

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2004-25

**The Importance of R&D for
Innovation : A Reassessment
Using French Survey Data**

J. MAIRESSE¹
P. MOHNEN²

Les documents de travail ne reflètent pas la position de l'INSEE et n'engagent que leurs auteurs.

Working papers do not reflect the position of INSEE but only the views of the authors.

¹ CREST-INSEE, 15 Boulevard Gabriel Péri, 92245 Malakoff Cedex, France et EHESS.

Email : mairesse@ensae.fr

² MERIT, Maastricht University, PO Box 616, 6200 MD Maastricht, Pays-Bas et CIRANO.

Email . p.mohnen@merit.unimaas.nl

The Importance of R&D for Innovation: A Reassessment Using French Survey Data⁺

Jacques Mairesse* and Pierre Mohnen**

Abstract - This paper compares the contribution of R&D to innovation in terms of the various innovation output measures provided by the third Community Innovation Survey (CIS 3) for French manufacturing firms and in terms of accounting for interindustry innovation differences.

Key words: R&D, Innovation, Patents. JEL classification: C35, L60, O31, O33

L'importance de la R&D pour l'innovation : Un réexamen sur données d'enquêtes pour l'industrie française

Jacques Mairesse et Pierre Mohnen

Résumé: Cette étude estime la contribution de la R&D à l'innovation sur la base des différents indicateurs disponibles dans la 3ème Enquête Communautaire sur l'Innovation (CIS3), faite par le SESSI sur la période 1998-2000 pour les entreprises industrielles françaises. Elle examine dans quelle mesure les résultats obtenus rendent compte des différences observées entre industries dans leur propension et intensité à innover.

Mots-clefs : R&D, Innovation, Brevets. JEL Classification: C35, L60, O31, O33

⁺ Paper written for a special issue of The Journal of Technology Transfer (vol. 30, n°1-2, January 2005), in memory of Edwin Mansfield, edited by A. Link and F.M. Scherer.

We are grateful to SESSI for access to the CIS 3 and EAE data for French manufacturing under strict confidentiality requirements, and for help in constructing the final working sample. We are particularly indebted to Elizabeth Kremp for her expertise and collaboration. This study builds directly on Mairesse-Mohnen (2001, 2002) and is part of an ongoing research on R&D, Innovation and Productivity. It has benefited from support by the EU Commission as part of the project on "Innovation and Employment in European Firms" (IEEF).

* CREST-INSEE, 18, boulevard Gabriel Péri, 92245 Malakoff Cedex, France, et EHESS.
Mail : mairesse@ensae.fr

** MERIT, Maastricht University, PO Box 616, 6200 MD Maastricht, Pays-Bas, and CIRANO. Mail: p.mohnen@merit.unimaas.nl

1. Introduction

In this paper we want to pay tribute to Ed Mansfield for his pioneering research on the economics of technological change. We reassess the importance that R&D plays in the innovation process using the latest French data from CIS 3, the third Community Innovation Survey. We have chosen a topic that would have been dear to Ed Mansfield since innovation, and in particular R&D, were at the center of his research for the greatest part of his career. He was also, as Mike Scherer (2005) mentions in his introductory paper to this issue, one of the first user (and producer) of survey data on R&D and innovation. Nowadays these survey data are collected in a systematic way for large samples in many countries.

Traditionally, the importance of R&D is evaluated by relating R&D and production (or cost, or profit) data, estimating the output elasticity or rate of return of R&D from an extended Cobb-Douglas production (or cost, or profit) function, where a stock of R&D knowledge enters as a separate input. Ed Mansfield himself contributed to this literature (Mansfield, 1965-a, 1965-b, 1980). Another way of evaluating R&D is by estimating the value attached to R&D investment by capital markets (see the original paper by Griliches, 1981 and recent work by Hall and Oriani, 2004). Instead of being related to measures of economic performance, R&D can also be related to innovation indicators through some kind of knowledge production function. The returns to R&D have been usually estimated in this line of work in terms of patent counts or innovation counts (see for instance the debate between Jaffe, 1989, and Acs, Audretsch and Feldman, 1992).

This last approach is the one we follow here. Instead of using count data, we use the five dichotomous indicators of innovation and patents, and the three shares in total sales of innovative and patent protected sales, which are provided by the third Community Innovation Survey. We can thus make the distinction between indicators related to product and process innovations, and among the former between indicators for products new to the firm only (but already known in the market) and products new to the market (new can also mean substantial modifications of existing products). And we can confront these indicators with indicators on patent applications during a given time period and on patent holdings, in a way reflecting the distinction between flows and stocks of patents.

Prior to the late eighties, innovation surveys were conducted in isolated ways. Mansfield, as already indicated, based much of his work on company survey data (e.g. Mansfield et al., 1977). SPRU set up a database of innovations back in the mid 1970s, which was explored among others by Pavitt, 1984. In the early 1990s these surveys became institutionalized, in least in Europe, with the advent of the Community Innovation Surveys (CIS), which followed the guidelines of the Oslo Manual (OECD, 1992 and 1996). Up to now, there exist three official waves of CIS (CIS 1 for 1990-1992, CIS 2 for 1994-1996 and CIS 3 for 1998-2000). A few countries, notably France, Germany and the Netherlands, had actually an innovation survey prior to CIS 1. The fourth round of innovation survey is presently underway. The present study is based on the data of CIS3 for French manufacturing.

We assess the impact of R&D on innovation separately for high-tech and low-tech industries, and in two ways. We first estimate the marginal effects of R&D on the various innovation output indicators, controlling for other co-determinant factors, and correcting for both the selectivity

and endogeneity of R&D itself (section 5). We then ask how much of the inter-industry difference in innovation performance can be imputed to R&D and to the other factors explicitly taken into account in the analysis and how much remains to be explained or could be attributed to innovativeness (section 6). Before that, we give necessary explanations on our data (section 2) and our model (section 3), and provide a short preliminary descriptive analysis (section 4).

2. Data

Our sample consists of French manufacturing enterprises that responded to the CIS 3 survey, covering the years 1998 to 2000, and that have also been surveyed by the EAE survey (“Enquête Annuelle d’Entreprise” or Annual Survey of Enterprises) in these three years.¹

The CIS 3 survey, like the previous ones, is structured in such a way that specific filter questions lead to the selection of firms which are innovators as opposed to non-innovators. Only the former have to answer the full questionnaire. Firms are first asked whether they have introduced in 1998-2000 a new product or a new process, or whether they have had any ongoing or abandoned activities to do so during this period. If they answer positively to one of these questions (about 60% in our sample), they are asked additional information about their innovation outcomes, their R&D expenditures in 2000, and other characteristics. If they answer no to all the filter questions (about 40 %), they are considered as non-innovators so to say, having to report chiefly on their size, group affiliation and industry of main activity.²

We are thus left with little information about the non innovating firms, and a severe selectivity problem. In particular we have the information on the R&D expenditures of the innovating firms for the year 2000, in case they have engaged in R&D, but not for the non innovating firms. We have thus been lead to consider that all R&D performers were innovators, and conversely that all non-innovators were non-R&D performers. By merging the French annual R&D survey for 2000 with CIS 3, we have been able to check that it was not far from being the case: only 2 % of the R&D firms in the annual survey declare they were non-innovators in CIS 3, while about 60 % of innovating firms in CIS 3 declare that they were R&D performers (on a continuous basis).

¹ The CIS3 survey is a mandatory survey which has been conducted in 2001 (and available in its final form for research purpose in 2003). We restrict ourselves to manufacturing industries, excluding food industries, which have been surveyed by SESSI (“Service des Études et Statistiques Industrielles”). The food industries and service industries have been surveyed by other agencies than SESSI (namely the statistical offices of the Ministry of Agriculture and the Ministry of Education and Research, and by INSEE). The SESSI survey concerns some 5500 manufacturing firms with 20 employees or more, which have been chosen randomly, using the business register based on legal units and according to the following stratified sampling design: all firms over 500 employees, 1/2 for firms from 100 to 499 employees, 1/4 for firms from 50 to 99 employees, and 1/8 for firms from 20 to 49 employees. The rate of response was 86%, corresponding to an overall coverage of 89% of the total turnover for the manufacturing sector in 2000. We have restricted our sample to the firms which have also answered the EAE surveys in 1998, 1999 and 2000, and have also eliminated a number of firms for which important variables or ratios had missing values or extreme outlier values, for example firms with an R&D to sales ratio above 50%.

²Note that the non-innovating firms have to report also on the last group of questions of the survey, concerning patents and other protection methods. As should be expected, only very few (less than 20 in our sample) declare that they have applied for a patent in 1998-2000, and a small minority (about 80) that they hold at least one patent in 2000. We have chosen to ignore this information for these few firms, since we could not treat them in our model as the innovator firms for which we have much more information.

Actually, in CIS 3 the innovating firms reporting R&D expenditures in 2000 are also asked whether they engaged in R&D continuously over the period 1998-2000 or only occasionally. While about 60 % of them answered that they did R&D continuously and 25 % occasionally, we preferred to restrict our attention here to the continuous R&D performers for two reasons. Ideally we would have like to have some measure of R&D stock. Since this is not possible without knowing past R&D expenditures for at least several years, we thought that the 2000 R&D flow number would be a much better proxy for the R&D stock for the continuous R&D performers only. The second reason has to do with the timing problem. Being an innovator in CIS 3 refers to the period 1998-2000 whereas R&D expenditures are known for the year 2000 only. By focusing on continuous R&D performers we avoid attributing innovation in 1998-2000 to firms that had no R&D activity prior to the year 2000. To simplify, from now on R&D will refer to continuous R&D.

Overall, CIS 3 provides five indicators of innovation *stricto sensu*. The first three are dichotomous (or propensity) indicators, respectively for process innovations, product innovations new to the firm (but not necessarily to the market), and product innovations new to the market. New products refer to the three year period 1998-2000. Products new to the firm are defined as substantially improved or entirely new. Products new to the market refer to a first appearance on the market and therefore to more fundamental innovations than products new to the firm but which can already exist on the market. The other two are quantitative (or intensity) indicators measuring for product innovators (of both types) the importance of their flow of innovations, in terms of the share of their total sales in 2000 accounted by the new products.

One problem that is often raised regarding these survey innovation indicators is their subjectivity. The definition of what is and what is not an innovation remains in the end up to the appreciation of the respondent. Therefore it is interesting to compare these indicators with the more objective indicators concerning patents. At least for patents an outside patent examiner decides on the suitability of granting a patent to an invention. CIS 3 asks three indicators on patents. Two are dichotomous on whether the firm has applied or not for at least a patent in the period 1998-2000, and whether or not the firm holds at least one patent in the year 2000. The third measures the importance of the stocks of patents again in terms of the share of sales in the firm total sales in 2000, which are protected by patents. It is thus analogous to the two product innovation intensity indicators, but corresponds to a stock measure rather than a flow measure.

3. Model

Innovation survey data are qualitative and some of them are censored. These features call for a proper econometric modeling using latent variables.

Our model consists of two groups of equations. First, firms decide on whether they perform R&D and, if so, by how much. Then, depending on the extent of their R&D and other factors, they achieve a certain innovation output.

R&D is thus modeled as a generalized tobit:

$$(1) \quad s_R^* = x_R^1 b_R^1 + u_R^1 \quad \text{and} \quad i_R^* = x_R^2 b_R^2 + u_R^2$$

with $s_R = 1$ and $i_R^* = i_R$ if $s_R^* > 0$, and $s_R = 0$ otherwise

where s_R is the observed dichotomous indicator for R&D doing and non R&D doing firms, i_R is the observed intensity of R&D (measured in log of R&D per employee) for R&D doing firms, s_R^* and i_R^* are the corresponding latent variables, x_R^i ($i=1,2$) are the explanatory variables in both equations and b_R^i their respective coefficients, and u_R^1 and u_R^2 follow a bivariate normal distribution with correlation coefficient r_R and standard errors 1 (for reasons of identification) and σ_R^2 respectively.

For each measure of innovation we have an innovation equation that depends on R&D. In the case of process innovation and patent application, the dependent variable is dichotomous. In that case we specify a probit model. For instance for process innovation,

$$(2) \quad s_{PC}^* = b_R i_R^* + x_{PC} b_{PC} + u_{PC}$$

with $s_{PC} = 1$ if $s_{PC}^* > 0$, and $s_{PC} = 0$ otherwise

where s_{PC} is the observed dichotomous variable doing a process innovation or not, s_{PC}^* is the corresponding latent variable, i_R^* is the latent R&D intensity variable (as predicted by equations (1) for all firms, that is for $s_R = 1$ and $s_R = 0$) and b_R is the corresponding coefficient of interest, x_{PC} are the other explanatory variables and b_{PC} their respective coefficients, and u_{PC} is a random variable with a standard normal distribution. A similar model holds for patent applications.³

If the innovation variable is filtered, as for the two types of product innovative sales and for patent-protected sales, the appropriate model is, as for R&D, a generalized tobit model, with a selection equation and an intensity equation, where the error terms of both equations are correlated and the explanatory variables are not necessarily the same. The intensity of innovation is measured respectively by the share of sales accounted for by new or substantially improved products subject to having introduced such a product on the market in the last three years; the share of sales accounted for by products new to the market, subject to having introduced such a product on the market in the last three years; and the share in sales of product protected by patents, subject to having had at least one valid patent in the last three years.

³ In implementing the estimation, we make a logit transformation on the share of innovative sales, so that the dependent variable would vary from minus infinity to plus infinity, i.e. over the same domain as the normally distributed random error term. To avoid infinite values, the sales shares corresponding to the various innovation indicators are truncated from below at 0.01 and from above at 0.99.

We explicitly treat R&D as endogenous and account for the selection of R&D performing firms. By doing this, we wish to achieve two things. On the one hand, determinants of innovation, like size, are decomposed into their direct effect on innovation and their indirect effect operating through R&D. On the other hand, common causal factors of R&D and innovation that are not included in our model either because of our lack of knowledge or a lack of appropriate data, do not bias our R&D coefficient estimates (to the extent that the explanatory variables in the R&D equations (1) are indeed exogenous).

In all equations we control for size, belonging to a French group, belonging to a foreign group, foreign exposure, and industry characteristics summarized by industry dummies. *Size* is measured by the number of employees (in logs). *Foreign groups* are distinguished from *French groups* by the location of their head office. A dummy indicates whether the *international market* is the main market. We have 10 *industry dummies* corresponding to one or more NACE (rev 1) 2-digit industries. Firms are assigned to these industries on the basis of their main activity.⁴

In the R&D selection equation we also control for demand pull, cost push, market share and diversification.⁵ The French CIS 3 questionnaire contains two questions regarding the importance of *demand pull* and *cost push* as reasons for innovating. These variables can take values of 0 to 3. We have decided to isolate strong cost push and demand pull perceptions by creating dummies that take the value one when cost push (resp. demand pull) take value 3. The *market share* is constructed as the market share each firm has in each of the 227 industry segments in the NAP (“Nomenclature des Activités et Produits”) 600 industrial classification, where each share is weighted by the relative importance of each market in the firm total sales. The *diversification* index is constructed as the inverse of the Herfindahl concentration index of firm sales in the different markets. The Herfindahl index takes a maximum value of 1 if sales are concentrated in one market and a minimum value of $1/n$, where n is the number of markets in which the firm is active, if sales are equally divided between all markets. The diversification index can thus be interpreted as the equivalent number of markets in which the firm would equally share its total sales. Both market share and diversification are introduced in logs. Those are basically the variables that we have for all firms (innovating or not) and that can to a large extent be considered as exogenous.

We have more explanatory variables available to explain the intensity of R&D, because R&D performers are by construction innovators. Thus we also include as explanatory variables in the R&D intensity equation whether R&D was done in *cooperation*, whether the firm was a beneficiary of *government support for innovation*, and whether it benefited from four sources of *information for innovation* (that can be considered as proxies of knowledge spillovers): the enterprise or the group, basic research institutions (universities and government labs), suppliers, and clients.

⁴ The ten industries we control for are non-metallic minerals (NACE14 and NACE 26), textiles (NACE 17-NACE 19), wood (NACE 20-NACE 22), chemicals (NACE 23 and NACE 24), plastics (NACE 25), basic metals (NACE 27 and NACE 28), machinery (NACE 29), electrical products (NACE 30-NACE 33), vehicles (NACE 34 and NACE 35), and industries not elsewhere classified (NACE 36).

⁵ These variables are defined and included in the R&D equations (1) as in Crépon-Duguet-Mairesse, 1998.

In order to better identify and estimate the R&D coefficients in the innovation equations, we chose to impose as many exclusion restrictions as we thought reasonable. In the innovation equations, we only include, besides R&D, our general control variables for industry, size, domestic and foreign groups, and foreign exposure. We thus consider that all other variables included in the R&D equations affect our innovation output indicators only indirectly through R&D.⁶

We estimate the model by asymptotic least squares (ALS), also known as minimum distance estimator (Gouriéroux, Monfort, Trognon, 1985). The idea is in a first step to estimate consistently the parameters of the reduced form, and then in a second step to estimate the parameters of the structural form by minimizing the distances between the estimated reduced form coefficients and their predictions from the structural form coefficients, weighted by the inverse of the variance-covariance matrix of the estimated reduced form coefficients. The relations between the reduced form and structural form coefficients are in our case simply obtained by substituting the expression of the latent variable for R&D intensity in equations (1) into the equations (2) for the propensity and intensity of our innovation output indicators. Since there is no feedback in our model from innovation output to R&D (the modeling of such a feedback would require sufficiently long panel data to model dynamic effects), the R&D equations are estimated consistently by a generalized tobit. The second step of the minimum distance estimator hence only concerns the coefficients of the innovation equation.⁷

4. Descriptive analysis

To have a first idea of the differences between the various measures of innovation, the characteristics of innovative and R&D firms and the interrelations between R&D and innovation, it is useful to look at some descriptive statistics. Table 1 presents the sample averages of the main dependent and explanatory variables on the total sample of 2253 firms, the sub-sample of 1399 innovative firms (firms that declare to have either introduced a new product or a new process or that were at some point during the 1998-2000 period trying to do so), and the sub-sample of 855 R&D performers. It is striking to notice that all variables (except of course the percentage of firms in low-tech sectors) have higher averages as we move from all firms to innovative firms and to R&D performers. A larger proportion of innovative firms (however they

⁶ Crépon-Duguet-Mairesse, 1998, have also excluded the market share and diversification variables but included the demand pull and cost push variables.

⁷ For R&D in the high-tech sectors, the likelihood function of the generalized tobit has a local maximum for a positive value of the correlation coefficient between the error terms of the two parts of the tobit model and a global maximum for a negative value, but the two maxima are not far apart. In that case, we have preferred to impose a priori the positive correlation coefficient in our estimation procedure, on the presumption that left out determinants affect the probability to do R&D and its intensity in the same direction. We had the same problem in the case of the generalized tobit reduced form equations for the new to the market product innovation indicators in the low-tech sectors. In the case of generalized tobit reduced form equations for the patent holdings indicators in both the high- and low-tech sectors, we also encountered a somewhat similar problem with a correlation coefficient equal to one at the maximum maximum of the likelihood function. In all these cases the next maximum corresponded to a very close value of the likelihood function and to a positive correlation coefficient, and we just chose to impose that correlation coefficient in the estimation.

are measured) is thus found among R&D performers than among other firms. In addition, among the product innovators and the patent holders, the share of total sales accounted for by new products or covered by patents is higher for R&D performers than for non R&D performers. The share of total sales covered by patents is also higher for innovators than for non innovators. Moreover, R&D performers are larger, more diversified, more internationally oriented, and so on than non R&D performers.

The eight innovation indicators that we consider behave differently. There are more product innovators of all sorts (i.e. true innovators and imitators combined) than process innovators: 80.7% of all innovating firms are new to the firm products innovators whereas only 55.8% of them are process innovators. Among the R&D performers, the relative proportion of product innovators is even slightly higher. True innovators, i.e. those that introduce products new to the market, are less frequent of course (the average drops from 80.7% to 45.5%), but their relative number increases if we move to R&D performers. These simple descriptive statistics suggest that R&D is more geared towards product innovations than process innovations, and is positively correlated to the introduction of products new to the market. The proportion of firms with new patent applications is similar to the proportion of firms with valid patents, which seems to indicate that over a three-year period there is persistence in patenting. What is also very interesting to notice is that the proportion of firms with new to the market product innovations is almost as high as the proportion of firms with patents. This does not imply of course that all new products are patented. Still, 72% of the firms that introduced new to the market products also applied for a patent (figure not shown in table 1).

5. Comparison of the marginal impact of R&D on various innovation measures

In this section we present the results of our econometric analysis in terms of the marginal effects of R&D intensity on the various innovation indicators. Based on previous experience (Mairesse and Mohnen, 2001) and on the large difference in the proportions of R&D performers in high-tech and low-tech sectors (see table 1), we have performed a separate analysis for the two corresponding clusters or sub-samples of firms, that is for the 899 firms in the electrical products, chemicals, machinery and vehicles industries, and for the 1354 firms in the textiles, wood, plastics, non-metallic mineral products, basic metals and not elsewhere classified industries.⁸

As appendix table 1 shows, the marginal effects of the explanatory variables in the R&D equations are quite different. The propensity to engage in continuous R&D increases significantly with size, foreign exposure, demand pull and cost push, more so in high-tech than in low-tech sectors. In the high-tech sectors it also increases with diversification and in the low-tech sectors with market share. The difference between high-tech and low-tech industries is even

⁸ The marginal effects for the two sub-samples are simply obtained as the sub-samples averages of the marginal effects computed for each firm. The marginal effects on the five indicators of innovation propensity are themselves equal for a given firm to the corresponding estimated coefficient multiplied by the predicted probability to innovate for this firm. Similarly, the marginal effects on the three indicators of innovation intensity are equal for a given firm to the corresponding estimated coefficient multiplied by $\exp(z)/(1+\exp(z))^2$, where z is the predicted logit transformation of innovation intensity for the firm (see Wooldridge, 2003).

more visible if we look at the determinants of R&D intensity. In the high-tech sectors, size, demand pull, market share and cooperation play no significant role. By contrast, in low-tech industries a hundred percent increase in size decreases R&D intensity by 31.2 percent, and demand pull, market share or cooperation are significantly positive and of a similar order of magnitude. R&D intensity is also affected by different information channels in the two sub-samples: information from basic research institutions increases R&D intensity by 29 percentage points in high-tech sectors, information from clients increases R&D intensity by 25.5 percentage points in low-tech sectors. All other information sources are not significant. The fact that the demand pull, cost push, diversification and market share effects, which are excluded from the innovation equations, are in most cases significant indicates that these variables should be good instruments to identify the R&D parameter in the innovation equation.

Table 2 compares the marginal effects of changes in R&D/employee and the number of employees on the different measures of innovation output. Size, if significant, has a small impact on the probability to innovate. In high-tech sectors, a one percent increase in size increases by 4.5 percentage points the probability of introducing a new process, by 6.6 percentage points the propensity to apply for a patent and by 4.8 percentage points the propensity to hold a valid patent; it has only a minor non-significant effect on new product introductions. In low-tech sectors, it increases the probability to be a product or a process innovator by 3 to 4 percentage points and patent applications (resp. patent holdings) by 9.4 (resp. 7.2) percentage points. Size has, in general, no significant impact on the share of innovative sales.

The effect of R&D intensity on innovation is stronger than size and more pronounced in the high-tech sectors. In the high-tech sectors a one percent increase in R&D/employee increases the probability to innovate by around 20 percentage points for all five dichotomous indicators, somewhat more for new to firm product innovation. Conditional on being innovative, a one percent increase in R&D intensity increases the share in sales of products new to the firm by 3.7 percentage points, the share in sales of products new to the market by about twice as much, and the share of patented products by 13 percentage points. The effect of an increase in R&D intensity is even more pronounced in the low-tech sectors. A one percent increase in R&D/employee increases the propensity to innovate in new to the firm products by about the same amount as in the high-tech sectors but it increases the propensity to introduce a product new to the market, a new process, as well as the probability to hold a patent by twice the amount in the high-tech sectors. In the low-tech sectors R&D increases the share in total sales of new to the firm products more than the share in sales of new to the market products (7.9 versus 2.2 percentage points), and it increases even more than in high-tech sectors the share of patented products (18.5 compared to 13 percentage points).

6. Inter-industry comparisons of innovativeness

In this section we compare innovation across industries, within the high-tech and low-tech sectors, by decomposing it into the contribution from R&D, size together with other explanatory factors (being part of a domestic or foreign group and foreign exposure), and “innovativeness”. The framework of this decomposition is laid out in Mairesse and Mohnen (2002). The idea is borrowed from growth accounting. Just as differences in output between two periods (years,

decades) or between two spatial units (firms, industries, countries) can be ascribed to differences in the inputs and to a residual that has been named total or multi-factor productivity (TFP or MFP), or simply productivity, likewise, differences in innovation output (in terms of new products, processes or patents) between two periods or spatial units can be ascribed to differences in the factors of innovation (notably R&D) and to a residual that we call “innovativeness”, or the unexplained ability to turn innovation inputs into innovation outputs.

Within the high-tech or low-tech cluster separately, we compare the inter-industry innovative performance along three measures (new to the firm product innovations, new to the market product innovations and patents) and two criteria (the probability to innovate and the intensity of innovation). As a general notation, let us write the innovation function as

$$(3) \quad y = f(R, Z)$$

where y can thus be measured in three ways and f can represent two types of functions (the propensity and the intensity of innovation). R represents R&D intensity and Z the other determinants that have been introduced as explanatory variables. Suppose we have two industries. Let y^A , y^B , and y^C represent the expected shares of innovative sales at respectively the mean values of the explanatory variables for industry A, industry B and the average industry within a cluster. By a linear approximation of each industry’s average expected share of innovative sales around the cluster average we have

$$(4) \quad \begin{aligned} y^A &= y^C + f_R^C (R^A - R^C) + f_Z^C (Z^A - Z^C) + e^A \\ y^B &= y^C + f_R^C (R^B - R^C) + f_Z^C (Z^B - Z^C) + e^B \end{aligned}$$

where f_x^E represents the gradient of $f(\cdot)$ with respect to x ($x=R, Z$) evaluated at the cluster average values for x , and x^C ($x=R, Z$) represents the cluster average of variable x .⁹ The first term of the approximation is the cluster average, the second term is the R&D effect, the third term is the effect of variable Z , and the last term, which includes the industry effect and the first-order approximation error, captures innovativeness.¹⁰ Notice that $e^E = (e^A + e^B) / 2 = 0$.

The results of the innovation accounting are presented numerically in appendix tables 2 (for the high-tech cluster) and 3 (for the low-tech cluster) and graphically in figures 1 to 4. Appendix tables 2 and 3 are organized as follows. In the first column we have the cluster average of an innovation measure, in column 2 the difference between an industry innovation and the cluster average attributable to size (group and foreign exposure) (the third term of equation (4)), in column 3 the difference attributable to R&D (the second term of equation (4)), in column 4 the sum of columns 2 and 3 (i.e. the explained difference in innovation between a given industry and the cluster average), in column 5 the expected innovation intensity (i.e. the cluster average of column 1 plus the explained effect of column 4), and in column 6 the difference between the industry innovation average of column 7 and the expected innovation intensity of column 5 (in

⁹ We give each industry equal weight.

¹⁰ If the $f(\cdot)$ function was linear, there would be no approximation error.

other words, the difference in measured innovation that we cannot at this stage attribute to anything else but innovativeness). Since the comparison is always with respect to the cluster average, the average differences which are reported in the last row add up to zero and the average expected industry measure of innovation is by construction equal to the observed cluster average.

In figures 1 to 4 we concentrate on new to the firm product innovations and patent holdings, the most subjective and the most objective of the three measures for which we have intensity of innovation measures¹¹. In figures 1 to 4, size and R&D are regrouped. We want to contrast the explained and unexplained portions of the interindustry innovation differences along the two innovation measures.

As the appendix table 2 show, R&D explains most of the inter-industry difference in the probability to innovate in the high-tech cluster. For instance, if we take the two industries with the largest and the smallest proportion of new to the firm product innovation, chemicals and vehicles, 11.9 out of a total of 13.9 percentage points difference is attributable to the R&D effect. R&D plays a less dominant role in the explanation of the intensity of innovation and in the low-tech sectors, but it still outweighs the contribution of the other explanatory variables. It is also worth noticing that enterprises are more likely to innovate than to patent, but the share in total sales due to new products is lower than the share in total sales protected by patents. For instance, in the high-tech cluster 65.6 percent of enterprises innovate in products new to the firm but only 45.2 percent of them hold patents. However, whereas only 16.5 percent of total sales represent new products, 28.6 percent of total sales are patent-protected. The same pattern holds for the low-tech cluster.

Given that these magnitudes differ between indicators, it is not appropriate to compare the size of the individual contributions between the product and the patent decompositions in figures 1 to 4. However, it is interesting to compare the directions of the explained and unexplained contributions and their relative magnitude for each innovation measure. With one exception (the not elsewhere classified industries in the low-tech cluster) the explained portions of the inter-industry differences in product innovation and patenting go in the same direction. With a few exceptions, the same can be said for innovativeness. If an industry is more innovative than the cluster average in terms of introducing new products, it is also more innovative in terms of patenting. The objective and subjective measures of innovation do not contradict each other. Apart from the propensity to innovate in high-tech sectors, in most other cases innovativeness is higher than the explained part of innovation. What is even more astonishing is that in some cases the unexplained part of innovation gets even larger after accounting for R&D. For instance in appendix table 2, we observe that enterprises in machinery have a higher share of new (to the firm or to the market) innovative sales than chemical firms and yet they do less R&D. The greatest part of the innovation remains to be explained. This conclusion is not very surprising, given the limited number of explanatory variables that we could introduce in this study.

¹¹The numerical appendix tables also present the decomposition for new to the market product innovations, which is closer to the decomposition for new to the firm products than for patents.

7. Conclusion

We have systematically confronted all the indicators of innovation output that are provided by the French CIS 3: the five dichotomous innovation indicators for the incidence of process innovation, product innovations new to the firm, product innovations new to the market, patent applications and patent holdings and the three censored continuous indicators measuring the shares in total sales of sales accounted for by products new to the firm or new to the market, and that of patent-protected sales.

R&D is positively correlated with all measures of innovation output, and, all other things equal, more correlated than size to innovation. Innovation is generally more sensitive to R&D in the low-tech sectors than in the high-tech sectors.

The innovation indicators may differ in terms of their meaning and objective quality. The notion of product or process innovation defined in the innovation surveys is often criticized for being rather subjective, up to the respondent's appreciation, as opposed for example to a patent that has been granted by a patent officer. We nonetheless find that both measures tell a consistent story when it comes to comparing the innovation performance across industries and attribute them to explanatory factors (like size and R&D) and the unexplained innovativeness. We thus tend to downplay the skepticism on the value and meaning of subjective survey data, as expressed for example in the quote by Zvi Griliches (1965) cited by Mike Scherer (2005) in his introductory paper to this issue, and to support the faith that Ed Mansfield showed in much of his work relying on survey data— or in Scherer's acute wording “asking the people who know”.¹²

¹² The quote to Griliches (1965) is the following: “Mansfield is reporting on a very large body of work, and he should be congratulated for working, almost alone, in this very important area and for producing so many interesting and significant results. All I can do in this context is to reiterate Mansfield's caveat that some of his conclusions rest on quite shaky data....To illustrate some of the difficulties of empirical research in this area, consider Mansfield's first model on R&D expenditures by firm. The important variables in this model...were derived from answers to a questionnaire or interview. They are not “data” in the usual sense...[What firms say] does not prove that profitability affects investment, it tells only that when asked to explain why they invest in research, these firms can provide a sensible answer”.

References

- Acs, Z., D. Audretsch and M. Feldman (1992), "Real effects of academic research: comment", *American Economic Review*, 82(1), 363-367.
- Crépon, B., E. Duguet and J. Mairesse (1998), "Research and Development, Innovation and Productivity: An Econometric Analysis at the Firm Level", *Economics of Innovation and New Technology*, 7(2), 115-158.
- Gouriéroux, C., A. Monfort and A. Trognon (1985), "Moindres carrés asymptotiques", *Annales de l'INSEE*, 58, 91-122.
- Griliches, Z. (1965), "Comment on Mansfield" in R. Tybout ed. *Economics of Research and Development*, Ohio State University Press, Columbus, 148-150.
- Griliches, Z. (1981), "Market value, R&D and patents", *Economics Letters*, 7, 183-187.
- Hall, B. and R. Oriani (2004), "Does the market value R&D investment by European firms? Evidence from a panel of manufacturing firms in France, Germany, and Italy", NBER working paper 10408.
- Jaffe, A. (1989), "Real effects of academic research", *American Economic Review*, 79(5), 957-970.
- Mairesse, J. and P. Mohnen (2001), "To be or not to be innovative: An exercise in measurement", *STI Review. Special Issue on New Science and Technology Indicators*, OECD, 27, 103-129.
- Mairesse, J. and P. Mohnen (2002), "Accounting for innovation and measuring innovativeness: An illustrative framework and an application", *American Economic Review, Papers and Proceedings*, 92(2), 226-230.
- Mansfield, E. (1965-a), "Rates of return from industrial research and development", *American Economic Review*, 55, 310-322.
- Mansfield, E. (1965-b), "The process of technical change" in R. Tybout ed. "Economics of Research and Development", Ohio State University Press, Columbus, 136-147.
- Mansfield, E. (1980), "Basic research and productivity increase in manufacturing", *American Economic Review*, 70, 863-873.
- Mansfield, E., J. Rapoport, A. Romeo, S. Wagner and G. Beardsley (1977), "Social and private rates of return from industrial innovations", *Quarterly Journal of Economics*, 77, 221-240.
- Organization for Economic Co-operation and Development (1992, 1996), *Oslo Manual*, Paris, 1st, 2nd edition.

Pavitt, K. (1984), "Sectoral patterns of technical change: towards a taxonomy and a theory", *Research Policy*, 13, 343-373.

Scherer F. M. (2005), "Edwin Mansfield: an appreciation", *The Journal of Technology Transfer*, to appear in a special issue in memory of Edwin Mansfield, edited by A. Link and F.M. Scherer, 30 (1-2), January 2005.

Wooldridge, J. (2003), *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.

**Table 1: Summary statistics: France, CIS 3, 1998-2000, Manufacturing
Samples of All firms (TOTAL), Innovative firms (INNO) and continuously R&D
performing firms (R&D)**

Variable	TOTAL Sample	INNO Sub-sample	R&D Sub-sample
Number of firms	2253	1399	855
% firms performing R&D continuously	37.9	61.1	100.0
R&D over sales (in %)	n.r.	n.r.	3.9
Log (R&D/Emp) (in 10³ euros per person)	n.r.	n.r.	1.0
% firms innovating with products new to the firm* (A)	50.1	80.7	89.8
% firms innovating with products new to the market* (B)	28.3	45.5	58.0
% firms with process innovation	34.7	55.8	57.5
% firms with at least one patent applied for in 1998-2000	32.4	48.2	63.5
% firms with at least one valid patent at the end of 2000* (C)	34.6	49.6	64.7
Share of new to the firm innovative sales for firms of type (A)	n.r.	12.9	14.5
Share of new to the market innovative sales for firms of type (B)	n.r.	5.9	7.7
Share of total sales covered by patents for firms of type (C)	9.7	13.1	16.7

n.r.: not available or non relevant. * Indicator corrected to 0 if corresponding share =0.
All means are unweighted.

Table 1 (cont'd): Summary statistics: France, CIS 3, 1998-2000, Manufacturing Samples of All firms (TOTAL), Innovative firms (INNO) and continuously R&D performing firms (R&D)

Variable	TOTAL Sample	INNO Sub-sample	R&D Sub-sample
% of firms ...			
Number of employees: Mean/Median	410/133	541/207	727/340
% of firms in high-tech industries	39.9	48.4	57.7
% of firms in low-tech industries	60.1	51.6	42.3
% of firms for which the most significant market is international	41.8	53.0	64.1
% of firms belonging to a French group	46.7	48.4	51.2
% of firms belonging to a foreign group	27.1	33.8	36.3
% of firms with strong demand pull	37.0	59.6	67.6
% of firms with strong cost push	12.8	20.7	24.3
Diversification index	1.17	1.19	1.24
Average market share (in %)	2.6	3.4	4.3
% of firms with government support for innovation	n.r.	32.4	37.0
% of firms collaborating in innovation	n.r.	45.5	56.4
% of firms with significant source of information for innovation from:			
Within the firm or group	n.r.	74.1	84.1
Basic research institutions	n.r.	50.3	62.5
Suppliers	n.r.	47.9	49.7
Clients	n.r.	71.8	79.8

n.r.: not available or non relevant. All means are unweighted.

Table 2: Average marginal effects of R&D intensity and size, by innovation indicator, in high-tech and low-tech industries

Innovation Variables	Products new to the firm		Products new to the market		Process innovation	Patent applications	Patent holdings	
	In %	Yes/no	Share in sales	yes/no	Share in sales	yes/no	yes/no	Share in sales
High-tech industries								
R&D intensity	26.5	3.7	21.6	7.6	19.5	19.9	20.4	13.0
	(2.0)	(1.4)	(2.2)	(1.2)	(2.5)	(2.3)	(2.1)	(2.9)
Size	0.5	-0.3	0.7	0.1	4.5	6.6	4.8	4.9
	(0.8)	(0.5)	(0.8)	(0.4)	(1.0)	(1.0)	(0.8)	(1.1)
Number of firms	899	593	899	372	899	899	899	411
Low-tech industries								
R&D intensity	27.7	7.9	45.5	2.2	34.3	28.5	39.8	18.5
	(1.5)	(1.3)	(3.3)	(0.9)	(2.3)	(2.7)	(3.0)	(2.1)
Size	3.1	0.1	3.8	0.1	3.5	9.4	7.2	1.4
	(0.7)	(0.5)	(1.5)	(0.3)	(1.2)	(1.4)	(1.5)	(0.9)
Number of firms	1354	536	1354	265	1354	1354	1354	283

All regressions include binary indicators for industry effects (4 in high-tech industries, 6 in low-tech industries), for the affiliation to a domestic or a foreign group, and for the importance of international market. R&D intensity is measured by R&D expenditures per employee (in logs), and size by the number of employees (in logs). All marginal effects are in percentage points: for example a 10% increase in R&D intensity has an average marginal effect of 2.6% on the propensity to innovate in new to the firm-products and of 0.4% on the corresponding intensity to innovate. Standard errors are in parentheses.

Figure 1

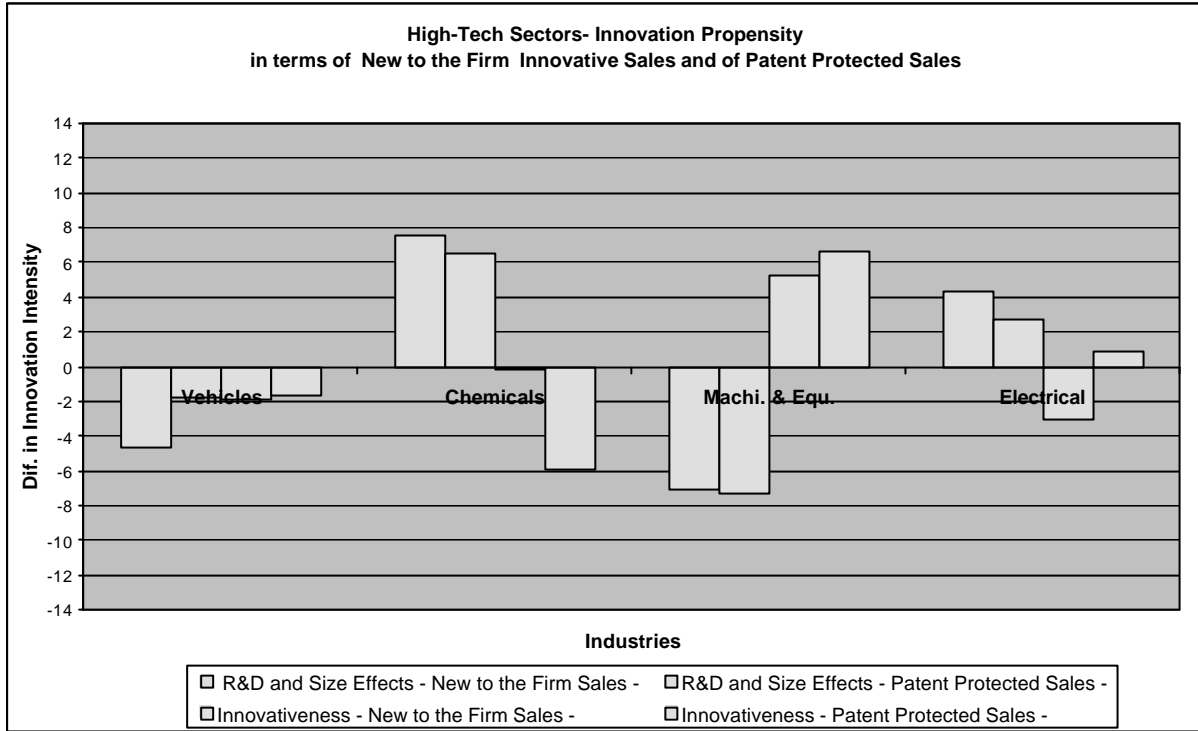


Figure 2

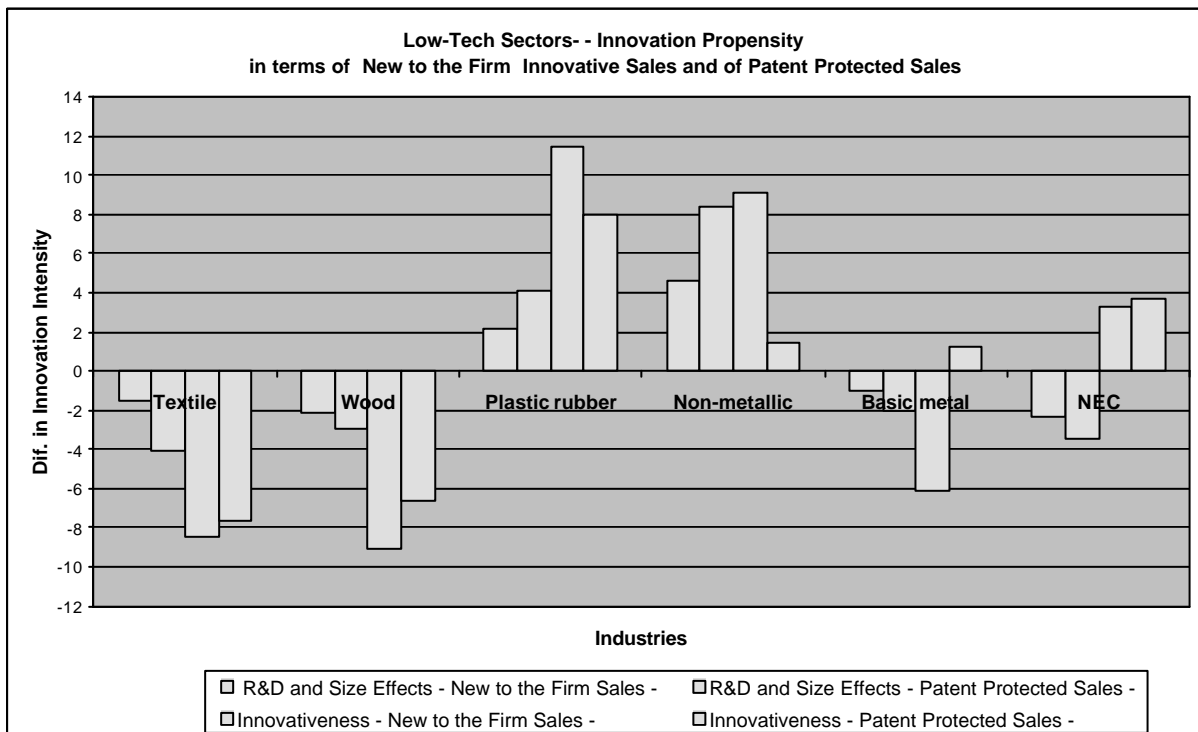


Figure 3

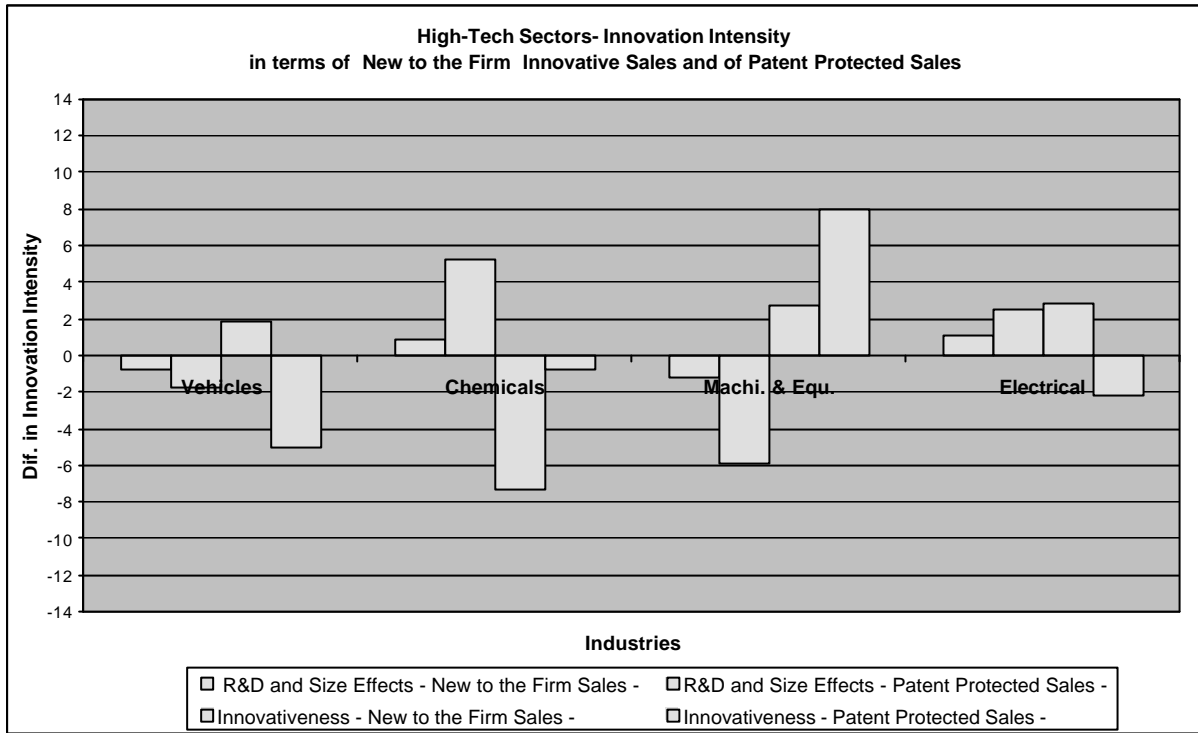
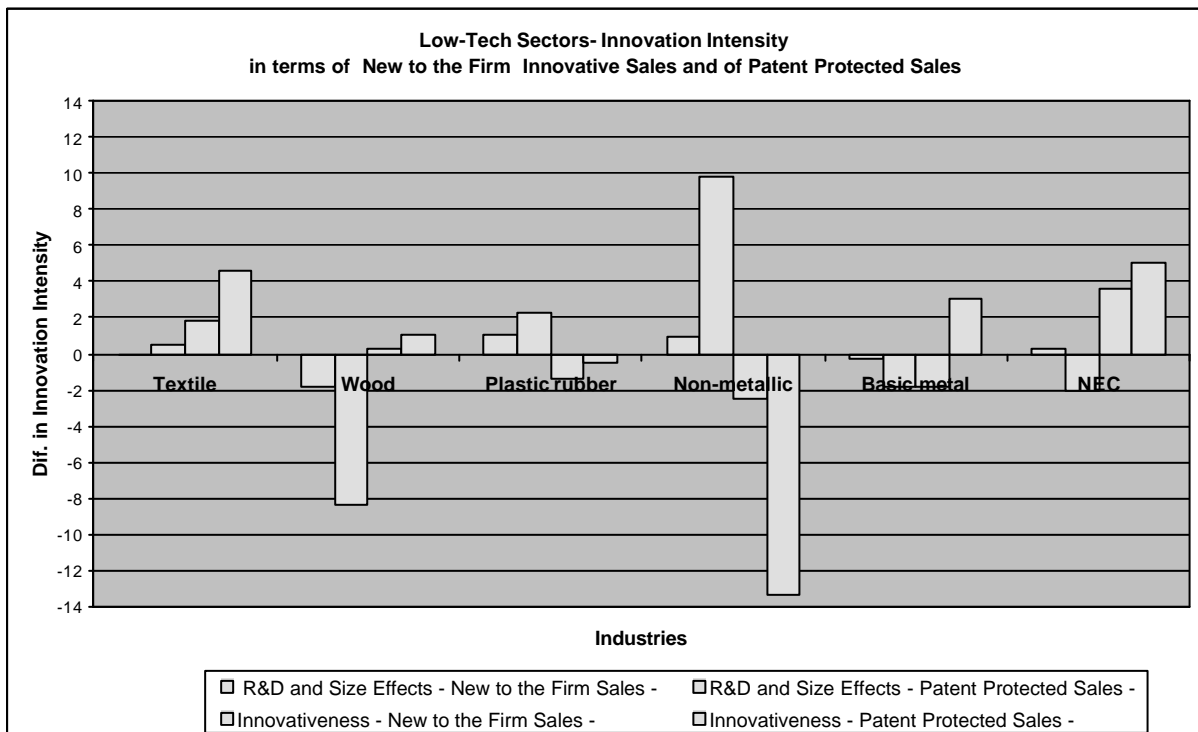


Figure 4



Appendix Table 1: Marginal effects of the determinants in the (generalized tobit) R&D equations – correcting for the selection of continuous R&D performing firms¹

Dependent variable	High-tech industries		Low-tech industries	
	Doing continuous R&D	Log of R&D/employee	Doing continuous R&D	Log of R&D/employee
Log of number of employees	6.8 (1.6)	0.7 (6.1)	4.3 (1.3)	-31.2 (10.8)
National group	4.5 (3.8)	-10.4 (18.4)	2.7 (2.5)	9.4 (20.6)
Foreign group	-0.6 (4.2)	-14.6 (20.1)	-1.7 (3.1)	15.7 (24.7)
International market	14.1 (2.8)	37.8 (12.5)	8.0 (2.0)	34.5 (13.4)
Demand pull	28.7 (2.7)	12.6 (12.9)	25.6 (2.0)	35.7 (16.6)
Cost push	23.1 (4.2)	60.9 (12.2)	16.7 (3.2)	43.4 (18.3)
Diversification	11.4 (5.5)	-16.0 (17.1)	6.0 (4.0)	7.1 (22.9)
Market share	0.9 (1.2)	7.6 (5.6)	4.0 (1.0)	29.6 (8.5)
Cooperation	-	14.7 (11.8)	-	33.2 (12.8)
Government support for innov.	-	8.3 (13.6)	-	-2.0 (12.8)
Information from within enter/group	-	17.5 (19.8)	-	18.1 (15.7)
Information from basic research inst.	-	29.0 (12.9)	-	11.1 (14.2)
Information from suppliers	-	-15.4 (10.8)	-	0.2 (12.3)
Information from clients	-	22.2 (18.5)	-	25.5 (12.5)
SE of regression	1	1.44 (0.13)	1	1.26 (0.08)
r	0.09*		0.55 (0.10)	
Number of firms	899	494	1354	361

All regressions include industry indicators. Standard errors are in parentheses. * imposed.

¹ All marginal effects are given in %.

Appendix Table 2: Accounting for inter-industry differences of innovation propensity and intensity in high-tech industries

Industry	Average innovation	Size related effects	R&D effects	Total Effects	Expected innovation	Innovation	Observed innovation
Innovation Propensity (in %) in terms of new to the firm innovative sales							
Vehicles	65.6	-0.1	-4.7	-4.7	60.9	-1.9	59.0
Chemicals	65.6	0.3	7.2	7.5	73.1	-0.2	72.9
Machinery	65.6	0.1	-7.2	-7.1	58.5	5.2	63.7
Electrical	65.6	-0.3	4.6	4.3	69.9	-3.0	66.9
Average industry	65.6	0.0	0.0	0.0	65.6	0.0	65.6
Innovation Intensity (in %) in terms of new to the firm innovative sales							
Vehicles	16.5	-0.2	-0.7	-0.8	15.7	1.8	17.5
Chemicals	16.5	0.1	0.8	0.9	17.4	-7.3	10.1
Machinery	16.5	0.0	-1.2	-1.2	15.3	2.7	18.0
Electrical	16.5	0.1	1.1	1.1	17.7	2.8	20.5
Average industry	16.5	0.0	0.0	0.0	16.5	0.0	16.5
Innovation Propensity (in %) in terms of new to the market innovative sales							
Vehicles	40.7	0.2	-3.8	-3.6	37.1	-2.6	34.5
Chemicals	40.7	0.2	5.9	6.1	46.8	-3.1	43.8
Machinery	40.7	-0.2	-5.9	-6.1	34.6	6.3	40.9
Electrical	40.7	-0.2	3.8	3.5	44.2	-0.7	43.6
Average industry	40.7	0.0	0.0	0.0	40.7	0.0	40.7
Innovation Intensity (in %) in terms of new to the market innovative sales							
Vehicles	14.5	0.1	-0.8	-0.8	13.8	2.9	16.7
Chemicals	14.5	0.0	1.1	1.1	15.6	-7.2	8.5
Machinery	14.5	0.0	-2.5	-2.6	12.0	4.4	16.4
Electrical	14.5	-0.1	2.3	2.2	16.8	-0.2	16.6
Average industry	14.5	0.0	0.0	0.0	14.5	0.0	14.5
Innovation Propensity (in %) in terms of patent protected sales							
Vehicles	45.2	1.8	-3.6	-1.8	43.4	-1.7	41.7
Chemicals	45.2	0.9	5.6	6.5	51.7	-5.9	45.8
Machinery	45.2	-1.8	-5.6	-7.3	37.9	6.6	44.5
Electrical	45.2	-0.9	3.5	2.7	47.9	0.9	48.8
Average industry	45.2	0.0	0.0	0.0	45.2	0.0	45.2
Innovation Intensity (in %) in terms of patent protected sales							
Vehicles	28.6	1.7	-3.4	-1.8	26.8	-5.0	21.8
Chemicals	28.6	1.2	4.0	5.2	33.7	-0.8	33.0
Machinery	28.6	-1.7	-4.2	-5.9	22.7	8.0	30.6
Electrical	28.6	-1.1	3.6	2.5	31.1	-2.2	28.9
Average industry	28.6	0.0	0.0	0.0	28.6	0.0	28.6

Note: Small discrepancies are due to rounding errors

Appendix Table 3: Accounting for inter-industry differences of innovation propensity and intensity in low-tech industries

Industry	Average innovation	Size related effects	R&D effects	Total Effects	Expected innovation	Innovation	Observed innovation
Innovation Propensity (in %) in terms of new to the firm innovative sales							
Textile	42.6	0.0	-1.4	-1.5	41.1	-8.5	32.6
Wood	42.6	1.0	-3.1	-2.1	40.4	-9.1	31.3
Plastic	42.6	-1.0	3.2	2.2	44.7	11.5	56.2
Non-metal	42.6	0.2	4.5	4.6	47.2	9.1	56.3
Basic metal	42.6	0.1	-1.0	-1.0	41.6	-6.1	35.5
Nec	42.6	-0.1	-2.1	-2.3	40.3	3.3	43.5
Average industry	42.6	0.0	0.0	0.0	42.6	0.0	42.6
Innovation Intensity (in %) in terms of new to the firm innovative sales							
Textile	13.1	0.2	-0.3	-0.1	13.0	1.8	14.7
Wood	13.1	0.0	-1.8	-1.8	11.3	0.3	11.6
Plastic	13.1	-0.2	1.3	1.1	14.2	-1.4	12.8
Non-metal	13.1	-0.3	1.2	0.9	14.0	-2.5	11.5
Basic metal	13.1	0.0	-0.4	-0.3	12.7	-1.8	10.9
Nec	13.1	0.2	0.1	0.3	13.4	3.6	17.0
Average industry	13.1	0.0	0.0	0.0	13.1	0.0	13.1
Innovation Propensity (in %) in terms of new to the market innovative sales							
Textile	20.4	0.2	-2.3	-2.1	18.3	-3.0	15.3
Wood	20.4	1.2	-5.1	-3.9	16.5	-3.9	12.6
Plastic	20.4	-1.4	5.3	3.8	24.3	4.1	28.4
Non-metal	20.4	-0.1	7.4	7.3	27.7	1.7	29.5
Basic metal	20.4	0.1	-1.7	-1.6	18.9	1.8	20.6
Nec	20.4	0.0	-3.5	-3.6	16.9	-0.7	16.1
Average industry	20.4	0.0	0.0	0.0	20.4	0.0	20.4
Innovation Intensity (in %) in terms of new to the market innovative sales							
Textile	10.8	-0.1	0.1	0.1	10.6	4.6	15.5
Wood	10.8	0.0	-1.2	-1.2	9.6	-2.1	7.5
Plastic	10.8	0.0	0.5	0.5	11.3	-1.0	10.3
Non-metal	10.8	0.1	0.6	0.7	11.5	-2.7	8.8
Basic metal	10.8	0.0	-0.1	-0.1	10.8	-1.6	9.2
Nec	10.8	0.0	0.1	0.1	10.9	2.8	13.7
Average industry	10.8	0.0	0.0	0.0	10.8	0.0	10.8

Note: Small discrepancies are due to rounding errors

Appendix Table 3 (cont'd): Accounting for inter-industry differences of innovation propensity and intensity in low-tech industries

Industry	Average innovation	Size related effects	R&D effects	Total Effects	Expected innovation	Innovation	Observed innovation
Innovation Propensity (in %) in terms of patent protected sales							
Textile	22.4	-2.1	-2.0	-4.1	18.3	-7.6	10.7
Wood	22.4	1.6	-4.5	-2.9	19.5	-6.6	12.9
Plastic	22.4	-0.5	4.6	4.1	26.5	8.0	34.5
Non-metal	22.4	1.9	6.5	8.4	30.8	1.4	32.1
Basic metal	22.4	-0.5	-1.5	-2.0	20.4	1.2	21.6
Nec	22.4	-0.4	-3.1	-3.5	18.9	3.7	22.6
Average industry	22.4	0.0	0.0	0.0	22.4	0.0	22.4
Innovation Intensity (in %) in terms of patent protected sales							
Textile	19.3	-1.0	1.6	0.5	19.8	4.6	24.5
Wood	19.3	0.4	-9.1	-8.4	10.7	1.1	11.8
Plastic	19.3	0.2	2.0	2.3	21.6	-0.5	21.1
Non-metal	19.3	0.7	9.1	9.8	29.1	-13.4	15.8
Basic metal	19.3	0.1	-1.9	-1.8	17.5	3.1	20.6
Nec	19.3	-0.4	-1.7	-2.1	17.2	5.0	22.2
Average industry	19.3	0.0	0.0	0.0	19.3	0.0	19.3

Note: Small discrepancies are due to rounding errors