Firms' age, process innovation and productivity growth

Elena Huergo[a,1], Jordi Jaumandreu[b,*]

[a] Dpto. Fundamentos del Análisis Económico I, Facultad de Ciencias Económicas y Empresariales,
Campus de Somosaguas, Universidad Complutense de Madrid, 28223 Madrid, Spain
[b] Dpto. Economía, Facultad de Ciencias Sociales y Jurídicas, Universidad Carlos III de Madrid,
C/Madrid 126, 28903 Getafe, Spain

Abstract

This paper looks directly at the impact of firms' age and (process) innovations on productivity growth. A model that specifies productivity growth as an unknown function of these variables is devised and estimated using semiparametric methods. Results show that firms enter the market experiencing high productivity growth and that above-average growth rates tend to last for many years, but also that productivity growth of surviving firms converges. Process innovations at some point then lead to extra productivity growth, which also tends to persist somewhat attenuated for a number of years.
© 2004 Elsevier B.V. All rights reserved.

1. Introduction

There is a vast amount of literature about the impact of technological activities on productivity, including an important tradition of empirical estimations of this effect using firm-level data (see Griliches, 1995, for a survey and Griliches, 2000, for example, for an updated assessment[2]). The standard form of these exercises has been the construction of a

---

* Corresponding author. Tel.: +34-916245737; fax: +34-916249329.
E-mail addresses: ehuergo@ccee.ucm.es (E. Huergo), jordij@eco.uc3m.es (J. Jaumandreu).
[1] Tel.: +34-913942408; fax: +34-913942561.
[2] See Hall and Mairesse (1995) and Klette (1996) for recent examples of this literature.

stock of knowledge capital, starting from R&D investment data, and its introduction as an additional input into the firms' production function. However, as Griliches (1979) already pointed out in his pioneering work, the relationship between productivity and the (constrained) weighted average of R&D expenditures embodies in a simplified way two very different and presumably complex processes: the production of innovations starting from R&D activities and the incorporation of these innovations to production.

The knowledge capital construction and specification imply a number of important constraints on the form of these processes (see Klette, 1996, for a discussion and the relaxation of a number of these constraints). This provides an important reason for looking more closely at every one of these processes. Moreover, recently, data on the innovative output of R&D activities have become increasingly available, opening the possibility of these types of analyses. For example, Crepon et al. (1998) constitute an interesting departure from the traditional modelling using innovation data.

This paper carries out an investigation focussed on the relationships between the introduction of innovations and the growth of productivity. It looks directly at the effects of innovation on total factor productivity growth, using (unbalanced) panel data on the age of more than 2300 Spanish manufacturing firms and their process innovations brought in during the period 1990–1998. The investigation is mainly intended to examine whether innovations really induce growth, the life span and time pattern of these productivity effects and the presumed heterogeneity associated with different frequencies of innovations. To answer these questions, it seems the effects of firm age must also be disentangled (in some sense, the first radical process innovation takes place with entry into a market). Conclusions contribute evidence on the effects of firms' innovative activity and have interesting implications for their modelling.

Productivity growth is measured by means of the (cost shares-based[3]) Solow residual, corrected for (possible) nonconstant returns to scale. To address the presumably highly nonlinear relationships between productivity growth, age and process innovation, we devise a specific semiparametric model. A central piece of this model is the flexible estimation of the expectation of productivity growth conditional on age and innovations, while controls enter linearly. The main advantage of this modelling is that we avoid imposing any a priori functional form constraint on the impact of the key variables of age and innovation. Estimation can then be claimed to be fully robust, in the sense of being free of specification bias, in the involved relationships. This seems a particularly appropriate way to learn about a subject on which there is little evidence and for which we can hardly guess the specific form of the relationships.

Estimates show that firms enter the market experiencing high productivity growth, and that above-average growth rates tend to last, although progressively weakened, for many years. The estimates also point out that productivity growth of surviving firms converges to different values according to activities and, on average, to almost 1.5% annually. Process innovations at some point then lead to some extra productivity growth, which tends to persist, although somewhat attenuated, for a number of years. If the introduction of process innovations then stops, however, innovation appears to be associated with an end to all productivity growth in the following years.

---

[3] The cost shares-based Solow residual is robust to the presence of market power. See, e.g., Hall (1990).

The rest of the paper is organised as follows. Section 2 is devoted to establishing the framework of measurement of productivity and Section 3 to estimating the impact of age and innovation. Section 4 deals with the data and variables and Section 5 presents the empirical results. Section 6 comments on implications and concludes. Appendix A describes the sample and gives the definition of the variables employed.

2. Measuring productivity growth

In this section, we describe the theoretical framework which relates productivity growth to firms' age and innovations and we derive the econometric model to estimate this relationship. Assume that firms in population are characterised by production functions of the type

$$Y_{it} = A_i(t)F_i(K_{it}, L_{it}, M_{it})$$

where Y denotes output, K is capital, L represents labour and M stands for materials. F_i are the (presumably specific) production relationships that link produced output to conventional inputs, and the factor $A_i(t)$ represents the level of efficiency reached by firm i. The way this factor enters the equation implies Hicks neutrality of productivity increases, and notation emphasises the idiosyncrasy and time dependence of the efficiency level. It can be interpreted, for example, as an unspecified form for the role of the traditional technological or knowledge capital variable used in the exercises aimed at measuring the productivity effects of technological activities. But it can also be understood simply as a completely unspecified efficiency level evolving over time.

We assume that total differentiation of production relationships yields the population equation

$$y_{it} = a_i(t) + \varepsilon_{K,it}k_{it} + \varepsilon_{L,it}l_{it} + \varepsilon_{,it}m_{it} + u_{it} \tag{1}$$

where the small letters employed to represent outputs and inputs denote logarithmic differences, the ε's stand for the respective input elasticities, the term $a_i(t)$ stands for the proportional change dA_i/A_i or productivity growth and u_{it} stands for a disturbance which we assume to have zero mean conditional on the included variables. We also assume that our available sample constitutes a random sample from this population.

An estimable model can be obtained in the following way. On the one hand, in order to investigate productivity growth characteristics, we specify the term $a_i(t)$, without loss of generality, as the sum of a firm and time idiosyncratic term and an unknown function (common to all firms) of the vector of the key variables z whose role we want to assess: $a_i(t)=a_{it}+a(z_{it})$. On the other hand, cost minimisation implies that input elasticities equal the products of scale elasticity and cost shares. Then, using s to denote cost shares and assuming common scale elasticity γ across firms, we can write:

$$y_{it} = a_{it} + a(z_{it}) + \gamma(s_{K,it}k_{it} + s_{L,it}l_{it} + s_{M,it}m_{it}) + u_{it}$$

Rearranging terms, this expression can be easily transformed into an equation that links the observable (cost shares-based) Solow residual, a correction for the scale

elasticity effect on productivity and the terms accounting for productivity growth. The model is

$$\theta_{it} = (\gamma - 1)v_{it} + a_{it} + a(z_{it}) + u_{it} \tag{2}$$

where θ represents the Solow residual, $(\gamma-1)$ is a parameter to be estimated involving scale elasticity γ and v stands for the weighted sum of input variations.

Let us comment and further develop specification (2). Firstly, the weighted input sum term is a correction for the productivity effects of variations in the scale of operation. While this term should become irrelevant under constant returns to scale, in practice, short-run input movements turn out to be associated to somewhat decreasing returns (probably due to the short-run fixity of some misspecified production aspects). Hence, its inclusion is important.

Secondly, the a_{it} term must be specified to account for any variation in productivity growth across firms which should be controlled for when studying the relationship of productivity growth with the variables z (in our case, age and innovation). We will specify it as a linear function of control variables. In particular, the empirical exercise performed here includes the variations in the firms' capacity utilisation (see Delgado et al., 1998), dummy variables to account for some sources of discrete changes in firms' efficiency levels (mergers, acquisitions, scissions) and time dummies as a way to pick up the influence of "macroeconomic" factors common to all firms (e.g., manufacturing cycle). In addition, the presumably high heterogeneity among activities makes it convenient to include sets of activity and firms' size dummies in order to control for any systematic differences in productivity growth.

Thirdly, we want the unknown function to represent average growth (i.e., includes the constant term of the regression). Consequently, we are going to specify the time, activity and firms' size dummies as picking-up differences from this average (constraining them to add up to zero; see Suits, 1984).

The linear nature of the controls makes it possible to write Eq. (2) in the slightly more compact form:

$$\theta_{it} = x_{it}\beta + a(z_{it}) + u_{it} \tag{3}$$

where $x=$(variable for scale correction, capacity utilisation, merger or acquisition dummy, scission dummy, time dummies, size dummies and industry dummies) and $a(\cdot)$ is the unknown function aimed at picking up the presumably highly nonlinear relationships between productivity growth, firms' age and innovations. The next section is devoted to the specification and estimation of these relationships.

3. Estimating the impact of age and innovation

We specify age as the argument of a first component $a^*(\tau)$ of the unknown function $a(\cdot)$. The impact of age can be guessed to be highly nonlinear and the variable age can take a relatively high number of values (we consider ages up to 40 years in the market; see Section 4 and Appendix A for details). In principle, several possibilities arise to estimate

this unknown function, even the estimation of the age effects by a system of dummies. We prefer, however, a "smoothing" estimator because of the high number of values and the likely low "volatility" of the age effect[4]. We use the kernel-based semiparametric estimator[5] analysed by Robinson (1988) and Speckman (1988) and applied, for example, in Stock (1989). Because τ and u are uncorrelated, from Eq. (3), we have

$$\theta_{it} - E(\theta_{it} \,|\, \tau_{it}) = [x_{it} - E(x_{it} \,|\, \tau_{it})]\beta + u_{it} \qquad (4)$$

and the semiparametric estimator of β is the ordinary least squares estimator (OLS) after replacing the conditional expectation functions by a nonparametric estimate. We use the kernel regression Nadaraya–Watson estimator[6]. Once the β parameters have been estimated, we can form the nonparametric regression model

$$\theta_{it} - x_{it}\hat{\beta} = a^*(\tau_{it}) + e_{it}$$

where the unknown function is estimated again applying the Nadaraya–Watson estimator (see Pagan and Ullah, 1999)[7]. This procedure makes it easy to estimate the variance of the function estimate and hence the computation of the confidence intervals[8]. This is important for inference because as firms' numbers become lower (or productivity growth more disperse) at some ages, the variance of the estimator becomes higher.

We use a normal kernel and all kernel estimates proved to be very sensible using the smoothing parameter[9] $h=n^{-1/5}$. Asymptotic variance of the estimated expectation is $V(\hat{a}^*) = \frac{1}{nh}\frac{\sigma^2(x)}{f(x)} \int K(s)^2 \mathrm{d}s$, where $K(\cdot)$ is the kernel density function, h is the smoothing parameter or bandwidth, n is the number of observations, σ^2 and $f(x)$ are the variance and density associated with each x and $s=(x_i-x)/h$. We estimate this variance using $\int K(s)^2 \mathrm{d}s = 0.2821$, estimating σ^2 by $\hat{\sigma}^2 = \sum_i \hat{e}_i^2 K(s_i)/\sum_i K(s_i)$, and employing the Rosenblatt–Parzen kernel estimator of $f(x)$.

Let us finally comment on an important property of this estimate. The estimated function constitutes an expectation conditional on surviving. As disappearance from the sample (by death or attrition) is likely to be correlated with low productivity growth, older (surviving) firms are likely to show better-than-average productivity growth.

---

[4] A dummies estimator, using a set of nonoverlapped dummies, would be equivalent to the semiparametric estimator based on the nonsmoothing estimation which merely averages the y values corresponding to each x value. To see this, notice that the coefficients on the set of dummies will simply be the averages of the corresponding y values after deducing the value of the controls (using the suitable previously estimated parameter). Delgado and Mora (1995) show the consistency of the nonsmoothing estimator and argue on the relative behaviour of the asymptotically equivalent nonsmoothing and kernel estimators with base on a series of Monte Carlo experiments.

[5] The kernel nonparametric estimator could obviously be replaced, with likely close results, by a nonparametric spline smoothing or even by a "series" expansion estimator (see, for example, Pagan and Ullah, 1999).

[6] See, for example, Wand and Jones (1995) or Pagan and Ullah (1999).

[7] The function could be alternatively recovered from the equality $\hat{a}^*(\tau)=\hat{E}(\theta_{it}|\tau_{it})-\hat{E}(x_{it}|\tau_{it})\hat{\beta}$.

[8] Variance will have an additional component coming from the substitution of the estimated parameters $\hat{\beta}$ for the true parameters β, but this component can be shown to be asymptotically unimportant.

[9] Estimates show a low sensitivity to the band width choice, but we use the slightly modified value of $h=1.2n^{-1/5}$ for the industry regressions.

The next questions to answer are whether process innovations introduced by firms induce extra productivity growth, with what intensity and for how much time. This implies the estimation of the effects of process innovations the year the innovation is introduced and the years to follow. To estimate these effects, it seems natural to complete the specification of the $a(\cdot)$ function by adding to $a*(\tau)$ the innovation impacts on productivity growth (at any point of life)[10]. Calling α_s the successive impacts of an innovation, from moment 0 (introduction) to T (the last considered lag), we have in whole $a(z) = a*(\tau) + \sum_{s=0}^{T} \alpha_s$. Given the time length of our sample, which allows us to specify only a limited number of lagged effects, we estimate them using a system of dummies[11].

This completion of the $a(\cdot)$ function can be estimated simply by also including in estimating Eq. (4) a system of dummies (each one minus its expected value given age) aimed at picking up the contemporaneous and delayed effects of innovations without any a priori constraint. Notice that, with this specification, the estimated effects will measure the productivity growth impact of innovation as the extra (or less) growth with respect to the average growth for noninnovating firms (after taking into account all the growth heterogeneity allowed by the control variables).

Among the different equivalent ways of specifying the impacts $\sum_{s=0}^{T} \alpha_s$, we choose the use of the system of dummies defined by the expression for the impacts $\alpha_0 d_{it} + \sum_{s=0}^{T} \delta_s tes_{it}$, where the artificial variables in the set $(d, te1, te2, \ldots teT)$ respectively indicate that an innovation has been introduced and the time elapsed since then. These variables are defined as follows:

$$d_{it} = \begin{cases} 1 & \text{if the firm brings in an innovation at time } k \text{ and } 0 \leq t - k \leq T \\ 0 & \text{otherwise} \end{cases}$$

$$tes_{it} = \begin{cases} 1 & \text{if } d_{it} = 1 \text{ and } t - k = s \text{ for } s \in [1, T] \\ 0 & \text{otherwise} \end{cases}$$

Estimates of α_s are given by $\hat{\alpha}_0$ and by $\hat{\alpha}_s = \hat{\alpha}_0 - \hat{\delta}_s$ when s takes the values from 1 to T.

Let us briefly comment in turn on some properties of this estimate. Firstly, in applying it, we must face the problem of left censoring. For firms which were born during the sample period, we can observe every innovation carried out until the final year of the sample, but for firms with a history previous to the initial year of the sample (or the firm's particular entering year in a few cases), we cannot determine the time elapsed since the last innovation, or even if there was such innovation at any point in time. Hence, there is a set of data points before the first innovation is observed for which, strictly speaking, we have no reliable value to attribute to our set of dummies. However, this set of data points is likely to include mostly observations of really noninnovating firms (the bulk of the firms for which we have not yet observed any innovation by the final year of the sample), or

---

[10]   Hence, we are implicitly employing an additive model.

[11]   In this case, we choose the use of a dummies estimator, which can be seen as an equivalent procedure to the employment of a nonsmoothing estimator (see the previous discussion).

scarcely innovating firms, after which a long period of time has elapsed when we observe the first innovation. We will experiment alternatively by dropping these observations from the sample and attributing to them an "absence of innovation" value.

Secondly, delays between innovations constitute a sample with a special type of selectivity problem. Firms' available time observations reach a maximum at the years covered by the sample (in our case, 9 years). Hence, again for censoring reasons, the probability of observing each delay value between innovations is lower the longer the delay is and zero for the length of the sample. Our dummy method of estimating the conditional expectation of productivity growth for each time elapsed is statistically robust to this sort of selectivity, although the estimates must be attributed a lower precision the higher the value of the time elapsed.

4. Data and variables

Estimations are carried out with an unbalanced panel data sample of more than 2300 firms surveyed during the period 1990–1998[12]. Details are provided in Appendix A. This sample is approximately representative of manufacturing, and hence inferences can be considered globally valid for the manufacturing population. Firms with fewer than 200 workers were sampled randomly by industry and size strata, retaining 5%, while firms with more than 200 workers were all requested to participate, and the positive answers represented more or less a self-selected 60%. The statistical methods applied here are robust to this type of sample mixture. In addition, the coefficients obtained for the size dummies confirm that very little or nothing linked to size remains to be explained.

To preserve representation, samples of newly created firms were added to the initial sample every subsequent year. At the same time, there are exits from the sample, coming from both death and attrition. The two motives can be distinguished and attrition was maintained to sensible limits. All the exercises performed here use all observations with complete data, independently of the available firm time observations. Hence, the sample includes, approximately in population proportions, surviving, entrant and exiting firms and experiences some decay over time due to attrition.

The available information allows us to compute the cost-based Solow residual, construct the control variables and fix the age of the firm according to the number of years it has been active in the marketplace, establishing a maximum category of 40 years or more (see Appendix A for details).

A process innovation is assumed to occur when the firm answers positively to the (innovation) question of whether it has introduced some significant modification of the productive process (affecting machines, organisation or both) along the year. The question appears in the questionnaire along with all the other R&D and innovation-related questions (e.g., product innovation) and is clearly separated from other sections on technology adoption and usage. Hence, it is likely to pick up rather precisely what firms consider

---

[12] The survey was sponsored by the Ministry of Industry, under the name "Encuesta sobre Estrategias Empresariales" (ESEE), and carried out at the Programa de Investigaciones Económicas of the Fundación Empresa Pública.

major innovative changes in their productive process and the frequency of these changes[13]. In fact, some of our results suggest that firms tend not to report changes which are already generalised in the market as innovative productive modifications.

The sample average relative frequency of process innovation is about 1/3 (34%). This implies that we expect a firm to introduce a process innovation every 3 years (see Table A.1). However, the sample also has a high proportion of firms which never innovate (about 42%) and a small proportion which innovate every year they are in the sample (15%). These values constitute two modes, and proportions of intermediate relative frequencies are slightly decreasing. The sample average relative frequency of innovation of the strictly uncensored sample (all the noninnovation-datable observations dropped) is, as expected, higher: about 1/2 (52%). The probability of introducing process innovations varies greatly by activities, sizes, and over firm ages. Huergo and Jaumandreu (in press) estimate this probability, showing how small size per se tends to reduce the probability of innovation, but also how entrant firms tend to present the highest probability of innovation. The oldest firms tend to present a somewhat lower probability, although some firms from intermediate to old ages present a high probability, which may be attributed to selection. Exiting firms are clearly associated to lower levels of preexit innovations[14].

5. Empirical results

This section presents the empirical exercise. Firstly, we report the results of estimating the age and the age/innovation models using the whole sample and the strictly non-censored sample. Then, we briefly comment on the results of estimating the age model by industries, using a disaggregation of manufacturing in 10 industries.

Table 1A reports the results of the estimations with the whole sample. The first estimate reports the results of a fully parametric estimation of the age model, where age enters the equation linearly. The second and third estimates report the results of semiparametric estimations of the age and age/innovation models. The third estimate adds to the function of age the set of dummies designed to account for the effects of innovation and their persistence over time. The fourth column reports the same estimate as the third, but applied to the strictly noncensored sample (68% of the previous data points). Table 1B reports the dummies' coefficients for the semiparametric estimate with the whole sample of the age/innovation effects model (third estimate; but, in fact, dummy coefficients remain fairly stable across estimates), and panels a–c of Fig. 1 depict the functions obtained in the semiparametric estimates, plotting the value of productivity growth as a function of firm age.

Controls turn out to give repeatedly robust and sensible results, which, in addition to their own interest, stress the validity of the framework employed. Firstly, the average

---

[13] Notice that an advantage of this type of output measure is that it avoids the well-known reporting problems associated with the coexistence of formal and informal innovative activities.

[14] All this agrees well with the standard findings on industry dynamics related to innovation; see, for example, Audretsch (1995).

Table 1A
Results from the estimation of models $\theta=x\beta+a^*(\tau)+u$ and $\theta=x\beta+a^*(\tau)+\alpha_0 d+\Sigma\delta_s tes+u$ (age and age/innovation effects models)

Dependent variable: (cost-based) Solow residual[a]			
Explanatory variables	Coefficients and *t*-ratios[b]		
	Age effects (parametric estimate)	Age effects (semiparametric estimate)	Age and innovation effects (semiparametric estimate)	Age and innovation effects (semiparametric estimate; uncensored sample)
Scale correction	−0.236 (−15.6)	−0.237 (−15.6)	−0.238 (−15.7)	−0.226 (−12.8)
Capacity utilisation	0.092 (7.6)	0.092 (7.6)	0.091 (7.5)	0.107 (7.1)
Merger	0.054 (4.4)	0.055 (4.4)	0.055 (4.4)	0.043 (3.0)
Scission	−0.071 (−2.8)	−0.072 (−2.8)	−0.071 (−2.8)	−0.079 (−3.0)
Time, size, industry dummies[c]	Included	Included	See Table 1B (cont.)	Included
Constant	0.030 (8.8)			
Age	−0.0003 (−2.2)	Fig. 1, panel a	Fig. 1, panel b	Fig. 1, panel c
Process innovation dummy			0.015 (4.3)	0.0003 (0.0)
Time elapsed 1 year			−0.008 (−1.4)	−0.007 (−1.2)
Time elapsed 2 years			−0.006 (−1.0)	−0.005 (−0.9)
Time elapsed 3 years			−0.007 (−1.1)	−0.008 (−1.1)
Time elapsed 4 years			−0.025 (−2.9)	−0.025 (−2.9)
Time elapsed 5 years			−0.019 (−1.7)	−0.021 (−1.8)
Time elapsed 6 years			−0.037 (−2.1)	−0.037 (−2.2)
Sigma squared	0.026	0.025	0.025	0.022
No. of firms	2356	2356	2356	1750
No. of observations	10,735	10,735	10,735	7293

[a] Sample period: 1991–1998.
[b] *t*-ratios computed using (unbalanced panel) robust standard error formulas.
[c] Eight time dummies, 6 size dummies and 18 industry dummies, with coefficients of each set constrained to add zero (Suits method).

elasticity of scale, estimated through the coefficient of the scale correction, is about 0.76. Secondly, firms' utilisation of capacity is important in explaining variations in productivity growth [10 percentage points of increase (decrease) in the utilisation of capacity imply nearly 1 percentage point of increase (decrease) in productivity growth]. Thirdly, mergers or acquisitions and scissions turn out to have a significant impact on productivity growth (the year following the fact). On average, this impact is positive for mergers or acquisitions and negative and stronger for scissions.

Let us briefly comment on the time, size and industry dummy coefficients. Recall that given the method used to specify these sets of dummy variables, dummy coefficients must be interpreted as giving percentage deviations from average growth.

Firstly, time dummies show how the industrial cycle determined a sharp average productivity decrease which reached bottom in 1993 and intense increases the following two recovering years. Secondly, interestingly enough, firm size dummies are not significant, which points to the absence of firm size patterns in the heterogeneity remaining to be explained beyond the model. This means that, with the determinants explicitly embodied,

Table 1B
Dummy coefficients of semiparametric estimate of the age and innovation effects model[a]

Dummies	Coefficient	t-ratio
Time dummies		
91	−0.002	(−0.4)
92	−0.006	(−1.3)
93	−0.028	(−5.9)
94	0.012	(2.8)
95	0.022	(5.5)
96	0.000	(0.0)
97	−0.009	(−2.4)
98	0.011	(3.2)
Size dummies (no. of workers)		
Less than 20	0.001	(0.2)
From 21 to 50	0.002	(0.8)
From 51 to 100	−0.003	(−0.7)
From 101 to 200	0.002	(0.7)
From 201 to 500	−0.001	(−0.3)
More than 500	−0.001	(−0.4)
Industry dummies		
Ferrous and nonferrous metals	0.016	(2.2)
Nonmetallic minerals	−0.004	(−0.9)
Chemical products	0.020	(4.6)
Metal products	−0.004	(−1.0)
Industrial and agricultural mach.	0.001	(0.3)
Office and data processing m.	0.011	(0.9)
Electrical and electronic goods	0.008	(2.0)
Vehicles, cars and motors	0.018	(3.4)
Other transport equipment	−0.032	(−2.3)
Meat and preserved meat	−0.001	(−0.2)
Food and tobacco	−0.012	(−3.3)
Beverages	−0.005	(−0.9)
Textiles and clothing	−0.001	(−0.3)
Leather and shoes	−0.005	(−0.8)
Timber and furniture	−0.015	(−2.7)
Paper and printing products	−0.003	(−0.7)
Rubber and plastic products	0.016	(3.2)
Other manufacturing products	−0.009	(−1.3)

[a] Third estimate in main panel.

we can account for all the differences in productivity growth apparently linked to size that emerge so often in empirical exercises. Thirdly, up to 1/3 of industries tend to show significant differences with respect to average productivity growth.

Let us comment on the central results. The second estimate shows a clear relationship between productivity growth and age (see Table 1A and Fig 1, panel a). Entrant firms present high productivity growth (beginning at about 5%) and, although decreasing as time goes by, average growth tends to be higher than average until firms reach about 8 years in the market. At this age, growth tends to stabilise by about 2% (the wandering of the curve is consistent with a higher variance as age becomes higher, but shows no clear trend). That
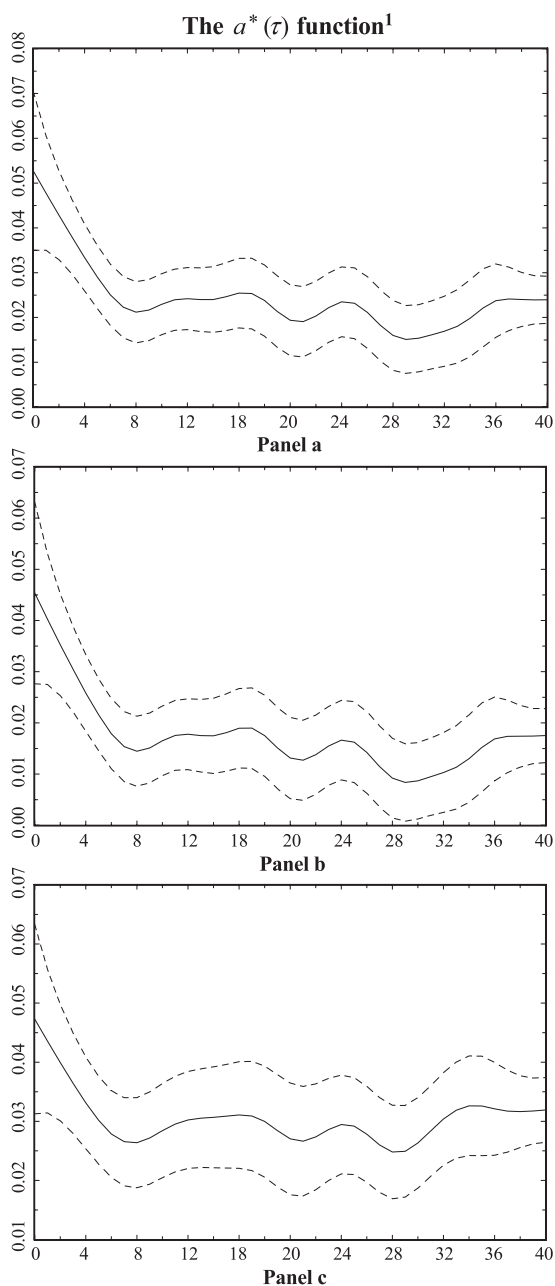
The $a^*(\tau)$ function[1]

Fig. 1. The $a^*(\tau)$ function. Values of the function $a^*(\tau)$ (productivity growth), estimated with semiparametric techniques, against age. Upper and lower 95% confidence bands.

is, productivity growth clearly tends to a "normal" rate, presumably different according to activities (recall the industry dummy effects), to which surviving firms converge after a number of years of rapid productivity growth. Recall that the estimation is conditional on observed lives, and the result consequently suggests that closing firms are the ones that fail to reach this "normal" rate.

Notice that the linear parametric estimate, reported as the first estimate, is clearly unable to approach properly the evidence contributed by the semiparametric estimate. Average productivity growth is evaluated at 3% (something more or less in the middle of the early growths), and age appears to reduce this rate by about 0.3% after 10 years and 0.6% after 20. The radical nonlinearity of the relationship is missed[15].

The third and fourth estimates introduce the set of dummies aimed at picking up the effects of innovation. The third estimate (Table 1A and Fig. 1, panel b) uses the whole sample. Introduction of innovation leaves the estimate almost unchanged, mainly affecting the constant implicit in the age function (notice that shape changes are minimal). The convergence value is now situated at almost 1.5%. At the same time, innovation shows a clear contemporaneous impact on productivity growth by about another 1.5%. Positive impacts seem to persist for 3 more years with a lower average value of 0.7% (estimated, however, with a high variance).

Positive impacts, however, also seem to be followed by 3 more years in which, if other innovation is not introduced, firm productivity growth is surprisingly reduced below average growth (by about 1.2% a year, which implies virtually no growth). It must be noticed that only a few firms among the firms which innovate will show such a time delay before introducing a new innovation (remember that innovators introduce on average an innovation in more than half of the years). But, for these firms, everything happens as if product innovation had moved future productivity growth ahead by 3 or 4 years.

When the same specification is estimated employing the strictly noncensored sample, the average values of the age function again move upwards (convergence value is now located at about 2.8%) and contemporaneous innovation turns out to have no impact on average rates, but the schedule of diminishing returns of innovations is estimated to be virtually the same. The reasons for this result are the following. On the one hand, this sample avoids wrongly attributing no-innovation values to data points which by their nature cannot be established without further nonavailable information. But, on the other hand, and by the same token, it constitutes a selected sample of entry, innovations and close data points, at which the value of innovation is already picked up by the implicit average productivity growth. Results with this subsample confirm for us, however, that the possible bias in the whole sample cannot be too important, and that the innovation lags schedule is reliable.

Then, estimate three offers a good picture of the average impact of innovation over the life of firms. According to this picture, process innovation clearly accelerates productivity growth during a number of years (1.5% the year of introduction of the innovation and a bit less for 3 more years), but productivity growth of the innovating firm also tends to fall below average and even fully stop the following years when new innovations are not introduced as

---

[15] Once the shape of the function is known, one could obviously try to adjust a suitable parametric function if needed.

well. This suggests that there is a common pace of growth as production improvements become ''common knowledge'' and are incorporated by every firm in the market, and that innovating firms are basically the firms which make a successful effort to move forward future productivity increases, creating the changes which will afterwards be generalised[16].

Table 2 reports semiparametric estimates of the age model for a disaggregation of manufacturing in 10 sectors. The number of data points for each estimation is obviously much lower and hence variance is higher, but these estimates let us assess to some extent the degree of heterogeneity involved in the estimates for all manufacturing. The lack of statistical precision makes the industry semiparametric estimates of the age/innovation effects model of no use here (there are too few observations on each innovation lag) and they are not reported[17].

Let us briefly comment on the results reported in Table 2. Control variables show quite plausible values. Returns to scale present values which range from 0.65 to 0.80, with the exception of the constant returns to scale case found for the transport equipment sector. Impacts of capacity utilisation on productivity growth range from a reasonable maximum with Transport equipment (6) to a minimum with the Food, drink and tobacco industry (7). At the same time, a simplified set of size dummies, which divides the samples according to the threshold of 200 workers, shows small coefficient values and not very much significance, and time dummies reveal some heterogeneity in sectors' productivity evolution over time.

The estimated age functions, whose precision depends heavily on the number of observations, are more difficult to interpret (separating the effects of heterogeneity and higher variance). However, in Fig. 2, we depict four particularly significant estimates [corresponding to the sectors of Chemical products (3), Office and electrical machines (5), Food, drink and tobacco and (7) Textile, leather and shoes (8)]. The number of observations is always relatively big and/or productivity growth tends to show more intraindustry homogeneity than in other sectors. Interestingly enough, among these sectors, there are two that are technologically intensive and two in which more mature traditional technologies tend to dominate. Panels of Fig. 2 suggest that in every industry, there is a starting period of high relative productivity growth, corresponding to the initial years of a firm's life, even if this productivity growth is very different and tends to decrease more or less steadily according to the sector (compare 3 and 5 to 7 and 8)[18]. The number of years of this initial period tends to be of 8 years. Then the results suggest that, judged by the observed growth in productivity, entrants reach maturity at around 8 years.

Some of the sectors seem to show a somewhat higher-than-average productivity growth for the oldest firms. We attribute this to the importance of selection, which implies that

---

[16] This interpretation gives a sensible explanation of why an occasional innovator, which do not innovate in the sequel, would cease to experiment productivity growth: it is just the only one in the market which cannot profit again from the improvement which has contributed.

[17] Nevertheless, they have been computed and do not show remarkable novelties.

[18] Given Fig. 2, it could be argued that a horizontal line can be drawn within the limits of the confidence bands and this would essentially mean that one cannot reject the hypothesis of no age effects. However, the hypothesis of no age effects, either for the whole expectation or part of it, should be contrasted by means of a test on the conditional function jointly implying the relevant estimated values. Here, the important thing is that the imprecisely estimated nonparametric part of the industry models does not contradict the main conclusions attained at the aggregate exercise.

Table 2
Industry semiparametric estimations of model $\theta = x\beta + a^*(\tau) + u$

Dependent variable: (cost-based) Solow residual[a]

Explanatory variables	Coefficients and t-ratios[b]									
	(1) Ferrous and nonferrous metals and metal products	(2) Nonmetallic minerals	(3) Chemical products	(4) Industrial and agricultural machinery	(5) Office and data-processing machines and electrical goods	(6) Transport equipment	(7) Food, drink and tobacco	(8) Textile, leather and shoes	(9) Timber and furniture	(10) Paper and printing products
Scale correction	−0.215 (−5.7)	−0.298 (−5.7)	−0.205 (−4.0)	−0.187 (−4.5)	−0.205 (−3.7)	−0.002 (0.0)	−0.299 (−6.4)	−0.328 (−10.9)	−0.243 (−4.7)	−0.349 (−5.9)
Capacity utilisation	0.059 (2.4)	0.143 (2.8)	0.098 (3.5)	0.104 (1.7)	0.067 (1.3)	0.142 (3.1)	0.056 (2.7)	0.089 (2.8)	0.070 (2.0)	0.068 (1.7)
Merger	−0.014 (−0.4)	0.049 (1.1)	0.067 (2.3)	−0.062 (−2.2)	0.096 (2.5)	−0.058 (−1.4)	0.071 (2.6)	0.061 (1.5)	−0.108 (−5.1)	0.130 (2.1)
Scission	−0.080 (−1.2)	0.038 (0.7)	−0.176 (−3.3)	−0.559 (−11.4)	−0.050 (−0.9)	−0.032 (−1.1)	0.062 (1.2)	−0.251 (−1.7)	−0.052 (−0.4)	0.029 (0.3)
Up to 200 workers	−0.003 (−0.9)	−0.007 (−1.5)	−0.004 (−1.3)	−0.011 (−2.0)	0.000 (0.0)	−0.001 (−0.3)	0.008 (2.1)	−0.001 (−0.2)	−0.016 (−2.2)	0.012 (1.9)
More than 200 workers	0.003 (0.9)	0.007 (1.5)	0.004 (1.3)	0.011 (2.0)	0.000 (0.0)	0.001 (0.3)	−0.008 (−2.1)	0.001 (0.2)	0.016 (2.2)	−0.012 (−1.9)
1991	−0.022 (−1.3)	−0.012 (−0.6)	0.013 (1.0)	0.002 (0.1)	0.007 (0.5)	−0.014 (−0.6)	0.004 (0.4)	0.012 (0.8)	−0.022 (−0.9)	−0.009 (−0.6)
1992	−0.006 (−0.5)	−0.033 (−2.2)	−0.004 (−0.4)	−0.026 (−1.2)	−0.013 (−0.8)	0.001 (0.1)	−0.002 (−0.2)	−0.001 (−0.1)	0.024 (1.0)	0.005 (0.3)
1993	−0.012 (−1.0)	−0.044 (−2.6)	−0.018 (−1.8)	−0.025 (−1.1)	−0.030 (−1.5)	−0.029 (−1.1)	0.009 (1.1)	−0.072 (−5.8)	−0.044 (−2.2)	−0.015 (−1.1)
1994	−0.003 (−0.3)	0.035 (2.1)	0.012 (1.2)	−0.001 (0.0)	0.022 (1.5)	0.014 (0.7)	−0.004 (−0.4)	0.056 (5.1)	−0.032 (−1.7)	−0.005 (−0.4)
1995	0.034 (2.8)	0.039 (2.2)	0.011 (1.2)	0.041 (2.1)	0.008 (0.6)	0.035 (1.5)	0.011 (1.5)	0.005 (0.4)	0.032 (2.2)	0.033 (2.8)
1996	−0.005 (−0.5)	0.012 (0.9)	−0.004 (−0.3)	0.002 (0.1)	0.023 (1.8)	0.009 (0.7)	−0.014 (−1.9)	−0.004 (−0.4)	0.021 (1.2)	−0.013 (−0.9)
1997	−0.006 (−0.6)	−0.005 (−0.4)	−0.031 (−3.0)	−0.018 (−1.2)	−0.019 (−1.5)	0.001 (0.0)	0.004 (0.5)	−0.003 (−0.4)	−0.024 (−1.3)	0.005 (0.4)
1998	0.021 (2.6)	0.008 (0.6)	0.021 (2.1)	0.025 (2.0)	0.002 (0.1)	−0.018 (−1.6)	−0.007 (−0.9)	0.009 (0.9)	0.045 (3.0)	−0.002 (−0.2)
Sigma squared	0.024	0.025	0.029	0.032	0.025	0.031	0.019	0.029	0.031	0.021
No. of firms	325	153	305	141	219	156	356	361	157	183
No. of observations	1321	752	1363	587	924	704	1795	1598	644	826

[a] Sample period: 1991–1998.
[b] t-ratios computed using (unbalanced panel) robust standard error formulas.
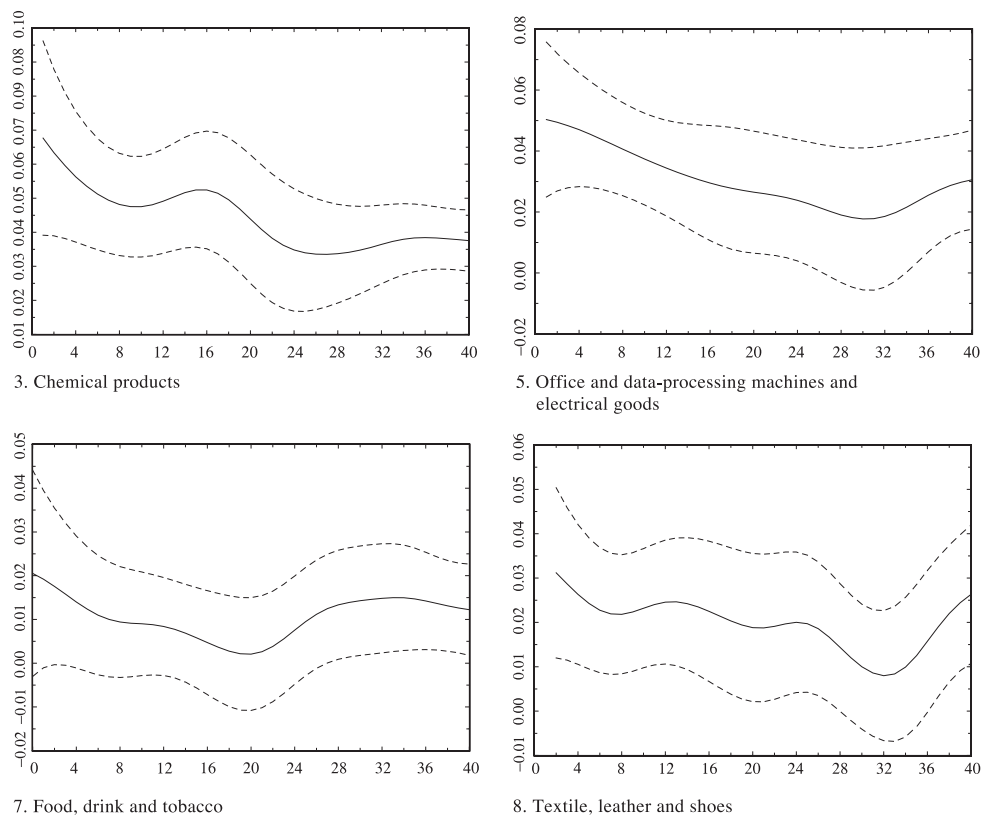
The a^* (τ) function by industries[1]

3. Chemical products

5. Office and data-processing machines and
   electrical goods

7. Food, drink and tobacco

8. Textile, leather and shoes

Fig. 2. The $a^*(\tau)$ function by industries. Values of the function $a^*(\tau)$ (productivity growth), estimated with semiparametric techniques, against age. Upper and lower 95% confidence bands.

among the group of the oldest firms, the share of the surviving high performers increases, which tends to displace the mean upwards. However, notice that these figures are depicted without controlling for innovation (we have already pointed out that the oldest firms are often also good innovators), and that the precision of the estimate is lower than in the aggregate.

6. Implications and concluding remarks

This paper has looked directly at the productivity growth impact of process innovations introduced by firms along their different ages. The main findings may be summarized as follows. Newborn firms tend to show higher rates of productivity growth which, as time goes by, tend to converge on average to common (activity-specific) growth rates. Process innovation clearly induces, however, extra productivity growth at any point in this process.

Extra growth tends to persist for a number of years, but it is followed by below-average productivity growth, very close to a halt, if innovation then stops. Let us briefly develop some interpretations and implications of these findings.

Firstly, how can average productivity growth of every firm, extra growth of innovators and below-average growth of firms that cease for a long time to perform as innovators be reconciled? Average growth is likely to be based on ingredients like incorporated technological change, learning over time and the absorption of "spillovers." We understand by absorption of spillovers the incorporation of productive ideas and methods, flowing more or less freely in the market, from the innovations and experience of others. These incorporations are unlikely to be classified as innovations by firms because they have been introduced years before in the market (and some of them are perhaps already partially commercialised). The same technical ideas and methods would in this way play several roles. In particular, they would be an important part of average productivity growth as well as the reason for the leading productivity growth of their (early) introducers as innovations, but they cannot definitely bring in new productivity growth to the introducers who have remained at the initial point. Productivity of these firms will have, in the end, grown by more or less the same amount as the productivity of any noninnovator. Our estimates then suggest an industry full of these "dynamic" spillovers, in which process innovations are generalised with some lag, bringing productivity growth to even the noninnovative firms.

Secondly, entrant firms are likely to derive their high rates of productivity growth from a mix, with unknown weights, of innovative processes and the course of learning. Notice that our estimation does not say a word about efficiency levels, but only indicates that new firms' productivity increases more rapidly. Then, in principle, we can attribute these productivity improvements either to the potential of their completely new processes brought into the market to jump away from the average efficiency levels or just the necessity to adapt quickly to them. Evidence on low surviving rates of entrant firms seems to point to the importance of the learning factor. But trying to disentangle the relative weight of these two effects seems a relevant question which cannot be answered with the present model.

Thirdly, the impact of process innovation seems to spread beyond what can be picked up by the simplest "knowledge capital" models. In these models, the perpetual inventory method of capital stock construction implies time productivity growth effects proportional to the contemporaneous net rates of R&D investment (investment over cumulated capital minus depreciation). The obtained evidence departs from this model in several aspects. The productivity growth impact takes place when a process innovation is introduced and it is spread over a number of years. But "knowledge capital" models possess the interesting feature of trying to weigh innovations by their value. This suggests the relevance of trying to advance in the modelling of innovation-specific investment weights and their dynamic effects.

Acknowledgements

the 27th EARIE Conference (Lausanne), the IO Workshop at the University of Vigo, the Workshop on Demography of Firms and Industries (Barcelona) and the Conference on Empirical Economics of Innovation and Patenting (Mannheim) for useful comments. Comments by two anonymous referees are also gratefully acknowledged. This research has been partially funded by the CICYT project SEC2000-0268. Elena Huergo also acknowledges financial support from the FEP.

Appendix A. Data appendix

All employed variables come from the information furnished by firms to the ESEE survey (see footnote 2). The employed sample results from dropping the observations for which the data needed to perform the exercise are incomplete. Composition of the unbalanced panel sample in terms of time observations is shown in Table A.1. The table also reports the frequency with which firms introduce process innovations. The columns "% of process innovations" are constructed by averaging across firms the relative frequencies or proportions of their time observations in which they report process innovations. The number of firms by size intervals are the following: up to 20 workers: 712; 21–50 workers: 555; 51–100 workers: 179; 101–200 workers: 203; 201–500 workers: 488; and more than 500 workers: 219. Fig. A.1 depicts the histogram of the variable age. Notice its bimodal character after grouping values at 40 years.

A.1. Detail on variables construction

Solow residual: Computed using the Tornqvist index $\theta = y - s_L l - s_K k - s_M m$, where the input measures are in log differences and the s weights for moment t are average cost shares for years t and $t-1$. Output and intermediate consumption real changes are obtained by deflating, respectively, (sales+inventory changes) and (raw materials and services purchases+energy and fuel costs). The price indices used are Paasche-type firm individual indices, constructed starting from the price changes on output and inputs reported by firms. Labour input changes are the changes in total effective hours of work (normal hours+o-

Table A.1
Number of firms, time observations and frequency of process innovation

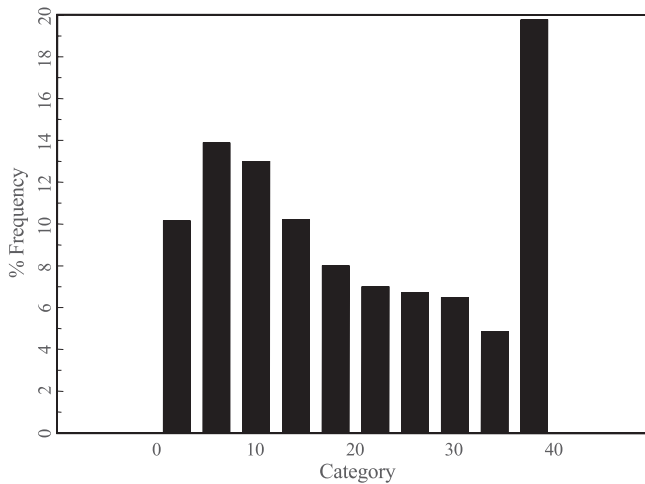| Time obs.	Total sample			Strictly uncensored sample		
	No. of firms	No. of observations	% of process innovations	No. of firms	No. of observations	% of process innovations
1	393	393	33.8	342	342	64.7
2	353	706	31.6	288	576	49.1
3	221	663	30.6	165	495	51.5
4	278	1112	31.1	197	788	48.5
5	159	795	37.9	169	845	48.1
6	180	1080	30.2	155	933	43.8
7	190	1330	35.3	155	1085	51.6
8	582	4656	38.2	279	2232	57.9
Total	2356	10,735	34.0	1750	7296	52.2

Fig. A.1. The histogram of age.

vertime-lost hours), and capital variations are computed from a measure of the stock of capital obtained starting from the firms' investments in equipment goods. Cost and cost shares are also computed using the labour cost per worker and a user cost of capital calculated as the firm's interest rate paid by long-run debt plus a sectoral estimate of equipment depreciation minus the rate of change of a capital goods price index.

Capacity utilisation: yearly percentage of utilisation of installed capacity reported by firms.

Merger/acquisition and scission variables: dummy variables that take the value 1 the year in which a merger/acquisition or a scission has taken place. When two observed firms merge, only the biggest is supposed to survive.

Size variables: dummy variables based on the average number of workers of the firm during the year.

Industry variables: 18 industry dummy variables classification (see Table 1B) which constitutes an adaptation of a standard NACE classification, and the 10 industries classification aggregates the previous one (to have a significant number of firms in each industry) in the following way: 1=1+4, 2=2, 3=3+17, 4=5, 5=6+7, 6=8+9, 7=10+11+12, 8=13+14, 9=15, 10=16. Firm numbers by industry appear in Table 2.

Age: computed from the difference between the current year and the constituent year reported by the firm; when this difference is higher than 40 years, we change it to a unique category of 40 or more years. This is the maximum life span with economic meaning in Spanish manufacturing circa 1998. Higher ages reported by firms are probably important in terms of prestige, but we assume that cannot have technological content. The unit surveyed is the firm, not the plant or establishment, and some closely related firms answer as a group. Constitution of groups and mergers implying major law changes including a new constituent year introduces a small number of ambiguous ages that we have respected. Ages distribution, given the character of the sample, is expected to be representative of the

ages distribution in manufacturing population. Quartiles of the ages distribution of firms in the sample for 1991 are 7, 17 and 31 years.

Process innovations: a process innovation is assumed to have occurred when the firm answers positively to the following request: "Please indicate if during the year 199× your firm introduced some significant modification of the productive process (process innovation). If the answer is yes, please indicate the way: a) introduction of new machines; b) introduction of new methods of organisation; c) both."

References

Audretsch, D., 1995. Innovation and Industry Evolution. MIT Press, Cambridge, Massachusetts.

Crepon, B., Duguet, E., Mairesse, J., 1998. Research, innovation and productivity: an econometric analysis at the firm level. Economics of Innovation and New Technology 7, 115–158.

Delgado, M.A., Mora, J., 1995. Nonparametric and semiparametric inference with discrete regressors. Econometrica 63, 1477–1484.

Delgado, M.A., Jaumandreu, J., Martín, A., 1998. Input cost, capacity utilization, and substitution in the short run. Spanish Economic Review 1, 239–262.

Griliches, Z., 1979. Issues in assessing the contribution of R&D to productivity growth. Bell Journal of Economics 10, 92–116.

Griliches, Z., 1995. R&D and productivity: econometric results and measurement issues. In: Stoneman, P. (Ed.), Handbook of Economics of Innovation and Technological Change. Blackwell, Oxford, pp. 52–89.

Griliches, Z., 2000. R&D, Education and Productivity. Harvard Univ. Press, Cambridge, Massachusetts.

Hall, R.E., 1990. Invariance properties of Solow's productivity residual. In: Diamond, D. (Ed.), Growth, Productivity, Unemployment. MIT Press, Cambridge, Massachusetts, pp. 71–155.

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

Huergo, E., Jaumandreu, J., 2004. How does probability of process innovation change with firm age? Small Business Economics (in press).

Klette, T.J., 1996. R&D, scope economies, and plant performance. RAND Journal of Economics 27, 502–522.

Pagan, A., Ullah, A., 1999. Nonparametric Econometrics. Cambridge Univ. Press, Cambridge.

Robinson, P.M., 1988. Root-*n* consistent semiparametric regression. Econometrica 56, 97–118.

Speckman, P., 1988. Kernel smoothing in partially linear models. Journal of the Royal Statistical Society, Series B 50, 413–446.

Stock, J.H., 1989. Nonparametric policy analysis. Journal of the American Statistical Association 84, 567–575.

Suits, D., 1984. Dummy variables: mechanics vs. interpretation. Review of Economics and Statistics 66, 177–180.

Wand, M.P., Jones, M.C., 1995. Kernel Smoothing. Chapman & Hall, London.