Laboratory of Economics and Management
Sant'Anna School of Advanced Studies

Piazza Martiri della Libertà, 33 - 56127 PISA (Italy)
Tel. +39-050-883-343 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://www.lem.sssup.it/

LEM
Working Paper Series

"Stacking" or "Picking" Patents? The Inventors' Choice Between Quality and Quantity

Myriam MARIANI*
Marzia ROMANELLI**

* Bocconi University, Milan, Italy
** University of Pisa, Italy

2006/06 February 2006

“STACKING” OR “PICKING” PATENTS?

THE INVENTORS’ CHOICE BETWEEN QUANTITY AND QUALITY

Myriam Mariani
Bocconi University
IEP and *Cespri*, Milano, Italy
myriam.mariani@uni-bocconi.it

Marzia Romanelli
University of Pisa, Italy
romanelli@sssup.it

February, 2006

Abstract

This paper studies the determinants of the *quantity* and *quality* of inventors’ patents. It uses a sample of 793 inventors drawn from the PatVal-EU dataset and the information on EPO patents that they contributed to inventing during the period 1988-1998. It explores three aspects of the inventors’ productivity: 1) the number of EPO patents that they produce; 2) their average quality; 3) the quality of the most valuable patents. By jointly estimating the three equations we find that the inventors’ level of education, employment in a large firm and involvement in large-scale research projects positively affect *quantity*. Yet, apart from the size of the research project, none of these factors directly influences the expected *quality* of the innovations. They do, however, indirectly, as we find that the number of innovations explains the probability of producing a technological hit (the maximum value). Also, there are no decreasing returns in the innovation process at an individual level, as the number of innovations that an inventor produces is not correlated with their average quality.

Acknowledgements. This paper is based on the data collected through the PatVal-EU project (Contract HPV2-CT-2001-00013) funded by the European Commission. Economic support was also provided by the Italian Ministry of Education (Miur, 2003-2005), by the Italian CNR (Promozione Ricerca 2004), by Bocconi University (RdB project, II tranche, 2005) and by the Econchange Project of the European Commission. We thank Bart Verspagen and Karin Hoisl for providing us with the patent citations, Grid Thoma for helping us with the EPO search, and the ZEW Research Institute for giving us access to the *SearchEngine* matching software. We are grateful to Manuel Trajtenberg for his helpful and encouraging comments, and to all participants in the ZEW Conference, Mannheim, 2005. The usual disclaimers apply.

1. Introduction

Innovation and human capital are key factors for the growth of firms and for economic growth more generally. Yet, little is known about the key actors in this process - the industrial inventors - and the determinants of their productivity.

Traditional contributions focus on scientists and use scientific publications as a measure of their research output (for an overview, see Stephan, 1996). They show that the distribution of the scientists' productivity is skewed (Lotka, 1926; de Solla Price, 1963; Allison and Stewart, 1974; Turner and Mairesse, 2006), and that age and vintage matter, with scientists becoming less productive as they get older (Oster and Hamermesh, 1998; Levin and Stephan, 1991). This holds after controlling for individual fixed effects that proxy for differences in motivation and ability. Our knowledge about industrial inventors is sparser. The difficulty in obtaining information about individual inventors has prevented previous research from performing systematic empirical studies on this matter. The existing evidence is based on small samples, specific industries and firms (e.g. Narin and Breitzman, 1995; Ernst et al., 2000).

By relying on the novel and detailed information from a large sample of European inventors (PatVal-EU, 2005), our paper explores the determinants of the productivity of industrial inventors in terms of *quantity* and *quality* of the innovations that they produce. In fact, inventor productivity may take various forms. While the number of patents that they develop is one form, the inventors often acquire visibility for the "value" of their innovations, and sometimes their reputation depends on one or several, major achievements (Jones, 2005). This calls for an indicator of the quality or importance of the innovations. We start by using the number of citations that the patents have received within five years of their publication date (i.e. forward citations). Alternatively, by combining different patent indicators, we extract a composite index – i.e. a common component – that proxies for the technological and economic importance of the innovations as in Lanjouw and

Schankerman (2004). We then study the determinants of the inventors' productivity measured as follows:

1) Number of patents that the inventors contributed to inventing and that were applied for at the European Patent Office (EPO) in the period 1988-1998;
2) Average quality of these innovations as measured by the average number of forward citations across each inventor's patents and, alternatively, by the average common component indicator;
3) Maximum quality of the patents invented by the individual inventor, i.e. the inventor's patent with the largest number of forward citations and, alternatively, with the highest level of indicator.

The empirical investigation uses a sample of 793 European inventors. Information on individual characteristics is drawn from the PatVal-EU survey that interviewed the inventors of 9,017 EPO patents with a priority date of 1993-1997. Information on all the patents that the 793 individuals contributed to inventing and that were applied for at the EPO in 1988-1998 is collected from the EPO database. We jointly estimate three equations at the inventor-level with 1)-3) above as dependent variables. Our covariates are individual, firm, industry, and country characteristics.

The results of the empirical analysis suggest an intriguing story about the driving forces of inventors' productivity. Individual and organisational characteristics affect *quantity* after controlling for countries and sectors. Specifically, the number of individual patents increases with the inventors' age and academic degrees, their involvement in large research projects, and employment in large firms that apply for many patents. Surprisingly, however, none of these factors produce a direct effect on the *quality* of innovations (both average and maximum). Only the scale of the research project leading to the patents has a small impact on *quality* when this is measured by the composite indicator.

Further investigation, however, reveals that there is an additional factor to take into account in order to explain *quality*. This is the *quantity* of the inventors' patents that positively affects the probability of developing a technological hit. Therefore, while *quantity* depends on a few systematic factors, *quality* is the output of a stochastic process, where inventor and firm characteristics enter only indirectly through *quantity*. This implies a sort of hierarchy in the effects that we studied. Individual characteristics or other factors affect the number of patents that the inventors produce, and this, in turn, affects (positively) their maximum quality. By contrast, apart from a weak effect of the scale of the research project, the average quality of an inventor's patents is not correlated with any of the factors that we control for. We also find no evidence of regression to the mean (i.e. no decreasing returns) in the innovation process at the level of the individual inventor.

To identify the effect of the number of patents on the expected average and maximum quality of the innovations we exploit the information contained in the variance-covariance matrix of the residuals of the system of three equations.

The paper is organized as follows. We first overview the background literature in Section 2. We then present the data, describe the estimation procedure, and show results of the empirical tests (Sections 3, 4 and 5). Section 6 summarizes the results and draws some conclusive remarks.

2. Background literature

The determinants of research productivity over a researcher's life cycle have been studied in the economic literature as well as in other disciplines. A pioneer work is Lotka (1926) who shows that research productivity is concentrated. Other authors confirm these findings and explain them with differences in the distribution of ability among scientists, and with the allocation of recognition and resources to the most productive individuals that make them even more productive – the "Matthew Effect" whereby an initial success entails increasing productivity and reputation (Merton, 1968; Allison and Stewart, 1974; Cole 1979; David, 1994).

Yet other authors show that age matters in many disciplines with older scientists becoming less productive (Dalton and Thompson, 1971; Goldberg and Shenhav, 1984). Levin and Stephan (1991), for example, examine the research productivity of scientists over their life cycle in six scientific areas, and find that it declines over time. Oster and Hamermesh (1998) follow the careers of 208 economists in the economic departments of 17 top research institutions who received PhD degrees between 1959 and 1983. They provide evidence that publishing diminishes with age. They also demonstrate the presence of persistent heterogeneity among individuals: the most productive economists early in their careers keep producing high-quality research (though at a lower rate) as they become older.[1] Turner and Mairesse (2006) explore the differences in productivity among French condensed matter physicists between 1986 and 1997 in terms of the number and quality of their publications. They find a strong impact of individual and institutional characteristics. For the same sample of scientists, Hall et al. (2005b) try to disentangle the impact of cohort, age and period effects on researcher productivity.

Existing evidence about industrial inventors is much more limited compared to academic scientists, and it is based on small-scale samples, specific industries and firms. Narin and Breitzman (1995) tested Lotka's inverse square law of productivity on a sample of inventors in the R&D departments of four companies in the semiconductor industry. Similarly, Ernst *et al.* (2000) studied the research productivity of inventors in 43 German companies, both in terms of quantity and quality of their innovations (see also Ernst, 1998 for a study at the firm level).[2] This literature confirms that the distribution of productivity among industrial inventors is skewed. However, given the difficulties of obtaining information at an individual level and of tracing the careers of the industrial inventors, the reasons behind these disparities were not investigated.

Our work contributes to a more profound study of the relationship between research productivity, age and other determinants of *quantity* versus *quality* of the inventors' innovations. *Quantity* is defined by the number of innovations that the inventors contributed to inventing. *Quality* is

[1] Cole (1979) finds that age is concavely related to the quantity and quality of scientists' productivity in a cross-section analysis, while he finds no relationship between age and productivity with longitudinal data.

[2] From a different point of view, Breschi *et al.* (2006) investigate the relationship between publishing and patenting by Italian academic inventors, and find a strong and positive relationship between the two research outputs.

measured by two indicators that the literature shows to be correlated with the importance and economic value of the innovations. Quite a few contributions highlight the positive relationship between patent value indicators and the actual ex-post value of the innovations as given by traditional accounting evaluation (Hall et al., 2005a).[3] Trajtenberg (1990) shows that there is a close association between patent counts weighted by forward citations and the social value of innovations in the Computer Tomography Scanner industry. Harhoff *et al.* (1999) demonstrate that the number of backward citations to previous patents and to the non-patent literature, and the number of forward citations received after the publication date is positively correlated with the value of the innovations. Schankerman and Pakes (1986) use patent renewal data to estimate the value of patent rights, while others use the number of countries in which the patent is applied for and the number of claims in the patent application as indicators of the value of the patents (see, for example, Putnam, 1996). Finally, patents that undergo opposition and annulment procedures are also shown to be more valuable (Harhoff and Reitzig, 2004).

We start by employing the number of forward citations received by the patents within five years of the publication date as a measure of their quality. Then, in order to solve some of the problems that arise with the use of patent citations and to check for the robustness of the results we follow the approach developed by Lanjouw and Schankerman (2004). We use multiple indicators of the patent value to construct a composite index that proxies for the technological and economic quality of the patents. As we shall see in the next Section, this common component indicator reduces substantially the measured variance in patent quality (Lanjouw and Schankerman, 2004) and is correlated with the actual monetary value of the patents (Gambardella *et al.*, 2005).

[3] On the limitations of patent indicators see Griliches (1990), Almeida and Kogut (1999), Alcacer and Gittelman (2004), Singh (2005).

3. Data sources and construction of variables

3.1. Data sources

Our major source of data is the PatVal-EU survey conducted in 2003-2004. The survey interviewed the inventors of 9,017 patents granted by the EPO with a priority date of 1993-1997, and located in France, Germany, Italy, the Netherlands, Spain and the United Kingdom. The PatVal-EU database provides critical information for our study on the age, education, career and affiliation of a large sample of EPO inventors (for details and descriptive statistics see Giuri, Mariani *et al.*, 2005; PatVal-EU, 2005). We complemented the PatVal-EU dataset with data on all the 1988-1998 EPO patents of the inventors in our sample. We chose this time window because the EPO data are not fully reliable before the end of the 1980s. However, an eleven-year window is large enough to capture a sizable portion of an inventor's career.

We focused on a sample of 793 PatVal-EU inventors that we selected by taking all the German, Italian, Dutch and British inventors that responded to the PatVal-EU questionnaire on patents invented in five technological classes – Information Technology, Chemical Engineering, Civil Engineering, Optics, and Biotechnology. We employed the ISI-INIPI-OST classification developed by the German Fraunhofer Institute of Systems and Innovation Research (ISI) together with the French patent office (INIPI) and the Observatoire des Sciences and des Techniques (OST). This classification is based on the International Patent Classification (IPC) and distinguishes between 5 macro and 30 micro technological classes. The advantage of this classification is that it translates the IPC classes, which are largely technological, into classes that mimic industrial sectors. We employed one "micro" ISI-INIPI-OST technological class for each of the five "macro" classes (respectively, Electrical Engineering, Process Engineering, Mechanical Engineering, Instruments, Chemicals and Pharmaceuticals). We checked the distribution of patents in these five "micro" classes compared to the whole PatVal-EU sample. We found that the characteristics of the inventors and the innovation processes in our five micro classes did not differ substantially from those of the macro classes from which they are drawn. In particular, Wald tests on the mean difference of some key variables between countries and technologies confirm that our sample is not biased in any particular direction. Only in Biotechnology is the share of inventors with a PhD significantly higher than the average of the macro class. In Optics the estimated economic value of the innovations is also higher than the average of the macro class.

We employed the database Delphion to collect all the patents of our 793 inventors, either applied for or granted by the EPO in 1988-1998. The search on Delphion was followed by a matching procedure to solve problems of homonymy that arise because quite a few inventors had the same first and last names as some inventors in our list, although they were not the same individuals. To eliminate homonymous inventors we employed matching software that ran on two tables: a searching table with information on the 793 PatVal-EU inventors and a reference table with all the potential EPO matching patents extracted from Delphion.[4] The match between the EPO patents and the inventors was performed by using a set of weighted criteria: last inventor's name, first name, second name, technological class. The matching software browsed the reference table and reported the probability for each patent to be a "good match" with the inventor in the searching table. At the end of this process we checked the list of inventors manually in order to remove the patents invented by homonymous inventors. The PatVal-EU information on the name of the applicant, the technological class, the extent of inventors' mobility helped this check as well. Finally, we searched on the internet to solve the remaining doubtful cases.

This searching and cleansing procedure left us with data on 4,376 patents invented by the 793 PatVal-EU inventors. For each patent we have information on the number of claims, the number of states in which the innovation is patented, the name and location of the applicant organisation, the IPC classes in which the patent was classified, and the number of forward and backward citations.

[4] Delphion is an on-line database released by Thomson Corporation. It collects all the EPO patent applications issued since 1979. We thank Grid Thoma for downloading the European patents, either applied or granted, in which the first and last names of the inventors corresponded to one of our 793 inventors. The matching software is *SearchEngine v. 5.751* developed at the ZEW Research Institute by Thorsten Doherr.

The PatVal-EU survey provided us with data on the inventor who contributed to developing the patent and the organisation in which he was employed. Table 1 describes the composition of our sample of 793 inventors.

[TABLE 1]

The share of women in the sample is very low: it is larger in Biotechnology and smaller in Civil and Chemical Engineering, with no significant differences across countries. The average age of the inventors is 45, with some variation across countries and technologies. Furthermore, the more a technology is science-intensive, the larger is the share of inventors with PhDs – e.g. in Optics and Biotechnology compared to Civil Engineering or to the overall share of 33.4%. In Italy only 5.1% of the inventors in the sample have a PhD degree.[5] Table 1 also shows that the average number of patents per inventor in the database is 5.5, with a peak for German and Italian inventors (7.4 and 6.4 respectively).[6]

3.2. Productivity measures and regressors

The aim of this paper is to study the determinants of the inventors' research productivity in terms of the *quantity* and *quality* of the innovations that they produce.

The number of patents that the inventors contributed to inventing in 1988-1998 is our *quantity* measure (NPAT). As far as *quality* is concerned, we cannot use the monetary value of the patents as given by the PatVal-EU survey because this information is provided only for the surveyed patents. It is not available for all the patents that make up our 4,376 sample. We therefore decided to employ

[5] These data are consistent with OECD data that show that the share of population between 25 and 64 years old with a tertiary education degree is 24% in Germany and the Netherlands, 28% in the UK, and 10% in Italy (OECD, *Education at a Glance*, 2005).
[6] The distribution of patents is shown in Figure 1 below.

two quality indicators that the literature on the measurement of patent value shows to be correlated with the technological and economic impact of the innovations.

The first one is the number of forward citations that a patent receives. A patent must cite all related prior patents, and the patent examiner eventually checks and changes them in order to ensure that all appropriate citations are included in the list. These citations identify the rights of the applicant, and they are a signal for the technological importance of a patent as a source of knowledge on which subsequent patents are built. We collected the total number of patents that cite the 4,376 patents in our sample, and focused on the citations received within 5 years of the patent publication date to avoid "truncation" problems (i.e. more recent patents are less cited). We use these citations to construct two inventor-level measures of patent quality.

The first one is the average value of each inventor's patents, and it is given by the average number of forward citations across the inventor's patents (AVCITE). Inventors, however, often acquire visibility for one or several major inventions (Jones, 2005; Zucker *et al.*, 1998) that are rare and lie at the very right-hand side of the patent value distribution. We decided to look also at the factors that explain the probability of inventing the "best" innovation amongst those produced by the individual inventors. We measure the technological hits by the highest number of forward citations across each inventor's patents (MAXCITE). We therefore start by estimating a system of three equations in which NPAT, AVCITE and MAXCITE are the inventors' productivity measures and the dependent variables in our regression model.

However, these quality measures based only on the number of forward citations have some limitations (see Hall *et al*, 2001, for a survey). For example, citations cannot be made to or by innovations that are not patented, thus underestimating the actual importance of some of them. Second, patents applied for in different years and technological classes differ in their propensity to be cited, leading to changes in the number of citations per patent that stem from factors other than the actual changes in the technological impact of the innovations. Finally, there is also a different propensity to be cited according to the type and size of the organisation that applies for the patent. For example, large firms might have larger portfolios of "citing patents" compared to smaller enterprises and universities, and this can affect the number of citations that their patents receive if self-citations (i.e. citations made by patents applied for by the same applicant) are included.[7]

In order to solve these problems, and to check for the robustness of the empirical results, we follow Lanjouw and Schankerman (2004) and use an alternative measure of patent quality based on a composite quality-adjusted patent index which proxies for both the technological and economic impact of the innovations. As discussed by Lanjouw and Schankerman, it reduces the variance in patent quality compared to employing only one of the traditional value indicators, which suggests that there is a large information gain from employing multiple patent characteristics. The index is also cleaned from differences among patents that depend on country, time and technological characteristics. Moreover, since only a fraction of the index is composed of forward citations, potential differences between applicants in the propensity to be cited are less severe. Finally, the monetary value of patents is highly correlated with a common component indicator similar to the one that we employ in this paper (See Gambardella *et al.*, 2005).

By controlling for some observed patent characteristics, we derive the common factor as the unobserved characteristic of a patent that influences the following three indicators:

Forward Citations. This is the number of citations that a patent receives within 5 years of the patent publication date (as described above).

Backward Citations. We also collected the number of prior patents cited by the 4,376 patents. Backward citations are an indicator of others working on similar research fields, and therefore they signal the importance of the technological area.

---

[7] This potential bias cannot be solved by simply dropping self-citations from the citation list. This is because their role in measuring the value of the innovation compared to independent cites is not clear. For example, some firms cite their own patents because they have large patent portfolios to cite, while others cite themselves because there are internal spillovers and cumulative processes of knowledge creation, or because they are exploiting technological trajectories in specialised niches.

Claims. A claim describes the features of the invention, and defines the property rights protected by the patent. The inventor and the patent applicant have an incentive to write as many claims as they can, but the examiner may require some of them to be dropped - the larger the number of claims, the broader and the greater the expected profitability of an innovation.

From our sample of 4,376 patents we retrieved the parameter estimates to construct the quality index (Q).[8] To do so we controlled for some observed characteristics of the patents: the nationality of the inventors, the publication year, and the primary micro-technological class in which the patents are classified. For each indicator k (k = 1,...3) of the pth patent we ran the following multiple-indicator model with one latent common factor:

$$y_{kp} = \beta_0 + \boldsymbol{\beta' x_p} + \lambda_k q_p + \varepsilon_{kp} \qquad (1)$$

where y_{kp} indicates the value of the kth indicator for the pth patent (in logs). The common factor is q with factor loadings λ_k, while x_p denotes the vector of observed controls. From the matrix of variance-covariance between the error terms of the three equations we derive the parameter estimates of the common factor model. The top part of Table 2 shows the correlation between the errors of the three equations. The bottom part shows the parameter estimates of the indicators that make up the quality index.

[TABLE 2]

[8] We also constructed the index by employing four patent indicators: the three described in the text and Family Size, i.e. number of countries in which the innovation is patented. However, the correlation coefficients between the measurement error of the Family size equation and the measurement errors of the Backward citations and Claims equations are close to zero and they are statistically insignificant. This solves in a straightforward way the over-identification problem that arises with four indicators, and suggests that for our sample of patents we can use three patent indicators to construct the index, which provides exact identification. See Lanjow and Schankerman (2004) for details.

By employing this patent-level quality index we constructed two other inventor-level measures of the average and maximum patent quality. The average quality is calculated as the average of the index across all the inventor's patents (AVQ). The "most important" innovation amongst those produced by the individual inventor is measured by the inventor's highest value of the common component indicator across his patents (MAXQ). We then estimate a second set of equations where NPAT, AVQ, MAXQ are the dependent variables.

Figures 1, 2 and 3 show the distribution of NPAT, AVCITE, MAXCITE, AVQ and MAXQ across the 793 inventors. Consistent with other contributions on the productivity of individual scientists and inventors, these Figures confirm that the distribution of the inventors' productivity is skewed with few inventors being very productive in terms of the *quantity* and *quality* of the innovations that they produce.

[FIGURE 1, 2 and 3]

Our regressors are inventor characteristics, characteristics of the organisation in which they work, country and technological dummies. Table 3 lists and defines the variables.

[TABLE 3]

Age, sex, and education are the inventor characteristics. The age of the inventors in 1995 is included in linear and in quadratic form (AGE and AGE2). This is to mimic existing work in the literature that the effect of age on the researchers' productivity may first increase and then decline with age (Cole, 1979). The educational background proxies for the inventors' unobservable ability and for the knowledge that they embody and that was assimilated in the different stages of their scientific training. We employ the highest degree of education of the inventors among the following four types: Secondary School (SecSc), High School (HighSc), University BSc or Master (Uni), PhD (PhD). We expect the level of education to be positively correlated with the research performance of the inventors. Not only is this because inventors with better ability and scientific knowledge are expected to be more productive, but also because education might be a signal that the inventors and the employer organisations use to search for a good "match" between the research potential of the former and the characteristics of the latter.[9]

The type of employer organisation and the number of patents that the applicant was granted by the EPO are the two organisation characteristics included in the regressions. As far as the number of patents granted by the EPO to the applicant organisations is concerned (PATSORG), we use the data provided by the PatVal-EU database.[10] PATSORG is a proxy for the experience to innovate and the propensity to patent of the applicant. Moreover, the development of innovations often requires extensive resources in terms of technical equipment, research laboratories, instruments, research assistants and complementary expertise. The type of organisation in which the inventor is employed partially proxies for the availability of such resources. We differentiate between large firms (LARGE), small and medium companies (SME) and public research organisations (GOV). Compared to SME and GOV, large firms might have the financial resources to engage in complex research projects, to produce a large number of innovations and to apply for patent protection for many innovations. They are also likely to be endowed with a large pool of heterogeneous and specialised researchers that are involved in a wide number of larger research projects compared to smaller enterprises. We expect these characteristics to be positively correlated with the productivity of the individual inventors, both in terms of quantity and maximum quality of the innovations. It is

[9] See, for example, the seminal contributions by Arrow (1973) and Spence (1973) on the consequences of imperfect information. Also, Moore (1911) argues that "Large establishments are able to carry out the work of selection [of more capable individuals] because in consequence of their large capital and better organisation, they offer opportunities for more capable individuals to reap the reward of their differential ability". Idson and Oi (1999) confirm that "the adoption of advanced technologies, employment of inherently more able individuals, and higher work standards go together to raise labour productivity [...]".

[10] We use the number of EPO patents granted to the organisation and included in the PatVal-EU survey as we are interested in the distribution of patents across applicants rather than in the absolute number. Given the sampling methodology that we employed in the PatVal-EU survey, these data provide a good approximation of such distribution. A long process of data cleaning for company names would be needed to use the whole EPO dataset for this purpose.

more difficult to predict the effect of these variables on the average quality of the innovations. For example, if large companies have the financial strength and the human resources to apply for patent protection not only for important innovations, but also for less valuable ones, this might produce a decline in the average quality of the patented innovations both at the company and at the inventor level. Therefore, the net effect of the size of the firm on the average quality will depend on the extent to which higher quality compensates for the number of low quality patents that are applied for. By controlling for the size and type of the employer organisation, we also separate the effect of the scale of the organisation and its patent propensity, which otherwise would both be reflected by the same PATSORG variable.

We also include in the regression model a project-level measure of the resources available to the inventors for developing the innovations. This is the size of the research project leading to the patents: for each patent in our database we collect the number of inventors involved in the development of the innovation, and for each inventor we compute the average size (NINV) of the research projects in which he participated. We expect a positive correlation between NINV and the *quantity* and *quality* of the research performed by the inventors, which would suggest that the investment in large-scale research projects leads to better and larger innovative output, and that research teams matter for the development of a large number of high quality innovations (see Andrews, 1979 and Lawani, 1986 for the relationship between researchers' productivity and collaboration in scientific research).

Finally, we control in all the regressions for the country of the inventors (COUNTRY) and for the macro technological classes in which the patents are classified (TECH). Table 4 provides the descriptive statistics of the variables.

[TABLE 4]

4. Specification and Estimation: Step 1

We start by estimating a set of three equations in a reduced form model by Seemingly Unrelated Regressions (SUR). We employ two specifications. The first one, which uses the quality measures constructed with the number of forward citations, jointly estimates the probability of NPAT, AVCITE and MAXCITE.

$$\begin{cases} NPAT_i = \boldsymbol{x}_i'\boldsymbol{\alpha}_1 + \varepsilon_{1i} \\ AVCITE_i = \boldsymbol{x}_i'\boldsymbol{\alpha}_2 + \varepsilon_{2i} \\ MAXCITE_i = \boldsymbol{x}_i'\boldsymbol{\alpha}_3 + \varepsilon_{3i} \end{cases} \qquad (2)$$

The subscript i denotes the inventor, while α_1 α_2 α_3 are the coefficients to be estimated for the impact of the x_i organisation and inventors' characteristics on the three productivity variables. All the variables are in logs. Table 5 presents the estimated results that can be interpreted as elasticities of the dependent variables to changes in the regressors. Since NPAT and MAXCITE are count variables, we also show the results of these two equations by using Negative Binomial regressions.[11]

[TABLE 5]

The estimated results in Table 5 suggest that inventors and firm characteristics produce a different impact on the expected *quantity* compared to the *quality* of the innovations. Specifically, they are positively correlated with the probability of producing a large number of innovations (NPAT), but many of them lose significance in the quality equations (AVCITE and MAXCITE).

[11] We use a zero-truncated negative binomial regression model for $NPAT_i$ because the data are strictly positive (Cameron and Trivedi, 1998). Our definition of "European inventor" implicitly requires that the individual contributed to inventing at least one EPO patent. This is a necessary condition for the inventors to be included in our sample and interviewed in the PatVal-EU survey.

This is the case of AGE. As the inventors grow older, the probability of producing a large number of patents increases, but after a certain point the relationship becomes negative.[12] The same AGE variable, however, does not affect the probability of inventing valuable innovations (both average and maximum). Being a MALE is positively correlated with NPAT, AVCITE and MAXCITE, though with a different level of statistical significance. However, the positive effect of MALE on the productivity measures might be due to the very low number of women in our sample and to the fact that, on average, male inventors can spend more time and effort in their job compared to women.

As expected, the academic degree of the inventors produces a positive effect on NPAT: inventors with a PhD have a higher probability of producing a large number of patents. The coefficient of PhD is positive and statistically significant both in the SUR and in the Negative Binomial regressions. Also the applicant characteristics matter for NPAT. As expected, the dummy LARGE is positive and statistically significant at the 0.01 level on the number of patents (it is significant at the 0.05 level in the Negative Binomial model). Also PATSORG is positive and statistically significant at 0.01. Finally, the size of the research project is positively correlated with NPAT, confirming that investment in large-scale research projects and collaboration among a large number of researchers is correlated with the inventors' productivity. The Negative Binomial estimation confirms these results.

When we turn to the factors that affect the expected quality of the innovations, however, Table 5 shows that, apart from MALE, inventors' personal characteristics are not correlated with the average and maximum quality measures. Not even the inventors' academic degree matters for producing high quality innovations.

Some firm characteristics affect the *quality* equations. The size of the research project (NINV) and the size of the firm (LARGE) are positively correlated with AVCITE, even though the statistical

[12] We are aware of the potential source of selectivity bias in our paper (Cole, 1979; Levin and Stephan, 1991). However, since the inventor's life cycle is not the primary focus of our analysis, we do not include any correction to predict the likelihood that the inventors are active in the innovation business.

significance of the coefficients is only 5% and 10% respectively. Employment in a LARGE firm, the size of the research project (NINV) and the number of patents granted to the organisation (PATSORG) increase the probability of inventing a technological hit (MAXCITE). In reading these results, however, it is worth keeping in mind that we are measuring quality in terms of the number of forward citations received by the patents. As we mentioned in Section 3.2, large firms own larger portfolios of "potentially" citing patents compared to smaller enterprises. These large patent portfolios open the way to self-citations, which, in turn, contribute to the total number of citations received by the patents invented in large companies quite independently of their actual economic and technological value.

To address the concerns associated with forward citations as a proxy for the importance of the innovations, and to check for the robustness of our results, we also employ the composite index. We estimate the same set of three equations with the logs of NPAT, AVQ and MAXQ as dependent variables. The results in Table 6 confirm that inventor and firm characteristics affect the probability of developing a large number of patents. Interestingly, however, when quality is measured by the composite index the effect of firm characteristics (LARGE and PATSORG) vanishes. These results confirm our suspicion that the larger number of forward citations to patents applied for by large companies compared to small firms and universities might be independent of the difference in the actual quality of the cited innovations. In Table 6, apart from MALE, only the size of the research project (NINV) is positive and statistically significant on the average and maximum quality of the inventor's patents.[13]

[13] In order to check for complementarity between DEG4 and LARGE we included the interaction between DEG4 and the type of employer organisation (LARGE, SME and UNI) in our system of equations. The estimated results do not change compared to those in Tables 5 and 6, and the impact of the interacted variables is statistically not significant. This suggests that firm and personal factors have an independent effect on NPAT. We also checked for the robustness of the results by replacing AGE and AGE2 with five age classes. We also dropped alternatively NINV, PATSORG and LARGE, with no significant changes in the results shown in Tables 5 and 6. Finally, we included three motivations for inventing that the inventors reported in the PatVal-EU survey: economic compensation, reputation and career advances. Career advances turned out to be positively and significantly correlated with the quality of the innovations. However, it is possible that due to the phrasing of the questionnaire, the inventors misunderstood the question and interpreted it as the ex-post reward from patenting rather than an ex-ante motivation, i.e. the inventors who consider career important are those who experimented it. In other words, career is endogenous with respect to patent quality. This is confirmed by the fact that when we drop it, the coefficient and statistical significance of AGE increase.

[TABLE 6]

These results are intriguing, as they suggest a peculiar story arising from our data. Firm and personal characteristics influence the number of innovations that the inventors produce but, apart from participation in large-scale research projects, they do not affect the quality of the innovations at an inventor level. This is interesting and puzzling at the same time. We therefore move one step further in order to understand whether the effect of inventor and firm characteristics on *quality* takes place indirectly through *quantity*. Our perception is that firm and inventor factors are correlated with *quantity*, which, in turn, explains *quality*. The next section explores this hypothesis.

5. Specification and Estimation: Step 2

This section builds empirical evidence for the story that we envision about the drivers of inventors' productivity. The story goes as follows. Highly educated inventors are employed in large companies that have the financial resources and research capabilities for developing a large number of innovations. This is suggested by the estimated results of the NPAT equation, and it would also be consistent with anecdotal evidence that shows that the inventors are often evaluated and rewarded according to the number of patents that they contribute to inventing rather than the quality of any one of them. This motivates the researchers to fill out their CV with a long list of innovations. For example, in 2004, Siemens AG honoured 13 inventors responsible for about 600 patents invented in the same year. Similarly, the WIPO Award scheme, launched in 1979, grants prizes to inventors for their "outstanding research activities and numerous patented inventions" (Hoisl, 2005).

In turn, we know that the *quantity* of innovations that the inventor produces might affect their *quality*. This is true in the case of a technological hit, since the expected value of the maximum as an ordered statistic increases with the number of trials (see Mood *et al*, 1974). This is also possible, with a negative sign, for the average quality of the innovations. Unless the quality of the best innovations compensates for the large number of lower value patents, we expect that there is regression to the mean in the average quality of the innovations as the number of attempts increases. We therefore include in the two *quality* equations an additional explanatory variable: the number of innovations (NPAT) that we expect to have a positive effect on MAXCITE (and alternatively on MAXQ). It is more difficult to predict the sign of the relationship between NPAT and AVCITE (and alternatively AVQ) as it depends on how much the "best" innovations compensate for the large number of lower quality patents.

Operationally, we propose a structure for our model in which NPAT enters the AVCITE and MAXCITE equations, and not vice-versa. Specifically:

$$\begin{cases} NPAT_i = \boldsymbol{x}_i'\boldsymbol{\alpha}_1 + \varepsilon_{1i} \\ AVCITE_i = \boldsymbol{x}_i'\boldsymbol{\alpha}_2 + \vartheta_2 NX_i + \varepsilon_{2i} \\ MAXCITE_i = \boldsymbol{x}_i'\boldsymbol{\alpha}_3 + \vartheta_3 NX_i + \varepsilon_{3i} \end{cases} \tag{4}$$

By substituting the NPAT equation in the AVCITE and MAXCITE equations, the system above can be rewritten as follows:

$$\begin{cases} NPAT_i = \boldsymbol{x}_i'\boldsymbol{\alpha}_1 + \varepsilon_{1i} \\ AVCITE_i = \left(\boldsymbol{\alpha}_2 + \vartheta_2\boldsymbol{\alpha}_1\right)\boldsymbol{x}_i' + \vartheta_2\varepsilon_{1i} + \varepsilon_{2i} \\ MAXCITE_i = \left(\boldsymbol{\alpha}_3 + \vartheta_3\boldsymbol{\alpha}_1\right)\boldsymbol{x}_i' + \vartheta_3\varepsilon_{1i} + \varepsilon_{3i} \end{cases} \tag{5}$$

Now, to retrieve θ_2 and θ_3 we need a variable that affects only NPAT with no influence on the two quality measures AVCITE and MAXCITE. Unfortunately, we do not have a structural model that justifies such exclusion restrictions. A possible solution is to estimate θ_2 and θ_3 from the variance-covariance matrix between the residuals of our three equations. To do so it is natural to allow the errors ε_2 and ε_3 of the two quality equations to be correlated. Still, we assume the correlation between ε_2 and ε_3 on the one hand, and ε_1 on the other hand equal to zero.[14] By doing this we obtain the estimates of θ_2 and θ_3 and their standard errors as shown in the upper part of Table 7.

[TABLE 7]

As expected, in the case of the maximum, the sign and statistical significance of θ_3 confirm our story that, at the individual level, the larger the number of trials, the higher the probability to invent a technological hit. Surprisingly, also θ_2 is positive and statistically significant, which would suggest that there are increasing returns in the average quality of the innovations as their number increases.

However, when we estimate θ_2 and θ_3 from the variance-covariance matrix between the residuals of the NPAT, AVQ, MAXQ equations as shown in the bottom part of Table 7, θ_2 is small and it is not statistically significant. This suggests that, when quality is measured by means of the composite indicator to which forward citations contribute only partially, there is no regression to the mean in the innovation process at the individual level. The quality of the "best innovations" compensates for the many lower quality patents that are produced during the process. However, there is not the increasing returns process that emerges when we use forward citations to measure the average quality of the patents. Still, θ_3 is positive and statistically significant, which confirms the results above: the number of patents that the inventors produce influences the probability of developing a technological hit.

[14] *De facto* this assumes that we control for all the observed factors that simultaneously affect the three productivity measures. This is clearly a hypothesis, although we really control for major individual and institutional factors. Moreover, by including NPAT in the two quality equations, shocks on NPAT do not enter the quality equations via the error term.

The last question concerns the direct effect of inventor and firm characteristics on the two quality measures once NPAT is included in the regressions. The Wald test for the α_2 and α_3 parameters shows that the direct effect of all our regressors is statistically not significant once NPAT is controlled for. In particular, the size of the firm (LARGE) completely loses its explanatory power in the two quality equations. The same applies to PATORG that was statistically significant at the 0.01 level (0.05 in the Negative Binomial regression) in the MAXCITE equation. The Wald test shows that it is not directly correlated with any of the quality indicators. Also the average size of the research projects in which the inventor is involved (NINV) does not produce any direct effect on the quality of the patents.

Similarly, the Wald test for the α_2 and α_3 parameters in the AVQ and MAXQ equations shows that, once NPAT is included as a regressor, the other factors do not directly affect the expected average and maximum quality of the innovations. Only NINV is still statistically significant at 0.05 on both AVQ and MAXQ, suggesting that investment in large-scale research projects and collaboration among a large number of researchers positively affect the productivity of inventors not only in terms of the number of innovations, but also in terms of their quality when this is measured with the composite index.[15]

In sum, our story is that highly educated inventors are employed in large firms that give them the opportunity (and incentive) to produce a large number of innovations and to apply for many patents.[16] In turn, the larger the number of attempts, the higher the probability of developing a valuable innovation: more ideas lead to better ideas. Moreover, our study shows that the number of

---

[15] Also MALE is positive and statistically significant at the 0.05 level on MAXCITE and at 0.10 on MAXQ, when we control for NPAT.

[16] We checked for the idea that larger firms create jobs that match with more productive individuals. Specifically, we estimated a Probit regression where the dependent variable is 1 if the employer organisation is LARGE and 0 if it is SME. Inventors' characteristics, country and technology dummies are the regressors. We found that the level of education matters for the inventors to be hired by a large company: the coefficients of Uni and PhD are positive and statistically significant at 0.01 and 0.05 respectively, supporting the idea of large firms employing highly educated inventors, and inventors using the level of education as a signal for their "ability" to the potential employer. This is also consistent with a recent contribution by Baumol (2005) who argues that large-firm research and development laboratories require highly educated personnel. He also provides the example of Procter & Gamble (one of the world's largest holders of US and global patents) that has developed a global research organisation with over 7,500 scientists, including 1,250 PhDs.

inventor's patents, together with investment in large research projects, are the only factors that, at the individual inventor level, impact directly on the maximum quality of innovations. Inventor and organisational characteristics are only indirectly important for *quality*, as they positively influence the number of inventors' patents. Finally, as the number of inventor's patents increases, we find no regression to the mean in the average quality. This suggests that the average quality of an inventor's patents is determined by stochastic factors.

Still, we do not know whether the effect of NPAT on the maximum quality is the result of a stochastic process where the larger the number of draws, the higher the probability of obtaining a success, or whether this is the output of a learning process through which knowledge accumulates and raises the probability of inventing a high quality patent. Albeit it is not within the scope of this paper to analyse systematically this issue, Figure 4 gives some insights into the "learning versus stochastic" process leading to high quality patents. It depicts the progression of the patent quality index (in logs) over time for a sub-sample of 39 inventors. Precisely, we selected all the male inventors with a PhD employed in large firms and with more than 10 EPO innovations applied in the period 1988-1998. We show the progression in patent quality (IND) for each of these inventors in the 11-year window.[17]

[FIGURE 4]

The horizontal axis measures the sequence of patents invented over time in the period 1988-1998, and the vertical axis shows the quality of each innovation by means of the composite quality index. Each line identifies an inventor. As a first approximation it is hard to extrapolate a specific monotonic trend in patent quality at the inventor level, either decreasing or increasing as the number of patents grows. Although these results are not conclusive, they would suggest that no learning

[17] For inventors with multiple patents in the same year we consider the patent with the highest quality index. It does not change much, however, if we take the average quality index of the multiple patents applied for in the same year.

process is at work in the relationship between quantity and quality, and that the positive correlation between NPAT and MAXCITE (and MAXQ) is the output of a stochastic process, where the number of trials is correlated with the probability of finding a maximum.

6. Conclusions

This paper studied the driving factors of inventors' productivity. It estimated the impact of individual and organisation characteristics on the expected *quantity* and *quality* of the innovations produced by the European inventors. In so doing it used information on the individual characteristics of 793 inventors drawn from the PatVal-EU database, which was compiled in 2003-2004 by interviewing the inventors of 9,017 European patents. For our 793 inventors we also collected information on 4,376 patents that they contributed to inventing and that were applied for at the EPO in 1988-1998.

Quantity is measured by the number of EPO patents that the inventors produced in 1988-1998. We constructed two measures of patent *quality* at the level of the individual inventor: the average quality of the patents that each inventor produced in 1988-1998; and the quality of the "best" patent that he invented in the same period. We computed the average and maximum quality by means of the number of forward citations that the patents received within 5 years of the publication date. Alternatively, we employed a common component index as developed by Lanjouw and Schankerman (2004) that proxies for the technological and economic impact of the patents.

We regressed our measures of patent quantity, average quality and maximum quality on inventors' personal characteristics (i.e. age, academic degree, gender) and on the characteristics of the organisation in which the inventors were employed while developing the innovations (large firms, small and medium enterprises, universities and other public research organisations), including the propensity of the organisation to apply for patent protection and the average size of the research project in which the inventors were involved. We employed seemingly unrelated regressions to estimate the effect of each variable on the probability of developing a large quantity of innovations as compared to producing high quality innovations. We used the matrix of variance-covariance between the residuals of the three equations to retrieve the estimated impact of the number of innovations developed by the inventor on their average and maximum quality. Our results indicate that the drivers of *quantity* differ from those that impact on *quality.*

Specifically, inventors with a high level of education and inventors employed in large firms are more likely to produce a large *quantity* of innovations. This might be because large firms provide the inventors with the financial resources and complementary expertise to develop and apply for many patents. Probably, this is also due to the fact that large firms provide the inventors with the incentives to produce a large number of innovations, as they use this information to identify and reward productive inventors.

Inventor and firm characteristics, however, do not impact directly on the expected *quality* of the innovations, neither average nor maximum. The impact of inventors' characteristics on quality is indirect through the number of innovations that skilled inventors produce. We find, indeed, that the number of patents developed by the individual inventor produces a positive and statistically significant effect on the *maximum quality* of his patents. In other words, more ideas lead to better ideas. Apart from the number of innovations, only the size of the research project produces a direct effect on the maximum quality.

Interestingly, our results also suggest that there is no regression to the mean at the inventor level: as the number of patents per inventor gets larger, the *average quality* of the innovations does not decrease. Finally, further investigation on the distribution of patent quality at the individual inventor level suggests that also the technological hit is not the outcome of a systematic learning process that takes place as the number of trials increases. As a matter of fact, our data is consistent with a story in which "fortune" leads to the inventor's best innovation, where fortune depends on the number of draws. An in-depth investigation of this issue is a topic for future research.

All in all, the inventors' productivity in terms of *quantity* of patents that they produce is driven by a few deterministic factors at the level of the individual and the employer organisation. This means that, on the basis of inputs availability, there is room for predicting such productivity. Differently, productivity in terms of patent *quality* is more uncertain and difficult to predict, as the probability of getting a technological hit increases with the number of trials and the average quality depends on stochastic factors. As a consequence, if inventors are given a certain amount of resources to use in the innovation process, they might choose to focus on the "less risky" strategy of producing *quantity* rather than concentrating on *quality*.

It is also worth noting that our findings do not diminish the importance of firm and individual characteristics in producing high-quality innovations. Although indirectly through *quantity*, they affect *quality*. This suggests, for example, that a key resource for fostering innovation is the availability and employing of highly educated researchers, which, in turn, depends on investment in high level and postgraduate education and training. Moreover, the importance of the inventors' education and the fact that their productivity distribution is much skewed reinforce the "retaining the best" policy (Narin and Breitzman, 1995): if ability is concentrated in a few key individuals, the organisations in which they are employed should do their best to retain them. This is most important if these individuals are responsible for both the *quantity* (directly) and *quality* (indirectly) of the innovations developed by the organisation, and if, at the inventor level, there is no regression to the mean in the average value of the innovations.

References

Alcacer, J., Gittleman, M., 2004. How do I Know what you Know? Patent Examiners and the Generation of Patent Citations, New York University, Working Paper.

Allison, P.D., Stewart, J.A., 1974. Productivity differences among scientists: evidence for accumulative advantage. American Sociological Review 39 (4), 596-606.

Almeida, P., Kogut, B., 1999. Localization of knowledge and the mobility of engineers in regional networks. Management Science 45 (7), 905-917.

Andrews, F.M. (Ed.), 1979. Scientific Productivity: The Effectiveness of Research Groups in Six Countries. Cambridge University Press, Cambridge.

Arrow, K., 1973. Higher education as a filter. Journal of Public Economics 2 (3), 193-216.

Baumol, W.J., 2005. Education and innovation: entrepreneurial breakthroughs versus corporate incremental improvements, in: Jaffe, A.B., Lerner, J., Stern, S. (Eds.), Innovation Policy and the Economy, NBER, pp. 33-56.

Breschi, S., Lissoni, F., Montobbio, F., 2006. The scientific productivity of academic inventors: new evidence from Italian data. Economics of Innovation and New Technology, forthcoming

Cameron, A.C., Trivedi, P.K., 1998. Regression Analysis of Count Data. Cambridge University Press, New York.

Cole, S., 1979. Age and scientific performance. American Journal of Sociology 84, 958-977.

Dalton, G.W., Thompson, P.H., 1971. Accelerating obsolescence of older engineers. Harvard Business Review 49, 57-67.

David, P., 1994. Positive feedbacks and research productivity in science: reopening another black box, in: Granstrand, O. (Ed.), Economics of Technology. North Holland, Amsterdam, pp. 65-85.

De Solla Price, D., 1963. Little Science, Big Science. Yale University Press, New Haven, CT.

Ernst, H., 1998. Industrial research as a source of important patents. Research Policy 27 (1), 1-15.

Ernst, H., Leptien, C., Vitt, J., 2000. Inventors are not alike: the distribution of patenting output among industrial R&D personnel. IEEE Transactions on Engineering Management 47(2), 184-199.

PatVal-EU, 2005. The Value of European Patents. Evidence from a Survey of European Inventors. Final Report of the PatVal-EU Project, DG Science & Technology, European Commission, Contract N. HPV2-CT-2001-00013. Brussels.

Gambardella, A., Harhoff, D., Verspagen, B., 2005. The value of patents, paper presented at the NBER Conference, The Economics of Intellectual Properties, Cambridge (MA) July, 19th 2005.

Giuri, P., Mariani, M., Brusoni, S., Crespi, G., Francoz, D., Gambardella, A., Garcia-Fontes, W., Geuna, A., Gonzales, R., Harhoff, D., Hoisl, K., Lebas, C., Luzzi, A., Magazzini, L., Nesta, L., Nomaler, O., Palomeras, N., Patel, P., Romanelli, M., Verspagen, B., 2005. Everything you always wanted to know about inventors (but never asked): Evidence from the PatVal-EU survey. LEM Working Paper No. 2005/20, Sant'Anna School of Advanced Studies, Pisa.

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

Goldberg, A.I., Shenhav, Y.A., 1984. R&D career paths: their relation to work goals and productivity. IEEE Transactions on Engineering Management 31(3), 111-117.

Hall, B., Jaffe, A., Trajtenberg, M., 2001. The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools, NBER working paper No. 8498.

Hall, B.H., Jaffe, A., Trajtenberg, M., 2005. Market value and patent citations. RAND Journal of Economics, 35(1), 16-38

Hall, B.H., Mairesse, J., Turner, L., 2005. Identifying Age, Cohort and Period Effects in Scientific Research Productivity: Discussion and Illustration Using Simulating and Actual Data on French Physicists, NBER Working paper no. 11739.

Harhoff, D., Narin, F., Scherer, F.M., Vopel, K., 1999. Citation frequency and the value of patented innovation. Review of Economics and Statistics 81, 511-515.

Harhoff, D., Reitzig, M., 2004. Determinants of opposition against EPO patent grants - the case of biotechnology and pharmaceuticals. International Journal of Industrial Organization 22, 443-480.

Hoisl, K., 2005. A Closer Look at Inventor Productivity. LMU Discussion Paper.

Idson, T.L., Oi, W.Y., 1999. Workers are more productive in large firms, *American Economic Review* 89(2): 104-108.

Jones, B. F., 2005. Age and Great Invention. NBER Working Paper No. 11359.

Lanjouw, J.O., Schankerman, M., 2004. Patent quality and research productivity: measuring innovation with multiple indicators. Economic Journal 114, 441-465.

Lawani, S.M., 1986. Some bibliometric correlates of quality in scientific research. Scientometrics 9 (1-2), 13-25.

Levin, S.G., Stephan, P.E., 1991. Research productivity over the life cycle: evidence for academic scientist. American Economic Review 81 (1), 114-132.

Lotka, A.J., 1926. The frequency distribution of scientific productivity. Journal of the Washington Academy of Science 16(2), 317-323.

Merton, R.K., 1968. The Matthew effect in science. Science 159 (January), 56-63.

Mood, A., Graybill, F.A., Boes, D.C., 1974. Introduction to the Theory of Statistics. McGraw-Hill.

Moore H.L., 1911. Laws of wages: An essay in statistical economics, New York: Augustus M. Kelley

Narin, F., Breitzman, A., 1995. Inventive productivity. Research Policy 24(4), 507-519.

OECD, 2005. Education at a Glance. OECD, Paris.

Oster, S.M., Hamermesh, D.S., 1998. Aging and productivity among economists. The Review of Economics and Statistics, 80(1), 154-156.

Putnam, J., 1996. The value of international patent rights, Ph.D. thesis, Yale University.

Schankerman, M., Pakes, A., 1986. Estimates of the value of patent rights in European countries during the post-1950 period. Economic Journal 96, 1052–1076.

Singh, J., 2005. Collaborative networks as determinants of knowledge diffusion patterns. Management Science, 51(5), 756-770

Spence, M.A., 1973. Job market signalling. Quarterly Journal of Economics 87(3), 355-374,

Stephan, P., 1996. The economics of science. Journal of Economic Literature 34 (Sept.), 1199-1235.

Trajtenberg, M., 1990. A penny for your quotes: patent citations and the value of innovations. RAND Journal of Economics 21 (1), 172-187.

Turner, L., Mairesse, J., 2006. Individual productivity differences in scientific research: an econometric study of the publication of French physicists. Annales d'Economie et de Statistique.

Zucker, L., Darby, M., Armstrong, J., 1998. Geographically localized knowledge: spillovers or markets? Economic Inquiry 36, 65–86.

Figures and Tables

Table 1. Gender, age, education and number of patents applied for by the 793 inventors (1988-1998). Distribution by technological class and country.

	% of female inventors	average age* yrs.	% of inventors with university BSc or Master	% of inventors with PhD	average # of patents in 1988-1998
Information Technology (145)	1.4	41 (9.5)	51.2	33.3	5.9 (9.3)
Optics (139)	5.0	42 (8.9)	34.1	46.4	6.4 (6.9)
Biotechnology (53)	9.8	43 (9.0)	15.4	48.1	4.6 (4.2)
Chemical Engineering (198)	1.0	46 (9.8)	39.8	38.3	6.1 (8.5)
Civil Engineering (258)	0.8	48 (9.1)	44.3	19.6	4.6 (5.8)
Germany (304)	2.9	47 (9.6)	46.9	33.4	7.4 (8.5)
Italy (119)	2.3	44 (10.8)	56.4	5.1	6.4 (10.3)
The Netherlands (162)	1.7	42 (7.5)	13.8	54.4	3.5 (3.9)
UK (208)	1.9	45 (10.1)	43.4	34.6	3.8 (4.3)
Total	2.3	45 (9.7)	40.6	33.4	5.5 (7.4)

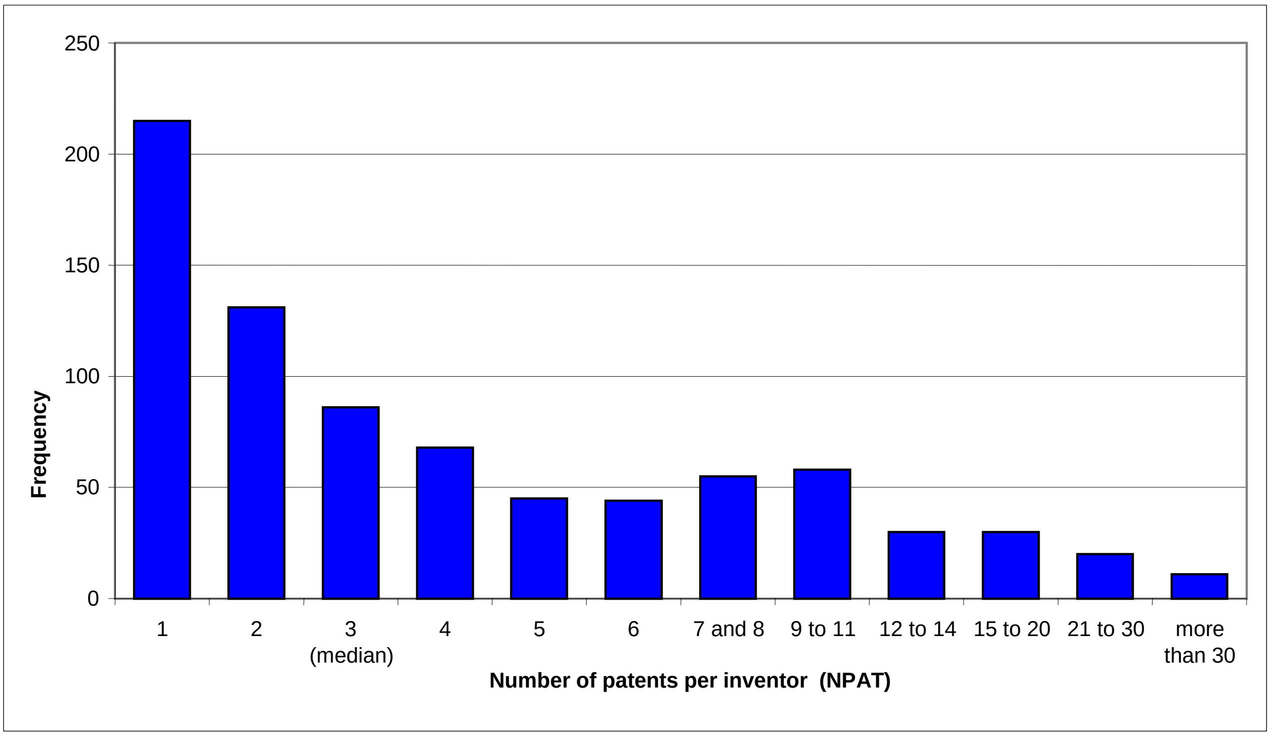
* The age of the inventors is calculated as 1995-date of birth. Standard deviations in parenthesis.
Note: The number of observations is in parentheses next to technological classes and countries.

Table 2. Correlation between the errors of the three equations and parameter estimates.

Errors of the three equations	Forward citations equation	Backward citations equation	Claims equation
Forward citations eq.	1.000		
Backward citations eq.	0.0282*	1.000	
Claims eq.	0.102***	0.027*	1.000
Parameter Estimates of the Common Factor model			
Variable (log)			
Forward citations (5 years after publication date)		0.21*** (0.082)	
Backward citations		0.04** (0.017)	
Number of claims		0.27*** (0.106)	

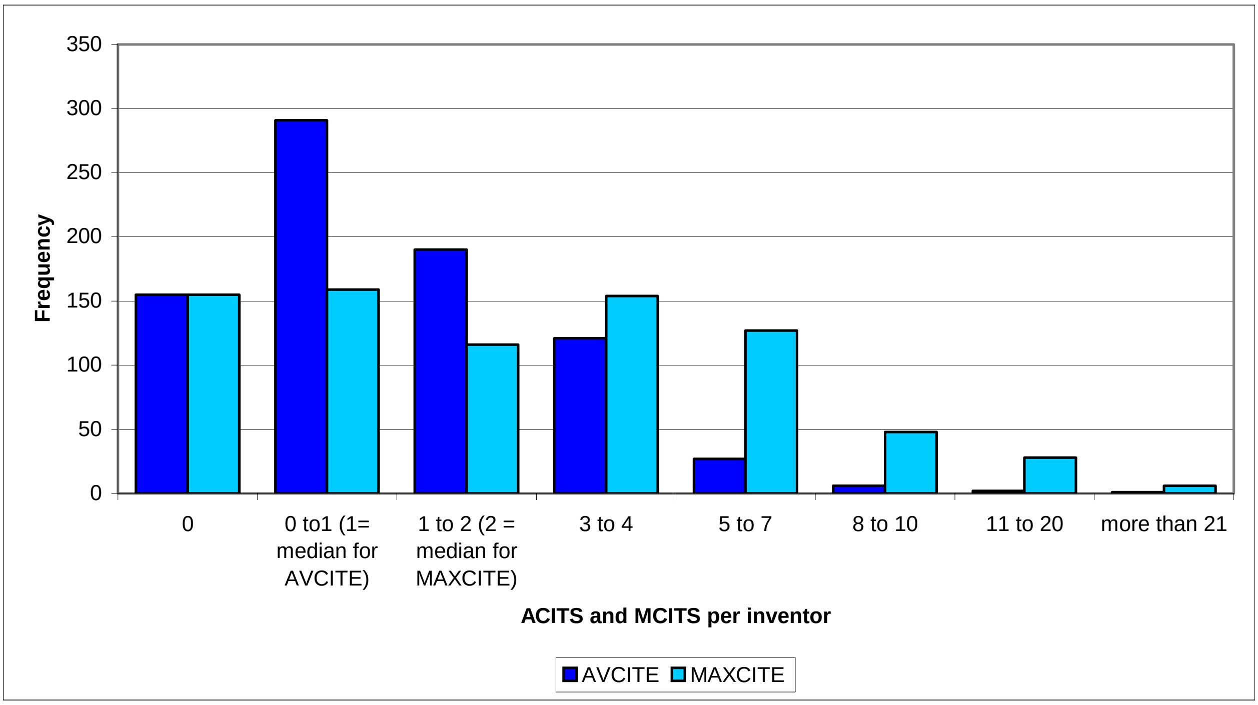
Note: To calculate the index we used dummies for the nationality of the inventors, the publication year, and the primary micro-technological class in which the patents are classified. Standard errors in parenthesis. Significant at: *0.1 level; **0.05 level; ***0.01 level.

Figure 1: Distribution of NPAT.



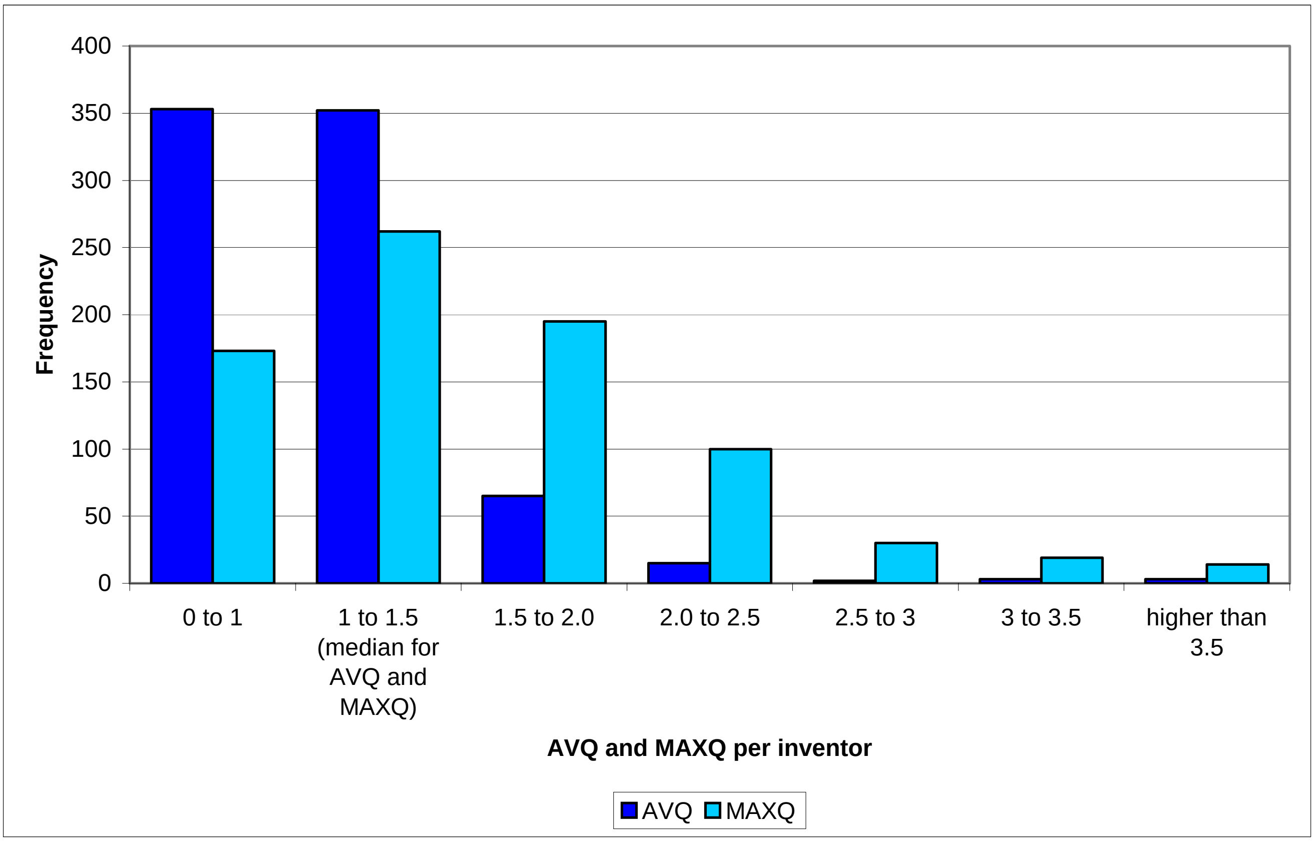
Note: # obs. 793

Figure 2: Distribution of AVCITE and MAXCITE.



Note: # obs. 793

Figure 3: Distribution of AVQ and MAXQ.



Note: # obs.: 793.

Table 3. List of variables.

	Productivity measures
NPAT	Number of patents invented by each inventor. Publication date 1988-1998 (Source: EPO)
AVCITE	Average quality of the patents invented by each inventor and applied for in 1988-1998. Measured by the average number of citations received by the inventor's patents within 5 years of the publication date (Source: EPO).
MAXCITE	Maximum quality of the patents invented by the inventor among those applied for in 1988-1998. Measured by the highest number of citations received within 5 years of the publication date across each inventor's patents (Source: EPO).
AVQ	Average quality of the inventor's patents applied for in 1988-1998. Measured by mean of common component index across each inventor's patents. (Source: Elaborations from EPO data)
MAXQ	Maximum quality of patents invented by the inventor among those applied for in 1988-1998. Measured by the highest common component index among each inventor's patents. (Source: Elaborations from EPO data)
	Inventors characteristics
AGE and AGE^2	Age of inventors = 1995-date of birth. (Source: PatVal-EU)
DEGREE	Dummy variable. Highest Academic degree of inventor at the time in which he developed the PatVal-EU innovation: Secondary School or lower (SecSc); High School (HighSc); University BSc or Master (Uni); PhD (PhD). (Source: PatVal-EU)
GENDER	Dummy variable. Male or Female inventor (Source: PatVal-EU)
NINV	Number of inventors who take part in development of a patent: average across each inventor's patents applied for in 1988-1998 (Source: EPO).
	Characteristics of employer organisation
EMPL	Dummy. Type of employer organisation: Large Firm (LARGE) with more than 250 employees; Medium and Small Firms (SME) with less than 250 employees; Universities and Government Institutions (GOV). (Source: PatVal-EU).
PATSORG	Number of patents granted to the applicant organisation in the PatVal-EU sample (Source: PatVal-EU)
	Controls
TECH	Dummy variables for the macro ISI-INIPI-OST technological classes in which the inventor's patents are classified: Electrical Engineering (ELENG), Process Engineering (PRENG), Mechanical Engineering (MECENG), Instruments (INST), Chemicals & Pharmaceuticals (CHEM). We used the technological class in which the majority of the inventor's patents falls.
COUNTRY	Dummy variables for the country of the inventor: Germany (GER), Italy (IT), the Netherlands (NL), UK (UK)

Source: PatVal-EU dataset and EPO

Note: The characteristics of employer organisations are provided by the PatVal-EU survey. We checked whether the patents in our 1988-1998 sample were applied for at the EPO during the period in which the inventors were employed in the organisation reported in the PatVal-EU survey. We found out that almost all our inventors did not change job during 1988-1998.

Table 4. Descriptive statistics.

	Productivity measures			
	Mean	S.D.	Min.	Max
NPAT	5.51	7.40	1	87
AVCITE	1.40	1.75	0	24
MAXCITE	3.30	3.82	0	31
AVQ	1.11	0.39	0.37	3.93
MAXQ	1.54	0.69	0.37	6.43
	Inventors characteristics			
AGE	45	9.71	18	73
DEGREE	3.26	0.82	1	4
GENDER	0.98	0.15	0	1
NINV	2.40	1.25	1	8.6
	Characteristics of the employer organisation			
LARGE	0.60	0.49	0	1
SME	0.30	0.46	0	1
GOV	0.11	0.32	0	1
PATSORG	33.44	72.95	1	286

Source: PatVal-EU dataset and EPO

Table 5. Estimates of SUR and Negative Binomial Regression. Dependent variables: NPAT, AVCITE and MAXCITE (Variables in logs).

Dependent variables: Log of Number of Patents and Forward Citations					
	SUR			NegBin	
	NPAT	AVCITE	MAXCITE	NPAT	MAXCITE
AGE	9.17*** (3.52)	-1.52 (2.66)	1.83 (3.26)	20.41*** (7.55)	3.31 (4.31)
AGE^2	-1.17** (0.47)	0.21 (0.35)	-0.22 (0.43)	-2.62*** (1.00)	-0.41 (0.57)
DEGREE: HighSc	0.06 (0.17)	-0.17 (0.11)	-0.17 (0.15)	-0.09 (0.33)	-0.29 (0.22)
DEGREE: Uni	0.12 (0.15)	-0.10 (0.10)	-0.03 (0.13)	0.22 (0.17)	-0.01 (0.19)
DEGREE: PhD	0.33** (0.16)	-0.10 (0.10)	0.02 (0.13)	0.59** (0.29)	0.01 (0.20)
MALE	0.54*** (0.21)	0.29* (0.14)	0.52*** (0.19)	0.77 (0.47)	0.72** (0.38)
NINV	0.35*** (0.07)	0.09** (0.05)	0.21*** (0.06)	0.51*** (0.15)	0.34*** (0.08)
LARGE	0.31*** (0.10)	0.12* (0.07)	0.23** (0.09)	0.36** (0.17)	0.28** (0.13)
SME	0.07 (0.10)	0.04 (0.07)	0.08 (0.10)	0.09 (0.12)	0.05 (0.13)
PATSORG	0.06*** (0.02)	0.02 (0.01)	0.05*** (0.02)	0.08*** (0.02)	0.05** (0.02)
Const.	-18.37*** (6.53)	3.19 (4.98)	-3.58 (6.07)	-40.56*** (14.44)	-6.61 (8.03)
# obs. 767					
Log Likelihood	-1360.69			-1866.81	-1729.84

Note: Robust Standard Errors in parentheses. Sample: 793 inventors. All regressions include dummies for country of inventors and for the macro ISI-INIPI-OST technological class of his patents.
Coefficient significant at: *0.1 level; **0.05 level; ***0.01 level.
Zero-truncated Negative Binomial regression for NPAT: $\ln(\alpha) = 0.45$ (1.08); R-squared = 0.20
Negative Binomial regression for MAXCITE: $\ln(\alpha) = -0.38$ (0.08); R-squared = 0.12

Table 6. Estimates of SUR and Negative Binomial Regression. Dependent variables: NPAT, AVQ and MAXQ (Variables in logs).

Dependent variables: Log of Number of Patents and Quality Index			
	SUR		
	NPAT	AVQ	MAXQ
AGE	9.17*** (3.52)	0.21 (1.57)	2.72 (1.82)
AGE^2	-1.17** (0.47)	-0.03 (0.21)	-0.36 (0.24)
DEGREE: HighSc	0.06 (0.17)	-0.07 (0.07)	-0.06 (0.08)
DEGREE: Uni	0.12 (0.15)	-0.03 (0.06)	0.00 (0.07)
DEGREE: PhD	0.33** (0.16)	-0.07 (0.06)	-0.01 (0.07)
MALE	0.54*** (0.21)	0.17 (0.10)	0.33** (0.12)
NINV	0.35*** (0.07)	0.07** (0.03)	0.15*** (0.03)
LARGE	0.31*** (0.10)	0.03 (0.04)	0.08 (0.05)
SME	0.07 (0.10)	0.04 (0.04)	0.05 (0.05)
PATSORG	0.06*** (0.02)	-0.01 (0.01)	0.01 (0.01)
Const.	-18.37*** (6.53)	-0.47 (2.93)	-5.36 (3.40)
# obs. 767			
Log Likelihood - 660.881			

Note: Robust Standard Errors in parentheses. Sample: 793 inventors. All regressions include dummies for inventors' country and macro ISI-INIPI-OST technological class.
Coefficient significant at: *0.1 level; **0.05 level; ***0.01 level.

Table 7. Parameter estimates: θ_2 and θ_3.

Computed with NPAT, AVCITE, MAXCITE as dependent variables	
θ_2	0.108*** (0.022)
θ_3	0.480*** (0.024)
Computed with NPAT, AVQ, MAXQ as dependent variables	
θ_2	0.016 (0.013)
θ_3	0.248*** (0.015)

Note: Robust Standard Errors in parentheses.
Coefficient significant at: *0.1 level; **0.05 level; ***0.01 level.

Figure 4: Progression of patent quality index (in logs) over time. Sample: male inventors with PhD who invented 10 or more patents in 1988-1998