but a large proportion of the observations will be equal to zero (especially in the case of diversification). Therefore, the observed decision variables have two components: a qualitative component that corresponds to the strictly positive and zero-valued observations, and a continuous component that describes the range of positive values. We can write the equation of interest, transformed in first differences so as to remove the time-invariant firm-specific effect, as

$$\Delta y_{jit}^* = \Delta W_{jit}' \theta_j + \Delta u_{jit} \quad (j = d, r) \quad (2)$$

where y_{jit}^* ($j = d, r$) denotes the dependent variable of interest (i.e., either d_{it} or r_{it} , respectively) in the absence of censoring, W_{jit} is the vector of covariates (which includes the lags of the two strategic variables), and θ_j is its corresponding vector of unknown coefficients that represent the corresponding effects of the covariates. The dependent variable that we observe can nevertheless be described as

$$y_{jit} = \begin{cases} y_{jit}^* & \text{if } I_{jit}^* > 0 \\ 0 & \text{if } I_{jit}^* \leq 0 \end{cases} \quad (j = d, r) \quad (3)$$

where I_{jit}^* determines the censoring (i.e., whether or not the firm does innovate/diversify), in accordance with the auxiliary equation

$$I_{jit}^* = Z_{jit}' \gamma_j + v_{jit} \quad (j = d, r). \quad (4)$$

In the case of the R&D intensity (diversification) decision, I_{jit}^* denotes the marginal net revenue of such a decision so that we will observe a firm innovating (diversifying) if the marginal net revenue of R&D intensity (diversification) is positive; otherwise, the firm will not innovate (diversify). In principle, both the right-hand-side variables (and thus the censoring mechanism) and the magnitudes of their effects need not coincide in the equation of interest that determines the positive amount of R&D (diversification) or in the auxiliary equation describing the marginal net revenue of the qualitative decision on whether or not to innovate (diversify).

The parameters of interest can be consistently estimated through the subsample of positive observations using a two-stage procedure (Heckman, 1978). When conditioning on the subsample of observations for which y_{jit} is strictly positive in two consecutive periods (i.e., conditioning on the event that $I_{jit}^*, I_{jit,t-1}^* > 0$), our

equation of interest can now be written as

$$\Delta y_{jit}^* = \Delta W_{jit}' \theta_j + \varphi_j \lambda_{jit} + \Delta v_{jit} \quad (j = d, r), \quad (5)$$

endogenous variables become valid instrument in the first-differences transformation of the original equations.

where $\varphi_j \lambda_{jit}$ is a selectivity correction term, which accounts for the sample selection bias that arise when conditioning on the subsample of positive observations (see Amemiya, 1984).⁸

If λ_{jit} were observed, we can introduce it as an additional variable in the equation of interest. However, λ_{jit} is unobserved but can be estimated for each non-censored observation through the auxiliary equation about the qualitative decision on whether or not to diversify (innovate), using a binary choice model. Afterwards, the predictions of the selectivity correction term, λ_{jit} , can be introduced in the corresponding equation of interest (5) as an additional regressor. Estimating without taking proper account of the possible selection bias can lead to inconsistent estimates if the selection is endogenous, if $\varphi_j \neq 0$.

Once a firm has decided to diversify or to spend in R&D, however, these decisions appear to be strongly entrenched. The transition probabilities for the events of diversifying or R&D spending are, respectively, 87% and 91% in our sample, indicating that, once we condition on non-censored observations, the selection bias is mostly time-invariant, so that it can be subsumed in the fixed effect, which disappears when we apply the first differences transformation. We would thus expect these estimates to be similar to those that use a selectivity correction term as an additional regressor.⁹

The proper econometric treatment of these problems will generally produce larger standard errors of the estimates, and therefore lower precision in comparison with other approaches that disregard these sources of endogeneity. Ignoring these problems, however, would lead to inconsistent, and therefore, meaningless estimates, regardless of their level of precision.

4. Data and variables

4.1. Data

The data set is an unbalanced panel of Spanish manufacturing firms, recorded in the database *Encuesta Sobre Estrategias Empresariales* (Survey on Business Strategies, ESEE hereinafter).¹⁰

⁸ Formally, denoting $E(\cdot)$ as the expectation operator conditional on the past values of R&D and related diversification and on the past and current values of the exogenous variables, we have that $E(\Delta u_{jit}) = 0$ for the whole sample, which is the main conditions for consistency of the OLS estimator. But if we condition on the subsample of observations for which y_{jit} is strictly positive in two consecutive periods (i.e., on the event that $I_{jit}^*, I_{jit,t-1}^* > 0$), then in general $E(\Delta u_{jit} | I_{jit}^*, I_{jit,t-1}^* > 0) = \varphi_j \lambda_{jit} \neq 0$, since the unobserved error terms u_{dit} and u_{rit} are not independent of the discrete choices on whether or not to diversify or innovate, respectively. Assuming normality, φ_j is the covariance between the error term in the equation of interest u_{jit} and the error term in the auxiliary equation for the decision to diversify or innovate, respectively, in two consecutive periods. Also, λ_{jit} , under normality, is given by the inverse Mills' ratio (see Heckman, 1978), based on the binary choice model for the decision to diversify or innovate in two consecutive periods.

⁹ An alternative two-stage approach, based on Bover and Arellano (1997) and Arellano et al. (1999), closely follows the Heckman two-stage procedure. Unlike the auxiliary probit equation for no censoring, we could consider the estimation of auxiliary equations for the censored variable y_{jit} and obtain predictions of the latent variable y_{jit}^* for the censored observations. Then, we could estimate the model with the full sample as if censoring had not taken place. We have also considered this alternative approach (yet the estimates are not reported here). Given that the proportion of censored observations is very large, however, particularly for related diversification, our predictions for y_{jit}^* were very imprecise and extremely sensitive to the choice of variables in the auxiliary equation.

¹⁰ The ESEE is produced by *Fundación Empresa Pública*, a public institute financed by the Spanish Ministry of Industry. The original dataset was designed with the aim of ensuring the representativeness of Spanish manufacturing firms. For this purpose, all companies with more than 200 employees were surveyed (and approximately 70% completed the survey), and smaller companies with more than 10 employees were selected on the basis of a stratified sampling. It is important to note that the unbalanced feature of the panel implies that the time period during which we observed the different companies was unequal, reflecting the fact that such companies may enter and exit from the survey (in the same way that companies appear

Table 2
Correlation matrix.

	R&D	Unrelated div.	Related div.	Firm size	Firm age (years)	Advertising exp.
Unrelated div.	0.1311 [§]					
Related div.	0.0616 [†]	-0.0086				
Firm size	0.2969 [§]	0.0241	-0.0266			
Firm age (years)	-0.0025	0.0124	0.0892 [§]	0.1331 [†]		
Advertising exp.	0.0085	-0.0756 [†]	-0.0219	-0.0532	0.2426 [§]	
ln(Sales)	0.0803 [§]	-0.0033	0.0225	0.5877 [§]	0.2251 [§]	0.2158 [§]

[†] Significance at the 10 percent level.

[†] Significance at the 5 percent level.

[§] Significance at the 1 percent level.

This database was previously used by Merino and Rodríguez (1997) and others in the analysis of corporate diversification. It contains annual information for a large number of Spanish companies whose main activity was manufacturing between 1990 and 2001.

Our final sample consisted of 513 non-energy manufacturing companies, with at least 25 employees, operating in at least two markets, and whose geographic scope was national or wider.¹¹ We also required that the company's situation was not substantially altered during the sample period. Therefore, we discarded mergers or splits, as well as companies whose main activity changed during the sample period (defined at the two-digit industry level). Based on these criteria, less than 4% of the observations were removed, including companies that were generally older and larger than the companies that remained in the sample, which carried out only organic growth strategies during the period analysed. Table A.1 shows the sample distribution of companies by two-digit industry and by size.

5. Variables and measures

We defined R&D intensity as R&D expenditure based on a proportion of a firm's sales, consistent with earlier related work, such as Baysinger and Hoskisson (1989) and Hitt et al. (1997), among others. We used the entropy index as a diversification measure (Jacquemin and Berry, 1979; Palepu, 1985), which is widely used in strategy research. The validity of this measure was analysed in several studies (Robins and Wiersema, 1995, 2003). This measure is defined as

$$d = \sum_s P_s \times \ln \left(\frac{1}{P_s} \right),$$

where the sub index s represents the s th product segment. The entropy index is thus a weighted average of the sales attributed to each firm's segment, P_s . The weights are the natural logarithm of the inverse of the segment's sales, $\ln(1/P_s)$. This entropy index is additive, so that it can be broken down as the sum of the indices of related and unrelated product diversification, dR and dU . The first index, dR , captures the diversification across four-digit SIC industries within each two-digit industry (with two-digit firm's sales as the reference). The second index, dU , captures diversification across two-digit industries (with overall firm's sales as the reference). For the diversification equation, we focus on related product diversification as the dependent variable.

and disappear in the economy). Restricting the sample of companies to be observed to the same time period would affect the randomness of the sample. Instead, to obtain sufficient information concerning firm dynamics, we only required sample companies to have full information available on all the relevant variables for at least five consecutive years between 1990 and 2001.

¹¹ The rationale for these last two restrictions is that, in general, the potential role of diversification (and probably innovation) is quite small in the case of companies focused on local markets and/or those that are relatively small.

We considered several control variables. First, we included the natural logarithm of firm employment to control for its size. The influence of firm size on diversification has been supported in several studies, such as Bettis (1981), Hill and Snell (1988), or Hoskisson et al. (1993), which reported a significant correlation. In the case of R&D intensity, however, the influence of firm size on R&D intensity remains controversial (King et al., 2003: 591).

Furthermore, in the level estimations, we included industry dummies (at the level of the firm's two-digit SIC major industry) that interacted with time dummies to control for heterogeneity in the firm environment, which is associated with the operating industry and can vary over time (see Bergh, 1995; McGahan and Porter, 1997, among others). For example, the aggregate level of industry R&D may affect both the relative R&D intensity at the firm level and their propensity to introduce new products (Baysinger and Hoskisson, 1989).

We also included age dummies, for which we chose the following age groups: less than 10 years, between 10 and 20 years, between 20 and 40 years, and older than 40 years. Age dummies capture the potential effect of the maturity of the firm, and therefore the accumulated know-how of the firm, on R&D intensity and diversification decisions. The direct use of age is precluded due to the unreliability of such a measure for older firms (Alonso-Borrego and Collado, 2002). As additional firm-level variables, we included sales (in logarithms) and advertising intensity (ratio of advertising expenditure to sales).¹²

The simple correlations among the main variables are reported in Table 2. The positive correlation between R&D intensity and related diversification is noteworthy. In addition, firm size appears positively correlated with R&D intensity but not with related diversification. Finally, age appears to be correlated with diversification but not with R&D intensity. These findings suggest that the maturity of the firm, rather than its size, is relevant to diversification. The opposite is true for R&D intensity. Nevertheless, the correlation analysis disregards the censored nature of the decision variables as well as the interaction between the potential explanatory variables.

In the upper part of Table 3, we report the means and standard deviations of the variables of interest for the full sample. Given the censored nature of our decision variables, however, the statistics were broken down depending on whether both R&D intensity and related diversification were zero or positive. We observed that zero R&D intensity and zero diversification are more likely when the firm is smaller and/or younger. There appears to be no link, however, between censoring in any of the two decision variables and sales. Furthermore, advertising expenditure seems to be higher for

¹² We also included firm characteristics and other variables. As firm characteristics, we considered whether the firm has industrial establishments abroad (to approximate international diversification) or belongs to an enterprise group. As strategic variables, we considered the firm market share in its main market (as a measure of its market power) as well as measures of firm liquidity and leverage. These latter variables, however, were not statistically relevant in our analysis. Therefore, the results obtained using these latter variables are not reported.

Table 3
Descriptive statistics: means and standard deviations of the relevant variables.

	R&D	Unrelated div.	Related div.	Firm size	Firm age (years)	Advertising exp.	ln(Sales)
<i>Full sample</i>	0.0130 (0.0263)	0.0518 (0.1601)	0.0569 (0.1711)	572.3 (1234.6)	32.96 (26.70)	2.28 (3.98)	15.13 (1.35)
<i>Subsamples</i>							
Related diversification = 0							
R&D = 0	0	0.0358 (0.1382)	0	255.6 (625.9)	29.11 (25.00)	1.70 (3.07)	14.41 (1.24)
R&D > 0	0.0190 (0.0301)	0.0554 (0.1652)	0	652.2 (1254.4)	33.46 (27.77)	2.66 (4.52)	15.46 (1.27)
All	0.0124 (0.0260)	0.0486 (0.1565)	0	537.1 (1122.1)	31.94 (26.90)	2.33 (4.09)	15.09 (1.36)
Related diversification > 0							
R&D = 0	0	0.0613 (0.1708)	0.3883 (0.2531)	344.4 (362.6)	37.12 (22.89)	1.14 (1.43)	15.30 (1.22)
R&D > 0	0.0231 (0.0308)	0.0793 (0.1874)	0.4939 (0.2038)	1024.5 (2197.9)	41.31 (24.41)	2.26 (3.31)	15.42 (1.29)
All	0.0173 (0.0285)	0.0748 (0.1829)	0.4675 (0.2210)	834.1 (1884.6)	40.26 (24.02)	1.98 (2.99)	15.39 (1.27)

Standard deviations in parentheses below.

innovating firms (i.e., those with positive R&D intensity), but average advertising expenditure is not significantly different between diversifying and non-diversifying firms. Finally, the probability of positive R&D intensity is higher when diversification is greater, and vice versa.

6. Results

To illustrate the aforementioned consequences of ignoring the potential endogeneity problems, we report alternative estimates based on different statistical assumptions. Table 4 shows the preliminary estimates with alternative treatments of censoring, whose consistency depends on the absence of firm-specific unobserved heterogeneity. Our preferred estimates, which take into account all the potential sources of endogeneity, are reported in Table 5. The alternative treatments of censoring in Table 4 range from the ordinary least squares (OLS) and truncated OLS estimates to the tobit and generalised selection models.¹³ In all these specifications, we previously removed any additional covariate that was statistically non-significant.

Not surprisingly, the coefficient of the lagged dependent variable was large and strongly significant in both equations, irrespective of the estimation method.¹⁴ We also observed that the alternative treatments of censoring induced striking differences in the estimated effects. In particular, the substantial differences between the tobit and generalised selection reveal the rejection of the tobit constraints, so that the factors determining the discrete decisions on whether to innovate or diversify seem to differ from the factors behind the magnitude of R&D intensity or diversification.¹⁵

¹³ Notice that OLS estimates with the full sample will be inconsistent due to the fact that a significant fraction of the observations for the dependent variables is censored at zero. By the same token, OLS estimates with the truncated sample of strictly positive observations (that is, uncensored observations) will also be inconsistent, as the sample selection bias, after excluding censored observations, is not taken into account.

¹⁴ When unobserved heterogeneity is ignored, the absolute value of the estimated coefficient of the lagged dependent variable is typically oversized.

¹⁵ Regarding the diversification equation, the likelihood of the tobit model did not converge so that their associated estimates are actually unreliable. This fact also provides evidence against the tobit constraints and supports the generalised

selection model where the sample selection term, given by the inverse of Mills' ratio, appears to be significant.

Table 5 shows the estimates based on a dynamic panel data estimation that explicitly take into account unobserved heterogeneity, together with the censoring problem of the dependent variables. We also provide standard errors appropriately corrected from potential finite-sample bias, as proposed by Windmeijer (2004). In both equations, we introduced the explanatory variables lagged one period, except for our two strategic variables, which we introduced the first two lags. For each equation, the first two columns show the unrestricted estimates (with and without the selectivity correction term).

Our preferred estimates, for which we removed the lags that were clearly non-significant, are shown in the last two columns.¹⁶ Although most qualitative findings from the earlier estimates in Table 4 remained, the magnitudes of the estimated coefficients were remarkably different. In particular, the estimated coefficients corresponding to the lagged dependent variable in both equations were substantially smaller, in accordance with the existence of unobserved firm heterogeneity. Therefore, when ignoring unobserved heterogeneity, we would be interpreting the spurious correlation between unobserved firm effects and the covariates as further persistence in the dependent variable.

In any case, the significance of the lagged dependent variables underlines the inertia, which we can attribute to the existence of significant learning effects, behind the decisions on R&D intensity and related diversification. Both in the R&D and the diversification equations, the selectivity correction term was significant at usual levels. In both cases, however, the results were very similar to the results of equations that exclude this term, and therefore, the assumption that sample selection bias is approximately time-invariant and can be subsumed into the firm-specific effect seems appropriate. We concentrate our discussion on the corresponding last columns for each equation.

In the diversification equation, the lagged endogenous variable at $t - 1$ and $t - 2$ was significant, and the coefficients decrease with the time lag. The large estimated coefficients show the existence of

¹⁶ Any variable kept constant over time after applying the fixed effects of transformation was not considered, because the fixed-effects of transformation would eliminate it. In particular, age effects were not taken into account.

Table 4
Preliminary estimates.

Variable	Diversification equation				R&D equation			
	OLS	Truncated OLS	Tobit	Generalised selection	OLS	Truncated OLS	Tobit	Generalised selection
R&D ($t-1$)	0.1504 [†] (0.0590)	0.3666 (0.2980)	1.3016 [§] (0.0000)	0.5756 [†] (0.2704)	0.5814 [§] (0.0890)	0.5593 [§] (0.0182)	0.6632 [§] (0.0181)	0.5669 [§] (0.0178)
Unrelated diversification ($t-1$)	0.0073 (0.0091)	-0.0118 (0.0548)	0.0989 [§] (0.0000)	0.0201 (0.0493)	0.0098 (0.0089)	0.0024 (0.0138)	0.0251 [†] (0.0122)	0.0040 (0.0132)
Squared Unrelated diversification ($t-1$)					-0.0175 (0.0145)	-0.0047 (0.0227)	-0.0380 [†] (0.0198)	-0.0071 (0.0218)
Related diversification ($t-1$)	0.8658 [§] (0.0082)	0.5488 [§] (0.0395)	1.7956 [§] (0.0000)	0.8764 [§] (0.0573)	0.0119 [†] (0.0060)	0.0184 [†] (0.0096)	0.0158 [†] (0.0083)	0.0187 [†] (0.0092)
Squared Related diversification ($t-1$)					-0.0104 (0.0082)	-0.0182 (0.0141)	-0.0120 (0.0118)	-0.0185 (0.0135)
Size	-0.0006 (0.0032)	-0.0065 (0.0251)	0.0104 (0.0000)	-0.0052 (0.0225)	0.0031 [§] (0.0008)	0.0043 [§] (0.0012)	0.0049 [§] (0.0011)	0.0045 [§] (0.0012)
Young firm (Y/n)	0.0023 (0.0047)	-0.0419 (0.0366)	-0.0315 [§] (0.0000)	-0.0273 (0.0327)	0.0011 (0.0011)	0.0027 (0.0018)	0.0008 (0.0016)	0.0025 (0.0017)
Advertising exp. ($t-1$)	-0.0002 (0.0004)	0.0022 (0.0031)	-0.0015 (0.0000)	0.0007 (0.0028)	0.0002 [§] (0.0001)	0.0002 (0.0002)	0.0004 [§] (0.0001)	0.0002 [†] (0.0001)
ln(Sales) ($t-1$)	0.0014 (0.0026)	0.0045 (0.0211)	0.0228 [§] (0.0000)	0.0135 (0.0188)	-0.0018 [§] (0.0006)	-0.0040 [§] (0.0010)	-0.0002 (0.0009)	-0.0037 [§] (0.0009)
Lambda				0.1498 [§] (0.0173)				0.0032 [†] (0.0014)
Number of observations	3508	375	3508	3508	3508	2248	3508	3508
R-squared	0.79	0.75			0.51	0.51		
Chi ²				1084.49				2552.45
Log-likelihood	3933.38	280.28	-343.13		8972.07	5375.46	4478.18	

The estimation method is indicated in the lower upper part of each column. None of the estimates reported in this Table are consistent in the presence of unobserved heterogeneity. The lag ($t-1$) is indicated to the left of each variable. Standard errors are in parentheses below each estimated coefficient. The variable *lambda* denotes the selectivity correction term in the generalised selection model. Time dummies, industry dummies and interactions between them are included in all equations. Chi² (whenever applicable) is a test of joint significance of all the covariates (excluding the constant). The joint significance test evaluates the hypothesis that the coefficients of related diversification and squared related diversification are jointly equal to zero.

^{*} Significance at the 10 percent level.

[†] Significance at the 5 percent level.

[§] Significance at the 1 percent level.

substantial inertia in related diversification. Neither the AR(2) nor the Hansen-Sargan test provided evidence against the specification. In addition, size (measured by the logarithm of employment) had a slightly negative effect, whereas advertising and sales were not significant. Interestingly, we found a one-to-one immediate effect of R&D intensity on related diversification so that an increase in R&D intensity by one percentage point implied a 0.01 increase in related diversification. This result clearly supports Hypothesis 1a, which establishes that R&D intensity positively influences related diversification.

In the R&D intensity equation, the lagged endogenous variable at $t-1$ and $t-2$ is significant, demonstrating that R&D dynamics are quite rich. In addition, related diversification has a positive effect on R&D, significant at the 10% level. Additionally, the estimated coefficient of square related diversification has the opposite sign, indicating that related diversification stimulates R&D at moderate levels of diversification, yet a positive effect decreases as the level of related diversification increases and may eventually become negative. This coefficient was also significant at the 10% level. This result supports Hypothesis 1b, whereby the influence of related diversification on R&D intensity assumes an inverted U-shaped function. We also implemented a test of joint significance of related diversification and its square, obtaining a p -value slightly below 20%, so that evidence was not fully conclusive. We attribute this result to the fact that our estimation approach provides robust estimates at the

expense of a larger variance-covariance matrix of the estimated coefficients.¹⁷

Regarding unrelated diversification, we found no significant effect on R&D intensity. Other additional variables, such as size, advertising, and sales, were not significant at any reasonable level. The AR(2) specification test for the existence of an error autocorrelation in the untransformed model did not provide evidence against the specification and the p -value was clearly above 20%. The Hansen-Sargan test, however, was not so conclusive.

It is worth noting that the two coefficients on diversification and squared diversification cannot be interpreted separately, as the effect of related diversification on R&D depends critically on the level of diversification. From the estimated numerical values of the coefficients of related diversification and squared related diversification, we computed the level after which additional increments in related diversification implied a negative effect of this variable on R&D intensity. This level is approximately 0.30, which seems quite high when compared with the full sample, for which the average value of related diversification is below 0.06. For those firms already diversifying, however, the related diversification

¹⁷ There are two potential causes of such a loss of precision: the transformation needed to remove the unobserved individual heterogeneity component and introduction of an estimated selectivity correction term rather than an actual (unknown) selectivity correction term.

Table 5
Dynamic panel data estimation.

Variable	Diversification equation				R&D equation			
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
R&D ($t-1$)	1.4252 [§] (0.5522)	1.3300 [§] (0.0536)	0.8116 [*] (0.4605)	1.0105 [§] (0.3939)	0.4058 [†] (0.2070)	0.4015 [†] (0.2029)	0.3936 [*] (0.2338)	0.3871 [*] (0.2311)
R&D ($t-2$)	-0.5347 (0.4184)	-0.4209 (0.3332)			0.3264 [†] (0.1666)	0.3241 [†] (0.1630)	0.3434 [*] (0.1944)	0.3435 [*] (0.1910)
Related diversification ($t-1$)	0.4622 [§] (0.1078)	0.4252 [§] (0.1020)	0.4480 [§] (0.1045)	0.4263 [§] (0.0991)	0.1049 (0.0892)	0.1071 (0.0918)	0.0742 [*] (0.0392)	0.0735 [*] (0.0420)
Related diversification ($t-2$)	0.2400 [§] (0.0711)	0.2168 [§] (0.0731)	0.2572 [§] (0.0738)	0.2198 [§] (0.0673)	-0.0447 (0.0837)	-0.0438 (0.0887)		
Squared Related diversification ($t-1$)					-0.1380 (0.0965)	-0.1433 (0.1013)	-0.1242 [*] (0.0878)	-0.1258 [*] (0.0716)
Squared Related diversification ($t-2$)					0.0318 (0.0811)	0.0301 (0.0904)		
Unrelated diversification ($t-1$)					-0.0098 (0.0186)	-0.0094 (0.0184)	-0.0111 (0.0156)	-0.0115 (0.0171)
Squared Unrelated diversification ($t-1$)					0.0136 (0.0210)	0.0133 (0.0210)	0.0153 (0.0219)	0.0153 (0.0213)
Size ($t-1$)	-0.0618 [†] (0.0310)	-0.0765 [§] (0.0298)	-0.0540 (0.0359)	-0.0768 [§] (0.0290)	0.0039 (0.0026)	0.0044 (0.0031)	0.0041 (0.0035)	0.0045 (0.0035)
Advertising exp. ($t-1$)	0.0001 (0.0065)	0.0003 (0.0061)	0.0008 (0.0072)	0.0003 (0.0054)	-0.0002 (0.0001)	-0.0001 (0.0001)	-0.0002 [*] (0.0001)	-0.0002 (0.0001)
ln(Sales) ($t-1$)	0.0318 (0.0251)	0.0412 (0.0271)	0.0270 (0.0309)	0.0433 (0.02634)	-0.0023 (0.0016)	-0.0027 (0.0019)	-0.0026 (0.0027)	-0.0031 (0.0027)
Selectivity correction		-0.0154 (0.0100)		-0.0158 [*] (0.0094)		-0.0012 [*] (0.0007)		-0.0012 [*] (0.0006)
Wald test of joint significance (% p -value)	572.4 (0.0)	574.2 (0.0)	406.5 (0.0)	532.3 (0.0)	684.7 (0.0)	724.1 (0.0)	670.6 (0.0)	694.7 (0.0)
Hansen-Sargan test (% p -value)	54.6 (84.1)	53.0 (85.8)	53.5 (88.4)	51.9 (89.7)	260.4 (3.6)	260.8 (3.1)	251.0 (9.6)	256.8 (5.4)
AR(2) test (% p -value)	-1.5 (14.0)	-1.4 (16.9)	-1.7 (9.0)	-1.6 (12.2)	-1.1 (24.9)	-1.1 (24.9)	-1.2 (23.3)	-1.2 (23.2)

The lag ($t-1$ or $t-2$) is indicated to the left of each variable. Standard errors in italics are below each estimated coefficient. Time dummies and age dummies were included in all estimations. The instrument set includes lags of the RHS variables at $t-3$, time, age and industry dummies. Due to the small sample size in the sample truncated on the basis of positive related diversification, and to avoid that the number of instruments that were too large with respect to the number of truncated observations, in the diversification equation, we exploit for each instrument one moment condition per lag instead of one moment condition per lag and per available year. In either case, the main results did not substantially change. The p -values indicate the significance level below which the null hypothesis is rejected. The Hansen-Sargan test evaluates the null hypothesis of validity of the over-identifying restrictions, and it is asymptotically distributed (under the null of validity) as a chi-square with as many degrees of freedom as the number of over-identifying restrictions. The AR(2) test is asymptotically distributed as a standard normal under the null hypothesis of no second order autocorrelation in the error term of the first-differences transformed model. This test is based on the fact that if the model is properly specified, the transformed error term should not exhibit second order autocorrelation.

^{*} Significance at the 10 percent level.

[†] Significance at the 5 percent level.

[§] Significance at the 1 percent level.

index was, on average, approximately 0.49. This result indicates that for a substantial proportion of diversifying firms, increasing levels of related diversification have a negative effect on R&D effort.

From the estimates shown in Table 5, we can compute the effects over time of a change in one strategic variable on the other. The effect after one year is simply the estimated coefficient of the corresponding variable at $t-1$. To compute the cumulative effect after additional years, we must consider the coefficients of the lagged dependent variables. Table 6 shows that the effects after one year are remarkably smaller than the full effect, which takes several years to complete.¹⁸ Together with the firm decision about either to diversify in further related activities or to undertake further effort in R&D activities, the firm may need to undergo organisational changes. As we have mentioned earlier, such changes are usually costly, so that the firm may have incentives to gradually distribute them over time rather than instantaneously implement the full change.

For example, it takes approximately two years for 50% of the effect to take place, whereas the full effect requires approximately 10 years to reach completion. Further, the sign and the size of the effect of diversification on R&D depend on the level of diversification. For this reason, we tabulated such effects for several levels of related diversification. Namely, for a firm that initially does not diversify, a 0.01 increase in related diversification will lead, on average, to an increase in R&D intensity by 0.07 percent after one year. The full effect is 0.26. For those firms whose diversification index is 0.20, a 0.01 increase in diversification would increase R&D, on average, by 0.02 after one year. The full effect (approximately 10 years later) is 0.09. The effect becomes null for diversification levels around 0.30 and becomes negative thereafter. Therefore, the complementarity between diversification and R&D occurs at moderately low levels of related diversification. Our results support Hypothesis 1b, whereby the effect of related diversification on R&D intensity is ruled by a U-shape pattern.

With regard to the effect of R&D on diversification, we found a one-to-one effect after one year, but the overall effect, after approximately 10 years, is almost three times the one-year effect. Therefore, we find empirical support for Hypothesis 1, which establishes a bidirectional and dynamic relationship between related diversification and R&D. In this dynamic context, the full effect of a change in R&D intensity on diversification and vice versa takes several years to reach completion.

¹⁸ Consider, for simplicity, the model $y_{jit} = \alpha_j + \rho_1 y_{jit,t-1} + \rho_2 y_{jit,t-2} + \delta_1 y_{kit,t-1} + X'_{jit} \beta_j + u_{jit}$. The immediate effect (after 1 year) of y_k on y_j is given by the coefficient Δ_{jj} . To compute the full effect, it is convenient to rewrite the model as $y_{jit} - \rho_1 y_{jit,t-1} - \rho_2 y_{jit,t-2} = \alpha_j + \delta_1 y_{kit,t-1} + X'_{jit} \beta_j + u_{jit}$, and the full effect is simply $\Delta_{jj} / (1 - \rho_1 - \rho_2)$.

Table 6
Distribution of effects of a unit change over time.

			Accum. effect							% of full effect
R&D → Diversification										
After	1	Year	1.01							35%
After	2	Years	1.66							58%
After	3	Years	2.09							73%
After	4	Years	2.36							83%
After	5	Years	2.53							89%
After	10	Years	2.82							99%
Full effect	2.86									
			Acum. effect for different diversification values							% of full effect
			$d_R = 0.0$	$d_R = 0.1$	$d_R = 0.2$	$d_R = 0.3$	$d_R = 0.4$	$d_R = 0.5$	$d_R = 0.6$	
Diversification → R&D										
After	1	Year	0.07	0.05	0.02	0.00	-0.03	-0.05	-0.08	27%
After	2	Years	0.13	0.08	0.04	0.00	-0.05	-0.09	-0.13	47%
After	3	Years	0.17	0.11	0.05	0.00	-0.06	-0.12	-0.18	61%
After	4	Years	0.20	0.13	0.06	-0.01	-0.07	-0.14	-0.21	72%
After	5	Years	0.22	0.14	0.07	-0.01	-0.08	-0.15	-0.23	79%
After	10	Years	0.26	0.17	0.08	-0.01	-0.10	-0.19	-0.28	96%
Full effect	0.27	0.18	0.09	-0.01	-0.10	-0.19	-0.29			

Calculations based on dynamic panel data estimates from Table 5. The full effect is $\beta/(1 - \rho_1 - \rho_2)$, ρ_1 and ρ_2 being the coefficients of the lagged endogenous variable at $t - 1$ and $t - 2$, respectively.

To evaluate the robustness of our results, we also considered alternative sets of instruments. In particular, due to the small sample size in the sample truncated on the basis of positive related diversification and to prevent the problem whereby the number of instruments was too large with respect to the number of truncated observations, we have exploited for each instrument one moment condition per lag instead of one moment condition per lag and per available year. In either case, the main results did not substantially change. In addition, we have also considered additional lags of the RHS at $t - 4$. In all these alternative estimates (not reported here), we found that the choice of instruments affected the precision of the estimates, but both the qualitative and quantitative results remain unchanged.

7. Conclusions

Over the last several decades, many studies have empirically analysed the link between corporate diversification and R&D intensity. However, they have provided contradictory evidence, particularly with regard to the effect of R&D on diversification. The existing contributions differ in the companies and countries of reference, and therefore, in their institutional environment and in the prevailing mode of firm growth. Earlier empirical work relied mostly on single cross-sections of data that focused on unidirectional effects of one strategic variable on the other; the potential feedback between both variables was not regarded. Finally, earlier contributions differ in the statistical assumptions, particularly in the treatment of potential sources of endogeneity. Our work, motivated by this inconclusive evidence about the mutual effects between R&D and diversification, contributes to this literature in several ways.

We posited a bidirectional relation between these two decisions and assessed it empirically by estimating a dynamic model with longitudinal data of companies that pursue internal growth. In particular, we used an unbalanced panel of Spanish manufacturing companies between 1990 and 2001. Our empirical approach takes into account firm-level unobserved heterogeneity, endogeneity of the right-hand-side variables, and the fact that the two strategic variables are subject to censoring. The results assert the dynamic nature of the relationship, on the one hand, and the existence of feedback effects between both strategic variables on the other hand.

In particular, we found an unambiguous positive effect of R&D on related diversification. There is a one-to-one effect of an R&D increase on related diversification after one year. Significant inertia in diversification results in an R&D increase on diversification over a 10-year period, which is considerably larger than the one-year effect. The effect of related diversification on R&D intensity was nonlinear. The effect is positive but decreasing up to a certain level of related diversification; eventually, it becomes negative when related diversification becomes high enough. Unrelated diversification, though, has no significant effect on R&D. Unlike unrelated diversification, related diversification reinforces the benefits of synergies (leading to an increase in technological resource endowments) and economies of scope (favouring a more efficient use of such resources). Our results suggest that diversifying in related activities, at least up to moderate levels, may increase the return to innovation in two instances. First, firms are more prompted to innovate to consolidate their position in the new activities. Second, firms become more efficient in R&D expenditure. We have also found that innovation activity creates incentives to the firms to branch out into new, related activities. Both strategic decisions exhibit a bidirectional and dynamic relationship, by which feedback effects are distributed over time. This result implies that the firms find a method to distribute changes over time in these two strategic variables, instead of adjusting instantaneously to their optimal levels because it is less costly to introduce such changes gradually. This perspective may have important research implications on the effects of diversification on performance.

Our results are also relevant from the managerial point of view. The results suggest the convenience of an integrated approach to the strategic decisions about increasing technological resources through R&D investment and about becoming involved in new related businesses. These two strategic decisions exhibit, in many instances, an important degree of complementarity. Although the planning of the different strategic decisions might depend on different decision-units within the firm, coordination among them can greatly improve the efficiency in decision-making. In this sense, the firm knowledge about the relationship between growth and technological resources can help it to design a coherent policy establishing the objectives of the different decision-units. It must also be pointed out that the effects of strategic decisions are not instantaneous, but they are distributed over several years. Therefore, time

planning must be taken into account for these two strategic decisions.

The main results also acknowledge the importance of performing a proper treatment of the potential sources of endogeneity in the empirical analysis, since disregarding these problems would lead to misleading conclusions. Unobserved heterogeneity should always be taken into account in empirical applications with firm-level data, since, no matter how many covariates we control, relevant unobserved characteristics of such firms potentially correlated with the existing covariates will generally remain. This problem is usually disregarded in the literature, with some exceptions. In many cases, this could be due to the lack of longitudinal data. We think that the proper econometric treatment of these problems may produce a relative loss of precision in the estimates. Nevertheless, this loss of precision should not lead one to disregard these problems because the alternative estimates would be inconsistent, and therefore meaningless, no matter how precise they might be.

Unlike most empirical work regarding this issue based on US companies, our analysis focused on firms that operate in settings where firms are more prone to organic growth. Thus, our results contribute to a better understanding about the link between related diversification and R&D in alternative institutional frameworks. The importance of the institutional framework was posed by Mayer and Whittington (2003).¹⁹ In comparison with US companies, which are the basis for most empirical work about diversification, European firms are smaller and more prone to internal growth. Therefore, the conclusions from US studies cannot necessarily be extrapolated to most European countries. Furthermore, this is the first work in which the link between diversification and R&D, focused solely on internal growth, has been analysed. The sole focus on internal growth further contributes to extending research in this area. Of course, our approach leaves aside strategies of external growth, firm splits, or cooperation modes between firms that may induce differences in a firm's strategies about diversification and R&D. These alternative strategies deserve further research and would provide more complete evidence about the link between diversification and R&D under different growth modes. These different strategies should be isolated, however, to elucidate this link in each particular case.

The method of measuring our two strategic variables is worth noting. Whereas the diversification measures, based on entropy indices, are commonly acknowledged in empirical work, there is no general agreement about the best measure of firm technological activity. These measures are often determined by the available data. In our case, we used R&D intensity to measure firm technological activity. Our conclusions may differ under other alternative measures: internal vs. external, or at the input vs. at the output level. We believe that it will be important to test our hypothesis under alternative measures of technological resources in future studies.

Our study highlights the importance of accounting for dynamics in the analysis of corporate diversification. The conceptual and empirical approach developed in this work, considering the dynamic effects of the relationship between resources and corporate diversification, will be an important contribution to future theoretical and empirical studies in the analysis of corporate diversification. In this context, a robust empirical approach acquires special importance. The econometric approach that we propose here may be a helpful benchmark for future research aimed at evaluating other impli-

¹⁹ The home country environment is characterized by features and institutions (Wan and Hoskisson, 2003) that determine both the allocations of technological resources and the diversification strategies of the companies.

Table A.1
Distribution of companies by industry and by size.

Industry (2-digit SIC code)	Size			
	Small	Medium	Large	All
Iron, steel and metal (22)	4	1	19	24
Building materials (24)	7	10	17	34
Chemicals (25)	14	13	33	60
Non-ferrous metals (31)	14	17	17	48
Machinery (32 + 33)	12	14	18	44
Electric materials (34)	12	6	20	38
Electronic (35)	4	5	11	20
Motor vehicles (36)	2	7	26	35
Ship building (37)	1	1	2	4
Other motor vehicles (38)	0	0	7	7
Precision instruments (39)	0	2	3	5
Non-elaborated food (41)	5	6	17	28
Food, tobacco & drinks (42)	3	2	15	20
Basic Textile (43)	5	10	16	31
Leather (44)	2	1	0	3
Garment (45)	15	7	8	30
Wood and furniture (46)	5	3	5	13
Cellulose and paper edition (47)	4	10	17	31
Plastic materials (48)	8	12	8	28
Other non-basic (49)	4	2	4	10
All	121	129	263	513

Source: Sample selected by the authors from *Encuesta Sobre Estrategias Empresariales (ESEE)*.

Firms are broken down by Size, in accordance with their average employment, as *Small* (up to 100 employees), *Medium* (between 100 and 250 employees) and *Large* (more than 250 employees).

cations of the resource-based view, such as the relationship between diversification decisions, resources, and corporate performance.

Acknowledgment

We thank two anonymous referees for their very helpful comments in the reviewing process. We are very grateful to Fundación Empresa Pública for providing the raw data. Both authors acknowledge financial support from the Spanish Ministry of Science and Innovation, Grants No. ECO2009-11165 and ECO2009-10358.

Appendix A. Distribution of companies by industry and by size

See Table A.1.

References

- Alonso-Borrego, C., Collado, M.D., 2002. Innovation and job creation and destruction: evidence from Spain. *Recherches Economiques de Louvain* 68, 149–168.
- Amemiya, T., 1984. Tobit models: a survey. *Journal of Econometrics* 24, 3–61.
- Anderson, T.W., Hsiao, C., 1982. Formulation and estimation of dynamic models using panel data. *Journal of Econometrics* 18, 47–82.
- Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58, 277–297.
- Arellano, M., Bover, O., 1995. Another look at the instrumental-variable estimation of error-components models. *Journal of Econometrics* 68, 29–52.
- Arellano, M., Bover, O., Labeaga, J.M., 1999. Autoregressive models with sample selectivity for panel data. In: Hsiao, C., Lahiri, K., Lee, L.F., Pesaram, M.H. (Eds.), *Analysis of Panels and Limited Dependent Variable Models*. Cambridge University Press, Cambridge.
- Baldwin, W.L., Scott, J.T., 1987. *Market Structure and Technological Change*. Harvard Academic Publishers, London.
- Baysinger, B., Kosnik, R.D., Turk, T.A., 1991. Effects of board and ownership structure on corporate R&D strategy. *Academy of Management Journal* 34, 205–214.
- Baysinger, B., Hoskisson, R.E., 1989. Diversification strategy and R&D intensity in large multiproducts firms. *Academy of Management Journal* 32, 310–332.
- Bergh, D.D., 1995. Problems with repeated measures analysis: demonstration with a study of the diversification and performance relationship. *Academy of Management Journal* 38 (6), 1692–1708.

- Bettis, R.A., 1981. Performance differences in related and unrelated diversified firms. *Strategic Management Journal* 2, 379–393.
- Bettis, R.A., Mahajan, V., 1985. Risk/return performance of diversified firms. *Management Science* 31 (10), 785–799.
- Blundell, R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87, 115–143.
- Bover, O., Arellano, M., 1997. Estimating dynamic limited dependent variable models from panel data. *Investigaciones Económicas* 21, 141–165.
- Burgelman, R.A., 1983. A process model of internal corporate venturing in the diversified major firm. *Administrative Science Quarterly* 28, 223–244.
- Canals, J., 1997. *Universal Banking*. Oxford University Press, Oxford.
- Capaldo, A., Cogman, D., Suonio, H., 2009. What's different about M&A in this downturn. *McKinsey Quarterly* 30 (Winter), 31–36.
- Chang, S.J., Singh, H., 1999. The impact of modes of entry and resource fit on modes of exit by multibusiness firms. *Strategic Management Journal* 20, 1019–1035.
- Chatterjee, B., 1990. Excess resources, utilization costs, and mode of entry. *Academy of Management Journal* 33, 780–800.
- Chatterjee, S., Singh, S., 1999. Are tradeoffs inherent in diversification moves? A simultaneous model for type of diversification and mode of expansion decisions. *Management Science* 45 (1), 25–45.
- Chatterjee, S., Wernerfelt, B., 1991. The link between resources and types of diversification: theory and evidence. *Strategic Management Journal* 12, 33–48.
- Chen, R., 1996. Technological expansion: the interaction between diversification strategy and organizational capability. *Journal of Management Studies* 33 (5), 649–668.
- Davis, R., Thomas, L.G., 1993. Direct estimation of synergy: a new approach to the diversity-performance debate. *Management Science* 39, 1334–1346.
- Delios, A., Beamish, P.W., 1999. Geographic scope, product diversification, and the corporate performance of Japanese firms. *Strategic Management Journal* 20, 711–727.
- Gedajlovic, E., Shapiro, D.M., 2002. Ownership structure and firm profitability in Japan. *Academy of Management Journal* 45 (2), 565–575.
- Grant, R.M., Jammine, A.P., 1988. Performance differences between Wrigley/Rumelt strategic categories. *Strategic Management Journal* 9, 333–346.
- Hamilton, B.H., Nickerson, J.A., 2003. Correcting for endogeneity in strategic management research. *Strategic Organization* 1 (1), 51–78.
- Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029–1054.
- Heckman, J.J., 1978. Dummy endogenous variables in a simultaneous equations system. *Econometrica* 46, 931–959.
- Helfat, C.E., 1994. Firm specificity in corporate applied R&D. *Organization Science* 5 (2), 173–184.
- Helfat, C.E., 1997. Know-how and asset complementarity and dynamic capability accumulation: the case of R&D. *Strategic Management Journal* 18, 339–360.
- Helfat, C.E., Eisenhardt, K.M., 2004. Intertemporal economies of scope, organizational modularity, and the dynamics of diversification. *Strategic Management Journal* 25, 1217–1232.
- Helfat, C.E., Raubitschek, R.S., 2000. Product sequencing: co-evolution of knowledge, capabilities and products. *Strategic Management Journal* 21 (10–11), 961–980.
- Hill, C.W., Hansen, G.S., 1991. A longitudinal study of the causes and consequences of changes in diversification in the U.S. pharmaceutical industry 1977–1986. *Strategic Management Journal* 12, 187–199.
- Hill, C.W., Snell, S.A., 1988. External control, corporate strategy, and firm performance in research-intensive industries. *Strategic Management Journal* 9, 577–590.
- Hill, C.W., Snell, S.A., 1989. Effects of ownership structure and control on corporate productivity. *Academy of Management Journal* 32 (1), 25–46.
- Hitt, M.A., Hoskisson, R.E., Ireland, R.D., Harrison, J.S., 1991. Effects of acquisitions on R&D inputs and outputs. *Academy of Management Journal* 34 (3), 693–706.
- Hitt, M.A., Hoskisson, R.E., Johnson, R.A., Moesel, D.D., 1996. The market for corporate control and firm innovation. *Academy of Management Journal* 39 (5), 1084–1119.
- Hitt, M.A., Hoskisson, R.E., Kim, H., 1997. International diversification: effects on innovation and firm performance in product-diversified firms. *Academy of Management Journal* 40 (4), 767–798.
- Holtz-Eakin, D., Newey, W., Rosen, H., 1988. Estimating vector autoregressions with panel data. *Econometrica* 56, 1371–1395.
- Hoskisson, R.E., Hitt, M.A., 1988. Strategic control systems and relative R&D investment in large multiproduct firms. *Strategic Management Journal* 9, 605–621.
- Hoskisson, R.E., Hitt, M.A., Johnson, R.A., Moesel, D.D., 1993. Construct validity of an objective (entropy) categorical measure of diversification strategy. *Strategic Management Journal* 14, 215–235.
- Hoskisson, R.E., Johnson, R.A., 1992. Corporate restructuring and strategic change: the effect on diversification strategy and R&D intensity. *Strategic Management Journal* 13, 625–634.
- Hoskisson, R.E., Cannella, A.A., Tihanyi, L., Faraci, R., 2004. Asset restructuring and business group affiliation in French civil law countries. *Strategic Management Journal* 25, 525–539.
- Itami, H., Numagami, T., 1992. Dynamic interaction between strategy and technology. *Strategic Management Journal* 13 (Special Issue), 119–135.
- Jacquemin, A., Berry, C., 1979. Entropy measure of diversification and corporate growth. *Journal of Industrial Economics* 27 (4), 359–369.
- King, D., Covin, J., Hegarty, W., 2003. Complementary resources and the exploitation of technological innovations. *Journal of Management* 29, 589–606.
- La Porta, R., Lopez-De-Silanes, F., Shleifer, A., Vishny, R.W., 1997. The *Journal of Finance* 52 (3), 1131–1150.
- Lecraw, D.J., 1984. Diversification strategy and performance. *Journal of Industrial Economics* 33, 179–198.
- Lunn, J., Martin, S., 1986. Market structure, firm structure and research and development. *Quarterly Review of Economics and Business* 26 (1), 31–44.
- McGahan, A.M., Porter, M.E., 1997. How much does industry matter, really? *Strategic Management Journal*, Special issue 18, 15–30.
- Mayer, M., Whittington, R., 2003. Diversification in context: a cross-national and cross-temporal extension. *Strategic Management Journal* 24, 773–781.
- Merino, F., Rodríguez, D.R., 1997. A consistent analysis of diversification decisions with non-observable firm effects. *Strategic Management Journal* 18, 733–743.
- Miller, D.J., 2004. Firm's technological resources and the performance effects of diversification: a longitudinal study. *Strategic Management Journal* 25, 1097–1119.
- Miller, D.J., 2006. Technological diversity, related diversification, and firm performance. *Strategic Management Journal* 27, 601–619.
- Montgomery, C.A., Hariharan, S., 1991. Diversified expansion by large established firms. *Journal of Economic Behavior and Organization* 15, 71–89.
- Palepu, K., 1985. Diversification strategy, profit performance and the entropy measure. *Strategic Management Journal* 6, 239–255.
- Penrose, E., 1995. *The theory of the growth of the firm*, 2nd ed. Oxford University Press, New York.
- Robins, J.A., Wiersema, M.F., 1995. A resource-based approach to the multi-business firm: empirical analysis of portfolio interrelationship and corporate financial performance. *Strategic Management Journal* 16 (4), 277–299.
- Robins, J.A., Wiersema, M.F., 2003. The measurement of corporate portfolio strategy: analysis of the content validity of related diversification indexes. *Strategic Management Journal* 24 (1), 39–59.
- Schoenecker, T.S., Cooper, A.C., 1998. The role of firm resources and organizational attributes in determining entry timing: a cross-industry study. *Strategic Management Journal* 19, 1127–1143.
- Scherer, F.M., 1984. Corporate size, diversification, and innovative activity. In: Scherer, F.M. (Ed.), *Innovation and Growth: Schumpeterian Perspectives*. MIT Press, Cambridge, pp. 222–238.
- Silverman, B.S., 1999. Technological resources and the direction of corporate diversification: toward an integration of the resource-based view and transaction cost economics. *Management Science* 45 (8), 119–124.
- Stimpert, J.L., Duhaime, I.M., 1997. Seeing the big picture: the influence of industry, diversification and business strategy on performance. *Academy of Management Journal* 40, 560–583.
- Tanriverdi, H., Venkatraman, N., 2005. Knowledge relatedness and the performance of multibusiness firms. *Strategic Management Journal* 26, 97–119.
- Teece, D.J., 1982. Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior and Organization* 3, 39–63.
- Teece, D.J., 1980. Economies of scope and the scope of the firm. *Journal of Economic Behavior and Organization* 1, 233–247.
- Teece, D.J., 1986. Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. *Research Policy* 15, 285–305.
- Tsang, E., 2000. Transaction cost and resource-based explanations of joint ventures: a comparison and synthesis. *Organization Studies* 21 (1), 215–242.
- Wan, W.P., Hoskisson, R.E., 2003. Home country environments, corporate diversification strategies, and firm performance. *Academy of Management Journal* 45 (1), 27–45.
- Williamson, O.E., 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. Free Press, New York.
- Windmeijer, F., 2004. A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics* 126 (1), 25–51.
- Yip, G.S., 1982. Diversification entry: internal development versus acquisition. *Strategic Management Journal* 3, 331–345.