Inventive Productivity and Patent Quality: Evidence from Italian Inventors

Schettino, Francesco and Sterlacchini, Alessandro and Venturini, Francesco

Department of Management Industrial Organization, Università Politecnica delle Marche

17 March 2008

Inventive Productivity and Patent Quality: Evidence from Italian Inventors

Francesco Schettino, Alessandro Sterlacchini and Francesco Venturini*

Abstract
By considering a regional sample of Italian inventors, this paper explores the factors behind the different individual performances in terms of number and quality of patents. Our reference population is composed of 570 inventors residing in the Marche region who, over the period 1991-2005, have contributed to 743 patent applications filed at the European Patent Office. Looking at the number of patents per inventor, a Lotka's distribution emerges suggesting that also for geographical areas inventive activities are highly concentrated in a few key inventors. To examine whether both the inventive productivity and quality are affected by individual and firm characteristics, we use the outcomes of a survey on 106 inventors. We find that the patent productivity is not influenced by individual characteristics but it is higher for the inventors working in teams and employed in large firms with greater patent portfolios. With respect to patent value we employ a composite index in which forward citations, claims and patent families are taken into account. Measured in this way, patent quality is significantly associated, along with the presence of an inventive team, with a set of individual features such as the inventors' experience and level of education. This suggests that inventions coming from individuals working in small firms or independently can be as valuable as those generated by inventors occupied in larger companies.

Keywords: Inventors, Inventive productivity, Patent quality.

JEL codes: O31, O34.

This draft: 17 March 2008

* Department of Management & Industrial Organization, Università Politecnica delle Marche, Ancona (Italy). Corresponding author: a.sterlacchini@univpm.it.

1. Introduction

Albeit imperfect indicators of innovative activities, patent data have been extensively used in empirical studies concerned with industries, firms, countries and regions with a view to explain their economic performances. To overcome a typical drawback of count data, the patents (or patent applications) attributed to a unit of observation should be weighted by their technological and economic value which can be approximated by some quality measures, such as forward citations. Thanks to the recent availability of comprehensive databases containing information on a number of patent characteristics, the economists of innovation have increasingly taken into account the qualitative dimension of patenting.

The above advancements, coupled with the recent explosion of patent applications and the mounting importance of Intellectual Property Rights, help explain why the economic analysis of patenting is thriving. In this favourable context, little attention has been devoted to the role played by inventors in spite of the fact that patented or patentable inventions derive from their ingenuity and creative work. If the majority of inventors were dependent workers of the companies applying for patents or holding patent rights, there will be little or no reason for taking them as units of observations. However, this is not the case. As documented in this paper, in a time span of 39 years, 58% of the applicants who have obtained a patent from the United States Patent and Trademark Office (USPTO) have got only one patent. In Europe, from 1991 to 2005, 2 out of 3 applicants at the European Patent Office (EPO) have sought patent protection only once. This suggests that a remarkable share of inventors is composed of entrepreneurs or owners of small firms, as well as individuals having nothing to do with extant business units. Moreover, some case studies have shown that also in large companies - holding a very high number of patents - most of their inventive outcomes derive from the work of a few key inventors. Accordingly, the distribution of patents among inventors is extremely skewed, with the key inventors of large companies contributing to a disproportionate share of patents. However, there is little evidence on whether the distribution of inventive outputs is equally skewed when the quality of patented inventions is taken into account.

In this paper we analyse the inventive productivity and patent quality of the inventors residing in an Italian region, the Marche, who have contributed to 743 patent applications filed at the EPO during the period 1991-2005. Looking at the number of patents per inventor, a Lotka distribution emerges suggesting that also for geographical areas inventive activities are highly concentrated in a few key inventors. To examine whether both the inventive productivity and quality are affected by individual and organizational characteristics, we use the outcomes of a survey on 106 inventors. We find that patent productivity is not influenced by individual characteristics but turns out to be higher for inventors working in teams and those employed in large firms with greater patent portfolios. With respect to patent value we employ a composite index in which forward citations, claims and patent families are taken into account. Measured in this way, patent quality is not significantly associated with the inventors' productivity but it is positively affected by the presence of an inventive team and a set of individual features such as the inventors' experience and level of education. This suggests that the distribution of patent quality among inventors can be less skewed than that of patent counts or, to put it another way, that inventions coming from individuals working in small firms or independently can be as valuable as those generated by inventors occupied in larger companies.

The paper is organized as follows. Section 2 examines the patents' distribution among inventors according to count data. Section 3 illustrates the main features of the inventors composing our sample and, on the basis of different measures, the quality of their patent applications. In section 4 we estimate a composite index of patent quality extracted from forward citations, claims and patent families. Section 5 presents the results of a regression analysis for both the number of patent applications per inventor and their average quality. Section 6 concludes.

2. The patents' distribution among inventors

This paper explores the patenting activities of a small subset of Italian inventors, those residing in the Marche region. Located in the east side of Central Italy, Marche hosts 1.48 million people and, compared to other European regions, has a low propensity to apply for EPO patents: in 2002, it recorded 55 EPO applications per million inhabitants while the average for the whole EU15 was 160. However, from 1991 to 2004 the number of applications submitted by Marche residents increased from 35 to 82, i.e. by 147%, while the EPO applications coming from EU residents rose by 90%. Thus, although the absolute and relative figures are not remarkable, Marche inventors have ascribed an increasing importance to European patents.

For the purposes of our study, we have considered the patent applications filed at the EPO during the period 1991-2005 (and published by December 2006) in which both the applicant and at least one inventor were residing in the Marche region[1]. In this way, we got a population of 570 inventors contributing to 743 patent applications.

Looking at the distribution of patents among inventors we found that the large majority of them (75%) contributed, in a time span of fifteen years, to one patent application only. At a first sight, we interpreted this prevalence of 'occasional' inventors as a peculiar feature of the Marche region (probably due to a low presence of high-tech industries and large firms; cf. Schettino and Sterlacchini, 2007). However, by examining how patents are distributed in other much broader contexts we discovered that this is not the case. In fact, as Table 1 shows, the prevalence of occasional inventors appears to be a generalised phenomenon.

Table 1 – Number and percentage of patent applications (EPO) or patent grants (USPTO) per inventor

	EPO-Marche		EPO-EU15*		USPTO**	
	No.	Perc.	No.	Perc.	No.	Perc.
1 patent	429	75%	427825	69%	908152	58%
2-4 patents	116	21%	149714	24%	454076	29%
5-9 patents	17	3%	31571	5%	125262	8%
10 or more patents	8	1%	12367	2%	78289	5%
Total inventors	570		621478		1575780	

*Own computations from OECD/EPO patent citation database. Period: 1991-2005.
**Trajtenberg (2005) with data taken from the NBER Patent Data File. Period: 1963-1999.

By using the OECD/EPO patent citations database[2] (which contains information for all the patent applications filed at the EPO since 1978; cf. Webb et al., 2005) we found that, over the period 1991-2005, 69% of the inventors residing in the EU15 countries took part to one patent application only. Thus, the Marche inventors seem less prolific than their European counterparts but the gap is not substantial[3]. Table 1 also presents the patents per inventor

[1] The few Marche inventors contributing to extra-regional inventions were neglected because, through a survey, we wanted to obtain information on both regional inventors and applicants.

[2] We are grateful to Colin Webb for providing us with the database.

[3] It must be added that the gap becomes negligible by excluding the patent applications in the technological fields of Chemicals & Pharmaceuticals and Electrical engineering in which the Marche inventors are by far less

granted by the USPTO, a figure extracted by Trajtenberg (2005) from the NBER Patent Data File (Hall et al., 2001). Along with the different reference period, the comparison with EPO applications is limited by the fact that, here, we are dealing with patent grants which are less likely to be obtained by occasional applicants. In spite of that, also in the United States the majority of inventors contributed to only one patent in a time span of 37 years. Albeit such a finding has not been sufficiently emphasised in the literature, as a stylized fact, most of the inventors try to get (or get) a patent once in a lifetime.

Obviously, this does not mean that patenting output is dominated by small entities. On the contrary, patent applications and grants are extremely concentrated. Suffice it to say that, at the EPO, 80% of applications come from fewer than 20% of applicants and, in 2006, the ten most active companies account for 12% of total applications (European Patent Office, 2007). But, how patents are distributed among inventors working in large companies? A seminal contribution in this direction was provided by Narin and Breitzman (1995) who found, also for these inventors, a remarkably skewed distribution ruled by a Lotka's law.

Looking at chemistry publications, Lotka (1926) discovered an inverse square law of academic productivity: the number of people producing n papers is proportional to $1/n^2$, i.e. for every 100 researchers with one paper, there are $100/2^2$, with two papers, $100/3^2$ with three, and so on. Shifting the analysis from academic to industrial research, Narin and Breitzman fitted the above law into the patents' distribution of inventors. Such a parallelism is justified by the same concept of *power law*, i.e. a mathematical representation of a *natural law*, which simply states that most of the brightest ideas are the products of a few outstanding brains[4]. This may be due to different motives such as 'experience' and 'prestige', although the main reason rests on the normal distribution of 'intelligence' (Ernst et al., 2000). However, with the partial exception of experience, many of these explanatory factors can be hardly quantified. Moreover, while for academics the role of individual prestige can be very important, the same cannot be assumed for inventors because, to be patented, an invention must fulfil a set of 'objective' requirements established by the patent law. This is not to say that, especially nowadays, to publish an article in a highly ranked journal is easier than to have a patent granted by the USPTO or the EPO.

---

specialized (cf. Schettino and Sterlacchini, 2007).

[4] With a view to find an (albeit partial) explanation, de Solla Price (1976) developed a probabilistic model of academic publications – the so-called *square root of elitism* – in which, for a researcher, the probability to publish an article increases with the number of papers she/he has already published (a sort of learning by doing or *Matthew's effect*).

Thus, the idea pushed forward by Ernst et al. (2000) that the distribution of patenting output should be even more concentrated than that of scientific publications is hardly acceptable.

Narin and Breitzman (ibid.) examined how the patents held by four large companies active in semiconductor technology (namely, AT&T, Matsushita, Fuji and Xerox) were distributed among the inventors occupied by them. Looking at the inventors' productivity, they found that, in each company, the top 1% of them was from 5 to10 times more productive than the average. Moreover, they estimated the logarithmic transformation of the following exponential equation:

$$Y=C/X^{\alpha} \qquad (1)$$

where X is the number of patents, Y denotes the number of inventors with X patents, C is a constant and α is the coefficient that, according to Lotka's law, should be equal to 2. For each company, Narin and Breitzman's findings were consistent with the above prediction, suggesting that the overall patenting performances of large firms are strongly dependent on the ingenuity and creativity of a few key inventors. Accordingly, companies should make every effort to retain and nurture them as a sort of golden eggs' gooses.

The crucial role played by a narrow set of very productive inventors has been re-stated by Ernst at al. (2000) who, by estimating the Lotka's equation for the patents of 43 German companies, find that the α coefficient is very close to 2. However, they also find that patenting activity (measured by the raw number of patent applications) is more highly concentrated than patenting performance (measured by adjusting the number of applications for their quality). Accordingly, these authors contend that "patenting quality tends to increase when the total number of patent applications decreases" so that "a higher patent quality can compensate for an inventor's low patenting activity" (ibid., p. 190). Since not all the patented inventions have the same value, this casts some doubts on the effectiveness of the Lotka's law in capturing an important dimension of the inventive process.

Apart from the above two, a few studies (often based on single companies) have examined the distribution of patents among individual inventors. Moreover, the functioning of the Lotka's distribution has been almost exclusively tested at company level. However, because such a distribution should capture a *natural law*, it should work also in other contexts. For instance, some authors have considered the scientists of different countries and found that Lotka's law is suitable for representing the distribution of their per-capita publications.

In the remaining of this section we perform the same test on the inventors residing in Marche, an administrative region whose extension has not been established according to economic or social criteria. For other types of analysis the above feature could be a shortcoming while, for studying how the inventors' productivity is distributed, it represents an advantage because this region can be viewed as a randomly chosen geographical area.

To test whether the Lotka's law applies to the patent applications attributed to Marche we considered, along with the 570 inventors residing in the region, also 97 extra-regional inventors mentioned in the same applications. Indeed, ignoring them would have artificially increased the productivity of regional inventors measured in terms of 'fractional patent counts' (see below).

Table 2 – Number of inventors contributing to patent applications of Marche applicants

Whole patent counts		Fractional patent counts (rounded)	
Patent applications	Inventors	Patent applications	Inventors
1	480	1	271
2	95	2	72
3	40	3	18
4	20	4	12
5	8	5	3
6	5	6	4
7	3	7	3
8	10	8	8
9	2	9	1
10	3	10	4
11	2	11	1
12	1	12	1
13	1	13	1
14	2	14	2
15	2	15	1
16	1	24	1
20	1		
34	1		

We then carried out two separate analyses for 'whole patent counts' and 'fractional patent counts' (cf. Table 2). In the first case, the whole invention is attributed to each inventor whose name appears on a patent application. In the second, instead, we work with fractional data by giving to each co-inventor an equal portion of the patent. It should be stressed that, in order for Lotka's inverse square law to make sense, fractional count have been rounded.

By using the data reported in Table 2, we run two estimations of the log-linear version of equation 1 ($lnY = c - \alpha\ lnX$) with 18 and 16 punctual observations for, respectively, whole and fractional patent counts. Table 3 shows that, in both cases, the estimated α is very close to that predicted by Lotka and the Wald test does not refuse the hypothesis that such a coefficient is equal to -2. It should be added that α turned out to be different from this value when, rather than among inventors, we run the regression across applicants. This confirms that the Lotka's law applies to the inventive performance of individuals and not to that of the organizations where they are employed.

Table 3 - Estimates of Lotka exponent

	α		Wald test (α = -2)	
	Estimate	Prob>| t |	F	Prob>F
Whole patent counts	–1.945	0.000	0.12	0.736
Fractional patent counts	–1.899	0.000	0.27	0.613

The above findings provide additional support to the *scale invariance* of the Lotka's distribution which appears to fit quite well into different contexts. Here, we have shown that this law can work also for the inventors contributing to the patent applications attributed to a given geographical area.

As already said, the above framework depicts the inventive activity as a sort of *natural phenomenon*. Although fascinating, it cannot be denied that such an explanation is not fully satisfactory for, at least, three reasons. First, contrary to that of academics, the ability of inventors can be partly captured by some individual features (such as the level of education and the recourse to more of less sophisticated sources of knowledge). Second, compared to academic publishing, patenting is a very expensive activity so that the characteristics of the organisations in which the inventors are occupied may exert a significant impact on their productivity. Third, as already mentioned, the Lotka's law performs well for patent counts while it may be less effective when patent quality is taken account (Ernst at al., 2000).

Accordingly, in what follows, we also explore the qualitative dimension of the inventive process.

3. Inventors' features and patent quality indicators

To assess whether some individual and organisational features are effective in explaining the different productivity of inventors, as well as the different quality of their inventions, we carried a survey by means of telephone interviews (see for details Schettino and Sterlacchini, 2007). The sample of Marche inventors who took part to the survey is composed of 106 individuals accounting for 19% of the original population of inventors and 42% of the patent applications attributed to regional applicant. As it emerges by comparing the two percentages, the sample was built with a view to over-represent the more productive inventors and, then, obtain an adequate variability of the phenomenon under investigation. This implies that, in terms of per-capita patents, the sample distribution is less skewed than that concerned with the population and, accordingly, the Lotka's law does not fit well into the sample.

For each inventor we obtained information on a number of personal features such as age, gender, level of education, professional position. As in the PatVal-EU survey (cf. Giuri et al., 2007), we also asked the inventors to assess the importance of the different sources of knowledge used in the inventive process (such as patent and scientific literature). Moreover, we distinguished independent inventors (13% of the sample) from firms' owners (30%) and dependent workers (57%). For the last two groups we got information on the size of the firms they were working in.

From the patent documents we were able to infer whether the inventors were working in teams or not. The presence of a team of inventors was identified when there were, at least, two or three individuals who jointly contributed to, respectively, three or two patent applications. By applying this criterion, we found that 24% of the Marche inventors worked in teams. It must be stressed that, in our sample, we did not find networks of inventors employed in different organisations, so that all the identified teams belong to specific firms.

Aside from inventive teams, the patent documents were mainly used to extract some information on patent quality (or value). In this connection, we collected data on family size, claims and forward and backward citations[5], whose descriptive statistics - referring to the 314 patent applications contained in our sample - are reported in Table 4.

Table 4 – Descriptive statistics of patent quality indicators: EPO applications of Marche inventors

	Mean	St. Dev.	Min	Max
Family size	7.46	5.39	1	34
Claims	12.93	9.79	1	58
Forward citations (adjusted)	0.97	1.77	0	8
Backward citations	3.27	2.70	0	15
Neutral backward citations	1.89	2.01	0	12
Bad backward citations	1.38	1.81	0	11

The extent of family size is given by the number of patents or applications – descending from the original one - obtained or filed by the applicant in different countries. As stressed by Putnam (1996), one of the first scholars using this indicator of patent quality, family size should reflect both the technological and economic importance[6] of the invention, simply because the overall cost of patenting rises with the number of countries in which a legal protection is sought. Lanjouw and Shankerman (2004) observe that more than two-thirds of US patentees do not seek protection outside their home markets and find that family size plays a relevant role for assessing the quality of computer patents only. The situation is quite different for European inventors and the reason is that the centralised procedure managed by the EPO becomes convenient when patent protection is sought for, at least, four countries adhering to the European Patent Convention. It is then not surprising that, in our sample, the average size

[5] Along with those employed in the present study, other indicators of patent quality have extensively been used in the empirical literature. Earlier works have employed data on patent renewals (Shankerman and Pakes, 1986) and patent scope (Lerner, 1994). The enduring payment of renewal fees signals that the patent has maintained some value over time. The patent scope, i.e. the number of technological classes to which it has been ascribed, is a proxy for the technological breadth of the invention.

[6] To a different extent, the indicators discussed in this section are good predictors of the economic value of the inventions. For a better assessment of the latter, however, direct suvey are needed (such as the recent PatVal-EU survey on European inventors; cf. Giuri et al., 2007). Harhoff et al. (2003) find that the economic value of patents (attributed by patent holders) is significantly correlated to a large set of quality indexes obtained from patent documents.

of patent families is greater than 7 (and that the patent applications with a family size lower than 4 account for only 12% of the total).

The claims included in the patent application identify the specific property rights that the patent should protect and their number approximates the 'scope' or 'width' of the invention (Hall et al., 2001). However, although the claims may refer to the both essential and secondary novelties embodied in the invention, such a distinction is not provided in the patent documents. Also for this reason, the number of claims per patent can be very high and, as Archontopoulus et al. (2007) have documented by observing EPO applications over 1980-2004, it has increased from 10 to 21. In our sample of patent applications the average number of claims is close to 13. On the one hand, it cannot be denied that applicants (often following the advice of patent attorneys) are prone to claim as much as possible in their applications. This casts some doubts on the effectiveness of claims to capture the inventions' quality. On the other hand, however, it must be taken into account that an excessive amount of claims may compromise the positive outcome of the examination process and, in any case, makes it more costly and time consuming.

The number of forward citations received by a patent (or patent application) is particularly linked to the technological rather than economic importance of the innovation. In fact, the more a patent is cited, the wider is the innovation basis it provides to future inventions (Harhoff et al., 1999; Hall et al, 2001). At the same time, however, the invention might not receive any monetary rewards for it.

In attributing forward citations to a patent one has to address the problem of "time truncation", i.e. the fact that more recent patents have, by definition, a lower probability to be cited. A solution proposed in the literature is that of considering patents with, at least, a window of five years from their application date (Lanjouw and Shankerman, 2004; Mariani and Romanelli, 2007)[7]. In our case, this would have meant to neglect all the patent applications filed after 2001. In order to preserve all the available observations, we adopted instead the "fixed-effects" approach proposed by Hall et al. (2001). Accordingly, the number of citations received by a patent in a given year is divided by the average number of citations received, in the same year, by all the patents of Marche inventors. Nonetheless, such a procedure gives rise to a staggering

[7] Hall et al. (2001) show that 60% of the citations recorded by the USPTO patents granted in 1990 is received after five years.

increase of citations for the few very recent applications that have been already cited, artificially over-stating the average quality of the corresponding inventors. To avoid it, we attributed a maximum of 8 citations to the most recent applications, which corresponds to the maximum number of actual (i.e. not adjusted) citations received in our entire sample. On the basis of this conservative procedure, in our sample the average number of (adjusted) forward citations per patent is slightly below one, with about one fourth of applications without any citations. As for the number of patents per inventor (cf. section 2), this finding is not exceptional. Schettino and Sterlacchini (2007) show that, in terms of forward citations per patent, the population of Marche inventors is not too different from that concerned with the whole Europe (0.20 versus 0.28 citations after three years)[8].

Finally, we also considered backward citations. The latter are important to trace the knowledge flows between patented inventions (Hall et al, 2001) also with a view to measure technological spillovers among industries and geographical areas (Criscuolo and Verspagen, 2006). However, the fact that backward citations are usually significantly correlated with forward citations does not imply that also the extent of the former can be viewed as a good proxy of patent quality. A large number of backward citations, in fact, could be a signal that the patent has a derivative nature and, as such, its novelty and inventive step could be questionable. This problem is particularly severe for EPO applications in which almost 90% of citations are added by the patent examiners rather than the applicants (cf. Criscuolo and Verspagen, 2006)[9]. For the patents issued by the USPTO the same share, though relevant, is significantly lower (63% according to Alcacér and Gittelman, 2005).

[8] This occurs when controlling for the technological distribution of patents which, in the Marche region, is strongly concentrated in the field of Mechanical engineering while being, in the EU15, much more homogenous among the different technological fields.

[9] "Under USPTO rules the inventor and his/her attorney are obliged to provide a (complete) list of reference describing the state of the art which are considered relevant to the patentability of the invention – the so called 'duty of candour'; the EPO has no similar requirements […]. As a result, in the EPO system, examiners tend to be responsible for the majority of patent citations." (Criscuolo and Verspagen, 2006, p. 4).

Table 5 – Backward citations by type

	Total EPO applications (1985-2000)*	EPO applications from Danish inventors (1978-1997)**	EPO applications from Marche inventors (1991-2005)***
Share of type X citations[(a)]	0.20	0.20	0.25
Share of type Y citations[(b)]	0.16	0.13	0.17
Share of bad backward citations (type X + type Y)	0.36	0.33	0.42
Share of neutral backward citations[(c)]	0.64	0.67	0.58

(a)= Citations added by EPO examiners concerned with particularly relevant documents if taken alone. The citations included into this category may seriously undermine the novelty and inventive step of the citing application.
(b)= Citations added by EPO examiners concerned with particularly relevant documents if combined with others.
(c)= Citations added by either applicants or EPO examiners which do not undermine the novelty and inventive step of the citing application.
* Source: Criscuolo and Verspagen (2006). **Source: Schneider (2007). ***Source: own computations from OECD/EPO patent citations database.

The different types of backward citations attached to EPO applications can be obtained from the OECD/EPO patent citations database. The most problematic citations added by EPO examiners - which can put at risk the granting of the patent[10] - are those conventionally classified with the X and Y codes (see Table 5) while other types of citations (including those originally inserted by the applicants) do not generate similar prejudices. Summing up types X and Y one can compute what we have termed 'bad' backward citations while all the other types can be grouped under the label of 'neutral' backward citations. Table 5 shows that, in our sample, the former account for 42% of total backward cites. This share is higher than that recorded, on average, by all the EPO applications (Criscuolo and Verspagen, 2006) and by those coming from Denmark (Schneider, 2007).

4. A composite index of patent quality

Since any quality indicator is likely to convey different pieces of information on the intrinsic value of patented inventions, Lanjouw and Shankerman (2004) have proposed a composite index able to maximise the information content embodied into the distinct measures

[10] Lazaridis and Van Pottelsbeghe de la Potterie (2007) contend that the interventions of EPO examiners are among the most important determinants of patent withdrawals (this occurs to about one third of total applications filed at the EPO). Accordingly, the report on prior art predisposed by examiners - from which the type of backward citations can be identified - may have the function of discouraging some of the applicants in continuing the examination process.

of patent quality discussed in the previous section. Accordingly, we estimate the following multiple-indicator model with one latent common factor:

$$y_{ki}=\mu_{ki}+\beta'\mathbf{X}_i+\lambda_k q_i+e_{ki} \qquad (2)$$

y_{ki} is the value in logs of the k indicator concerned with the ith patent. As quality measures we considered patent families, claims, forward citations and backward citations: the latter, however, are broken down into 'bad' cites which, according to the previous discussion, should signal less valuable inventions and 'neutral' cites whose impact on patent quality cannot be established a priori. $\mathbf{X}_i$ is a set of characteristics observed for each patent. q_i is a latent factor, assumed to be distributed as a standard normal, representing the unobserved features common to all the quality indicators. The loading factor λ_k identifies the degree of correlation between the original indicators and the common factor. Finally, e_{ki} denotes an idiosyncratic component. The main assumption of the above model is that the variability of each quality measure is jointly generated by those of the common factor and the residual disturbances.

As in Hall et al. (2007), we estimated q_i through a two-step procedure. First, we built a system where each indicator of patent value was regressed on two observable characteristics, i.e. the year of application and the main IPC technology class, using the seemingly unrelated regression. In the second step, we extracted the common factor from the residuals of such auxiliary regression by means of maximum likelihood. It must be stressed that the latent factor model was estimated on the full set of patent applications (314 observations). Subsequently, we computed the mean score attributable to each inventor.

In building the common quality factor, we first examined the correlation existing among the separate indicators. The top panel of Table 6 illustrates two clear points. On the one hand, there is a weak statistical association among patent families, claims and forward citations. On the other, 'neutral' and, to a lower extent, 'bad' backward citations are positively correlated only with forward citations while a negative association arises between the former and the number of claims and between the latter and family size. As a consequence, the inclusion of both kinds of backward citations in the latent factor model will probably polarize the variance of the composite index.

Table 6 - Correlation matrixes between patent quality indicators

A. Original variables (in logs)

	Family size	Claims	Forward citations	Neutral back cit.	Bad back cit.
Family size	1.000				
Claims	0.058	1.000			
Forward citations	0.069	-0.026	1.000		
Neutral backward citations	-0.048	-0.127**	0.332**	1.000	
Bad backward citations	-0.233**	-0.017	0.156**	0.101*	1.000

B. Residuals of the multiple-equation system

	Family size	Claims	Forward citations	Neutral back cit.	Bad back cit.
Family size	1.000				
Claims	0.106*	1.000			
Forward citations	0.038	0.027	1.000		
Neutral backward citations	-0.183**	-0.074	0.272**	1.000	
Bad backward citations	-0.218**	-0.018	0.162**	0.120**	1.000

*= Significant at 0.1 level of confidence. **= Significant at 0.05 level of confidence.

Table 7 – Composite value indexes: correlation with the original variables

	Value index 1	*Value index 2*	*Value index 3*
Family size	0.386	-0.183	-0.218
Claims	0.275	-0.074	0.162
Forward citations	0.099	0.272	-0.018
Neutral backward citations		1.000	
Bad backward citations			1.000

To account for such a problem, we decided to estimate three common quality factors. The first, as in Lanjouw and Shankerman (2004), employs patent families, claims and forward citations only (*Value index1*). The other two factors also include, along with the previous indicators, 'neutral' backward (*Value index2*) and 'bad' backward citations (*Value index3*). The bottom panel of Table 6 reports the correlation among the residuals of the first-step regression. It describes the statistical relations that remain in the original indicators once the effects exerted by the application year and the main technological class are removed. In this case, both kinds of backward citations appear to be negatively associated with family size.

The exact correlation between the composite quality indexes and the first-step residuals is provided by the loading factors displayed in Table 7 which clearly shows that, in our sample, backward citations are inadequate proxies of the value of patented inventions. In fact, the second and third composite indexes are negatively associated with some of the original quality measures and, as such, can be hardly interpreted as proxies of patent value. Instead, *Value index1* is positively correlated with all the original measures, although particularly with family size and claims and, to a lesser extent, with forward citations.

It should be stressed that our first and preferred indicator of patent quality is identical to that employed by Lanjouw and Shankerman (2004) in their econometric analysis of research productivity and firms' market value. Originally, in computing a composite index, these authors also used backward citations but, subsequently, they considered only forward citations, claims and patent families. Mariani and Romanelli (2007) obtain a synthetic index of patent quality and, as in our study, attach it to a sample of 793 European inventors. Along with claims, they make use of both forward and backward citations but neglect patent families which, in the European context, should be quite effective in signalling the value of patented inventions (cf. section 3). Moreover, looking at their estimate of the common quality factor it emerges that the latter is mainly influenced by forward citations and claims while backward citations play a little role. Instead, as we have clearly shown, in our sample backward citations are not only weakly correlated with the other quality measures but those classified as 'bad' can be negatively associated with patent quality.

5. Exploring the determinants of inventive productivity and quality

This section investigates econometrically the determinants of the inventors' productivity, measured both in quantitative and qualitative terms. We first assess what are the main explanatory factors of the two measures of inventive performance taken separately. In a second step, we check whether there is a mutual self-enforcing effect between the two productivity indicators, i.e. if the average number of applications affects patent quality (and viceversa).

For the first step, the equations to be estimated are the following:

$$NPAT_i = \alpha_0 + \alpha' \mathbf{X}_i + e_i \qquad (3)$$

$$AVQ_i = \beta_0 + \beta' \mathbf{X}_i + u_i \qquad (4)$$

NPAT is the number of patent applications, in logs, of the *i*th inventor, taken as a quantitative indicator of her/his inventive productivity. *AVQ* is the average level of patent quality attributable to each inventor on the basis of the score obtained by her/his patent applications. Such a score is the value of the composite index of patent quality extracted from the number of forward citations, claims and patent family (see, in the previous section, *Value index 1*). The vector $\mathbf{X}_i$ includes a set of explanatory variables, common to both specifications, representing some key characteristics of either the inventor or the firm where she/he is employed.

With respect to the firms' features we considered two dummy variables for large and medium-sized firms (with, respectively, more than 250 and between 50 and 249 employees) and the extent of patent portfolios measured by the log of patent applications that the firm (or, in case of an independent inventor, the individual) filed at the EPO during the examined period. Among the individual features, there are two dummy variables for whether the inventor is a male and possesses a tertiary level of education (university degree or PhD). Moreover, as a proxy of experience, we included the log of the inventor's age, both in linear and quadratic form. Somewhat in between individual and firm characteristics, we also inserted a dummy variable for the inventors who worked in teams. Moreover, we considered whether the inventor ascribed a great importance (identified by a score of 5 or 4 in a Likert scale ranging from 1 "not important" to 5 "very important) to the use of two sources of codified knowledge: patent literature (i.e. the search in patent databases) and scientific literature. Finally, as further controls, we included in both equations seven dummy variables for the main IPC technological classes in which the patent applications of the inventors are classified[11].

[11] The technological classes that have been considered are the following: Electrical-electronic technologies; Instruments; Chemicals, Drugs & Biotechnology; Material processing; Thermal processes; Other process engineering; Consumer goods; Mechanical engineering. The first three labels correspond to the first three of the 5 macro-technological fields of the IPC classifications. The remaining five labels derive from a re-combination of

The equations (3) and (4) have been estimated as a simultaneous system through the Seemingly Unrelated Regression (SUR) method by making a small sample adjustment for computing the covariance matrix of the residuals. It should be stressed that, in our first estimation, the set of explanatory variable is the same for the two equations so that the results are equivalent to those provided by OLS. Table 8 reports the regression results apart from those concerned with the dummies for technological fields which are omitted for the sake of brevity.

Table 8 - Determinants of inventors' productivity and patent value: SUR estimates.

	(1) Ln Number of patent	System (1)-(2) (2) Composite value index	System (1)-(3) (3) Composite value index
Ln Number of patents			-0.020
			(0.077)
Large firm	0.317*	-0.127	-0.121
	(0.191)	(0.139)	(0.141)
Medium-sized firm	0.110	0.053	0.055
	(0.133)	(0.096)	(0.097)
Ln Patent portfolio	0.145*	0.000	0.003
	(0.080)	(0.058)	(0.059)
Team	0.432**	0.291**	0.300**
	(0.162)	(0.117)	(0.122)
Ln Age	9.328	11.843**	12.029**
	(7.019)	(5.088)	(5.152)
$(\text{Ln Age})^2$	-1.172	-1.519**	-1.542**
	(0.895)	(0.648)	(0.656)
Higher education	-0.130	0.174**	0.172**
	(0.116)	(0.084)	(0.085)
Gender (male)	0.014	0.324**	0.324**
	(0.216)	(0.157)	(0.157)
Patent literature	0.192	0.054	0.058
	(0.131)	(0.095)	(0.096)
Scientific literature interacted with medium-sized firm	-0.116	-0.298**	-0.301**
	(0.176)	(0.128)	(0.128)
Constant	-18.077	-23.629**	-23.990**
	(13.751)	(9.968)	(10.090)
Log-likelihood		-89.59	-89.59
AIC		251.2	253.2

Notes: Each specification includes a set of dummies for IPC technological classes. Standard errors are in brackets. *= Significant at 0.1 level of confidence. **= Significant at 0.05 level of confidence.

the 30 technological fields in which the IPC classification is usually broken down.

Starting from the inventors' productivity (column 1), the presence of an inventive team within the firms exerts a positive and very significant impact. In addition, the number of patents per inventor is higher for larger firms and rises with the size of their patent portfolios. Both coefficients, however, are significant only at a 10% level of confidence and this is probably due to the fact that the two variables are highly correlated (in fact, the Spearman's correlation coefficient is 0.698). Since none of the individual features considered influences the dependent variable in a significant manner[12], the inventors' productivity appears to be explained by organizational characteristics only.

Moving to the quality of the inventors' patent applications (column 2), the size of the firms and that of their patent portfolios are no longer significant. This suggests that the inventors working in small entities or independently do not seek international protection for less valuable inventions. This is not surprising because, due to the difficulties and costs of European patenting, also 'small applicants' must be selective in deciding what invention is worth to be patented.

Although they tend to be more present in larger firms with wider patent portfolios, inventive teams exert an autonomous, positive and significant impact on patent quality. This suggests that it is not the size of the firms or their propensity to patent but the size of their inventive activities (proxied by the number of people who are formally recognised as 'inventors') which allow them to attain more valuable inventions. Thus, the presence of an inventive team is effective in rising both the quantity and quality of inventions. The latter, however, is also enhanced by individual features. Male inventors and those with higher education contribute to more valuable patent applications. Also experience matters, as witnessed by the fact that the patent quality rises with the age of inventors. However, the relationship is inverted U-shaped so that the oldest inventors have no advantages over the younger ones. Finally, while the importance attributed to patent literature is not significant, a negative impact on patent quality is found when the inventors working in medium-sized firms ascribe a great importance to scientific literature. This finding is probably due the peculiar occupational position of the few inventors who have attained a PhD.

[12] Almost identical results arise when equation (3) is estimated by means of a negative binomial regression (which is particularly adequate for analysing count data like patents).

In a second step we also checked whether there is cross-fertilization between the quantitative and qualitative performance of each inventor. For this purpose, we first inserted *NPAT* among the explanatory variables of *AVQ* (cf. equations 3 and 4) and, then, estimated the system by means of the SUR method. The third column of Table 8 shows that the patent quality attributed to each inventor is not significantly associated with the number of patents per inventor[13], while all the effects already stressed in commenting the previous estimation are confirmed. We have also tested the opposite relation, i.e. whether the patent quality may spur the amount of patent applications per inventor. However, when inserted into the *NPAT* equation the composite index of patent quality turns out to be not significant and the results reported in column 1 of Table 8 are confirmed[14].

Our results are not consistent with those provided by Mariani and Romanelli (2007) for a sample of European inventors. These authors show that patent quality is a sort of purely random phenomenon being not significantly influenced by a set of explanatory factors similar to those employed in the present study. Moreover, in contrast to the same study, we do not find that a larger number of patents increases the probability of developing more valuable inventions. The above discrepancies, however, are not surprising. Our sample, although slightly biased towards more productive inventors, is dominated by 'small applicants': only 36% of the inventors took part to more than two patent applications over a period of 15 years (1991-2005). The sample used by Mariani and Romanelli (2007), instead, over-represents the European inventors working in large firms as witnessed by the fact that 56% of them contributed to more than two patents in a time span of 11 years (1988-1998). Accordingly, by considering a larger share of inventors working in small firms or independently, our study shows that their patent quality is significantly affected by some individual features.

[13]This precludes the utilization of three-stage least squares in order to assess the consistency of the *NPAT* coefficient in the second equation.
[14] These results are not reported for the sake of synthesis but are available from the authors upon request.

6. Summary and concluding remarks

As we have documented in section 2, over a period of 15 years, the large majority of European inventors contributed to only one patent application at the EPO. On the other hand, 80% of patent applications at the EPO are coming from 20% of applicants and the ten most active companies account for 12% of total applications. Moreover, from the studies reviewed in the same section, we know that a disproportionate share of the numerous patents held by large companies is coming from a few key inventors. Combining these pieces of evidence, it is not surprising that the distribution of patents among inventors is extremely skewed and consistent with that identified by Lotka in 1926 for scientific publications.

In this paper we have considered the population of inventors residing in an Italian region (the Marche). Its borders were established long time ago according to historical, cultural and, above all, administrative criteria so that, for the purpose of our study, the Marche can be viewed as a randomly chosen geographical area. We found that also in this area the distribution of patent applications among inventors is consistent with the 'inverse square law' introduced by Lotka.

Having established that, the next step has been that of looking for some explanations of the above phenomenon. In effect, the extent of patenting activities among inventors is likely to be significantly influenced by some measurable characteristics which may refer either to the same inventors or the organizations they are working for. Moreover, an extremely skewed distribution emerges by looking at patent counts while the results could be different when patents are weighted according to their technological and economic value. As a consequence, the search for some possible determinants of the inventors' productivity and patent quality is worth to be pursued.

By using the findings of a survey concerned with a sample of Marche inventors, this paper has provided an attempt in this direction. Starting from the quantitative dimension, we found that the number of patent applications per inventor does not depend upon any individual characteristics but it is positively influenced by organizational variables such as the presence of an inventive team, the firms' size and the extent of their patent portfolios. The average quality of the patent applications of each inventor has been approximated by a composite index extracted from the numbers of forward citations, claims and patent families. Backward citations are not included in the above index because, in our sample, they are negatively correlated with the other quality measures and, accordingly, are likely to be imperfect proxies of patent value. The regression analysis across inventors show that patent quality is positively affected by different individual features such as age, level of education and gender while, among the organizational characteristics, only the presence of an inventive team exerts a significant positive impact. The fact that patent quantity and quality are rather independent phenomena is confirmed by a further estimate showing that the inventors' productivity does not significantly influence the average quality of their inventions.

In conclusion, the skewed distribution of inventive productivity is not only a product of a *natural law* but also depends upon the size of the firms the inventors are working in, as well as the extent of collective work undertaken to attain an invention. Along with quantity, the working within an inventive team also increases the quality of patent applications. The latter, however, is also enhanced by the inventors' experience and level of education. This helps explain why occasional inventors generate inventions that are as valuable as those coming from the most prolific ones.

References

Alcacér, J. and M. Gittelman (2005) Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations, *Review of Economics and Statistics*, 88(4): 774-779.

Archantopolous, E., D. Guellec, B. van Pottelsberge de la Potterie and N. van Zeebroeck (2007) When Small is Beautiful: Measuring the Evolution and Consequences of the Voluminosity of Patent Applications at the EPO, *Information Economics and Policy*, 19: 103-132.

Criscuolo, P. and B. Verspagen (2006) *Does it matter where patent citations come from? Inventor versus examiners citations in European patents*, revised versions (July 2006) of the ECIS Working Papers 05.06, Eindhoven Centre for Innovation Studies, Eindhoven University of Technology-

de Solla Price, D. (1976) A General Theory of Bibliometric and Other Cumulative Advantage Process, *Journal of the American Society for Information Science,* 27:292-306.

Ernst, H., C. Leptien and J. Vitt (2000) Inventors are not alike: the distribution of patenting output and industrial R&D personnel, *IEEE Transaction on engineering management*, 47:184-199.

European Patent Office (2007) *Top applicants in 2006*, www.epo.org/ about-us/ office/ statistics/ top-applicants-2006.html

Giuri, P. et al. (2007) Inventors and Invention Processes in Europe: Results from the PatVal-EU Survey, *Research Policy*, 36:1107-1127

Hall B., A. Jaffe and M. Trajtenberg (2001) *The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools*, NBER Working Paper 8498.

Hall, B., G. Toma and S. Torrisi (2007) *The Market Value of Patents and R&D: Evidence from European Firms*, NBER Working Paper 13426.

Harhoff, D., F. Scherer and K. Vopel (2003) Citations, family size, opposition and the value of patent rights, *Research Policy*, 32:1343-1363.

Harhoff, D., F. Narin, F. Scherer and K. Vopel (1999) Citation Frequency and the Value of Patented Inventions, *The Review of Economics and Statistics*, 81: 511-515.

Lanjouw, J.O. and M. Schankerman (2004) Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators, *Economic Journal*, 114:441-465.

Lazaridis G. and B. Van Pottelsberghe de la Potterie (2007) The Rigour of EPO's Patentability Criteria: An Insight into the "Induced Withdrawals", *World Patent Information*, 29: 317-326.

Lerner, J. (1994) The Importance of Patent Scope: An Empirical Analysis, *RAND Journal of Economics*, 25: 319-333.

Lotka A. J. (1926) The Frequency Distribution of Scientific Productivity *Journal of the Washington Academy of Sciences*, 16:317-323.

Mariani, M. and M. Romanelli (2007) 'Stacking' and 'Picking' Inventions: The Patenting Behavior of European Inventors, *Research Policy*, 36:1128-1142.

Narin, F. and A. Breitzman (1995) Inventive Productivity, *Research Policy*, 24: 507-519

Putnam, J. (1996) *The value of international patent rights*, Ph.D. Thesis, Yale University, Yale.

Schankerman, M. and A. Pakes (1986) Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period, *Economic Journal*, 96: 1052-1077.

Schettino, F. and Sterlacchini A. (2007) *European Patenting and the Size of Inventors,* Quaderno di Ricerca n. 308, Dipartimento di Economia, Università Politecnica delle Marche.

Schneider, C. (2007) *The Determinants of Patent Applications Outcomes – Does Experience Matter?,* MPRA Paper No. 3359.

Trajtenberg, M. (2005) *The mobility of inventors and the productivity of research*, Presentation made at the NBER Summer Institute Conference on "Economics of Intellectual Property" , July 19, 2005; www.nber.org/confr/2005/si2005/prlprg.html.

Webb, C., H. Dernis, D. Harhoff and K. Hoisl (2005) *Analysing European and International Patent Citations: A Set of EPO Patent Database Building Blocks*, OCDE Science, Technology and Industry Working Paper 2005/9, OCED Publishing.