Firms' age, process innovation and productivity growth

Elena Huergo*
Universidad Complutense de Madrid

Jordi Jaumandreu**
Universidad Carlos III de Madrid

January, 2002

Abstract

This paper looks directly at the impact of a firm's 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 a number of years.

JEL Classification: D24, L6, O3.
Keywords: Productivity, process innovation, Solow residual, firm age.

We are grateful to M.Delgado and J.Ruiz for detailed comments on previous versions of this paper, and to D.Audretsch, J.C.Fariñas, R.Griffith, N.Janz, J.Mairesse, G.Reid and M.Vivarelli as well as the audiences at the 27th EARIE Conference (the University of Lausanne), the IO Workshop at the University of Vigo and the Workshop on Demography of Firms and Industries (the Universidad de Barcelona) for useful comments. This research has been partially funded by the CICYT project SEC2000-0268. Elena Huergo also acknowledges financial support from the Fundación Empresa Pública.

*Dpto. Fundamentos del Análisis Económico I. Facultad de Ciencias Económicas y Empresariales. Campus de Somosaguas. 28223 Madrid. Spain. Tel.: 34 913942408. Fax: 34 913942651.e-mail: ehuergo@ccee.ucm.es
Dpto. Economía. Facultad de Ciencias Sociales y Jurídicas. C/ Madrid 126, 28903 Getafe. Spain. Tel.: 34 916245737. Fax: 34 6249329. e-mail: jordij@eco.uc3m.es

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.[1]) The standard form of these exercises has been the construction of a 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 the 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, and Crepon, Duguet and Mairesse (1998) for a departure from the traditional modelling using innovation data). This provides an important reason for looking more closely at every one of these processes.

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 2,300 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

[1] See Hall and Mairesse (1995) and Klette (1996) for recent examples of this literature.

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) Solow residual, corrected for (possible) non-constant returns to scale and using other necessary controls. To pick up the presumably highly non-linear relationships with age and process innovation, we devise and use a semiparametric model and techniques of estimation. 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 to a yearly 1.25% on average. Process innovations at some point then lead to some extra productivity growth, which tends to persist, somewhat attenuated 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 the following years. This suggests that process innovations are very much in a race to bring future growth.

The rest of the paper is organised as follows. Section two is devoted to establishing the framework of measurement of productivity and Section three to the way to estimate the impact of age and innovation. Section four deals with the data and variables and Section five presents the empirical results. Section 6 comments on implications and concludes. A Data Appendix 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 firms are characterised by production functions of the type

$$Y_{it} = A_i(t)F_i(K_{it}, L_{it}, M_{it}) \qquad (1)$$

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.

Total differentiation of equation (1) yields

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

where the small letters employed to represent output and inputs denote logarithmic differences, the ε's stand for the respective input elasticities, and the term $a_i(t)$ stands for the proportional change $\frac{dA_i}{A_i}$ or productivity growth. In order to investigate productivity growth characteristics, we will specify this last term, without loss of generality, as the sum of a firm and time idiosyncratic term and a function (common to all firms) of the vector of variables z whose role we want to assess:

$$a_i(t) = a_{it} + a(z_{it})$$

Now we must obtain an estimable model. Cost minimisation implies that input elasticities equal the products of scale elasticity and cost shares (robustness with respect to market power implies the use of cost shares, see e.g. Hall (1990)). Then, using s to denote cost shares and assuming a common scale elasticity γ across firms we can write:

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

Rearranging terms, this expression can be easily transformed into an econometric model that links the observable (cost shares based) Solow residual, a correction for the scale elasticity effect on productivity, and productivity growth. The model is

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

where θ represents the Solow residual, (γ-1) is a parameter to be estimated involving scale elasticity, v stands for the weighted sum of input variations, and e is a zero mean disturbance which we will assume uncorrelated across firms and time.

Let us briefly comment on 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 will include the variations in the firms' capacity utilisation, 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. As we are interested in the unknown function representing average growth (including the constant term of the regression), 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 (2) in the slightly more compact form:

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

Finally, we leave a(.) as an unknown function aimed at picking up the presumably highly non-linear 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 take the age of the firm, τ, as the number of years it has been in the market, establishing a maximum category of 40 years or more (see the Data Appendix for details). The impact of age can be guessed to be highly non-linear, and hence difficult to pick up with standard linear regression methods. Consequently, we specify age as the argument of the function $a(\cdot)$ and we use an estimator of the function based on the semi-parametric estimation of (3). Robinson (1988) and Speckman (1988), among others, have proposed estimators of this type. Assuming τ and e uncorrelated, from (3) we have

$$\theta_{it} - E(\theta_{it} \mid \tau_{it}) = \left[x_{it} - E(x_{it} \mid \tau_{it})\right] \beta + e_{it}$$

and the semiparametric estimator of β is the ordinary least squares estimator (OLS) after replacing the conditional expectation functions by some nonparametric estimate.[2] Once the β parameters have been estimated, we can recover an estimate of the $a(\tau)$ function:

$$\hat{a}(\tau) = \hat{E}(\theta_{it} \mid \tau_{it}) - \hat{E}(x_{it} \mid \tau_{it})\hat{\beta} \tag{4}$$

Let us cite two important properties of this estimate. Firstly, 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. Secondly, to implement the estimator, we compute the conditional expectations using nonparametric regressions,

[2] We use the kernel regression Nadaraya-Watson estimator; see, for example, Wand and Jones (1995).

and estimate (4) is hence derived from a linear combination of nonparametric regressions. As firm numbers become lower at some ages, the variance of the estimator becomes higher and the (implicit) confidence intervals are broader.[3]

The next questions to answer are whether process innovations introduced by firms along their lives induce extra productivity growth, in 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 this, we will enlarge the specification of the $a(\cdot)$ function by adding the innovation effects on productivity growth (at any point of life). Calling α_s the successive impacts of an innovation, from moment 0 (introduction) to T (the last considered lag), we have the function $a(\tau)+\sum_{s=0}^{T}\alpha_s$.

Consistent estimation of the $\alpha's$ may be achieved in our context by the simple enlargement of the econometric model with the terms $\alpha_0 d_{it}+\sum_{s=0}^{T}\delta_s tes_{it}$,[4] where the artificial variables on the set $(d, te1, te2, ...teT)$ indicate that an innovation has been introduced and the time elapsed since then, respectively. These variables are defined as follows:

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

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

[3] Variance of estimated expectations is $V(\hat{m})=\frac{1}{nh}\frac{\sigma^2(x)}{f(x)}\int K(s)^2\,ds$, where *K(.)* is the kernel density function, h the smoothing parameter and σ^2 and *f(x)* the variance and density associated with each *x*.

[4] This may be understood as an estimator of the type studied by Delgado and Mora (1995).

Estimates of α_s are given by $\hat{\alpha}_0$, and $\hat{\alpha}_s = \hat{\alpha}_0 - \hat{\delta}_s$ 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 are 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 non-innovating firms (the bulk of the firms for which we have not yet observed any innovation by the final year of the sample), or 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 and zero for the length of the sample. However, our non-parametric 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 2,300 firms surveyed during the period 1990-1998.[5] Details are provided in the Data Appendix. This sample is approximately representative of manufacturing, and hence inferences can be considered globally valid for this ambit. 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 (see the Appendix for details). A process innovation

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

is assumed to occur when the firm answers positively to the 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 major innovative changes in their productive process and the frequency of these changes.[6]

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 three years. 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 that are slightly decreasing. The sample average relative frequency of innovation of the strictly uncensored sample (all the non innovation-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 (2002) 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 as well as the oldest firms tend to present a somewhat lower probablity. Exiting firms are clearly associated to lower levels of pre-exit innovations.[7]

[6] 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.

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 ten industries.

Table 1 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, with age entering the equation as the argument of an unknown function. The third estimate adds to this function 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, applied to the strictly non-censored sample (68% of the previous data points). Table 1 (cont.) 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, b and c of Figure 1 depict the functions obtained by means of the semiparametric estimates of the model, 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 interest- stress the validity of the framework employed. Firstly, the average

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

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 two following 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, we can account for all the differences in productivity growth apparently linked to size that emerge so often in empirical exercises. Thirdly, some industries (1/3) 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 1 and Fig 1, panel a). Entrant firms present high productivity growth (more than 5%) and, although decreasing as time goes by, average growth tends to be higher than average until firms reach 8 years in the market. At this age, growth tends to stabilise by about 2% (the wandering of the curve denotes higher variance as age becomes higher, but shows no clear trend). That is, productivity growth tends clearly 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 fully 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 non-linearity of the relationship is missed.

The third and fourth estimates introduce the set of dummies aimed at picking up the effects of innovation. The third estimate (Table 1 and Figure 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 1.25%. At the same time, innovation shows a clear contemporaneous impact on productivity growth by about 1.5%. Positive impacts seem to persist for three years with a lower average value of 0.7% (estimated, however, with a high variance), but they also seem to be followed by three more years in which average firm productivity growth is reduced by about 1.2% a year. All happens as if product innovation moved ahead future productivity growth by three or four years.

When the same specification is estimated employing the strictly non-censored sample, the average values of the age function again move upwards (convergence value is now located at about 2.8%), 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 nature cannot be established without further non-available information. But, on the other, 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 three more years), but productivity growth also tends to fall and even fully stop the following years if new innovations are not introduced as well. This sensibly suggests that innovating firms basically made an effort to move forward future productivity increases, which other firms will reach at a different pace.

Table 2 reports semiparametric estimates of the age model for a disaggregation of manufacturing in ten sectors. The number of data points for each estimation are obviously 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 of manufacturing. The lack of statistical precision makes the industry semiparametric estimates of the age/innovation effects model less useful here (there are too few observations on each innovation lag) and they are not reported.[8]

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 on productivity of capacity utilisation range from a reasonable maximum with transport equipment to a minimum with the food, drink and tobacco industry. 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.

Figure 2 depicts the industry age functions obtained by means of semiparametric estimates. Panels of Figure 2 confirm that with every industry there is a starting period of high relative productivity growth, corresponding to the initial years of a firm's life, at which productivity growth tends however to decrease steadily. The number of years of this initial period varies between 8 and 12 according to the firm's activity. After this, the higher variance associated with the smaller number of firms makes it difficult to summarize the patterns (particular observations acquire an important weight). Panels show, however, some interesting features. In some sectors, the first years of productivity growth decay seem to inscribe in a more long term towards lower productivity growth rates. But, at the same time, there are sectors in which firms of intermediate age and even the oldest firms show high productivity growth rates, the likely output of selection.

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

6. Implications and concluding remarks

This paper has looked directly at the productivity growth impact of process innovations introduced by firms along the different ages of the firm. 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 tends, however, to induce extra productivity growth at some point in this process. Growth tends to persist for a number of years, and is followed by a halt if innovation then stops. Let us briefly develop some interpretations and implications of these findings.

Firstly, as we have already emphasized, process innovation appears to be an effort of firms to bring today productivity growth which tomorrow will be generalized to all firms. This is the likely reason for the halt in productivity growth observed in firms introducing innovations followed by a large delay without innovations. Innovators seem to reach in advance the productivity improvements that other firms will obtain with some lag at a more regular pace. Notice that this picture agrees very well with what we would expect from an industry full of "dynamic" spillovers, in which process innovations were systematically generalized with some lag, bringing productivity growth to even the non-innovative 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, 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 quickly adapt 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 posses the interesting feature of trying to weigh innovations by their value. This suggests the relevance of trying to advance in the modeling of innovation-specific investment weights and their dynamic effects.

Data Appendix

All employed variables come from the information furnished by firms at the survey ESEE (see footnote 1). 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. Figure A.1 depicts the histogram of the variable age. Notice its bimodal character after grouping values at 40 years.

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 + overtime - 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 computed also 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 one 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 1 (cont.)) 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 firms closely related answer as a group. Constitution of groups and mergers implying major law changes including a new constituent year introduce 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 199x 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.

Crepon, B., E. Duguet, and J. Mairesse (1998), "Research, innovation and productivity: An econometric analysis at the firm level," *Economics of Innovation and New Technology*, 7, 115-158.

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

Delgado, M. A., and J.Mora (1995), "Nonparametric and semiparametric inference with discrete regressors," *Econometrica* , 63, 1477-1484.

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 P. Stoneman (ed.), *Handbook of the economics of innovation and technological change*, Blackwell, Oxford.

Griliches, Z. (2000), R*&D, Education and Productivity,* Harvard University Press.

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

Huergo, E. and J. Jaumandreu (2002), "How does probability of process innovation change with firm age?," mimeo.

Klette, T.J. (1996), "R&D, scope economies, and plant performance," *RAND Journal of Economics* 27, 502-522

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.

Suits, D. (1984), "Dummy variables: mechanics vs. interpretation," *Review of Economics and Statistics* 66, 177-180

Wand, M.P. and M.C. Jones (1995), *Kernel Smoothing*, Chapman & Hall.

Table 1

Results from the estimation of models $\theta = x\beta + a(\tau) + e$ and $\theta = x\beta + a(\tau) + \alpha_0 d + \sum \delta_s tes + u$

(Age and age/innovation effects models)

Dependent variable: (Cost based) Solow residual[1]

Explanatory variables	Coefficients and t-ratios[2]							
	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[3]	Included		Included		See Table 1 (cont.)		Included	
Constant	0.030	(8.8)						
Age	-0.0003	(-2.2)	Figure 1, panel a		Figure 1, panel b		Figure 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	2,356		2,356		2,356		1,750	
No. of observations	10,735		10,735		10,735		7,293	

[1] Sample period: 1991-1998. [2] t-ratios computed using (unbalanced panel) robust standard errors formulas.

[3] 8 time dummies, 6 size dummies and 18 industry dummies, with coefficients of each set constrained to add zero (Suits method)

Table 1 (cont.)
Dummy coefficients of
semiparametric estimate of the age and innovation effects model[1]

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 non-ferrous metals	0.016	(2.2)
Non-metallic 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)

[1] Third estimate in main panel.

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

Dependent variable: (Cost based) Solow residual[1]

	Coefficients and t-ratios[2]									
Explanatory variables	1. Ferrous and non-ferrous metals and metal products		2. Non-metallic minerals		3. Chemical products		4. Industrial and agricultural machinery		5. Office and data-processing machines and electrical goods	
Scale correction	-0.215	(-5.7)	-0.298	(-5.7)	-0.205	(-4.0)	-0.187	(-4.5)	-0.205	(-3.7)
Capacity utilisation	0.059	(2.4)	0.143	(2.8)	0.098	(3.5)	0.104	(1.7)	0.067	(1.3)
Merger	-0.014	(-0.4)	0.049	(1.1)	0.067	(2.3)	-0.062	(-2.2)	0.096	(2.5)
Scission	-0.080	(-1.2)	0.038	(0.7)	-0.176	(-3.3)	-0.559	(-11.4)	-0.050	(-0.9)
Up to 200 workers	-0.003	(-0.9)	-0.007	(-1.5)	-0.004	(-1.3)	-0.011	(-2.0)	0.000	(0.0)
More than 200 workers	0.003	(0.9)	0.007	(1.5)	0.004	(1.3)	0.011	(2.0)	0.000	(0.0)
1991	-0.022	(-1.3)	-0.012	(-0.6)	0.013	(1.0)	0.002	(0.1)	0.007	(0.5)
1992	-0.006	(-0.5)	-0.033	(-2.2)	-0.004	(-0.4)	-0.026	(-1.2)	-0.013	(-0.8)
1993	-0.012	(-1.0)	-0.044	(-2.6)	-0.018	(-1.8)	-0.025	(-1.1)	-0.030	(-1.5)
1994	-0.003	(-0.3)	0.035	(2.1)	0.012	(1.2)	-0.001	(0.0)	0.022	(1.5)
1995	0.034	(2.8)	0.039	(2.2)	0.011	(1.2)	0.041	(2.1)	0.008	(0.6)
1996	-0.005	(-0.5)	0.012	(0.9)	-0.004	(-0.3)	0.002	(0.1)	0.023	(1.8)
1997	-0.006	(-0.6)	-0.005	(-0.4)	-0.031	(-3.0)	-0.018	(-1.2)	-0.019	(-1.5)
1998	0.021	(2.6)	0.008	(0.6)	0.021	(2.1)	0.025	(2.0)	0.002	(0.1)
Sigma squared	0.024		0.025		0.029		0.032		0.025	
No. of firms	325		153		305		141		219	
No. of observations	1321		752		1363		587		924	

[1] Sample period: 1991-1998.
[2] t-ratios computed using (unbalanced panel) robust standard errors formulas.

Table 2 (cont.)
Industry semiparametric estimations of model $\theta = x\beta + a(\tau) + e$

Dependent variable: (Cost based) Solow residual[1]

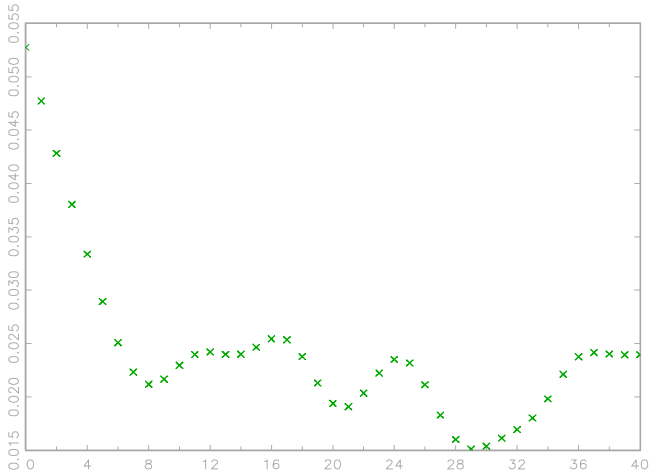
Explanatory variables	6. Transport equipment		7. Food, drink and tobacco		8. Textile, leather and shoes		9. Timber and furniture		10. Paper and printing products	
	Coefficients and t-ratios[2]									
Scale correction	-0.002	(0.0)	-0.299	(-6.4)	-0.328	(-10.9)	-0.243	(-4.7)	-0.349	(-5.9)
Capacity utilisation	0.142	(3.1)	0.056	(2.7)	0.089	(2.8)	0.070	(2.0)	0.068	(1.7)
Merger	-0.058	(-1.4)	0.071	(2.6)	0.061	(1.5)	-0.108	(-5.1)	0.130	(2.1)
Scission	-0.032	(-1.1)	0.062	(1.2)	-0.251	(-1.7)	-0.052	(-0.4)	0.029	(0.3)
Up to 200 workers	-0.001	(-0.3)	0.008	(2.1)	-0.001	(-0.2)	-0.016	(-2.2)	0.012	(1.9)
More than 200 workers	0.001	(0.3)	-0.008	(-2.1)	0.001	(0.2)	0.016	(2.2)	-0.012	(-1.9)
1991	-0.014	(-0.6)	0.004	(0.4)	0.012	(0.8)	-0.022	(-0.9)	-0.009	(-0.6)
1992	0.001	(0.1)	-0.002	(-0.2)	-0.001	(-0.1)	0.024	(1.0)	0.005	(0.3)
1993	-0.029	(-1.1)	0.009	(1.1)	-0.072	(-5.8)	-0.044	(-2.2)	-0.015	(-1.1)
1994	0.014	(0.7)	-0.004	(-0.4)	0.056	(5.1)	-0.032	(-1.7)	-0.005	(-0.4)
1995	0.035	(1.5)	0.011	(1.5)	0.005	(0.4)	0.032	(2.2)	0.033	(2.8)
1996	0.009	(0.7)	-0.014	(-1.9)	-0.004	(-0.4)	0.021	(1.2)	-0.013	(-0.9)
1997	0.001	(0.0)	0.004	(0.5)	-0.003	(-0.4)	-0.024	(-1.3)	0.005	(0.4)
1998	-0.018	(-1.6)	-0.007	(-0.9)	0.009	(0.9)	0.045	(3.0)	-0.002	(-0.2)
Sigma squared	0.031		0.019		0.029		0.031		0.021	
No. of firms	156		356		361		157		183	
No. of observations	704		1795		1598		644		826	

[1] Sample period: 1991-1998.
[2] t-ratios computed using (unbalanced panel) robust standard errors formulas.

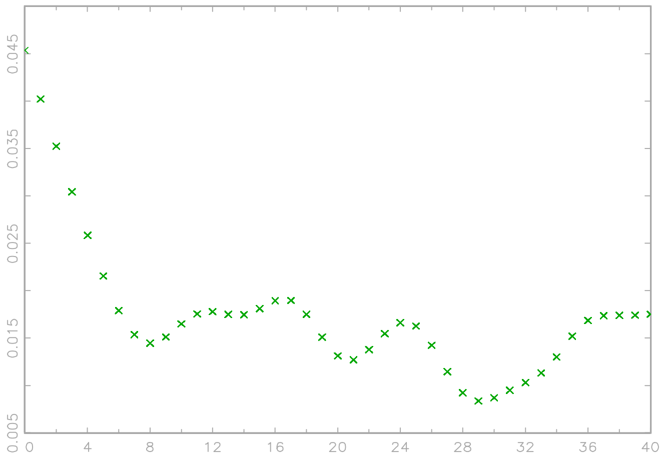
Table A.1
Number of firms, time observations and frequency of process innovation

Total sample				Strictly uncensored sample			
Time obs.	No. of firms	No. of observations	% of process innovations	Time obs.	No. of firms	No. of observations	% of process innovations
1	393	393	33.8	1	342	342	64.7
2	353	706	31.6	2	288	576	49.1
3	221	663	30.6	3	165	495	51.5
4	278	1,112	31.1	4	197	788	48.5
5	159	795	37.9	5	169	845	48.1
6	180	1,080	30.2	6	155	933	43.8
7	190	1,330	35.3	7	155	1,085	51.6
8	582	4,656	38.2	8	279	2,232	57.9
Total	2,356	10,735	34.0	Total	1,750	7,296	52.2

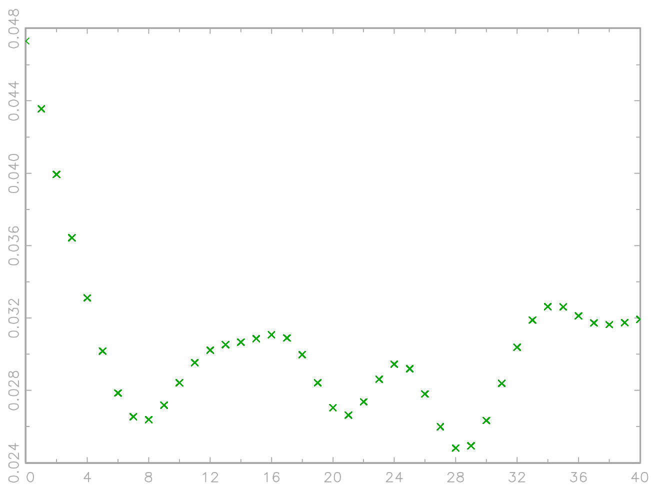
Figure 1
The $a(\tau)$ function[1]



Panel a



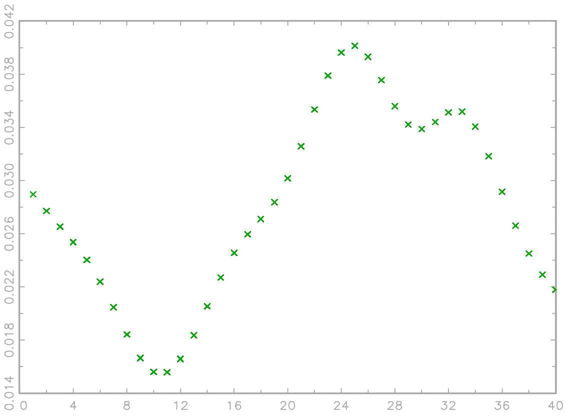
Panel b



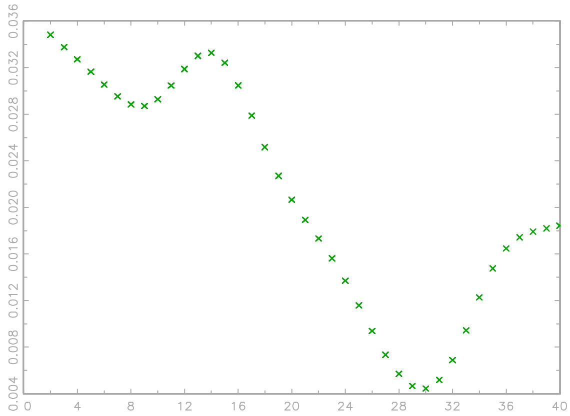
Panel c

[1] Values of the function $a(\tau)$ (productivity growth), estimated with semiparametric techniques, against age.

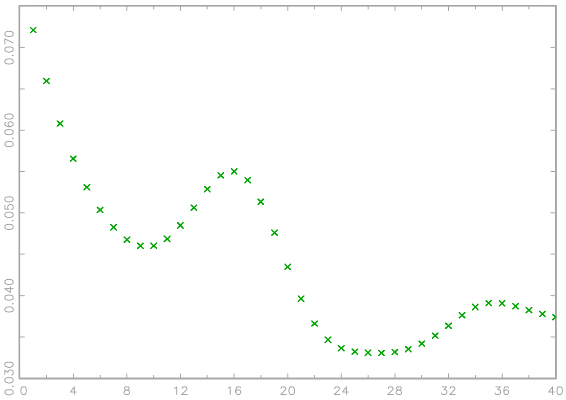
Figure 2
The $a(\tau)$ function by industries[1]



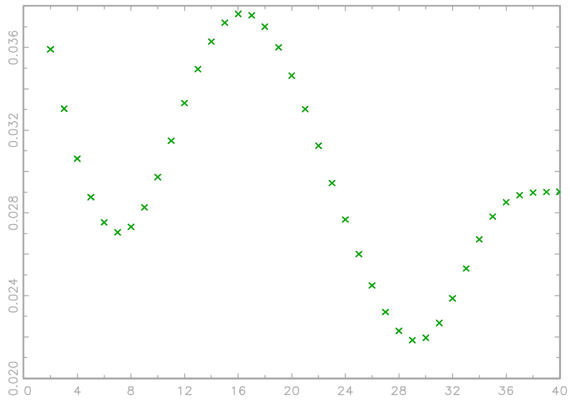
Sector 1



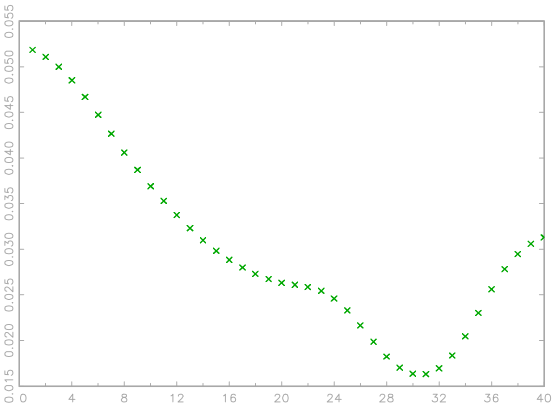
Sector 2



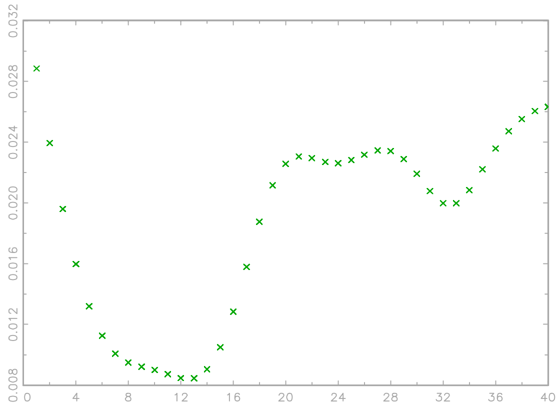
Sector 3



Sector 4

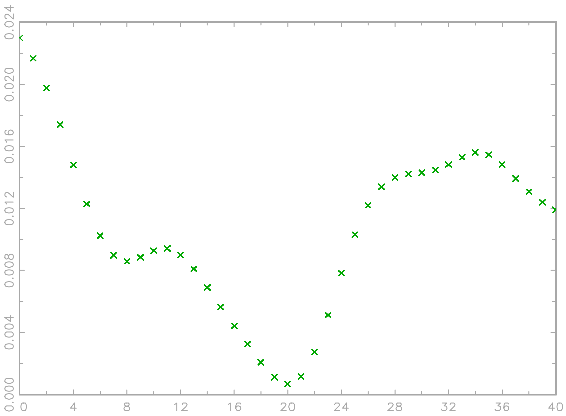


Sector 5

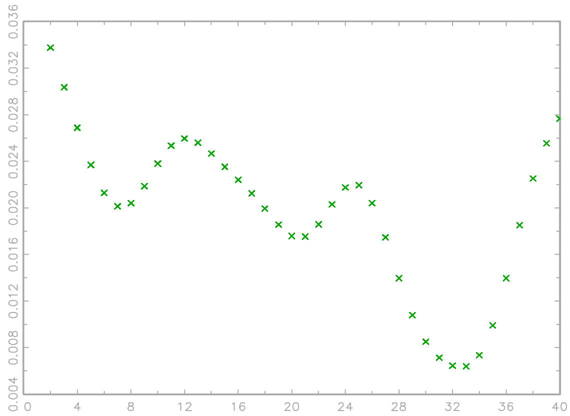


Sector 6

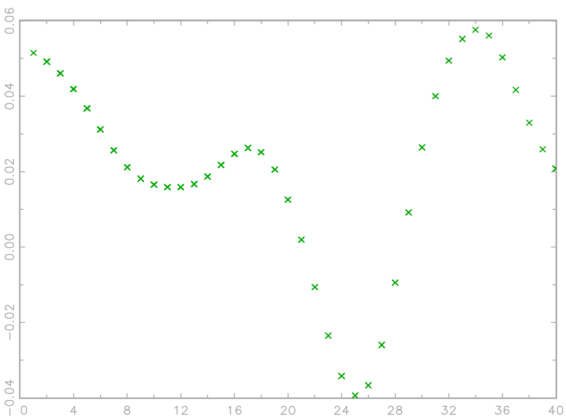
Figure 2 (cont.)
The $a(\tau)$ function by industries[1]



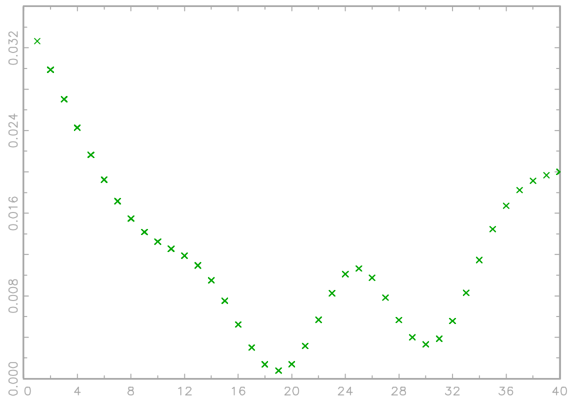
Sector 7



Sector 8



Sector 9



Sector 10

[1] Values of the function $a(\tau)$ (productivity growth), estimated with semiparametric techniques, against age.

Figure A.1
The histogram of age