UNIVERSIDAD CARLOS III DE MADRID

working papers

Working Paper 07-82
Economic Series (49)
December 2007

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 916249875

Corporate diversification and R&D intensity dynamics[†]

César Alonso-Borrego (Universidad Carlos III de Madrid)
Francisco Javier Forcadell (Universidad Rey Juan Carlos)[*]

Abstract

We study the dynamic bidirectional relationship between firm R&D intensity and corporate diversification, using longitudinal data of Spanish manufacturing companies. Our empirical approach takes into account the censored nature of the dependent variables and the existence of firm-specific unobserved heterogeneity. Whereas we find a positive linear effect of R&D intensity on related diversification, the evidence about the effect of related diversification on R&D intensity takes the form of an inverted U. Hence, the effect of related diversification on R&D intensity is positive but marginally decreasing for moderate levels of related diversification, but such effect can turn out negative for high levels of related diversification. Additionally, the consequences of the dynamic relation are that the effects are substantially larger in the long-run than in the short-run.

Keywords: Diversification; R&D Intensity; Dynamics; Organic Growth; Endogeneity; Panel Data.

.

JEL Classification: L22, M11, C33, C34.

† We acknowledge research funding from the Spanish Ministerio de Educación y Ciencia, Grant No. SEJ2006-05710/ECON (C. Alonso-Borrego), and SEJ2007-67496/ECON and SEJ2005-08805/ECON (F.J. Forcadell).

* Corresponding Author. Department of Management, Universidad Rey Juan Carlos, Paseo de los Artilleros s/n, 28032 Madrid, Spain. E-mail: francisco.javier.forcadell@urjc.es.

1. Introduction

The relationship between R&D intensity and corporate diversification has received considerable attention in empirical research on strategic management (e.g. Chatterjee and Wernerfelt, 1991; Chen, 1996; Silverman, 1999). While there is a pervasive evidence of a linear and positive effect of related diversification on R&D intensity, the empirical evidence about the effect of R&D intensity on diversification is mixed (see Table 1). While most empirical work has concentrated in unidirectional relationships, there is a lack of evidence about the potential feedback between diversification and R&D intensity.

In this paper, we attempt to shed some light on the contradictory findings concerning the link between corporate diversification and R&D intensity. We posit a dynamic bidirectional hypothesis between corporate diversification and R&D intensity, and evaluate such relationship at the empirical level. The dynamic nature of such relationship is sustained by the concepts of synergies and economies of scope. This bidirectional link emphasizes 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 looked at the experience of U.S.-based companies (Wan and Hoskisson, 2003). However, the institutional environment in which firms operate determines their dominant growth mode. In the US, diversification takes place mainly through external growth, but this is not the case in Europe (Mayer and Whittington, 2003), and particularly in Spain. In those contributions which consider both organic growth and external growth, the effects of these alternative growth modes are not distinguished (see Table 1). Regarding this, it would be worth to isolate the effect of organic growth by means of a sample of companies which only make use of this growth mode.[1]

We evaluate 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 allows to consider the dynamic features of R&D intensity and diversification decisions, allowing for lagged effects distributed over time and thus distinguishing between short-run and long-run effects. For the purpose of characterizing the simultaneous decisions regarding R&D intensity and diversification and the potential feedback between them, we estimate a bivariate VAR for R&D intensity and related diversification, augmented by additional covariates. In order to obtain robust results, we also control for two potential sources of endogeneity: censoring and unobserved firm heterogeneity. The failure to account for any of these two sources of endogeneity can provide misleading implications.

Our empirical results provide evidence in favor of the bidirectional relationship between corporate diversification and R&D intensity, in two extents. First, a rise in R&D intensity has a positive effect on related diversification. Second, related diversification influences R&D intensity in the form of an inverted-U shaped function, what implies that for moderate levels of diversification there is positive effect of related diversification on R&D intensity, yet marginally decreasing, so that this effect is lesser when the level of diversification is greater (and could be eventually offset when the degree of diversification is high enough). The dynamic nature of the relationship implies that short-run and long-run effects are different, the latter being significantly greater. This result points out that time is needed for generating and exploiting synergies and economies of scope.

2. Theory and Hypothesis

Since the last decades, there is a steady interest in the link between diversification strategy and R&D intensity, which has been empirically analyzed in several studies, which we have summarized in Table 1. However, to our knowledge, empirical studies have adopted a unidirectional approach, which provide conflicting evidence. On the one hand, the common finding about the effect of diversification on R&D intensity is that such effect is positive. Among the few exceptions, we should mention Hill and Hansen (1991) or Schnoenecker and Cooper (1998), referred to very particular industries, and Miller (2004), where, in an external growth framework, diversification affects R&D intensity negatively. On the other hand, with regard to the effect of R&D intensity on diversification, evidence is odd, finding a positive (Davis and Thomas, 1993; Hoskisson and Johnson, 1992) or negative influence (Hitt et al 1996; Stimpert and Duhaime, 1997).

In our view, there are several reasons which can be behind 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 be remarkably different. Second, most studies do not distinguish between growth modes, and the decision about growth mode influences on both corporate diversification and R&D intensity. Third, assumptions about the time schedule by which the effects take place may affect the results. Fourth, there are differences in the methodological approach that can affect conclusions. The different approaches that studies shown in Table 1 have undertaken are very likely to affect empirical findings. Nonetheless, and more importantly, no study analyzes the bidirectional relationship, which acknowledges the potential feedback between these two strategic decisions. We believe that it is necessary to go further and propose a new hypothesis posing a dynamic bidirectional relationship between both strategic variables.

Insert Table 1 about here

There are two kinds of arguments that support the link between related diversification and R&D intensity. From the point of view of production, super-additive value synergies among the firm's businesses (Davis and Thomas, 1993) derived from input complementarities, so that the output value of a multiproduct firm exceeds the sum of individual outputs. On the other hand, from the point of view of costs, economies of scope, because of sub-additive production costs derived from resource sharing between businesses (Teece, 1982) imply that joint production costs are lower than the sum of individual production costs. Nevertheless, as Tanriverdi and Venkatraman (2005) point out, the concepts of synergy and economies of scope have been used synonymously, because the cost function and the production function are alternative views of the firm's decision problem. While this line of reasoning is concerned with related diversification, no synergies and economies of scope among the technological activities are expected in the case of unrelated diversification.

In a diversified firm, the generation of synergies and economies of scope from the relationship between related corporate diversification and R&D intensity implies two benefits: more options to exploit R&D intensity; and a more efficient use of R&D intensity. Related diversification provides the firm broader options in which technological resources can be exploited, so that investment in R&D might be increased so as to reach wider technological areas. In turn, the larger scope of technological resources may lead to further related diversification so as to optimize the exploitation of available technological resources. A more efficient use of technological resources implies that the relative amount of technological resources may decrease with the number of related businesses because of the appearance of economies of scope (Baysinger and Hoskisson, 1989; Miller 2004). Thus, after a certain level of related diversification is achieved, lower additional investment in R&D is needed in order to compete in new businesses. As a consequence of the combined effect of synergies and economies of scope, as business relatedness reaches a certain degree, R&D saving might offset the synergies effect because of economies of scope.

According to the previous related literature, we can suggest that as R&D investment increases, the level of related diversification will increase as well. By the same token, as related diversification rises, R&D intensity increases because of the existence of synergies. However, whenever the degree of related diversification is high enough, the effect of economies of scope may eventually offset the effect of synergies. Therefore, the effect of related diversification on R&D intensity may be nonlinear, characterized by an inverted U-shaped function, where the effect of related diversification on R&D intensity is positive but marginally decreasing, and might be even negative for high levels of related diversification. Such non-linearity may also help to reconcile the empirical evidence of earlier research on this topic (Table 1).

Still, two additional elements of the relationship between R&D intensity and corporate diversification need to be considered: the growth mode and the time schedule in the relationship. There are different modes through which firms can grow, internal or organic, external, and cooperative (Tsang 2000). As we mentioned earlier, most studies included in Table 1 do not distinguish between external and organic growth, and there is no study that analyzes firms experiencing organic growth exclusively. We believe that the absence of control for growth mode may lead to unclear conclusions concerning R&D intensity and the related diversification link. The choice among these growth modes is related to the minimization of the transaction costs associated with these different governance structures (Williamson, 1975), which, in addition to firm characteristics, depend on the institutional environment in which firms are involved. Furthermore, growth modes affect how diversification is extended (Busija et al., 1997; Chang and Singh, 1999; Chatterjee and Singh, 1997) and how diversification and R&D intensity are related (Miller, 2004)[2]. Then, the choice of companies under study entails non neutral implications about the empirical findings. External growth is the dominant growth mode among US companies, in which most empirical work have focused. However, external growth (through mergers and acquisitions) entails diversification decisions which can be affected by agency problems. Yet this is not usually the case for companies located in most European countries, which are dominated by organic growth (Mayer and Whittington, 2003). Therefore, the evidence provided by European companies can help to identify the effect of organic growth. One of the advantages of concentrating in companies which are primarily involved in organic growth is that the diversification strategy can be isolated from other considerations. Also, the fact that most firms undertake organic growth makes conclusions based on firms for which this growth mode is the dominant, at least qualitatively, general for almost all firms.

The major concern in this paper is to acknowledge the dynamic nature of the relationship between R&D intensity and corporate diversification. These two decisions have a strategic and intertemporal nature because their consequences have a lasting influence in the organization and the performance of the firm. The intertemporal nature has been acknowledged, among others, by Helfat and Eisenhardt (2004), in the context of related diversification decisions. In other words, these decisions have persistent implications, so that their relevance exceeds the instantaneous or short-run consequences. The implicit framework within which the firm determines the target levels of its strategic variables is a dynamic intertemporal decision model in which the firm takes into account their effect on current and future performance. In addition, such decisions are constrained by the fact that variations in these strategic variables entail costs to the firm which affect the time length within the firm will adjust to the new target levels of such variables. Particularly, if the optimal targets in the strategic variables imply large changes, it may be more profitable to distribute such changes along a longer time until reaching the target level, instead of making the full change instantaneously, what would entail higher costs to the firm. Learning effects and organizational issues (Mayer and Whittington, 2003; Grant and Jammine, 1988; Bergh, 1995) are among the primary reasons which can induce such costs. In economics literature, these sorts of costs are labeled as "adjustment costs" (Helfat and Eisenhardt, 2004). Empirically, this is reflected by the fact that lagged values of the strategic variables will appear among the covariates of the equations characterizing such strategic decision. As a consequence, the short-run or instantaneous effect of a change in any covariate will differ from the long-run effect (the total accumulated effect). Thus, we would expect that the R&D intensity influence on related diversification (and vice versa) will be more intense in the long run than in the short run. These dynamic considerations provide further rationale for the potential feedback between R&D intensity and diversification. In addition, the dynamic considerations of the relationship have the additional implication that the effect of changes in any covariate can take place with a certain delay. Consequently, the initial effect of corporate diversification on R&D intensity (and vice versa) may take place with a certain time lag after the other strategic variable has changed. Moreover, this effect may be distributed over time, so the duration of the total effect can be long.

Hypothesis

> *Hypothesis 1*. In a context of organic growth, R&D intensity and related diversification show a bidirectional relationship, in which the total effect takes several years to be achieved, so that the long-run or total effect is greater than the short-run or instantaneous effect.

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

Hypothesis 1b. And on the other hand, related diversification influences R&D intensity through an inverted-U shaped function. In other words, although the effect of related diversification on R&D intensity is initially positive, it is marginally decreasing, so that it may turn out negative when the level of related diversification is high enough.

3. Econometric Issues

The dynamic hypothesis that we propose requires us to characterize the simultaneous decisions of R&D and diversification in the empirical analysis. Our concern for the importance of dynamic feedback has already been expressed in earlier work, in the sense that the lack of accounting for such feedback effects (because of the use of cross-sectional data) is behind the failure to establish a clear relationship between diversification and R&D. In his review of empirical work, Bergh (1995:1697) claims that change over time has not been accounted for empirically. Hill and Hansen (1991:187) state that cross-sectional studies are unable to determine the underlying causal relationship.

In order to allow for the potential feedback between R&D intensity and diversification decisions, we posit a dynamic model for the joint decisions of R&D intensity and related diversification, conditional on further covariates. The availability of longitudinal data allows us to characterize the potential dynamic effects. In particular, our model consists on a bivariate augmented VAR (vector auto-regression) 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, 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,

$$y_{jit} = \alpha_j + \rho_{j1} y_{ji,t-1} + \rho_{j2} y_{ji,t-2} + \delta_{j1} y_{ki,t-1} + \delta_{j2} y_{ki,t-2} + \beta' x_{jit} + v_{jit}$$
$$j,k = \text{related diversification, R\&D intensity } (j \neq k)$$

where v_{jit} is the error term, and x_{jit} is the vector of further covariates. 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 effect of a change in any of the right-hand-side variables will differ in the short and in the long run due to the presence of lagged values of the dependent variable on the right-hand side of each equation. Specifically, the short-run effects can be substantially lower than the long run-effects.

Among the contributions that exploit panel data to carry out empirical work, we should refer to Merino and Rodríguez (1997), Gedajlovic and Shapiro (2002), and Miller (2006). The first one studies the discrete decision on whether or not to diversify, being particularly concerned with the inconsistency bias of the estimates because of disregarding unobserved firm heterogeneity, for which they propose a proper econometric treatment that requires panel data. The second one studies the role of agency problems in profitability, taking into account dynamic effects. The third one finds a positive relationship between technological diversification and performance, using a random effects model.

The variables that we are considering are strategic, in the sense that 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 that we wish to consider jointly in order to obtain appropriate estimates that guarantee robust results. The first one arises from the fact that the dependent variables in each equation (R&D intensity and diversification) are censored. The second one is the existence of firm-specific unobserved heterogeneity, which is potentially correlated with the right-hand-side variables. In contrast with the random effects approach, which disregards such potential correlation, we will consider fixed effects models. For the sake of clarity, we will first describe these two problems separately.

The censoring problem arises from the fact that the decision variables can be positive, but for a large proportion of the observations will be exactly equal to zero (very especially in the case of diversification). Therefore, the observed decision variables have two components: a qualitative one corresponding to the strictly positive and zero-valued observations, and a continuous one describing the range of positive values. The proper treatment of this kind of dependent variables requires to characterize the firm's discrete decision on whether to innovate (respectively, diversify) or not, and the firm's decision on the amount of R&D intensity (resp., diversification).

Let our model of interest (omitting the indices *t* and *j* for the sake of simplicity) be

$$y_i^* = x_i'\beta + u_i \tag{2}$$

where y_i^* denotes the variable of interest in the absence of censoring, x_i is a vector of covariates and β is its corresponding vector of unknown coefficients which represent the respective effect of the covariates. For each observation, the dependent variable can be described as

$$y_i = \begin{cases} y_i^* & \text{if } I_i^* > 0 \\ 0 & \text{if } I_i^* \leq 0 \end{cases} \tag{3}$$

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

$$I_i^* = z_i'\gamma + v_i \tag{4}$$

In the case of R&D intensity (respectively, diversification) decision, I_i^* denotes the marginal net revenue of such decision, so that we will observe a firm innovating (resp. diversifying) if the marginal net revenue of R&D intensity (resp. diversification) is positive; otherwise, the firm will not innovate (resp., diversify). In principle, neither the right-hand-side variables (and thus the censoring mechanism) nor the magnitudes of their effects need to coincide in the equation of interest which determine the positive amount of R&D (resp., diversification) and in the auxiliary equation describing the marginal net revenue of the qualitative decision on whether to innovate (resp. diversify) or not.

The parameters of interest can be consistently estimated using the subsample of positive observations by means of a two-stage procedure (Heckman, 1978). Formally, the augmented equation for the variable of interest for the subsample of positive observations can be written as

$$y_i^* = x_i'\beta + \delta\hat{\lambda}_i + v_i \qquad (5)$$

where $\hat{\lambda}_i$ is an additional variable which captures the estimated bias for the ith observation due to the truncation of the sample. Such bias is based on the estimates of the auxiliary equation about the qualitative decision on whether to innovate (resp., diversify) or not.

The problem of unobserved heterogeneity arises from the existence of relevant unobserved firm characteristics 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 in the presence of unobserved heterogeneity. Using the subscript t to denote time, and in the absence of censoring, we can write the model for longitudinal data as

$$y_{it}^* = x_{it}'\beta + \eta_i + u_{it},$$

where η_i is an unobserved random variable which captures time-invariant differences across firms. The problem resembles an omitted variable problem. The availability of the panel makes it possible to apply a transformation that removes the unobserved heterogeneity term. Particularly, subtracting from the above equation the same equation lagged one period (which also includes the unobserved heterogeneity term), and denoting Δ as the difference operator, we find that

$$\Delta y_{it}^* = \Delta x_{it}'\beta + \Delta u_{it},$$

where the unobserved heterogeneity term η_i has been removed, under the assumption of time invariance, because of the first-differences transformation, but the parameter vector of interest β has remained unchanged after such 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. We can therefore apply a generalized method of moments (GMM) estimator (see Hansen 1982, Arellano and Bond 1991). In particular, our estimation approach consists of a system-GMM estimator proposed by Arellano and Bover (1995) and developed by Blundell and Bond (1998).

The combination of these two problems, censoring and unobserved heterogeneity, makes things even more complicated. Our approach is based on Bover and Arellano (1997), and Arellano, Bover and Labeaga (1999). In the first place, we will deal with the censoring problem. For this purpose, we will consider, separately for each year, the estimation of auxiliary equations for y_{it}^{*} in order to obtain predictions of y_{it}^{*} for the censored observations. These auxiliary equations will be estimated by means of the aforementioned two-stage procedure. We can construct afterwards an estimate of the variable y_{it}^{*} in the absence of censoring, $\tilde{y}_{it}^{*}$, by substituting the zeros by their corresponding predictions and estimate the model with the full sample as if censoring had not taken place. Although the proper econometric treatment of these problems may reduce the precision in the estimates, disregarding these problems would lead to inconsistent, and therefore meaningless, estimates, no matter how precise they might be.

3. Data and Variables

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)[3]. This database has been used previously by Merino and Rodríguez (1997), among 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 consists of 513 non-energy manufacturing companies which satisfy the following criteria. Firstly, the nature of the company cannot be substantially altered in the sample period, so that we discard mergers or splits, as well as changes in the main activity (defined at the two-digit industry level). Consequently, and in accordance with the hypothesis that we posit, our sample contains those companies which carry out only organic growth strategies, and those companies that undertake portfolio restructuring are disregarded. Secondly, we require the companies to operate in at least two markets whose geographical scope must be national or wider. Finally, the companies are required to have a minimum of 25 employees[4]. In Table A.1 we show the sample distribution of companies by two-digit industry and by size. It is important to notice that the unbalanced feature of the panel implies that the time length during which we observe the different companies is unequal, reflecting the fact that such companies may enter and exit from the survey (in the same way that companies appear and disappear in the economy). Restricting the sample of companies to be observed along the same time period would affect the randomness of the sample. Instead, we only require sample companies to have full information available on all the relevant variables for at least five consecutive years, between 1990 and 2001, in order to obtain sufficient information concerning firm dynamics.

Variables and measures

We define R&D intensity as the proportion in firm sales of its R&D expenditure, in line with earlier related work, such as Baysinger and Hoskisson (1989), and Hitt et al. (1997), among others. We employ the entropy index as diversification measure (Jacquemin and Berry, 1979; Palepu 1985), which is of common use in strategic management literature (see, e.g., Baysinger and Hoskisson, 1989). This measure is thus defined as

$$d = \sum_i \left[P_i \times \ln\left(\frac{1}{P_i}\right)\right],$$

where the subindex i represents the ith product segment. The entropy index is thus a weighted average of the sales attributed to each firm's segment, P_i, the weights being the natural logarithm of the inverse of the segment's sales, $\ln(1/P_i)$. The entropy index of total diversification can be broken down as the sum of two components: the index of related product diversification, d_R, and the index of unrelated product diversification, d_U. Whereas related product diversification, d_R, captures the diversification across four-digit SIC industries within each two-digit industry (with firm's sales in the two-digit industry as reference), unrelated product diversification, d_U, captures diversification across two-digit industries (with total firm sales as reference). For the diversification equation, and given the reasons discussed previously, we will concentrate on related product diversification as the dependent variable.

We consider several control variables. First, we include the natural logarithm of total firm sales in order to control for firm size. Regarding R&D intensity, the influence of firm size on R&D intensity is still a matter of controversy (King et al., 2003: 591). On the other hand, several studies such as those by Bettis (1981), Hill and Snell (1988), or Hoskisson et al (1993), among others, have found a significant correlation between firm size and diversification.

Furthermore, earlier studies have found spillover effects of R&D at the industry level, demonstrating that the average level of industry R&D affects the relative firms' R&D intensity as well as their propensity to introduce new products (Baysinger and Hoskisson, 1989). In addition, other factors associated with the firm's environment, defined as the industry in which the firm operates, can affect the outcome of our variables of interest. To control for such effects, we include as additional regressors dummies representing each firm's two-digit SIC major industry. Since industry effects can vary over time, we also consider time dummies as well as industry dummies interacted with time dummies (see Bergh, 1995; McGahan and Porter, 1997, among others).

In order to consider 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, we have also included age dummies, for which we have chosen the following age groups: less than 10 years, between 10 and 20 years, between 20 and 40 years, and older than 40 years. We believe that these qualitative dummies are preferable to the use of the continuous variable of the age of the firm, due to the unrealibility of such a measure for older firms (Alonso-Borrego and Collado, 2002).

Finally, we have also considered the potential effect of several covariates, whose omission could bias the results: first, the firm marketing decision using the percentage of advertising expenditure with respect to sales; second, the logarithm of firm sales. We have already evaluated the relevance of some other variables, such as whether the firm has industrial establishments abroad (which is the best approximation that we have found in our dataset to international diversification), or whether the firm belongs to an enterprise group, as well as the firm market share in its main market (as a measure of its market power), and measures of firm liquidity and leverage. However, these latter variables did not appear to be statistically relevant in our analysis. Therefore, the results including them have not been reported.

Among the significant correlations reported in Table 2, we should highlight the positive correlation of R&D intensity with related diversification. Furthermore, the logarithm of sales is also positively correlated with the two decision variables, R&D intensity and related diversification. In addition, firm size presents a positive correlation with R&D intensity, but not with related diversification. The opposite is true for firm age: we only find a positive correlation with the diversification decision. Hence, the preliminary descriptive analysis suggests that it is the maturity of the firm, rather than its size, what matters for 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. Concerning this latter point, in order to capture the partial effects on the decision variable of every explanatory variable keeping constant the others, we must proceed with a multivariate regression analysis.

Insert Tables 2 and 3 about here

In the upper part of Table 3, we report the means and standard deviations of the variables of interest for the full sample. However, given the censoring of R&D intensity and related diversification, we broke down these statistics depending on whether both R&D intensity and related diversification are zero or positive. We observe that zero R&D intensity and zero diversification are more likely when the firm is smaller and/or younger. However, there does not appear to be any link between censoring in any of the two decision variables and sales. Furthermore, advertising expenditure seems to be higher for innovating firms (i.e., those with positive R&D intensity), but average advertising expenditure is not significantly different for diversifying and non-diversifying firms. Finally, the probability of positive R&D intensity is higher when diversification is greater, and viceversa.

4. Results

In order to illustrate the consequences of ignoring the potential econometric problems that we have mentioned in the previous section, we report alternative estimates based on different statistical assumptions. Specifically, we will first provide preliminary estimates whose consistency depends on the absence of firm-specific unobserved heterogeneity in Table 4, and our preferred estimates which take into account all the potential sources of endogeneity in Table 5.

In Table 4 we present the preliminary estimation results, assuming that all the potential heterogeneity among firms is solely explained by its observed industry affiliation. Among these, we report estimates that consider different treatments of the censoring problem, ranging from the OLS and truncated OLS estimates (that ignore censoring) to the tobit and generalized selection models (that make different assumptions concerning how censoring is determined)[5]. We have dropped from the specifications further covariates which were clearly unimportant from a statistical point of view. Time and industry dummies were included in all estimations and found to be jointly significant.

Insert Table 4 about here

Both in the R&D intensity and in the diversification equation, we find a large and significant coefficient of the lagged dependent variable. However, behind the magnitude of the coefficient of the lagged dependent variable, there may be a lack of control for firm-specific unobserved heterogeneity.[6] The fact that the estimates of the tobit and generalized selection models appear substantially different suggest that the constraints behind the tobit specification are not satisfied. In particular, the factors determining the discrete decisions on whether to innovate or diversify are different from the factors that determine the continuous decisions concerning the magnitude of R&D intensity or diversification.[7] This result supports the generalized selection model, in which the sample selection term, given by the inverse of the Mills' ratio, appears to be significant.

It is worth noticing that the estimates from Table 4 are not robust in the presence of time-invariant unobserved firm heterogeneity correlated with the explanatory variables. In what follows, we will concentrate on results based on a dynamic panel data estimation which explicitly takes into account unobserved heterogeneity, together with the censoring problem of the dependent variables. Our reported estimates 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 the two lagged dependent variables, for which we initially introduced the first and second lag. The reported results only include the specification with those lags of the dependent variable that were found significant.

Our final results for R&D intensity and related diversification are presented in Table 5.[8] The first noticeable result is that, although the main qualitative findings remain, the magnitudes of the estimated coefficients are remarkably different. In particular, the estimated coefficients corresponding to the lagged endogenous variable in both equations are now much smaller. This result highlights the existence of unobserved firm heterogeneity which produced upward biased coefficients for the lagged endogenous variable in the earlier non-robust estimates. The significance of the lagged dependent variables underlines the inertia behind the R&D intensity and diversification decisions, which we can attribute to the existence of significant learning effects.

Insert Table 5 about here

In the diversification equation, although the estimated coefficient of R&D intensity at *t-2* is not significant at the five percent level, it is so at the ten percent level. We attribute the lower precision of the estimates in the diversification equation to the sharper incidence of censoring in the dependent variable. We can assert that the effect of R&D intensity on related diversification is positive, supporting hypothesis

1a. Nevertheless, this effect takes place with a two-period lag. This interesting result points out that the influence of R&D intensity takes longer, given that we found a very small and non-significant coefficient for the first lag of R&D intensity (not reported here for the sake of brevity). We also find that related diversification appears to have much inertia, as shown by the large coefficient associated with its lagged value. Whereas the AR(2) specification test does not provide evidence against the specification (the p-value being clearly above ten percent), the Hansen-Sargan test is not so conclusive.

In the R&D intensity equation, we confirm the positive (and significant) estimated effect of related diversification. Moreover, the fact that the coefficient of square related diversification has the opposite sign points out that, although related diversification stimulates R&D, such positive effect decreases as the level of related diversification rises. This result is coherent with hypothesis 1b, by which the influence of related diversification on R&D intensity comes through an inverted U-shaped function. Nevertheless, the coefficient on squared related diversification is significant at the 10 percent level. We attribute this lack of precision to the fact that this estimation procedure provides robust estimates at the expense of larger standard errors[9].

In any case, we achieve a much greater estimated short-run effect of related diversification than in the earlier estimates. Regarding unrelated diversification, we find no significant effect on R&D intensity. Finally, neither size, nor advertising expenditure, nor sales were significant at any reasonable significance level. Concerning the two specification tests that we have carried out, we do not find evidence against the specification, given that the p-value of both imply that we do not reject the corresponding null hypotheses at any usual significance level.

From the estimated numerical values of the coefficients of related diversification and squared related diversification, we could compute the critical level after which additional increments in related diversification would imply a negative effect on R&D intensity. This critical level of related diversification is above 0.56, which is quite high, given that the average value of related diversification is lower than 0.06 in the full sample and about 0.47 for the subsample of observations with strictly positive values of related diversification.

From the estimates in Table 5, we can compute the long-run marginal effects taking into account the model dynamics, that we in Table 6. It can be seen that such long-run effects are much greater than the short-run effects.[10] Specifically, an increase in related diversification of 0.01, will increase R&D intensity in the long run, on average, by 0.05 percent. On the other hand, if R&D intensity rises by 1 percent, related diversification will increase in the long run, on average, by 0.005. Although this last magnitude is substantially large, it is nonetheless much smaller, and clearly more plausible, than the long-run marginal effect implicit from the estimates reported in Table 5, which were inconsistent in the presence of firm-specific unobserved heterogeneity.[11] In fact, notice that after one year only 60 percent of the effect of diversification on R&D intensity takes place, and less than 40% of the total effect of R&D intensity on diversification takes place after two years. In fact, it takes four years to produce 95 per cent of the effect of a change in diversification, and eight for 95 per cent of the effect of a change in R&D intensity. Therefore, we find support for the hypothesis 1, which states that the total effect between R&D intensity and diversification takes several years to be achieved.

Insert Table 6 about here

----

5. Discussion and Conclusions

This paper is mainly concerned with the potential dynamic and bidirectional relation between corporate diversification and R&D intensity. Empirical studies in strategic management literature have typically adopted a unidirectional approach, with mixed evidence. In our paper, we try to fill this gap by positing a bidirectional relationship between related diversification and R&D intensity in an organic growth context, and providing empirical evidence, based on a longitudinal sample of Spanish manufacturing firms between 1990 and 2001.

Our findings emphasize the dynamic nature of the relationship between related diversification and R&D intensity, as shown by the significant coefficient of the corresponding lagged dependent variables.

This will imply remarkable differences between the long-run and the short-run effect of changes in these two strategic decisions of the firm, as postulated in our hypothesis. Our evidence supports the hypothesis that the effect of related diversification on R&D intensity takes on the shape of an inverted U, so that the effect of related diversification would be positive but at a marginally decreasing rate for moderate degrees, and eventually the sign of this effect could be reversed if the degree of related diversification is high enough. It is possible to attribute this shape to the generation and exploitation of economies of scope and synergies. On the other hand, we find a positive effect of R&D intensity on related diversification, but this (short-run) effect occurs with a delay of two years, implying that it takes some time for the effect to appear. This evidence supports some previous literature, which associates this effect to the existence of adjustment costs.

The evidence concerning the bidirectional effects between corporate diversification and R&D intensity extends the earlier research in this area by revealing the need to consider the dynamics of diversification, and the causes and consequences of corporate diversification strategy in terms of resources. Unlike unrelated diversification, related diversification reinforces the benefits of synergies (leading to an increase in technological resources endowments) and economies of scope (favoring a more efficient use of such resources). Consequently, the value of the related diversified firms may increase as result of such synergies and economies of scope. This perspective may entail important research implications on the effects of diversification on performance.

Our empirical findings in favor of the dynamic and bidirectional relationship have consequences about the distribution over time of the mutual effects between corporate diversification and R&D intensity, as well as the effect of further covariates. The interrelation between these two strategic variables follows a distributed lag structure, implying that the effect of a change in any variable takes some time. In other words, the change in the strategic variable after a change in any covariate is implemented gradually. Additionally, the effect of a change in any of the right-hand-side variables differs in the short and in the long run due to the presence of lagged values of the dependent variable on the right-hand side of each equation. In particular, we have found that the short-run effects are substantially lower than the long run-effects.

Given the magnitude of the estimates, we have also found that the time needed for the total effect on diversification of a change in R&D intensity to be completed is substantially greater than the time needed for the overall opposite effect to occur. Specifically, the effects do not occur instantaneously, and it takes several time periods for the overall effects to be completed. The ideas behind these findings should be acknowledged if the effects on performance are analyzed.

The main results also confirm the importance of ensuring a proper treatment of the potential sources of endogeneity in the empirical analysis, since failure of accounting for endogeneity will lead to inconsistent estimates, and therefore the conclusions could be misleading. The most obvious source of endogeneity is the endogenous character of the relationship between related diversification and R&D intensity, which could be circumvented by means of instrumental variable procedures. The two other sources, even though less obvious and less dealt with in empirical work, are the censored nature of the two decision variables, and the existence of firm-specific unobserved heterogeneity. Unobserved heterogeneity should always be taken into account in empirical applications with firm-level data, given the fact that, no matter how many covariates we can control, relevant unobserved characteristics of such firms potentially correlated with the existing covariates will always remain. To the extent of our knowledge, this problem have usually been disregarded in this literature, with some exceptions, although we understand that in many cases this could be due to the lack of longitudinal data. We understand that the proper econometric treatment of these problems may entail a relative loss of precision in the estimates. Nevertheless, this loss of precision should not lead to disregard these problems, because the alternative estimates would be inconsistent, and therefore meaningless, no matter how precise they might be.

Our empirical study is confined to a sample of Spanish firms that only undergo organic growth. With regard to the country specificity of the sample, Mayer and Whittington (2003) suggest that the relationship between diversification and performance is influenced by the economic environment of the country in which the company operates[12]. We should affirm that this idea can be extended to the context of our study. Nevertheless, Mayer and Whittington (2003) suggest that, although the theory of diversification strategies has been motivated largely from the experience of US companies, it can be generalized to Western European environments. Thus, we are prone to believe that our empirical results could be widely applicable within the scope of organic growth.

A sample composed of companies undergoing exclusively organic growth has allowed us to isolate the effect of the growth mode on the diversification strategy, but this also implies a limitation. On the one hand, it is very likely that strategies of external growth or cooperation play a differential role in the relationship under study, possibly opening up an important line of future research. On the other hand, restructuring processes, which have also been disregarded given the nature of our sample, may also modify the conclusions concerning such a relationship. Furthermore, the conclusion by Miller (2004) stating that the companies that diversify are less innovative is not applicable to our sample, since this conclusion is based on a sample of companies carrying out external growth. On the contrary, it is possible to argue that firms that follow a pattern of related diversification are more efficient in the use of the R&D investments (to the extent that they generate economies of scope) up to a certain degree of related diversification. Organic growth leads to more related diversification than external growth (Miller, 2004).

The evidence as to the functional form of the relationship has been provided with R&D intensity as an innovation measure, but it may differ for other measures of technological resources. Consequently, it is worth enquiring into how the relationship would be affected when alternative measures of technological resources, either internal or external measures, either input or output measures, are used. This suggests another subject of future investigation: attempting to test our hypothesis with alternative measures of technological resources.

In summary, the results provide important evidence as to the bidirectional relationship between corporate diversification and R&D intensity in an organic growth context. Our study highlights the importance of accounting for dynamics in the analysis of corporate diversification. Given the potential importance of the conceptual and empirical approach developed in this study, considering dynamics effects in the relationship between resources and corporate diversification, is likely to be fruitful in future theoretical and empirical studies in the analysis of corporate diversification. We believe that our study provides a robust approach for the empirical analysis of strategic decisions in general, and of corporate diversification in particular. In this regard, the econometric approach that we propose here may be a helpful benchmark for future research aimed at evaluating other implications of the RVB, such as the relationship between diversification decisions, resources and corporate performance.

REFERENCES

Alonso-Borrego, C., M.D. Collado. 2002. Innovation and job creation and destruction: Evidence from Spain. Recherches Economiques de Louvain 68 149-168.

Arellano, M., S. Bond. 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., O. Bover. 1995. Another look at the instrumental-variable estimation of error-components models. Journal of Econometrics 68 29-52.

Arellano, M., O. Bover, J.M. Labeaga. 1999. Autoregressive models with sample selectivity for panel data. C. Hsiao, K. Lahiri, L.F. Lee, M.H. Pesaram, eds. Analysis of panels and limited dependent variable models. Cambridge University Press, Cambridge.

Baldwin, W.L., J.T. Scott. 1987. Market structure and technological change. Harwood Academic Publishers, London.

Baysinger, B., R.D. Kosnik, T.A. Turk. 1991. Effects of Board and Ownership Structure on Corporate R&D Strategy. Academy of Management Journal 34 205-214.

Baysinger, B., R.E. Hoskisson. 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., V. Mahajan. 1985. Risk/Return Performance of Diversified Firms. Management Science 31(10) 785-799.

Blundell, R., S. Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87 115-143.

Bover, O., M. Arellano. 1997. Estimating dynamic limited dependent variable models from panel data. Investigaciones Económicas 21 141-165.

Busija, E.C., H.M. O'Neill, C.P. Zeithaml. 1997. Diversification strategy, entry mode and performance: Evidence of choice and constrains. Strategic Management Journal 18 321-328.

Chang, S.J., H. Singh. 1999. The impact of modes of entry and resource fit on modes of exit by multibusiness firms. Strategic Management Journal 20 1019-1035.

Chatterjee, S., S. Singh. 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., B. Wernerfelt. 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., L.G. Thomas. 1993. Direct estimation of synergy: A new approach to the diversity-performance debate. Management Science 39 1334-1346.

Delios, A., P.W. Beamish. 1999. Geographic scope, product diversification, and the corporate performance of Japanese firms. Strategic Management Journal 20 711-727.

Gedajlovic, E., D.M. Shapiro. 2002. Ownership structure and firm profitability in Japan. Academy of Management Journal 45(2) 565-575.

Geringer, J.M., S. Tallman, D. Olsen. 2000. Product and international diversification among Japanese multinational firms. Strategic Management Journal 21 51-80.

Grant, R.M., A.P. Jammine. 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., K.M. Eisenhardt. 2004 Intertemporal economies of scope, organizational modularity, and the dynamics of diversification. Strategic Management Journal 25 1217-1232.

Hill, C.W., G.S. Hansen. 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., S.A. Snell. 1988. External control, corporate strategy, and firm performance in research-intensive industries. Strategic Management Journal 9 577-590.

Hill, C.W., S.A. Snell. 1989. Effects of ownership structure and control on corporate productivity. Academy of Management Journal 32(1) 25-46.

Hitt, M.A., R.E. Hoskisson, R.D. Ireland, J.S. Harrison. 1991. Effects of acquisitions on R&D inputs and outputs. Academy of Management Journal 34(3) 693-706.

Hitt, M.A., R.E. Hoskisson, R.A. Johnson, D.D. Moesel. 1996. The market for corporate control and firm innovation. Academy of Management Journal 39(5) 1084-1119.

Hitt, M.A., R.E. Hoskisson, H. Kim. 1997. International Diversification: Effects on Innovation and Firm Performance in Product-Diversified Firms. Academy of Management Journal 40(4) 767-798.

Holtz-Eakin, D., W. Newey, H. Rosen. 1988. Estimating vector autoregressions with panel data. Econometrica 56 1371-1395.

Hoskisson, R.E., M.A. Hitt. 1988. Strategic control systems and relative R&D investment in large multiproduct firms. Strategic Management Journal 9 605-621.

Hoskisson, R.E., M.A. Hitt, R.A. Johnson, D.D. Moesel. 1993. Construct validity of an objective (entropy) categorical measure of diversification strategy. Strategic Management Journal 14 215-235.

Hoskisson, R.E., R.A. Johnson. 1992. Corporate restructuring and strategic change: The effect on diversification strategy and R&D intensity. Strategic Management Journal 13 625-634.

Jacquemin, A., C. Berry. 1979. Entropy measure of diversification and corporate growth. Journal of Industrial Economics 27(4) 359-369.

Lecraw, D.J. 1984. Diversification Strategy and Performance, Journal of Industrial Economics 33 179-198.

McGahan, A.M., M.E. Porter. 1997. How much does industry matter, really? Strategic Management Journal, Special issue 18 15-30.

Mayer, M., R. Whittington. 2003. Diversification in context: A cross-national and cross-temporal extension. Strategic Management Journal 24 773-781.

Merino, F., D.R. Rodríguez. 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., S. Hariharan. 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.

Palich, L.E., L.B. Cardinal, C.C. Miller. 2000. Curvilinearity in the diversification-performance linkage: An examination of over three decades of research. Strategic Management Journal 21 155-174.

Schoenecker, T.S., A.C. Cooper. 1998. The role of firm resources and organizational attributes in determining entry timing: A cross-industry study. Strategic Management Journal 19 1127-1143.

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., I.M. Duhaime. 1997. Seeing the big picture: The influence of industry, diversification and business strategy on performance. Academy of Management Journal 40 560-583.

Tanriverdi, H., N. Venkatraman. 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.

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., R.E. Hoskisson. 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.

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	
	Schoenecker and Cooper (1998)	Unrelated	-	No distinction	USA	Computer and PC industries
	Silverman (1999)	Related	+	No distinction	USA	Detailed level of technological resources analysis
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, Kosnik and Turk (1991)	Related Unrelated	n.s. n.s.	No distinction	USA	
	Davis and Thomas (1993)	Related	+	No distinction	USA	
	Hill and Snell (1989)	Unrelated	n.s.	No distinction	USA	R&D per employee
	Hitt, Hoskisson, Ireland and Harrison (1991)	Unrelated	-	External	USA	
	Hitt, Hoskisson, Johnson and Moesel (1996)	No distinction	-	External	USA	Diversification as control variable
	Hitt, Hoskisson and Kim (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	

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***

* denotes significance at the 10 percent level; ** denotes significance at the 5 percent level; *** denotes significance at the 1 percent level.

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

Standard deviations in italics

TABLE 4. Preliminary estimates

Variable	R&D equation				Diversification equation			
	OLS	Truncated OLS	Tobit	Generalized selection	OLS	Truncated OLS	Tobit	Generalized selection
R&D (t-1)	0.5814***	0.5593***	0.6632***	0.5669***	0.1504**	0.3666	1.3016***	0.5756**
	(0.0890)	(0.0182)	(0.0181)	(0.0178)	(0.0590)	(0.2980)	(0.0000)	(0.2704)
Unrelated diversification (t-1)	0.0098	0.0024	0.0251**	0.0040	0.0073	-0.0118	0.0989***	0.0201
	(0.0089)	(0.0138)	(0.0122)	(0.0132)	(0.0091)	(0.0548)	(0.0000)	(0.0493)
Squared Unrelated diversification (t-1)	-0.0175	-0.0047	-0.0380*	-0.0071				
	(0.0145	(0.0227	(0.0198	(0.0218				
Related diversification (t-1)	0.0119**	0.0184**	0.0158*	0.0187**	0.8658***	0.5488***	1.7956***	0.8764***
	(0.0060)	(0.0096)	(0.0083)	(0.0092)	(0.0082)	(0.0395)	(0.0000)	(0.0573)
Squared Related diversification (t-1)	-0.0104	-0.0182	-0.0120	-0.0185				
	(0.0082)	(0.0141)	(0.0118)	(0.0135)				
Size	0.0031***	0.0043***	0.0049***	0.0045***	-0.0006	-0.0065	0.0104	-0.0052
	(0.0008)	(0.0012)	(0.0011)	(0.0012)	(0.0032)	(0.0251)	(0.0000)	(0.0225)
Young firm(Y/n)	0.0011	0.0027	0.0008	0.0025	0.0023	-0.0419	-0.0315***	-0.0273
	(0.0011)	(0.0018)	(0.0016)	(0.0017)	(0.0047)	(0.0366)	(0.0000)	(0.0327)
Advertising exp. (t-1)	0.0002***	0.0002	0.0004***	0.0002**	-0.0002	0.0022	-0.0015	0.0007
	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0004)	(0.0031)	(0.0000)	(0.0028)
ln(Sales) (t-1)	-0.0018***	-0.0040***	-0.0002	-0.0037***	0.0014	0.0045	0.0228***	0.0135
	(0.0006)	(0.0010)	(0.0009)	(0.0009)	(0.0026)	(0.0211)	(0.0000)	(0.0188)
Lambda				0.0032**				0.1498***
				(0.0014)				(0.0173)
Number of observations	3508	2248	3508	3508	3508	375	3508	3508
R-squared	0.51	0.51			0.79	0.75		
Chi2				2552.45				1084.49
Log-likelihood	8972.07	5375.46	4478.18		3933.38	280.28	-343.13	
Joint significance test (p-value)	0.025	0.027	0.007	0.069				

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 in italics below each estimated coefficient. The variable *lambda* denotes the selectivity correction term in the generalized selection model. Time dummies, industry dummies and interactions between them 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.
* denotes significance at the 10 percent level; ** denotes significance at the 5 percent level; *** denotes significance at the 1 percent level.

TABLE 5. Dynamic panel data estimation

Variable	Equation	
	R&D	**Diversification**
R&D (t-1)	0.3955***	
	(0.1208)	
R&D (t-2)		0.1783**
		(0.1006)
Unrelated diversification (t-1)	-0.0082	0.0419
	(0.0247)	(0.0616)
Squared Unrelated diversification (t-1)	0.0250	
	(0.0559)	
Related diversification (t-1)	0.0355**	0.6417***
	(0.0177)	(0.0912)
Squared Related diversification (t-1)	-0.0316*	
	(0.0198)	
Size	0.0000	-0.0357
	(0.0051)	(0.0221)
Advertising exp. (t-1)	-0.0003	-0.0023
	(0.0005)	(0.0022)
ln(Sales) (t-1)	0.0031	0.0082
	(0.0045)	(0.0059)
Hansen-Sargan test	233.03	106.97
(p-value)	0.20	0.03
AR(2) test	0.15	-1.50
(p-value)	0.88	0.13

The lag (t-1 or t-2) is indicated to the left of each variable.
Standard errors in italics below each estimated coefficient.
Time dummies were included in all estimations.
Both the Hansen test and the AR(2) test are specification tests that help to evaluate the validity of the estimates. 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) as a chi-square with as many degrees of freedom as the number of over-identifying restrictions. In our context, this test of over-identifying restriction can be viewed as a test for instrument validity. If the instruments used in GMM estimation are valid, then the Hansen-Sargan test ought to be statistically equal to zero.
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 cannot show second order autocorrelation.
* denotes significance at the 10 percent level; ** denotes significance at the 5 percent level; *** denotes significance at the 1 percent level.

TABLE 6. Long run vs. short run effects

	Related Diversification → R&D	**R&D → Related Diversification**
Long run effect	**0.053**	**0.498**
Short run effect	0.032	0.178
Percentage of the short run effect with respect to the overall (long run) effect	60.4	40.4
Number of approximate periods after 95% of the overall effect has taken place*	4	8

Calculations based on dynamic panel data estimates from Table 5.
*This calculation is done taking into account that the accumulated effect after *s* periods is given by $\beta(1-\rho^{s})/1-\rho$, whereas the total long run effect is $\beta/(1-\rho)$. Therefore, the number of time periods after which 95% of the effect has taken place is $s=\ln(1-0.95)/\ln\rho$, which has been rounded to the nearest integer. Notice that it should also be taken into account that the variable for which the effect is being considered is lagged one period.

APPENDIX. DISTRIBUTION OF COMPANIES BY INDUSTRY AND BY SIZE

Table A1. 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).

[1] Growth modes affect the extent of diversification (Busija, O'Neill and Zeitthaml 1997, Chang and Singh 1999, Chatterjee and Singh 1997), and are associated with R&D intensity (Miller 2004).

[2] Those firms which diversify through organic growth use a highly specific set of resources, as for example, technological resources. On the other hand, organic growth usually encourages related diversification. Those firms with inferior technological resources will be prone to unrelated diversification towards external growth (Miller 2004).

[3] The ESEE is produced by *Fundación Empresa Pública*, a public institute financed by the Spanish Ministry of Industry. The original data set was designed with the aim of ensuring the representativeness of Spanish manufacturing firms. For this purpose, all the companies with more than 200 employees were surveyed (and, accordingly with the information provided by those responsible for the data set, about 70% completed the survey), and smaller companies with more than 10 employees were selected on the basis of a stratified sampling.

[4] The rationale for these last two restrictions is that, in general, the potential role of diversification (and probably innovation) is quite scarce in the case of companies focused on local markets and/or without a sufficient size.

[5] Notice that OLS estimates with the full sample will be inconsistent because of 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, since the sample selection bias, after excluding censored observations, is not taken into account.

[6] It is well known that when ignoring unobserved heterogeneity, the estimated coefficient of the lagged dependent variable is typically oversized in absolute value.

[7] Regarding the diversification equation, the likelihood of the tobit model did not converge, so that their associated estimates are completely unreliable. Behind this fact, there exists evidence that the constraints that the tobit model imposes do not hold.

[8] Any variable which kept constant over time after applying the fixed effects transformation was not considered, since the fixed-effects transformation would drop it. In particular, age effects could not be taken into account.

[9] Such a loss of precision is due to two reasons: the transformation needed to remove the unobserved individual heterogeneity component as well as the use of predictions of the latent dependent variable instead of its (unobserved) true value in order to take into account censoring in the dependent variable.

[10] Letting β be the short-run effect, and ρ the coefficient of the lagged dependent variable, the long-run effect is calculated as $\beta/(1-\rho)$.

[11] In particular, since the coefficient of lagged R&D in the generalized selection estimates was around 0.58 (which appeared to be unbelievably large), the implied long-run effect was about 5.

[12] 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.