R&D-EXPERIENCE AND INNOVATION SUCCESS*

Pilar Beneito[a], María Engracia Rochina Barrachina[b], Amparo Sanchis Llopis[b]

ABSTRACT

This paper analyses the role of firms' R&D-experience in their innovative success using a representative sample of Spanish firms for the period 1990-2002. Using count data models and within an innovation production function approach, we investigate the influence of firms' R&D-experience in the achievement of innovation results. To estimate R&D-experience, partially unobserved, we estimate a duration model and use the obtained results and a non-parametric procedure to impute R&D-experience when unobserved. We obtain that R&D effectiveness increases along the R&D history of the firm.

Keywords: innovation, accumulation of knowledge, R&D-experience, duration models, count data models.

JEL Classification: O30, O34, C23, C10.

*The authors acknowledge financial support from IVIE, and grants SEJ2005-05966, SEJ2005-08783-C04-01 and GVACOMP2007-132.

[a] **Corresponding author**: Pilar Beneito, Universidad de Valencia, Facultad de Economía, Departamento de Análisis Económico, Avda. de los Naranjos s/n, 46022 Valencia (Spain); telephone: 0034 963828223, e-mail address: pilar.beneito@uv.es.

[b] Universidad de Valencia and *ERI-CES*.

"*Learning is the product of experience. Learning can only take place through the attempt to solve a problem and therefore only takes place during activity.*" (Arrow, 1962, p.155).

"*I advance the hypothesis here that technical change in general can be ascribed to experience, that it is the very activity of production which gives rise to problems for which favourable responses are selected over time*." (Arrow, 1962, p.156).

1. INTRODUCTION

Among the still scarce empirical literature on the dynamics of firms' innovation behaviour, very recently a new strand of analysis has focused on *persistence*, or patterns of continuity, in innovation activities and results at the firm-level. From a theoretical point of view, there are three main hypotheses suggesting why firm innovative behaviour should exhibit persistence. First, innovative persistence may result from the existence of sunk costs associated with the performance of R&D activities (Sutton, 1991). These costs arise from the establishment of an R&D department, the purchasing of specific assets, and/or the hiring and training of specialized workforce, and may lead firms to spread innovation expenditures over a period of time, causing persistence. The second hypothesis states that "success-breeds-success": past innovations raise the probability to innovate again, i.e., innovative success generates profits that may be reinvested in future R&D activities (Mansfield, 1968, Stoneman, 1983). Finally, a third hypothesis refers to the existence of *dynamic increasing returns* in innovation. This hypothesis, coming from the evolutionary theory, emphasises the cumulative nature of the learning process (Rosenberg, 1976; Nelson and Winter, 1982): the generation of knowledge is based on previous knowledge and affects future research. This stream of literature considers that innovations are the result of a process of accumulation of firms' specific competencies (Rosenberg, 1976). In particular, by investing in R&D projects, firms develop abilities in the form of knowledge, both scientific and informal know-how, which may be used to develop further innovations at consecutive times. According to this view, firms benefit from *dynamic increasing returns* in the form of learning-by-doing, learning-to-learn or scope economies in the production of innovations (Cohen and Levinthal, 1989).

The existing empirical studies on innovation persistence have mainly focused on the analysis of patterns of continuity in the achievement of innovations, that is, the analysis of persistence in innovation output, usually measured as the number of patents and/or major innovations (Geroski *et al.*, 1997; Crépon and Duguet, 1997; Malerba and Orsenigo, 1999; Cefis and Orsenigo, 2001; Cefis, 2003, Raymond et al., 2006). The analysis of innovation persistence in the realization of innovation activities, or input persistence, is still very scarce (Máñez et al., 2004, 2006, and Peters, 2005).

How do the theoretical and the empirical approaches to innovation persistence match? On the one hand, it seems that the hypothesis of sunk costs associated with R&D investments should imply *input persistence*, that is, the observation of current R&D expenditures will be followed by the observation of future R&D expenditures. In particular, the hypothesis of sunk costs has been tested by Máñez et al., 2004, who analyzed innovation input persistence using firms' engagement in R&D activities.

On the other hand, if the "success-breeds-success" hypothesis suggests that past innovations raise the probability to innovate again, we should observe some degree of persistence in the achievement of innovations by firms. This approach has been assessed using innovation output measures, such as the number of patents or the number of major product or process innovations (see, e.g., Flaig and Stadler, 1994, 1998).

Finally, the hypothesis of *dynamic increasing returns* has been used to justify the empirical analysis of both innovation input and innovation output persistence. However, we consider that the empirical testing of this hypothesis is more complex and can not be solely based on the analysis of persistence of one of these two sides of innovation. As pointed out before, the hypothesis of *dynamic increasing returns* is based on the idea that the learning process is cumulative, and that innovations are the result of a process of accumulation of firms' specific competencies, both scientific and informal know-how, which determine their capabilities to achieve successful innovations. Therefore, we argue that the empirical hypothesis to be tested in this case is whether or not the continuity or persistence in the development of R&D activities affects the probability of successful innovation results.

Following Arrow (1962) and evolutionary theorists, in this paper we hypothesize that experience in the process of accumulation of technological knowledge (learning) is one of the main sources of dynamic increasing returns in innovation. Innovation experience could be defined either as experience in the consecution of innovations, or as experience in the realisation of innovative efforts, such as performance of R&D activities. We consider innovation experience as experience accumulated through the performance of R&D activities, irrespective of the achievement of innovation results in a given period, since the process of knowledge accumulation may even result in that "failure-breeds-success". We argue that the time dimension of the cumulative process of R&D knowledge, that is, technical skills and learning-by-doing accumulated through time, may not be properly measured by the standard inputs considered in the empirical literature, such as R&D expenditures or R&D capital stock. Our hypothesis is that R&D-experience, understood as time devoted to the performance of R&D activities, is a key driver in the innovation success: we consider that the effect of R&D in the achievement of innovations depends on

R&D-experience, that is, on the period of time during which the firm has been engaged in R&D activities.

The aim of this paper is not only to address whether or not R&D-experience affects the firms' innovation success, but also to measure R&D-experience effectiveness, that is, to determine to what extent the rate at which R&D investments yields innovation output depends upon firms' accumulated R&D-experience.

To the best of our knowledge, there is a lack of empirical evidence that explicitly deals with the measure of R&D-experience effectiveness, although the importance of continuity in R&D efforts has been stated in some empirical works. As an example, West and Iansiti (2003) assert that experience and experimentation are significantly correlated with innovation performance in the semiconductor industry, mainly because experience reduces search time and search costs. Their empirical setting (using a sample of 29 research projects in the semiconductor industry in the US) relies on a dummy variable indicating if the project members have previous experience in technology selection decisions. Still, a direct measure of R&D-experience is absent.[1]

The lack of empirical evidence in measuring R&D-experience returns is likely to be also due to data restrictions: innovation surveys do not usually report, in a retrospective way, the number of years the firm has been carrying out innovative activities. In this paper we use a representative panel sample of the population of Spanish manufacturing firms for the period 1990 to 2002, drawn from the *Encuesta sobre Estrategias Empresariales* (ESEE, henceforth). Although we also lack retrospective information on firms' R&D histories, we pay special attention to the empirical estimation of our measure of R&D-experience. For

[1] More indirectly, the work of Griffith, et. al. (2006) suggests the importance of R&D-experience by including in their estimation of firms' productivity an indicator for whether the firm answered yes to conducting R&D continuously. The same kind of indicator is used by Raymond, et al. (2006) to explain innovative sales.

this purpose, we first analyse firms R&D patterns in order to determine the duration of firms' R&D spells, i.e. periods of time during which firms perform R&D activities in a continuous way. The estimation of this duration model allows us to identify firm and industry characteristics affecting R&D histories, information which, in turn, is used to make an estimation of the R&D-experience measure. Once we have estimated the R&D-experience of firms, we proceed to estimate, within the framework of an innovation production function and using count data models, the influence of firms' accumulated R&D-experience on their R&D innovative effectiveness. In order to do this, we treat R&D-experience as a moderator variable and investigate how it influences the impact of R&D capital on firm innovation success.

Our work is also connected with that strand of the empirical literature that has focused on the analysis of the relationship between firms' R&D input (measured as R&D capital stock, R&D expenditures, or as the ratio of R&D expenditures to sales or revenues) and innovation output (measured, e.g., in terms of patents or productivity). In particular, the relationship between innovation, R&D and patents has been surveyed by Griliches (1990), who reports a robust R&D-patents relationship at the firm level.[2] More recently, the availability of CIS surveys has given rise to a number of empirical works that also analyse the innovative performance of firms by relating innovation inputs to innovation outputs[3]. However, these empirical studies do not explicitly take into account the possibility that the

[2] Among the most well known works are those of Schmookler (1966, ch. 2), Scherer (1965), Bound et al. (1984), Hausman, Hall and Griliches (1984), Hall, Griliches and Hausman (1986), Pakes and Griliches (1984), Scherer (1983) and Acs and Audretsch (1989). See also Henderson and Cockburn (1993), Branstetter (1996) and Crépon *et al*. (1998).

[3] Some of these works are Klomp and van Leeuwen (2001) for the Netherlands, Smith and Sandven (2001) for Norway, Lööf and Heshmati (2001) for Sweden, or Mairesse and Mohnen (2005) and Kremp and Mairesse (2004) for France.

effectiveness of the innovation inputs changes as firms accumulate experience in the performance of their innovation activities.

Therefore, this paper is the first attempt to empirically address, in a direct and explicit way, the role of firms R&D-experience in their innovation success, and this is the main contribution of this paper to the existing literature. To anticipate our results we obtain that, after controlling for R&D capital stock and other firms' individual heterogeneity, firms' R&D effectiveness rises with R&D-experience, that is, with the accumulation of technical skills and knowledge that emerge as firms invest in R&D over time. In addition to past R&D-experience, we also find that the performance of informal innovation activities, and the technological intensity of the industry in which the firm operates, are significant determinants in the achievement of innovations.

The rest of the paper is organised as follows. In section two we present the empirical model and the econometric procedure, where we outline the empirical framework we use throughout the paper. Section 3 presents the data. Section 4 is devoted to the estimation of firms' R&D-experience. Section 5 describes the estimation of the innovation production function. Finally, section 6 concludes.

2. EMPIRICAL MODEL AND ECONOMETRIC PROCEDURE

Our main hypothesis to be tested relies on the idea that the effectiveness of R&D activities may vary with the R&D-experience of the firm, that is, with the accumulation of knowledge that takes place along with the research effort that is undertaken. Technical skills and learning-by-doing accumulated with time may not be properly measured by the standard R&D inputs considered by the empirical literature that has tried to explain the factors underlying the achievement of innovation results. In this paper we attempt to measure the extent to which this R&D-experience matters in determining the effectiveness of R&D activities. Our approach is based on the concept of an *innovation production function* that may, in a very general form, be expressed as follows

$$N_{it} = f(x_{it}, \beta) \quad (1)$$

where *i* refers to the firm and *t* to the time period, N_{it} stands for any chosen indicator of innovation outcomes and x_{it} represents the vector of innovation inputs in the equation. A usual component of x_{it} are R&D inputs, quite often measured by R&D capital. Our innovation production function will differ from the standard one in that the effectiveness of R&D capital is specified as a function of the R&D-experience of the firm. In particular, the parameter vector β may be decomposed as

$$\beta = [\beta_1(E_{it}), \beta_2] \quad (2)$$

where β_1 is the parameter that measures the "*innovative effectiveness*" of the R&D input, E_{it} stands for firms' R&D-experience, and β_2 stands for other inputs' parameters. Therefore, the effect of R&D in the achievement of innovation outcomes depends on R&D-experience, that is, on the time the firm has been engaged in R&D activities.

The econometric approach to estimate the parameters in (1) is conditioned by the kind of data used to measure innovation success, that is, the output of the innovation process (N_{it}). By far, the measure used more frequently is the number of patents registered by the firm. In this paper, two alternative measures of innovation output will be used: the number of patents registered, and the number of product innovations introduced by the firm during the period under analysis. These two measures share two common features: both of them are event counts (non-negative integers) for unit *i* during time period *t*, and in any given year many firms do not register patents or do not introduce innovations.

It is standard in the literature to assume that the Poisson distribution is a reasonable description for count data. According to the Poisson process, research results are the outcome of an unknown number of Bernoulli trials with a small probability of success. The basic Poisson probability specification is

$$\Pr(N_{it} = n_{it}) = f(n_{it}) = \frac{e^{-\lambda_{it}} \lambda_{it}^{n_{it}}}{n_{it}!} \qquad (3)$$

We may model the single parameter of the Poisson distribution function, λ, as a function of our explanatory variables, *x*, and parameters, β, in the standard fashion[4]

$$\lambda_{it} = \exp(x_{it}\beta) \qquad (4)$$

It is easily shown that

$$E[N_{it} | x_{it}] = Var[N_{it} | x_{it}] = \lambda_{it} = \exp(x_{it}\beta) \qquad (5)$$

so that λ_{it} represents the arrival rate of innovations per firm per year and also the expected number of innovation outcomes per firm per year. Taking logs in (5) we get

$$\log E[N_{it} | x_{it}] = \log \lambda_{it} = x_{it}\beta \qquad (6)$$

If the explanatory variables are used in logs, the estimated β are the elasticities of the expected number of innovations with respect to these variables. We will consider $x_{it} = (R_{it}, E_{it}, z_{it})$ where R_{it} is knowledge or R&D capital (derived from the flow of real R&D investments),[5] E_{it} is the firm's R&D-experience, and z_{it} stands for an index of other inputs and control variables.

In our case, we assume that expression (5) takes the form

---

[4] Note that λ_{it} is a deterministic function of x_{it} and the randomness in the model comes from the Poisson specification for N_{it}.

[5] For a discussion on the use and construction of the R&D stock measure (the so-called R&D capital), see, for example, Hall and Mairesse (1995). The use of the stock measure has, at least, two advantages as compared with the use of R&D flows: it avoids making assumptions about distributed lags while being somewhat equivalent to imposing a geometric lag structure, and it prevents from having to drop much of the data in order to have a given number of lags in the R&D spending pattern of firms. Details about how we construct this stock are given in Table 1.

$$\lambda_{it} = A(t) R_{it}^{\beta_1(E_{it})} \exp(z_{it}\beta_2) \tag{7}$$

that is, the estimated function has a direct proportionate relationship between the R&D capital and innovation counts moderated by a multiplicative set of variables hypothesized to shift the distribution of expected innovation results. The impact of R&D capital on the rate of innovation is assumed to be a function of the R&D-experience of the firm. This function may be non-linear, so in order to allow for a non-linear relationship we assume the following quadratic form

$$\beta_1(E_{it}) = \alpha_0 + \alpha_1 E_{it} + \alpha_2 E_{it}^2 \tag{8}$$

Formally, β_1 is defined as the percentage change in innovation output generated by one percent change in R&D capital. Thus, this elasticity represents the effectiveness of R&D capital, moderated by R&D-experience, in obtaining innovation outputs, such as product innovations or patents. Note that α_0 would be the standard elasticity parameter if R&D-experience would not matter for R&D success. In addition, α_1 captures the impact of firm's R&D-experience on R&D effectiveness and α_2 is the change in the impact of firm's R&D-experience on R&D effectiveness.

Substituting expression (8) into (7) gives

$$\lambda_{it} = A(t) R_{it}^{(\alpha_0 + \alpha_1 E_{it} + \alpha_2 E_{it}^2)} \exp(z_{it}\beta_2) \tag{9}$$

and, taking logs,

$$\begin{aligned} \log \lambda_{it} &= \log A(t) + (\alpha_0 + \alpha_1 E_{it} + \alpha_2 E_{it}^2) \log R_{it} + z_{it}\beta_2 = \\ &\log A(t) + \alpha_0 \log R_{it} + \alpha_1 E_{it} \log R_{it} + \alpha_2 E_{it}^2 \log R_{it} + z_{it}\beta_2 \end{aligned} \tag{10}$$

We use equation (10) to examine the effect of firm's R&D-experience on R&D effectiveness. If the estimate of α_2 is significantly positive (negative), then the relationship between R&D effectiveness and firm's R&D-experience approximates to a "U-type" (inverted "U-type") relationship. However, if the estimate of α_1 is significantly different from zero but the estimate of α_2 is not significant, then firm' R&D effectiveness is a monotonically increasing or decreasing function of firm's R&D-experience.

In order to proceed further we need to address an important limitation of our measure of firms' R&D-experience. The problem we face is an empirical one: our data set does not include (similar to most of innovation surveys) retrospective information about the R&D history of firms, that is, when a firm enters the survey we do not have information on how many years this firm has been undertaking R&D activities. To see this problem more clearly we can have a look at Figure 1 (which will be explained in more detail in section 4.1). In this figure, the horizontal axis shows the passage of time, and the length of each horizontal line shows the time spent on performing R&D activities. If the year 1990 represents the first year a firm is observed, and the firm reports R&D investment for that year, we do not have information on whether the firm has been investing in R&D during previous years. This lack of retrospective information brings about a serious limitation to our possibility of measuring the R&D-experience of this type of firms. To deal with this problem, we implement a procedure to estimate such (*left*) *censored* R&D-experiences. Our procedure is developed in three consecutive steps: first, we identify the factors underlying the length of firms' R&D-spells, that is, the number of uninterrupted periods of R&D activities. To this end, we estimate a discrete time duration model.[6] The essence of duration models is to analyse the length of time that an individual spends in a relevant state (in our case, the length of time a firm performs R&D activities) before experiencing the exit from that state to another state (in our case the cease in R&D activities).

[6] The nature of our data set lead us to consider time as a discrete variable, not because it is intrinsically discrete but because the data in the survey is provided on a yearly basis.

Secondly, once we have the estimates associated with the factors explaining spells duration, we proceed to predict the total duration of spells that are still in progress at the end of our sample period (*right-censored* spells). At this stage, we should obtain values for two key statistical functions in duration analysis: the hazard function, which in discrete-time is defined as the probability of transition out of a state at each discrete point in time *t* given survival up to that point, and the discrete-time survival function, $S(t) = \Pr[T > t]$, which is the probability that the duration of the spell is higher than *t*. The mean or *expected duration* of a spell is the sum (in the case of discrete time) of the survival function evaluated at survival time 1 up to the maximum survival time (when the survival function reaches the value of 0).

Finally, we focus on the duration of R&D spells that were in progress at the moment the firms were firstly observed in our sample (the *left-censored* spells). For them, we cannot proceed as for right-censored spells because the key statistical measures in a duration model are *conditional on past information* until *t*, and for left-censored spells we do not have this necessary past information. Thus, for this type of spells we carry out a matching approach to impute them R&D spells durations non-parametrically. In this final step we use the information about both completed spells durations and the estimated total durations for right censored spells. Through non-parametric regression (kernel regression) we impute to each left censored spell a duration equivalent to a weighted sum of complete durations (either originally observed or estimated, as it is the case for right censored spells), with weights based on similarities in the firm and industry characteristics used in the duration model.

To sum up, our empirical procedure will proceed as follows: first, we estimate a duration model to identify firm and industry characteristics that affect R&D durations; secondly, we predict expected durations for right censored spells; thirdly, we use information on spells duration for observed complete and estimated (right censored) spells to impute durations, by non-parametric regression methods, to left-censored spells. Finally, once a measure of our R&D-experience variable is available, we estimate a panel count data model for our *innovation production function*. Section 4 presents and explains in more depth these empirical steps.

3. DATA

The data are drawn from the ESEE, a representative annual survey of Spanish manufacturing firms carried out since 1990.[7] The sampling procedure of the ESEE is the following. In the base year, 1990, firms were chosen using a selective sampling scheme with different participation rates depending on firm size. All firms with more than 200 employees (large firms) were requested to participate and the participation rate reached approximately 70% of the number of firms in the population. Firms that employed between 10 and 200 (small firms) were randomly sampled by industry and size strata, holding around 5% of the population.[8] Important efforts have been made to minimise attrition and to annually incorporate new firms with the same sampling criteria as in the base year so that the sample of firms remains representative of the Spanish manufacturing industry over time.

The sample used in this paper covers the period 1990-2002. To this sample, we have applied the following selection criteria. First, we have dropped out from the data those firms which do not respond to the questionnaire in some of the panel years, as well as those firms that have experienced any ownership change process such as a merger or absorption. Secondly, we have removed firms with missing observations in the R&D variables, and

[7] See http://www.funep.es for a more detailed description of the ESEE.

[8] Firms with less than 10 employees in 1990 were not included in the survey.

selected those which have undertaken R&D activities at least during one of the observed periods. As a result, we are endowed with a sample of 6,627 observations, corresponding to 671 firms.[9]

4. THE ESTIMATION OF R&D-EXPERIENCE

4.1. R&D Duration Model

Duration models analyse the length of time that an individual spends in a relevant state before experiencing the transition to another state. In the case of the study of R&D activities, it consists in the analysis of the period of time for which a firm uninterruptedly performs R&D activities. The unit of observation in this section is the R&D spell, defined as the number of uninterrupted years a firm performs R&D activities. Figure 1 presents our observation window (period of time for which we follow firms R&D patterns), corresponding to the period 1990-2002, and it provides visual and simplified information about the sample distribution, number and types of R&D spells. The total number of R&D spells in our sample is 985 spells.

[Insert Figure 1 about here]

We refer to censored R&D spells as those spells for which we do not observe their initial and/or final date, that is, those spells for which we do not know their exact length. We denote with T_e (elapsed duration) the length of time from the beginning of the spell still

[9] The number of observations in the empirical applications outlined in sections 4 and 5 may vary due to additional missing data in the variables used. The details about the data used in each of these applications, as well as some descriptive statistics on the variables of interest in each case, will be given in the corresponding sections below.

in progress at the time the firm is incorporated to the survey, to the year of incorporation.[10] We denote with T_o (observed duration) the observed spell duration over the observation window, and T_r (remaining duration) the length of time from 2002 to the end of the R&D spell. The lines in Figure 1 represent the different types of R&D spells firms may exhibit. The actual duration of the spell, T^*, is measured by the length of the line. There are four categories of R&D spells in relation to censoring. The first relates to completed spells (not censored), representing the 36.5% of the total number of sample spells, for which their full length is known and no problem of either right or left censoring arises ($T^*=T_o$). The second category corresponds to right censored spells (22.2% of the total), for which T_r is not observed. Given that the observation window is finite, some spells are still in progress at the end of the period analyzed, and therefore, these spells are only partially observed (we only observe T_o from the actual duration $T^*=T_o+T_r$). In the estimation of duration models the likelihood contribution of right censored spells can be easily handled.

The third category refers to left censored spells (20.4% of the total), for which T_e is not observed. As an example, consider the year 1990 (starting date of the survey). In this year, firms were asked to declare whether or not they invested in R&D activities. However, due to the lack of retrospective information on R&D activities performed by the firm previously to 1990, the starting date of the spells that were in progress in 1990 is unknown. Attempts to correct for left-censoring in empirical applications of duration models are rare, mainly because it is a very complex issue due to the lack of information on the value of the survival time, i.e., the previous duration of the spell (this is a kind of initial conditions

[10] To make simpler the typological representation of spells in Figure 1, we made elapsed duration to coincide with unobserved periods before the year 1990. However, for firms incorporated to the survey later than 1990, the elapsed duration period reaches the corresponding year of incorporation.

problem). The standard approach to handle this problem in the estimation of duration models is to discard all left-censored spells.

Finally, the fourth category corresponds to left-and-right censored spells (20.8% of the sample), for which both T_e and T_r are unobserved. This censoring case presents the same complexity than left censoring and it is usually handled in a similar way. We follow the standard approach and in the estimation of our duration model we only use those spells that are either complete or right censored (which sum up to 58.7% of total spells). The inclusion in the estimation of the duration model of the observed durations of left censored spells and left-and-right censored spells would have lead to a well known underestimation of the duration of the R&D spells.

To choose the appropriate econometric model for estimation we should treat time either as continuous or discrete. In general, it is assumed that the transition out from one state may occur at any particular instant in time, thus the stochastic process generating durations occurs in *continuous* time. However, as pointed out by Jenkins (2004), survival time is not necessarily a continuous variable and, as in our case, it is not intrinsically discrete but it is observed in discrete intervals. In the ESEE the data is recorded yearly and, therefore, we do not have information on the exact date at which a firm starts or ends an R&D spell, we only know (at most) the starting or the exiting year.

Further, we have to choose a family of duration models, that is, to choose between accelerated failure time models and proportional hazard duration models. We select a proportional hazard model, where the baseline hazard function (controlling for duration dependence) depends only on survival time t and multiplies an exponential component that incorporates the explanatory variables (covariates) and it is not a function of survival time t. The reason to select a proportional hazard model is threefold. First, it is the most widely used formulation in duration analysis. Secondly, it allows for a nice interpretation of the coefficients in the hazard function. That is, in proportional hazard models the estimated coefficients are interpreted as the proportionate response of the hazard to a unit change in a given covariate. However, this nice interpretation is lost when the model incorporates individual unobserved heterogeneity, in which case it is called a mixed proportional hazard model. Thirdly, it can be easily extended to include time varying covariates and unobserved individual heterogeneity, and it allows treating time both in continuous and discrete terms.

Therefore, the implemented duration model is a discrete time proportional hazard model that aims at capturing the particular nature of the dataset (the econometric procedure to estimate this model is explained in Appendix A). This duration model is estimated using both completed and right censored spells. We start up for estimation with 579 spells (1666 observations). After deleting observations for which some of the relevant variables were missing, we end up with 1653 observations corresponding to 569 spells. The specification of the model includes a number of variables that are considered to be relevant in determining the continuity of the performance of R&D activities. In addition, given that the duration model is a first step in order to obtain the parameter estimates that will be used for prediction purposes in the next sections, we have avoided the inclusion of highly time varying variables and/or variables with a clearly increasing or decreasing trend.

Table 1 presents a definition of all the variables used in the estimation. Among these variables we have included industry dummies to control for technological opportunities, appropriability conditions and spillovers at the industry level. We have also included a dummy variable indicating whether the firm sells in foreign markets, which may capture economic opportunities and competitive pressure; the age and size of the firm; its ownership structure; its R&D intensity and R&D workforce ratio; and, finally, a measure of regional and local spillovers.

[Insert Table 1 about here]

Table 2 shows the estimation results for the discrete time proportional hazard *cloglog* model. We find evidence of unobserved individual heterogeneity given that the hypothesis of the unobserved heterogeneity variance component (σ^2) being equal to zero is rejected at a 7% significance level. Furthermore, once controlling for unobserved individual heterogeneity, the duration dependence parameter is not significantly different from zero.

[Insert Table 2 about here]

For continuous variables, the interpretation of the estimated coefficients is as follows. A positive (negative) coefficient means that if the corresponding covariate increases, the hazard risk of ending the spell rises (decreases), and so, the expected duration of the R&D spell decreases (rises). The interpretation is analogous for sets of dummy variables, but in this case we do not generally refer to increases of the covariates but to the way in which the hazard (duration) is affected when the firm belongs to different components of each set of dummies.

According to our results, there are only two industries, *Leather and shoes* and *Motors and cars*, showing a differential longer R&D spell duration. Exporting firms experience longer R&D spells. This may indicate that firms in more competitive markets have greater incentives to undertake R&D activities in a continuous way in order to maintain market competitiveness and high quality standard products required by international markets (Kleinschmidt and Cooper, 1990, and Kotable, 1990). Firms' age increases the probability of experiencing longer R&D spells in a non linear manner. It is especially remarkable the effect on spell length for firms between 40 to 50 years old. For firms with more than 50 years the effect of age on duration decreases considerably and also the significance level with which this coefficient is estimated. These results are consistent with Huergo and Jaumandreu (2004) who also found a non linear effect of age and the probability to innovate.

Regarding the relationship between firm's size and R&D investments, our results confirm that R&D spells of larger firms have lower chances of ending. Arguments related to superior firm internal capabilities associated with size, such as exploitation of economies of scale and scope, larger market size, lower risk, higher appropriability conditions, financial means, etc., are the usual arguments to support a positive association between firm size and innovative activities in general. The coefficients corresponding to the two included size groups (the excluded one is the group with less or equal to 100 employees) are negative and significant, justifying then a lower ending risk and consequently a longer spell duration. However, the impact of firm size on the length of the R&D spell is not linear, as the comparison of both coefficients suggests that R&D spells of firms with more than 200 employees (size200) endure better survival prospects than firms between 100 and 200 employees.[11]

Firms that are not legally organized as a limited liability corporation have shorter R&D spells. This result is consistent with the hypothesis that these firms are relatively more risk averse (as compared to managed firms) and thus less willing to undertake risky investments such as R&D activities (Love *et al.*, 1996).

In relation to R&D intensity and the nature of the R&D investments, we have included two different measures. The first is the yearly ratio of R&D expenditure over sales and the second the yearly ratio of R&D employees over total number of employees in the firm. The greater these two ratios, the more the firm is expected to perform R&D activities in a continuous way. Both measures may be capturing the extent of sunk costs in which firms incur when undertaking R&D projects (Cohen and Klepper, 1996). According to our results, those firms in medium/high R&D intensity industries enjoy R&D spells with longer

[11] We obtain that, in absolute value, the negative coefficient of size100200 is significantly smaller than the coefficient of size200.

survival prospects, as compared to those firms in low R&D intensity industries (the coefficient for medium/high R&D intensity is negative and significant at 1% level). As regards the ratio of R&D specialized workforce, which may also capture technological opportunities, we find a very strong effect in decreasing the risk of ending an R&D spell, contributing then to explain longer spells duration. This variable has appeared to be the best one in capturing the internal nature of the R&D activities.

Finally, the literature on R&D has stressed the importance of spillovers on the decision to innovate. We find evidence of regional spillovers increasing the R&D spell duration. Local spillovers do not seem to be relevant, and industrial spillovers cannot be separately identified in the estimation from the industry dummies.

To conclude this section, we evaluate the goodness of fit of the duration model in Table 2. Provided that the hazard function with discrete time has the interpretation of a conditional probability (which lies between 0 and 1), we use for this purpose the information on predicted hazards. For each firm on each of the survival years of a given spell, we use the predicted value of the hazard to classify the firm as a firm ending the spell in a particular survival year or continuing the spell at least one more year. Given that the hazard is defined as the probability of ending the spell in year *t* provided the spell has lasted until *t-1*, a firm observation in a given spell is classified as continuing the spell when the predicted hazard for abandoning it is lower than 0.5, and as finishing the spell when this predicted hazard is higher than 0.5. This classification rule allows classifying correctly 70.7% of the exit/no-exit statuses for firms' observations along spells.

4.2 Out- of-sample Prediction for Right Censored Spells

Once the parameters from the duration model have been estimated, we are interested in computing the average duration of right censored R&D spells for firms with different characteristics. To do this, we need to know the shape of the survival function. As Jenkins (2004) has noticed, there are not typically closed form expressions for the mean in discrete time models, requiring then numerical solutions. In general,

$$E\left(T_i^*\right) = \sum_{k=1}^{J} S_i\left(k\right) \tag{11}$$

where J is the maximum survival time. The corresponding discrete time survival function (see in appendix A the *cloglog* hazard function given in A1 and A6) is

$$\begin{aligned} S_i\left(k\right) &= \left[1-h_1\left(x_{i1}\right)\right]\cdot\left[1-h_2\left(x_{i2}\right)\right]\ldots\left[1-h_k\left(x_{ik}\right)\right] = \\ &\exp\left\{\sum_{s=1}^{k} \ln\left[1-h_s\left(x_{is}\right)\right]\right\} \end{aligned} \tag{12}$$

Therefore, for right censored spells we may obtain the mean expected duration of the spell by predicting the hazard rates, given the values of the covariates and the value of k (survival time) in the relevant spell years. This allows generating survival function values per spell-years, and aggregating them until the maximum survival time (spell-year in which the survival probability is zero). This is what Jenkins (2004) calls "out of sample prediction", which requires from the model a parametric specification of the baseline hazard to be able to project to the future, with the model estimates, for right censored spells.

In our sample there are 219 right censored spells with observed durations from 1 to 13 years.[12] The distribution of observed durations for these spells can be found in Table 3. For all these spells we are going to calculate the value of the survival function from survival time 1 to survival time 200 (survival time that guaranties that for all the right censored spells the survival function value reaches 0). For the observed survival periods, the value of that function is calculated with the parameter estimates in the duration model applied to the value of the explanatory variables of any given firm in that survival time period. For the

[12] The 13 years observed spell length for right censored spells corresponds to firms that were born in 1990, and already in this year and in all the subsequent years, including the year 2002, claimed to invest in R&D.

non-observed survival periods in the future, we fix the values of the explanatory variables at their values in the observed final year (2002 for all of them), with the exception of the variable *log(t)* (log of the survival time) that before taking logs it is increased by one each considered extra year of the spell. Among other things, the need to project to the future for right censored spells was already conditioning the type of variables to be included in the first step estimation (duration model). We tried to capture main characteristics of the firms without the inclusion of highly time varying variables and/or variables with a clearly increasing or decreasing trend. The only exception was for the variable survival time itself, which value should increase by one each spell period. We did the full procedure for all the right censored spells, which graphical representation may be found in Figure 2.

[Insert Table 3 about here]

[Insert Figure 2 about here]

Finally, we imputed as the total spell duration for a right censored spell the already observed number or years plus the expected duration remaining afterwards. That is, for instance, for right censored spells which observed duration is of 13 years we apply the formula in (11) to get as expected spell duration

$$E\left(T_i^*\right) = T_o + \sum_{k=(T_o+1)}^{200} S_i\left(k\right) = 13 + \sum_{k=14}^{200} S_i\left(k\right) \tag{13}$$

(see in Figure 2 the remaining survival function values after the vertical line at survival time 13 years). The distribution of predicted durations for the right censored spells can be found in Table 4.

[Insert Table 4 about here]

4.3. Non-Parametric Prediction for Left and Left-and-Right Censored Spells

In order to impute predicted spell durations for those spells that are either left or left-and-right censored we proceed as follows. In the case of no censored spells and right censored spells we have either the actual spell duration or the (previously obtained) predicted spell duration, respectively. Thus, for each spell in these two spell categories we can associate its spell duration with a given value of $\beta_0 + x_{ij}\beta$ in (A6) in Appendix A.[13] From here, and according to the differences in the values of $\beta_0 + x_{ij}\beta$ for the aforementioned two spell categories with respect to the values of $\beta_0 + x_{ij}\beta$ for spells that are left or left-and-right censored, we can calculate by non-parametric regression (kernel regression) the prediction for the spell duration of left or left-and-right censored spells. The intuition behind the method is to predict the missing spell duration of any left and left-and-right censored spell by weighting the known (or predicted) spell durations for no censored and right censored spells, according to the corresponding differences in the linear index of characteristics in (A6), that is, $\beta_0 + x_{ij}\beta$. Then, the spell duration we are seeking will be a weighted average of other spells durations, with higher weights for spells that are close in terms of the value of $\beta_0 + x_{ij}\beta$, and lower weights for spells that are far in terms of this value. The weighting function is going to be a kernel function that is a probability density function which formula will be given in Appendix B.

We first calculate the value of the index $\beta_0 + x_{ij}\beta$ associated with each spell. For the matching step, the x_{ij} are not taken at any particular survival time (*j*) value, but taken as fixed during the spell and equal to its mean value over the observed years of the corresponding spell. Furthermore, the parameters in the index are the parameter estimates in

[13] Given that the individual unobserved heterogeneity component is unknown, the value of $u_i \equiv \ln(v_i)$ in (A6) is settled to zero, since the random term v_i has unit mean.

the duration model. Consequently, our working index is $\hat{\beta}_0 + x_i\hat{\beta}$, which can be reduced to $x_i\hat{\beta}$ given that $\hat{\beta}_0$ is a common constant to all spells.

For left and left-and-right censored spells, the conditional expectations $E\left(T_i^* \middle| x_i\hat{\beta}\right)$ are replaced by non-parametric estimators $\hat{E}\left(T_i^* \middle| x_i\hat{\beta}\right)$, such as kernel estimators. In order to compute the $\hat{E}\left(T_i^* \middle| x_i\hat{\beta}\right)$ values, in our application we will use the so-called Nadaraya-Watson kernel regression function estimator (details for this estimator are given in Appendix B).

In our sample there are 197 left censored spells, and 205 spells which are left-and-right censored spells. The distribution of observed durations for these spells can be found in Table 5. Observed left censored durations are more concentrated in the low part of the durations' distribution than observed durations for left-and-right censored spells. For all these spells we are going to calculate the expected value of the spell duration by the kernel regression method just described above.

[Insert Table 5 about here]

As we have already stated, we use the information related to the total spell length of observed complete spells and the one predicted for right censored spells (a total of 569 spells, of which 350 are complete and 219 are right censored). The total durations' distribution for these spells can be found in Table 6. For the left and left-and-right censored spells, which observed durations are denoted by $T_{o,i}$, we use for the implicit matching procedure in the non-parametric regression (kernel regression) those observed complete and predicted right censored spells with duration equal or higher than $T_{o,i}$. The corresponding number of matching spells with $T_j^* \geq T_{o,i}$ are included in the first column of Table 5.

Finally, the distribution of predicted durations for the left and left-and-right censored spells can be found in Table 7.

[Insert Table 6 about here]

[Insert Table 7 about here]

5. ESTIMATES OF THE INNOVATION PRODUCTION FUNCTION

Using the results of the previous section we estimate the innovation production function. Recall from section 2 our estimating equation (10), which takes the form

$$\log \lambda_{it} = \log A(t) + \alpha_0 \log R_{it} + \alpha_1 E_{it} \log R_{it} + \alpha_2 E_{it}^2 \log R_{it} + z_{it}\beta_2 \qquad (14)$$

Our R&D-experience variable (E_{it}) is constructed as the sum of the number of years the firm has been investing in R&D in the past. For firms undertaking R&D activities the first year they are observed, this past history of R&D investments is estimated following the procedure in section 4. Control variables in z_{it} include informal innovation-related activities carried out by firms, firm size, and the type of industry in which the firm operates according to the degree of technological intensity (see the Table 1 for details). Additionally, a time trend and its squared value substituting for *log A(t)* are included.

In the estimation we follow the econometric approach pioneered by Hausman, Hall, and Griliches (1984), (HHG from now onwards). Their work develops and adapts statistical models of counts in the context of panel data to analyze the relationship between patents and R&D expenditures.

Using the Poisson specification as our starting point, and following HHG, we estimate the model under three alternative distributional assumptions: the existence of overdispersion in the data, the existence of random firm specific effects, and the existence of fixed firm specific effects potentially correlated with the regressors.

One limitation of the Poisson model is the assumption that the variance of N_{it} equals its mean (see equation 5), which neglects the possible existence of 'overdispersion' in the data. In the presence of such overdispersion, though the estimated parameters will be consistent, their standard errors will typically be under-estimated, leading to spuriously high levels of significance. After an initial estimation of the Poisson model, we shall consider the possibility of such overdispersion. If the results indicate the presence of such overdispersion in the data, we will proceed to the estimation of a Negative Binomial (NB) regression model. The NB is an extension of the Poisson regression model which allows the variance of the process to differ from the mean. One way for the model to arise is as a modification of the Poisson model in which λ_{it} is re-specified as

$$\log \lambda_{it} = x_{it}\beta + \varepsilon_{it} \tag{15}$$

where $exp(\varepsilon_{it})$ has a gamma distribution with mean 1 and variance α. This is a natural form of 'overdispersion' in that the overdispersion rate is given by

$$\frac{Var[N_{it}]}{E[N_{it}]} = 1 + \alpha \; E[N_{it}] \tag{16}$$

If the results render an estimate for α different from zero, we will be rejecting the Poisson model against the NB model.

Both the Poisson and NB models may be changed to allow for fixed and random effects. Apparently, these extensions mirror the panel data models for the linear regression model. For the fixed effects case the model takes the form

$$\log \lambda_{it} = \mu_i + x_{it}\beta \;\; (+\varepsilon_{it} \text{ for the NB model}) \tag{17}$$

where μ_i is the coefficient of a binary variable indicating membership to the *i*-th group. The difference with linear regression panel data models is that now the model cannot be fit by least squares using deviations from group means. Instead, a conditional maximum likelihood approach is used which removes μ_i from (17). The random effects model is

$$\log \lambda_{it} = x_{it}\beta + \upsilon_i \quad (18)$$

where v_i is a random effect for the i-th group such that e^{v_i} has gamma distribution with parameters (θ,θ). For the NB model, it is assumed that e^{v_i} is distributed as gamma with parameters (θ_i , θ_i), which brings in a model with a parameter that varies across groups. Moreover, it is assumed that θ_i /(1+θ_i) is distributed as beta (r, s). An estimate for s statistically different from zero indicates a variance to the mean ratio, that is, a value for the overdispersion ratio, different from one (the Poisson case, see HHH, 1984, pp. 372-373). The approach for the NB model is to integrate out the random effect and estimate by maximum likelihood the parameters of the resulting distribution. To evaluate the convenience of estimating a random effects version of the model, a likelihood ratio (LR) test will be performed testing the random effects model versus a pooled estimation of the model. Subsequently, a Hausman's (1978) specification test is used to compare the estimated vectors under the random and fixed effects versions of the model, that is, to test fixed versus random effects.

Before turning to the econometric results, we present in Table 8 the descriptive statistics for two firm size groups (firms with 200 employees or less, and firms with more than 200 employees), according to the sample procedure of the ESEE. The first column shows intervals of years of R&D-experience. For instance, the first interval "1-3 years" corresponds to firms that are either in their first, second or third year of R&D-experience. This R&D-experience is calculated for each observed period as the sum of past years with positive R&D spending, using, at this descriptive stage, only the observed data of firms with no problems of left censoring. Thus, what we report in this table are averages of the number of product innovations, the number of patents and the R&D-to-sales ratio that firms achieve each year when they are in their 1st to 3rd year of R&D-experience, in their 4th to 6th year of R&D-experience, and so on.

[Insert Table 8 about here]

A first comparison between the two size groups suggests that large firms have, on average, longer R&D-experience: the percentage of firms in the first interval is above 61% in the case of small firms, whereas this percentage is about 49% in the case of large firms. Consequently, the percentage of observations in the higher intervals is higher in the case of large firms. This could be indicating that R&D-experience is positively correlated with firm size, which is consistent with the well established empirical finding of a positive correlation of firm size with the probability of performing R&D activities.

As regards to the average number of product innovations that firms achieve yearly, figures in Table 8 indicate that they rise with R&D-experience. For the group of small (large) firms this average number ranges from 0.83 (0.73) in the first three years of R&D-experience to 1.42 (1.68) in the highest observed interval of R&D-experience (10^{th}-13^{th} years). In the case of the average number of patents, similar patterns are observed, although for the group of large firms there is a decline between the second and the third interval, which is recovered in the last interval. Thus, at a descriptive level, the data in our sample show that firms tend to achieve more innovation results as they accumulate years of R&D-experience. Finally, the average R&D-to-sales ratio also shows a positive relationship with R&D-experience. This ratio goes from 1.82 to 2.75 in the case of small firms, and from 0.95 to 1.76 in the case of large firms.

Therefore, both the average number of our measures of innovation results and the R&D effort made by firms seem to increase with firms' R&D-experience. However, we cannot at this stage discern whether we are simply observing the well established and intuitive result that higher R&D efforts lead to higher number of innovation results, or whether it is the case that R&D effectiveness rises with R&D experience, that is, whether each "euro" spent in R&D activities is more effective in achieving innovation results if combined with higher R&D-experience. In order to test this last hypothesis, which is our main objective in this paper, we turn to the analysis of our econometric results.

The econometric results from estimation for both product innovations and patents are reported in Tables 9 and 10, respectively. All regressions include our R&D-capital variable and its interactions with R&D-experience and with squared R&D-experience. Additionally, our estimation equations include a set of dummy variables accounting for other informal innovation related activities carried out by firms (scientific and technical services, quality control, imported technology, marketing, design, and other). It has been argued that a considerable amount of firms' innovation output may be the result of these informal innovation activities undertaken by firms (Sirrili, 1987). Furthermore, a set of control variables such as firm size (in the form of six size dummies), three dummy variables that indicate whether or not the firm belongs to a low, medium or high technological industry and, finally, a time trend and its square have been included. Details about the construction and definition of these variables are given in Table 1.

Columns 1 and 2 in tables 9 and 10 report the pooled regressions results both under the Poisson and the NB distributional assumptions, respectively.[14] In both tables, the parameter capturing overdispersion is statistically significant at conventional levels, indicating the rejection of the Poisson against the NB model (see columns (2) in both tables). Moreover, in column (3), the NB random effects model is estimated and tested against the NB pooled model. In this case, the LR test leads to the rejection of the NB pooled model. Finally, column (4) reports the fixed effects NB model, and the

[14] The number of observations included in each table differs slightly due to missing data in the dependent variable. It is also noticeable that the fixed effects estimation drops out part of the sample because the conditional method in this case needs at least one observation per firm on the dependent variable to be different from zero along the whole period.

corresponding Hausman test rejects the null of no-correlated fixed effects. Our econometric sequence, therefore, suggests the choice of the estimates of the NB fixed effects model for both product innovations and patents as measures of innovation output.

A first result in Table 9 is that both R&D-capital and the interaction of R&D-capital with R&D-experience have positive estimated signs, while the sign of the interaction of R&D with squared experience is negative. These results would be suggesting that the relationship between R&D effectiveness (here expressed in elasticity form) and R&D experience is of an inverted U-type. These results arise regardless of the distributional assumptions considered in the estimation, although the coefficients are somewhat lower in the panel estimation, that is, in columns (3) and (4). If we take the (statistically significant) results in column (4), the corresponding R&D-elasticity would be of a magnitude of $0.043 + 0.008 \cdot E_{it} - 0.0003 \cdot E_{it}^2$.

These results indicate that our measure of R&D effectiveness is different for firms with different R&D-experience, and that the effectiveness of R&D rises with R&D-experience, although at a decreasing rate. For instance, in our sample, and for a value of 7 years undertaking R&D activities (corresponding approximately to the median of the sample distribution) the value of the elasticity would be of 0.084, that is, by about a 66% larger than the elasticity of a firm that has been undertaking R&D for only one year. Moreover, the maximum value of the estimated elasticity, 0.096, corresponds to an R&D-experience of about 13.3 years, and beyond that value the estimated elasticity decreases.

[Insert Table 9 about here]

Figure 3 illustrates the R&D-capital elasticities for product innovations and patents. The R&D-capital elasticity for product innovations is represented in Graph 1. As already stated, our estimated elasticity reaches its maximum value between the 13^{th} and 14^{th} year of R&D-experience, and decreases for further years of R&D-experience. However, not all points depicted in Graph 1 are equally probable in our sample, and, in particular, 90% of the distribution is below 14 years of experience.

[Insert Figure 3 about here]

If we turn now to Table 10, we observe somewhat different results when patent counts are taken as our indicator of innovation output. Our preferred results are also those from the NB fixed effects estimation but, in this case, the coefficient of the interaction term of R&D-capital with R&D-experience is not statistically significant, whereas the coefficient of the interaction term of R&D-capital with squared R&D-experience turns out to be positive and statistically significant. The formula for the calculus of the R&D-elasticity would be now $0.071 + 0.0003 \cdot E_{it}^2$. This R&D-elasticity is monotonically increasing with R&D-experience. However, for low levels of R&D-experience, the effect of the interaction term above is very small, and it becomes more noticeable as R&D-experience rises. This result is illustrated in Figure 3, Graph 2, where the (positive) slope of the curve rises with R&D-experience. For a value of 8 years of R&D-experience, which represents approximately the median of the sample distribution, the value of the elasticity is about 26.5% higher than the elasticity of a firm that has been undertaking R&D for only one year. We also observe, for instance, that a firm with 14 years of experience has an elasticity which is about 82% higher than the elasticity of a firm with only one year of R&D experience. Therefore, we obtain that the longer the R&D-experience the higher the value of the elasticity, possibly indicating that it is required a lengthy R&D-experience to benefit from dynamic economies of scale, but that, once accumulated the necessary knowledge, further R&D efforts pay more and more in terms of patents. Thus, our results indicate that the effectiveness of R&D-capital changes along the R&D history of the firm, and that the results may differ depending on the indicator of innovation results.

[Insert Table 10 about here]

The inverted U-shape of the R&D effectiveness for product innovations could be related to a decrease in technological opportunities of the life cycle of the firms' product. However, the joint consideration of the results for product innovations and for patents may be interpreted in a more suggestive manner: it might be the case that as firms accumulate experience in the development of their R&D activities, they are able to select more successfully those innovation projects with more innovative content and thus, more likely to be translated into patents. This would explain the joint occurrence of a decreasing R&D effectiveness in terms of product innovations but an increasing R&D effectiveness in terms of patents as R&D-experience grows.

The interpretation above would be in line with Beneito (2006), who has highlighted how the joint consideration of alternative indicators of innovation output may enrich the understanding of R&D performance. Unfortunately the available data do not allow us to distinguish the innovative and/or economic content of each of the obtained innovations the firm declares to have introduced in a given year. Instead we only know the number of innovations obtained, as well as the number of patents registered. With such a detailed data the offered interpretation could be tested more accurately.

Other complementary results in Tables 9 and 10 that deserve some attention are those related to informal innovation activities. In our preferred specification, all the dummy variables capturing these informal activities have turned out to be robustly significant. In the case of product innovations, all kinds of informal activities contribute to the achievement of product innovations, whereas importing technology and marketing is negatively correlated with the number of patents obtained by the firm. This negative and significant sign of 'assimilation of imported technologies' in the case of patents might be indicating that the more orientated the firm's technological strategy is towards the import of (already existing) technologies, the lower the propensity to patent innovations. Informal innovation activities exhibit in our sample a positive correlation with formal R&D activities, raising the estimated R&D-elasticity if they are excluded from the estimation.[15] This point is remarkable in our sample because of two reasons. On the one hand, in the case of the Spanish industry, with a considerable percentage of firms of small and medium size, these informal R&D activities may be important for their innovation effectiveness. On the other hand, empirical work in this area does not typically include this information in the R&D patents relationship, a point that, among others, may help to explain the lower obtained magnitude of our R&D elasticities.

6. CONCLUSIONS

In this paper we have tested the hypothesis that, due to knowledge cumulativeness, the period of time during which firms perform R&D activities, which we call R&D-experience, is a key determinant of the firms' innovation success. We have argued that the temporal dimension captured by R&D-experience goes beyond the effect of R&D investments. In particular, we have tested the hypothesis that the effect of R&D-capital stock in the achievement of innovations depends on R&D-experience, that is, the number of years the firm has been performing R&D activities. By doing so, this paper has been an attempt to contribute to a better understanding of the nature of the cumulative process of learning and the importance of experience in the achievement of innovations.

We have investigated the role of firms' R&D-experience in the achievement of innovations using a representative sample of Spanish manufacturing firms (ESEE) for the period 1990-2002. We first have analysed firms R&D patterns in order to determine the duration of firms' R&D spells. For those spells for which, due to data restrictions, the

[15] These results are available from the authors upon request.

starting year is unknown (left censored spells), we have implemented a three steps procedure. First, we have estimated a duration model to identify firm and industry characteristics affecting R&D durations; secondly, and as a necessary intermediate step, we have used the duration model results to predict expected durations for right censored spells; thirdly, the information on complete spells and estimated right censored spells has been used to non-parametrically impute durations to left-censored spells. Once we have estimated the R&D-experience of firms as described above, we have proceeded to estimate, within the framework of a knowledge production function and using count data models, the influence of firms' accumulated R&D-experience on their R&D innovative effectiveness.

Our empirical analysis has indicated that, after controlling for R&D-capital stock and other firms' individual heterogeneity, firms' R&D effectiveness rises with R&D-experience, that is, with the accumulation of technical skills and knowledge that emerge as firms invest in R&D in a continuous way over time. However, the relationship between R&D-effectiveness and R&D experience is somewhat different depending on the innovation output we consider (product innovations or patents). The results suggest an inverted U-type relationship between R&D effectiveness and R&D-experience in the case of product innovations, and a monotonic increasing relationship between them in the case of patents. Our preferred interpretation for such a result is that, possibly, as firms accumulate experience in the development of their R&D activities, they are able to select more successfully those innovation projects with more innovative content and thus, more likely to be translated into patents. Further investigation with other data sources and more detailed information could reinforce this hypothesis. Finally, and in addition to past innovation experience, the performance of informal innovation activities and the technological intensity of the industry in which the firm operates have also been found to be important determinants in the achievement of innovations.

These findings may contribute to a better understanding of the cumulative process of learning and the importance of R&D-experience in the effectiveness of R&D investments, and may be a guide for policy makers in the design of policy measures to be implemented in order to stimulate the production of R&D knowledge. In particular, given that R&D-experience matters for innovation, our results suggest the convenience of implementing measures aimed at inducing firms to engage in R&D activities in a continuous way. Among these measures, a technological policy planed within a medium run perspective, or measures designed with the aim of creating a stable institutional framework, could help firms to persistently perform innovative activities.

References

Acs, Z.J., & Audretsch, D.B. (1989) 'Patents as a measure of innovative activity', Kyklos, 42(2), 171-80.

Allison, P.D., 1984. Event history analysis: Regression for longitudinal data. Sage:Newbury Parca. CA.

Arrow, K.J., 1962, "The economic implications of learning by doing", The Review of Economic Studies, 29, 3, 155-173.

Beneito, P., 2006, 'The innovative performance of in-house and contracted R&D in terms of patents and utility models', *Research Policy,* 35, 502-517.

Bierens, H. J., 1987, 'Kernel estimators of regression functions', in *Advances in Econometrics*, Fifth World Congress, Volume I, Econometric Society Monographs, No. 13, ed. T. F. Bewley, Cambridge University Press.

Bound, J., Cummins, C., Griliches, Z., Hall, B.H., & Jaffe, A. (1984) 'Who does R&D and who patents?' in Z. Griliches (eds.), *R&D, Patents and Productivity*, Cambridge: University of Chigaco Press, 21-54.

Branstetter, L., 1996, 'Are knowledge spillovers international or intranational in scope?. Microeconometric evidence from the U.S. and Japan', NBER Working Paper Series, No. 5800.

Cefis, E., 2003, 'Is there persistence in innovative activities?', *International Journal of Industrial Organization,* 21, 489-515.

Cefis, E. & Orsenigo, L., 2001, 'The persistence of innovative activities. A cross-countries and cross-sectors comparative analysis', *Research Policy,* 30, 1139-1158.

Cohen, W.M. & Klepper, S., 1996, 'Firm size and the nature of innovation within industries. The case of process and product R&D', *Review of Economics and Statistics,* 78, 232-243.

Cohen, W.M. & Levinthal, D.A., 1989, ‘Innovation and learning: the two faces of R&D’, *The Economic Journal*, 99, 569-596.

Crépon, B. and Duguet, E., 1997, ‘Estimating the innovation function from patent numbers: GMM on count panel data’, Journal of Applied Econometrics, 12, 243-263.

Crépon, B., Duguet, E., & Mairesse, J., 1998, ‘Research, innovation and productivity: an econometric analysis at the firm level’, NBER Working Paper Series 6696.

Flaig, G. and Stadler, M., 1994, ‘Success breeds success. The dynamics of the innovation process’, *Empirical Economics*, 19, pp. 55-68.

Flaig, G. and Stadler, M., 1998, ‘On the dymanics of product and proc ess innovations. A bivariate random effects probit model, *Jahrbücher für Nationalokönomie und Statistik*, 217, 401-417.

Geroski, P.A., Van Reenen, J. & Walters, C.F., 1997, ‘How persistently do firms innovate?’, *Research Policy*, 26, 33-48.

Griffith, R., Huergo, E., Mairesse, J. & Peters, B., 2006, ‘Innovation and Productivity across four European Countries’ *Oxford Review of Economic Policy*, 22 (4), 483-498.

Griliches, Z., 1990, ‘Patent statistics as economic indicators: a survey’, *Journal of Economic Literature*, 28, 1661-1707.

Hall, B.H., Griliches, Z., & Hausman, J., 1986, ‘Patents and R&D: is there a lag?’, *International Economic Review* 27(2), 265-283.

Hall, B.H. & J. Mairesse, 1995, ‘Exploring the relationship between R&D and productivity in French manufacturing firms’, *Journal of Econometrics*, 65, 263-293.

Hausman, J.A., 1978, ‘Specification tests in econometrics’, *Econometrica*, 46, 1251-1271.

Hausman, J., Hall, B.H. & Z. Griliches, 1984, ‘Econometric models for count data with and application to the patents-R&D relationship’, *Econometrica*, 52(4), 354-383.

Henderson, R., & Cockburn, I., 1993, ‘Scale, scope and spillovers: the determinants of research productivity in the pharmaceutical industry’, NBER Working Paper Series 4466.

Huergo, E. & Jaumandreu, J., 2004, ‘How does the probability to innovate change with firms’ age?’, *Small Business Economics,* 22, 193-207.

Jenkins, S.P., 1995, ‘Easy estimation methods for discrete-time duration models’, *Oxford Bulletin of Economics & Statistics,* 57, 129-138.

Jenkins, S. P., 2001. ‘Discrete time proportional hazards regression’, *Stata Technical Bulletin* 39, 22-32.

Jenkins, S.P., 2004, ‘Survival Analysis. Unpublished manuscript’, Institute for Social and Economic Research, University of Essex, Colchester, UK.

Kleinschmidt, E.J. & Cooper, R.B., 1990, ‘The performance impact of an international orientation on product innovation’, *European Journal of Marketing* , 22, 56-71.

Klomp, L. & van Leeuwen, G., 2001, ‘Linking innovation and firm performance’, *International Journal of the Economics of Business,* 8 (3), 343-364.

Kotable, M., 1990, ‘The relationship between offshore sourcing and innovativeness of US multinational firms: an empirical investigation’, *Journal of International Business Studies,* 21, 623-638.

Kremp, E., & Mairesse, J., 2004, ‘Knowledge management, innovation and productivity: a firm level exploration based on French manufacturing CIS3 data’, *NEBR Working Paper* 10237, Cambridge, MA.

Lööf, H. & Heshmati, A., 2001, ‘Knowledge Capital and Performance Heterogeneity: A Firm Level Innovation Study’, *International Journal of Production Economics,* 0, 1–25.

Love, J.H., Ashcroft, B. & Dunlop, S., 1996, ‘Corporate structure, ownership and the likelihood of innovation’, *Applied Economics,* 28, 737-746.

Mairesse, J. & Mohnen, P., 2005, ‘The Importance of R&D for Innovation: A Reassessment Using French Survey Data’, *The Journal of Technology Transfer*, 30(1–2), 183–197.

Malerba F, and Orsenigo L., 1999, ‘Technological entry, exit and survival: an empirical analysis of patent data’, *Research Policy* 1999; 28; 643-660.

Mansfield, E., 1968, *Industrial Research and Technological Innovation: An Econometric Analysis*, New York.

Máñez, J.A., Rochina-Barrachina, M, Sanchis-Llopis, A. & Sanchis-Llopis, J.A., 2004, ‘The role of sunk cost in the decision to invest in R&D’, mimeo.

Máñez, J.A., Rochina-Barrachina, M, Sanchis-Llopis, A. & Sanchis-Llopis, J.A., 2006, ‘Why do some firms persistently perform R&D activities?’, mimeo.

Meyer, R., 1990. ‘Unemployment insurance and unemployment spells’ *Econometrica*, 58, 757-782.

Nelson R. & Winter, S., 1982, *An Evolutionary Theory of Economic Change*, Cambridge, Harvard University Press.

Pakes, A., & Griliches, Z., 1984, ‘Patents and R&D at the firm level: a first look’ in in Z. Griliches (eds.), *R&D, Patents and Productivity*, Cambridge: University of Chigaco Press, 55-72.

Peters, B., 2005, ‘Persistence in innovation: stylised facts and panel data evidence’, ZEW Discussion Paper 05-81, Mannheim.

Prentice, R.L., Gloecker, L.A., 1978, ‘Regression analysis of grouped survival data with application to breast cancer data’, *Biometrics*, 34, 57-67.

Raymond, W., Mohnen, P., Palm, F. & Schim van der Loeff, S., 2006, ‘Persistence of innovation in Dutch manufacturing: is it spurious?’, UNU-Merit Working Paper 2006-011, Maastricht.

Rosenberg, N., 1976, *Perspectives on Technology*, Cambridge University Press, Cambridge.

Smith, K., & Sandven, T., 2001, ‘Innovation an economic performance at the enterprise level’, *Innovation and Enterprise Creation: Statistics and Indicators*, European Commission, Luxembourg, 106-116.

Scherer, F. M., 1965, ‘Corporate inventive output, profits and growth’, *Journal of Political Economy* 73(3), 240-297.

Scherer, F. M., 1983, ‘The propensity to patent’, *International Journal of Industrial Organization*, 1, 107-28.

Schmookler, J., 1966, *Invention and economic growth*, Cambridge: Harvard University Press.

Sirrilli, G., 1987, ‘Patents and inventors: an empirical study’, *Research Policy*, 16(2-4), 157-174.

Stoneman, P., 1983, *The Economic Analysis of Technological Change*, New York.

Sutton, J., 1991, *Sunk Costs and Market Structure*, Cambridge, Massachusetts: The MITT Press.

West, J. and Iansiti, M., 2003, “Experience, experimentation, and the accumulation of knowledge: the evolution of R&D in the semiconductor industry”, *Research Policy*, 32, 809-825.

APPENDIX A. ***The duration model: econometric methodology.***

Time intervals in our data set are of one year. Thus, the interval boundaries are the positive integers *j=1, 2, 3, 4,…*, and the interval *j* is $\left(j-1, j\right]$. For estimation, one R&D spell can either be complete ($c_i = 1$) or right censored ($c_i = 0$). A censored R&D spell *i* with length *j* intervals contributes to the likelihood function with the discrete time survival function (the probability of survival until the end of interval *j*):

$$S_i\left(j\right) = \Pr\left(T_i > j\right) = \prod_{k=1}^{j}\left(1-h_{ik}\right) \quad \text{(A1)}$$

where $T_i = \min\left\{T_i^*, C_i^*\right\}$, and T_i^* is some latent failure time and C_i^* some latent censoring time for spell *i*, and $h_{ik} = \Pr\left(k-1 < T_i \leq k \middle| T_i > k-1\right)$ is the discrete hazard (the probability of ending the spell in interval *k* conditional to the probability of survival up to the beginning of this interval). A complete spell *i* in the *j* interval contributes to the likelihood with the discrete time density function (the probability of ending the spell within the *j* interval):

$$f_i\left(j\right) = \Pr\left(j-1 < T_i \leq j\right) = S\left(j-1\right) - S\left(j\right) = \frac{h_{ij}}{1-h_{ij}}\prod_{k=1}^{j}\left(1-h_{ik}\right) \quad \text{(A2)}$$

Using (A1) and (A2), the log likelihood function for the sample of spells is:

$$\log L = \sum_{i=1}^{n} c_i \log\left(\frac{h_{ij}}{1-h_{ij}}\right) + \sum_{i=1}^{n}\sum_{k=1}^{j}\log\left(1-h_{ik}\right) \quad \text{(A3)}$$

Allison (1984) and Jenkins (1995, 2004) show that (A3) can be rewritten as the log likelihood function of a binary dependent variable y_{ik} with value one if spell *i* ends in year *k*, and zero otherwise:

$$\log L = \sum_{i=1}^{n}\sum_{k=1}^{j}\left[y_{ik}\log h_{ik} + \left(1-y_{ik}\right)\log\left(1-h_{ik}\right)\right] \quad \text{(A4)}$$

This allows discrete time hazard models to be estimated by binary dependent variable methods and time-varying covariates to be incorporated.

Following Prentice and Gloeckler (1978), we assume that h_{ik} is distributed as a complementary log-log (*cloglog*) function to obtain the discrete time representation of an underlying continuous time proportional hazard: [16]

$$c\log\log\left[1-h_j\left(x_{ij}\right)\right]\equiv\log\left(-\log\left[1-h_j\left(x_{ij}\right)\right]\right)=\beta_0+x_{ij}\beta+c\left(j\right)$$
$$\Rightarrow h_j\left(x_{ij}\right)=1-\exp\left[-\exp\left(\beta_0+x_{ij}\beta+c\left(j\right)\right)\right] \qquad \text{(A5)}$$

where $c\left(j\right)$ is the baseline hazard parametrically specified as in a Weibull ($c\left(j\right)=\left(q-1\right)\ln\left(j\right)$),[17] and x_{ij} are explanatory variables (covariates), which may be time-varying (although constant within intervals).

Incorporating unobserved heterogeneity, the *cloglog* model in (A5) becomes

$$h_j\left(x_{ij}\right)=1-\exp\left[-\exp\left(\beta_0+x_{ij}\beta+\left(q-1\right)\ln\left(j\right)+u_i\right)\right] \qquad \text{(A6)}$$

where $u_i\equiv\ln\left(\nu_i\right)$, and ν_i originally enters the underlying continuous hazard function $h\left(t,x_{it}\right)=h_0\left(t\right)\exp^{\beta_0+x_{it}\beta}\nu_i$ multiplicatively. It is standard to assume that ν is Gamma distributed with unit mean and variance σ^2, to be estimated from the data (Meyer, 1990).[18]

---

[16] Given that we are interested in a proportional hazard specification with duration data observed discrete but with an underlying continuous time generating process, the complementary log-log function is the most appropriate one.

[17] If q>1, that is, if (q-1)>0, there is positive duration dependence, what means that survival time increases the risk of ending the R&D spell. If q<1, that is, if (q-1)<0, there is negative duration dependence, what means that survival time decreases the risk of ending the R&D spell. If q=1, that is if (q-1)=0, there is not duration dependence.

[18] An up-to-date Stata program drawn up by S. Jenkins that implements the cloglog with gamma distributed unobserved heterogeneity is available from http://www.bc.edu/RePEc/bocode/p or it can also be obtained, inside the Stata program, by typing “ssc install pgmhaz8”. An initial version of the program was presented in Jenkins (2001).

Not controlling for unobserved individual heterogeneity may cause two problems. First, the degree of negative (positive) duration dependence in the hazard (the parameter estimate for (q-1)) is over-estimated (under-estimated). This is the result of a selection process. For instance, with negative duration dependence, individuals with high ν-value finish the spell more rapidly. Then, as time goes by, a higher proportion of individuals with low values of ν remain in the spell, which implies a lower hazard. Secondly, positive (negative) β parameters are under-estimated (over-estimated).

APPENDIX B. *The kernel regression function estimator: the Nadaraya-Watson model.*

According to the Nadaraya-Watson kernel regression function estimator the corresponding non-parametric regression function estimator of $\hat{E}\left(T_i^*\middle|x_i\hat{\beta}\right)$ is

$$\hat{E}\left(T_i^*\middle|x_i\hat{\beta}\right)=\frac{\sum_{j=1,\&\,j\neq i}^{N} T_j^*\cdot K\left[\left(x_i\hat{\beta}-x_j\hat{\beta}\right)/c_N\right]}{\sum_{j=1,\&\,j\neq i}^{N} K\left[\left(x_i\hat{\beta}-x_j\hat{\beta}\right)/c_N\right]} \tag{B1}$$

which are *leave-one-out* kernel estimators constructed without T_i^* being used in estimating $\hat{E}\left(T_i^*\middle|x_i\hat{\beta}\right)$. This is convenient both theoretically and for our particular application of the method, given that we do not observe T_i^* for left and left-and-right censored spells.

In order to apply this method, one needs to choose the kernel function K and a particular bandwidth parameter c_N. We implement a univariate second order bias reducing kernel of Bierens (1987), corresponding to a normal probability density function of the form:

$$K_{1,2}\left(\frac{x_i\hat{\beta}-x_j\hat{\beta}}{c_N}\right)=\frac{\exp\left(-\frac{1}{2}\left(\frac{x_i\hat{\beta}-x_j\hat{\beta}}{c_N}\right)^{'}\Omega^{-1}\left(\frac{x_i\hat{\beta}-x_j\hat{\beta}}{c_N}\right)\right)}{\sqrt{2\pi}\cdot\sqrt{\det(\Omega)}} \qquad \text{(B2)}$$

where Ω is a positive definite matrix . We specify $\Omega=\hat{V},$ where $\hat{V}$ is the sample variance matrix; that is, $\hat{V}=\frac{1}{N}\sum_{j=1}^{N}\left(x_j\hat{\beta}-\overline{z}\right)^{'}\left(x_j\hat{\beta}-\overline{z}\right)$ with $\overline{z}=\frac{1}{N}\sum_{j=1}^{N}x_j\hat{\beta}$. Thus, we obtain

$$K_{1,2}\left(\frac{x_i\hat{\beta}-x_j\hat{\beta}}{c_N}\right)=\frac{\exp\left(-\frac{1}{2}\left(\frac{x_i\hat{\beta}-x_j\hat{\beta}}{c_N}\right)^{'}\hat{V}^{-1}\left(\frac{x_i\hat{\beta}-x_j\hat{\beta}}{c_N}\right)\right)}{\sqrt{2\pi}\cdot\sqrt{\det\left(\hat{V}\right)}} \qquad \text{(B3)}$$

We will now focus on the problem of the bandwidth selection. We need the convergence rate of $\hat{E}\left(T_i^*\middle|x_i\hat{\beta}\right)$ to the true value to be faster enough. According to Bierens (1987), the best uniform consistency rate for a univariate kernel of order two is obtained for $c_N=c\cdot N^{-1/6}$, and so we use this form in the estimation.[19] Therefore, the bandwidth selection problem is reduced to choosing the constant c. In our application, the constant part of the bandwidth was chosen to be equal to 1. There was no serious attempt at choosing c optimally, but we avoided values which could entail extreme bias or variability.

[19] If we were focused on convergence in distribution, the optimal rate would have been obtained by setting $c_N=c\cdot N^{-1/5}$.

Figure 1: Sample distribution, number and types of R&D spells

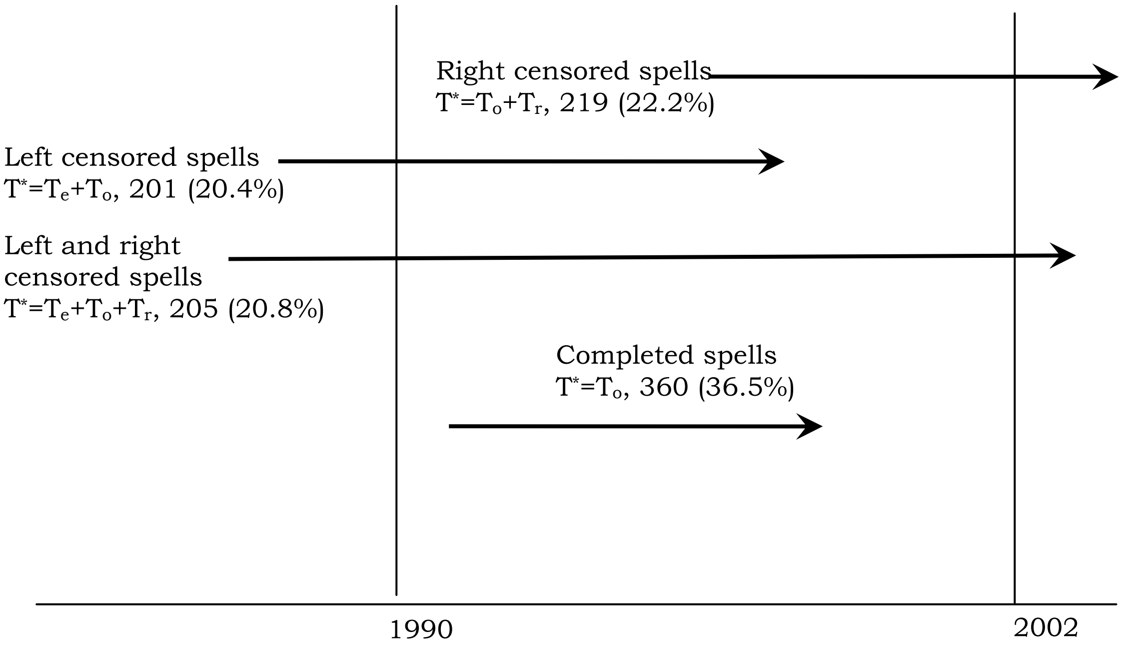


Figure 2: Out of sample prediction of the survival function

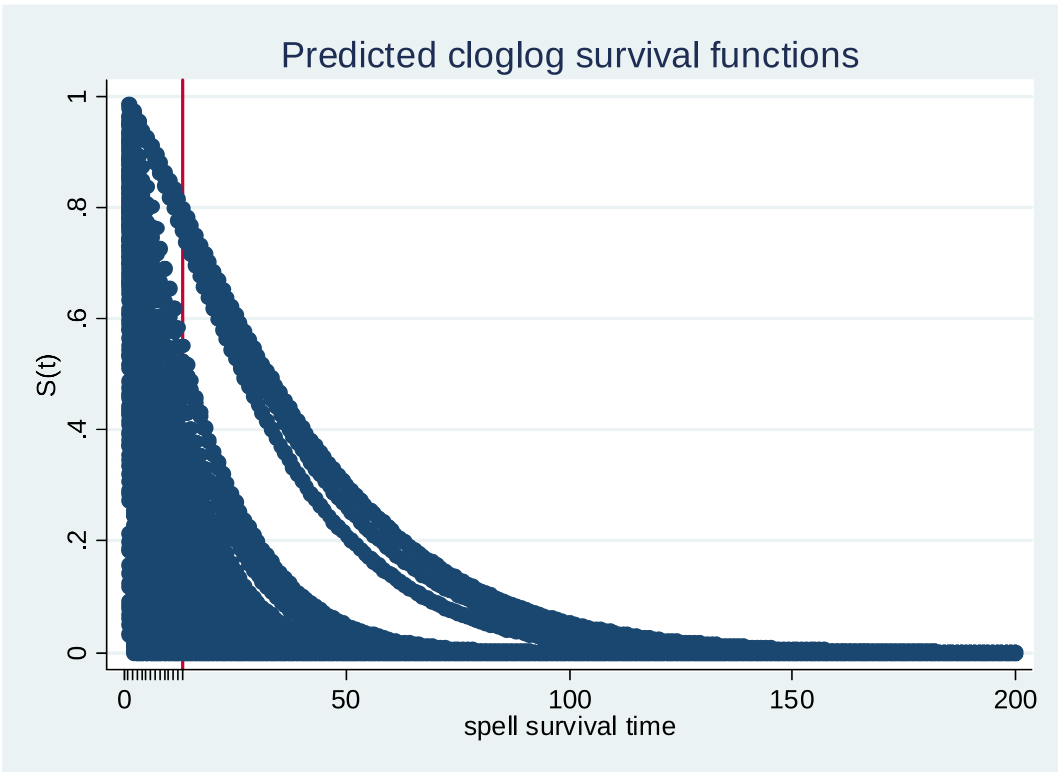
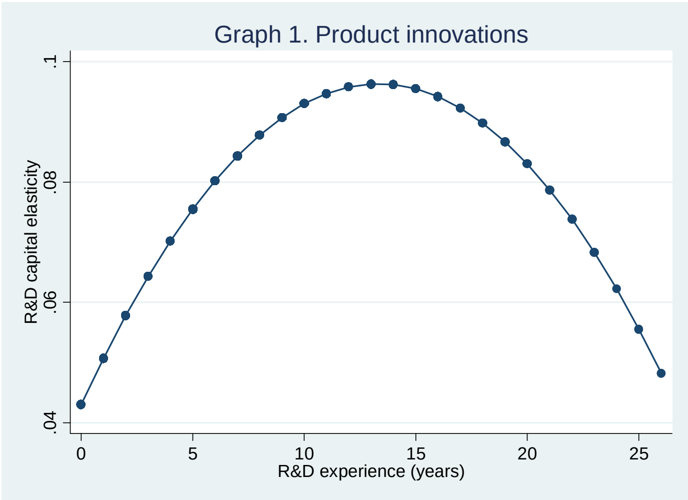


Figure 3. R&D-capital elasticities



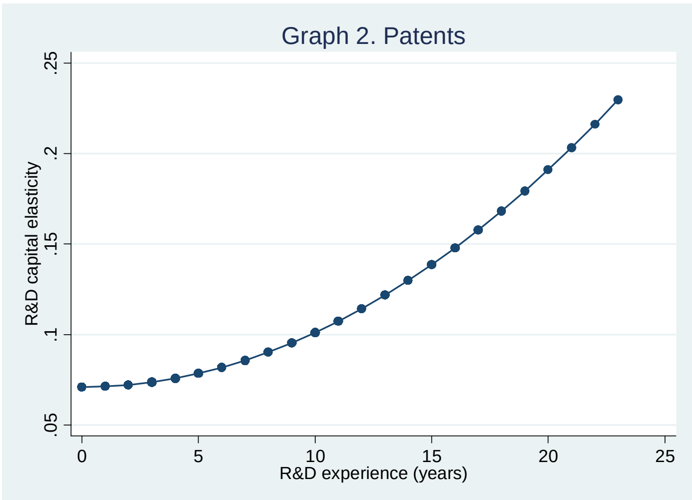


Table 1. Variable definitions for the duration model

Ln(t)	Ln of spell duration in years
Industry dummies	Industry dummies accounting for 20 industrial sectors of the NACE-93 classification. See Table 2 for the classification of industries.
International market	Dummy variable taking value 1 if the geographic limits of the firm main market are foreign or both national and foreign, and 0 otherwise.
Age5	Dummy variable taking value 1 if the firm mean age during the spell is smaller or equal than 5 years, and 0 otherwise.
Age510	Dummy variable taking value 1 if the firm mean age during the spell is greater than 5 and smaller or equal than 10 years, and 0 otherwise.
Age1020	Dummy variable taking value 1 if the firm mean age during the spell is greater than 10 and smaller or equal than 20 years, and 0 otherwise.
Age2030	Dummy variable taking value 1 if the firm mean age during the spell is greater than 20 and smaller or equal than 30 years, and 0 otherwise.
Age3040	Dummy variable taking value 1 if the firm mean age during the spell is greater than 30 and smaller or equal than 40 years, and 0 otherwise.
Age4050	Dummy variable taking value 1 if the firm mean age during the spell is greater than 40 and smaller or equal than 50 years, and 0 otherwise.
Age50	Dummy variable taking value 1 if the firm mean age during the spell is greater than 50 years, and 0 otherwise.
Size100	Dummy variable taking value 1 if the firm number of workers is smaller or equal than 100, and 0 otherwise.
Size100200	Dummy variable taking value 1 if the firm number of workers is greater than 100 and smaller or equal than 200, and 0 otherwise.
Size200	Dummy variable taking value 1 if the firm number of workers is greater than 200, and 0 otherwise.
No Corporate	Dummy variable taking value 1 if the firm is not a limited liability corporation, and 0 otherwise.
Med/High R&D intens.	Dummy variable taking value 1 if the firm's R&D intensity (R&D expenditure to sales ratio) belongs to the second and third thirds of the sample R&D intensity distribution, and 0 otherwise.
R&D workers ratio	Ratio of R&D workers to total number of workers.
Regional spillovers	Ratio of firms that perform R&D in the same region but outside the corresponding two digit NACE-93 industry.
Local spillovers	Ratio of firms that perform R&D in the same region and the same two digit NACE-93 industry.

Table 1 (continued). Variable definitions for the count data model

Patents	Number of patents registered during the year both in Spain and abroad.
Product innovations	Number of product innovations introduced by the firm during the year.
Scientific/technical services	Dummy variable taking value 1 if the firm has undertaken services of scientific and technical information, and 0 otherwise.
Quality control	Dummy variable taking value 1 if the firm has undertaken works of normalisation and quality control, and 0 otherwise.
Imported technology	Dummy variable taking value 1 if the firm has undertaken efforts to assimilate imported technologies, and 0 otherwise.
Marketing	Dummy variable taking value 1 if the firm has undertaken marketing studies orientated to the commercialisation of new products, and 0 otherwise.
Design	Dummy variable taking value 1 if the firm has undertaken design activities, and 0 otherwise.
Other	Dummy variable taking value 1 if the firm has undertaken other informal innovation activities, and 0 otherwise.
Low technological sector	Dummy variable taking value 1 if the firm belongs to a low technological intensity sector: meat products; beverages; textiles; leather and shoes; wood; paper; printing; non metallic minerals; metallic products; furnitures; other manufacturing goods.
Medium technological sector	Dummy variable taking value 1 if the firm belongs to a medium technological intensity sector: food and tobacco, rubber and plastic; metallurgy; machinery and mechanical equipment; electronic.
High technological sector	Dummy variable taking value 1 if the firm belongs to a high technological intensity sector: chemical products; office machines; electronic; other transport material.
E: R&D-experience	Number of years the firm has been investing in R&D in the past. For firms undertaking R&D activities the first year they are observed, this past history of R&D investments is estimated following the procedure in section 4.
Size1	Dummy variable that equals 1 if the number of employees of the firm is below or equal to 20, and 0 if otherwise.
Size2	Dummy variable that equals 1 if the number of employees of the firm is above 20 and below or equal to 50, and 0 if otherwise.
Size3	Dummy variable that equals 1 if the number of employees of the firm is above 50 and below or equal to 100, and 0 if otherwise.
Size4	Dummy variable that equals 1 if the number of employees of the firm is above 100 and below or equal to 200, and 0 if otherwise.
Size5	Dummy variable that equals 1 if the number of employees of the firm is above 200 and below or equal to 500, and 0 if otherwise.
Size6	Dummy variable that equals 1 if the number of employees of the firm is above 500, and 0 if otherwise.
K: R&D-capital	The knowledge capital derived from the firm's R&D investment follows the historical or perpetual inventory method: $K_{it} = (1-\delta) K_{it-1} + R_{it-1}$ where δ is the rate of depreciation, *K* is the R&D-capital stock and *R* are real R&D expenditures (current R&D has been deflated using industrial prices for the whole manufacturing industry). To calculate the R&D-capital according to the equation above we need an initial value for *R* to start the recursion. We use for that purpose the information about the number of years the firm has been undertaking R&D activities, that is the firm's R&D-experience, that comes from the estimation and predictions explained in section 4. By backwards induction, the sequence of past R&D expenditures can be imputed till the first year of R&D activities, when the initial R&D-capital stock is equal to zero. The R&D-capital is defined for a depreciation rate of 15 percent and a pre-sample growth rate of real R&D investment equal to the mean growth rate for the firms which perform R&D activities and are observed during the sample period, that is g = 4,5%.

Table 2. Maximum likelihood estimates for the discrete time proportional hazard model

cloglog with Gamma individual unobserved heterogeneity and Weibull type duration dependence

	Coefficients[a]	*p*-value
Ln(t)	0.305	0.72
Food and tobacco[b]	-1.189	0.23
Beverages	-1.673	0.15
Textiles	-1.385	0.21
Leather and shoes	-2.747*	0.08
Wood	-2.579	0.11
Paper	-1.739	0.16
Printing	-0.395	0.70
Chemical products	-1.983	0.11
Rubber and plastic	-1.876	0.19
Non metallic miner	-1.618	0.17
Metallurgy	-0.077	0.94
Metallic products	-0.858	0.38
Machin. and mech. eq.	-1.376	0.23
Office machines	-1.026	0.57
Electronic	-2.137	0.12
Motors and cars	-2.244*	0.09
Other transp. material	-0.756	0.51
Furniture	-1.789	0.20
Other manufact. goods	-1.580	0.30
International market	-0.383*	0.06
Age510	-1.076**	0.03
Age1020	-1.012**	0.05
Age2030	-1.732**	0.03
Age3040	-1.865**	0.03
Age4050	-2.882**	0.04
Age50	-1.631*	0.06
Size100200	-0.763*	0.10
Size200	-0.989***	0.01
No Corporate	0.636*	0.09
Med/High R&D intens.	-0.714***	0.01
R&D workers ratio	-4.860**	0.04
Regional spillovers	-1.571**	0.05
Local spillovers	-0.235	0.61
Intercept	3.666	0.11
Log likelihood	-717.008	
N. of observations	1653	
N. of spells	569	
Test for unobserved individual heterogeneity	LR test of Gamma variance=0 Chibar2(01)= 2.152 *p*-value =0.07	

[a] (***), (**), and (*), means statistically different from zero at the one, five, and ten-percent significance level.
[b] The meat industry is the reference category.

Table 3. Distribution of observed durations (T_o) for right censored spells

T_o	Number of spells	%
1	61	27.85
2	31	14.16
3	19	8.68
4	20	9.13
5	18	8.22
6	12	5.48
7	13	5.94
8	10	4.57
9	7	3.20
10	10	4.57
11	5	2.28
12	11	5.02
13	2	0.91
Total	219	100

Table 4. Distribution of predicted durations $\left(E\left(T_i^*\right)\right)$ for right censored spells.

T_o	Number of spells	%
2	27	12.33
3	31	14.16
4	24	10.96
5	26	11.87
6	17	7.76
7	21	9.59
8	13	5.94
9	15	6.85
10	7	3.20
11	10	4.57
12	5	2.28
13	11	5.02
14	3	1.37
16	2	0.91
17	1	0.46
19	2	0.91
21	1	0.46
33	1	0.46
37	1	0.46
41	1	0.46
Total	219	100

Table 5. Distribution of observed durations (T_o) for left and both left/right censored spells

T_o (Number of matching spells $\left(T_j^* \geq T_{o,i}\right)$)[a]	Left and both left/right censored spells: Number of spells	%	Left censored spells: Number of spells	%	Both left/right censored spells: Number of spells	%
1 (569)	59	14.68	59	29.95		
2 (381)	40	9.95	35	17.77	5	2.44
3 (274)	100	24.88	33	16.75	67	32.68
4 (208)	30	7.46	25	12.69	5	2.44
5 (161)	16	3.98	16	8.12		
6 (121)	29	7.21	8	4.06	21	10.24
7 (100)	10	2.49	6	3.05	4	1.95
8 (75)	2	0.50	2	1.02		
9 (61)	6	1.49	4	2.03	2	0.98
10 (45)	5	1.24	4	2.03	1	0.49
11 (38)	4	1.00	4	2.03		
12 (28)	7	1.74	1	0.51	6	2.93
13 (23)	94	23.38			94	45.85
Total	402	100	197	100	205	100

[a] For left and left-and-right censored spells, which observed durations are denoted by $T_{o,i}$, we use for the implicit matching procedure in the non-parametric regression (kernel regression) those observed complete and predicted right censored spells with duration equal or higher than $T_{o,i}$.

Table 6. Distribution of observed complete durations and predicted right censored durations

T_o	Number of spells	%
1	188	33.04
2	107	18.80
3	66	11.60
4	47	8.26
5	40	7.03
6	21	3.69
7	25	4.39
8	14	2.46
9	16	2.81
10	7	1.23
11	10	1.76
12	5	0.88
13	11	1.93
14	3	0.53
16	2	0.35
17	1	0.18
19	2	0.35
21	1	0.18
33	1	0.18
37	1	0.18
41	1	0.18
Total	569	100

Table 7. Distribution of predicted durations $\left(E\left(T_i^*\right)\right)$ for left and both left/right censored spells

T_o	Number of spells	%
2	8	1.99
3	23	5.72
4	41	10.20
5	32	7.96
6	47	11.69
7	49	12.19
8	28	6.97
9	28	6.97
10	14	3.48
11	7	1.74
12	11	2.74
13	8	1.99
14	55	13.68
15	16	3.98
16	12	2.99
17	10	2.49
19	3	0.75
20	1	0.25
25	3	0.75
26	2	0.50
29	2	0.50
32	2	0.50
Total	402	100

Table 8. Descriptive statistics on R&D-experience and innovation results.

Intervals of R&D-experience (years)	Firms with ≤ 200 employees				Firms with > 200 employees			
	N.obs. (%)	Average number of product innovations in each year	Average number patents registered in each year	$\frac{R\&D}{sales}$	N.obs. (%)	Average number product innovations in each year	Average number patents registered in each year	$\frac{R\&D}{sales}$
1 – 3 years	381 (61.75 %)	0.83	0.05	1.82	163 (48.95 %)	0.73	0.51	0.95
4 – 6 years	149 (24.25 %)	1.05	0.05	1.77	88 (26.43 %)	0.76	0.69	1.39
7 – 9 years	68 (11.02 %)	1.04	0.08	1.90	52 (15.62 %)	1.11	0.34	1.44
10 – 13 years	19 (2.90 %)	1.42	0.09	2.75	30 (9.0 %)	1.68	0.60	1.76
Total	617				333			

Table 9. Estimates of the Innovation Production Function.

	PRODUCT INNOVATIONS							
	Poisson (pooled)		Neg. Bin. (pooled)		Neg. Bin. (random eff.)		Neg Bin. (fixed eff.)	
	(1)		(2)		(3)		(4)	
log K	.064**	(.003)	.064**	(.016)	.047**	(.011)	.043**	(.011)
log K × E	.014**	(.7e-03)	.010**	(.003)	.007**	(.002)	.008**	(.002)
log K × E^2	-.8e-03**	(.4e-05)	-.6e-04**	(.2e-04)	-.3e-04**	(.1e-04)	-.3e-04**	(.1e-04)
size2	.805**	(.023)	.470**	(.117)	.015	(.091)	.036	(.098)
size3	.576**	(.027)	.489**	(.150)	.013	(.115)	.008	(.125)
size4	.259**	(.028)	.240*	(.141)	.221**	(.109)	.188	(.120)
size5	-.089**	(.026)	-.021	(.124)	-.270**	(.101)	-.325**	(.112)
size6	-.428**	(.034)	.007**	(.168)	-.035	(.129)	-.037	(.144)
cient./tecnic. services	.234**	(.016)	.048	(.088)	.234**	(.056)	.237**	(.059)
quality control	-.731**	(.015)	-.558**	(.090)	.165**	(.057)	.163**	(.060)
imported tech.	.250**	(.017)	.128	(.105)	.087	(.061)	.130**	(.064)
marketing	.171**	(.016)	.207**	(.094)	.235**	(.059)	.213**	(.061)
design	.900**	(.015)	.764**	(.086)	.267**	(.056)	.181**	(.058)
other	-.001	(.049)	-.160	(.307)	.545**	(.161)	.540**	(.166)
med. tech. sectors	-.865**	(.020)	-.619**	(.095)	.207**	(.072)	.194 **	(.077)
high tech. sectors	-.563**	(.020)	-.309**	(.110)	.455**	(.081)	.481**	(.087)
trend	-.082**	(.008)	.059	(.047)	.024	(.029)	.028	(.030)
trend^2	.003**	(.5e-03)	-.003	(.003)	-.002	(.002)	-.003	(.002)
intercept	.459**	(0.036)	.023	(.187)	-2.18**	(.129)	-2.11**	(.135)
N. obs (N.firms)	6464 (670)		6464 (670)		6464 (670)		5094 (510)	
log likelihood	-53383.9		-10058.8		-8977.2		-6451.1	
parameter ≠ 0 indicates overdispersion			7.860** (0.209)		1.246** (0.112)			
LR test pooled vs. random effects.					1965.57 p-value: 0.000			
Hausman test of correlated fixed effects							89.27 p-value: 0.000	

Standard errors in parenthesis. ** significant at 1% level; * significant at 5% level

Table 10. Estimates of the Innovation Production Function.

	PATENTS							
	Poisson (pooled)		Neg. Bin. (pooled)		Neg. Bin. (random eff.)		Neg Bin. (fixed eff.)	
	(1)		(2)		(3)		(4)	
log K	.139**	(.010)	.056**	(.016)	.088**	(.021)	.071**	(.022)
log K × E	.005**	(.001)	.008	(.005)	-.003	(.003)	-.003	(.003)
log K × E^2	-.5e-05	(.5e-05)	-.1e-04**	(.2e-04)	.3e-04*	(.16e-04)	.3e-04**	(.1e-04)
size2	.693**	(.106)	.550**	(.211)	.151	(.202)	.019	(.232)
size3	1.024**	(.111)	1.096**	(.260)	-.081	(.263)	-.202	(.300)
size4	.599**	(.110)	.853**	(.238)	.300	(.232)	.078	(.268)
size5	1.300**	(.099)	1.455**	(.213)	.266	(.213)	.017	(.250)
size6	.674**	(.109)	1.156**	(.266)	.169	(.248)	-.080	(.284)
cient./tecnic. services	.573**	(.039)	.298	(.161)	.361**	(.111)	.348**	(.118)
quality control	-.078*	(.044)	-.160	(.145)	.258**	(.113)	.286**	(.121)
imported tech.	-.303**	(.040)	-.288	(.176)	-.237**	(.114)	-.215*	(.119)
marketing	.364**	(.038)	.525**	(.155)	-.267**	(.107)	-.352**	(.111)
design	.363**	(.037)	.929**	(.138)	.371**	(.104)	.224**	(.109)
other	-.309*	(.176)	-.525	(.510)	1.159**	(.359)	1.35**	(.381)
med. tech. sectors	-.199**	(.051)	-.116	(.152)	.179	(.146)	.204	(.160)
high tech. sectors	.597**	(.044)	.391**	(.187)	.211	(.149)	.182	(.158)
trend	-.273**	(.021)	-.235**	(.078)	-.155**	(.052)	-.134**	(.053)
trend^2	.011**	(.001)	.011**	(.005)	.006*	(.003)	.004	(.003)
intercept	-2.75**	(.123)	-2.44**	(.321)	-1.98**	(.261)	-1.579**	(.185)
N. obs (N.firms)	6627	(671)	6627	(671)	6627	(671)	2261	(219)
log likelihood	-9415.49		-3437.0		-3007.4		-1870.4	
parameter ≠ 0 indicates overdispersion			17.801** (0.946)		0.256** (0.029)			
LR test pooled vs. random effects.					863.72 p-value: 0.000			
Hausman test of correlated fixed effects							82.27 p-value: 0.000	