

# CONCORDANCE AND COMPLEMENTARITY IN IP INSTRUMENTS\*

Marco Grazzi<sup>1</sup>, Chiara Piccardo<sup>2</sup>, and Cecilia Vergari<sup>3</sup>

<sup>1</sup>Department of Economic Policy, Università Cattolica del Sacro Cuore

<sup>2</sup>Department of Economics, University of Verona

<sup>3</sup>Department of Economics, University of Bologna

November 30, 2018

## Abstract

This work investigates the relationship between proxies of innovation activities, such as patents and trademarks, and firm performance in terms of revenues and growth. By resorting to the virtual universe of Italian manufacturing firms we provide a rather complete picture of the innovation activities of Italian firms, in terms of patents and trademarks, and we study whether the two instruments for protecting Intellectual Property (IP) exhibit complementarity or substitutability. In addition, and to our knowledge novel, we propose a measure of concordance (or proximity) between the patents and trademarks owned by the same firm and we then investigate whether such concordance appears to exert any effect on performance.

**Keywords:** Trademarks, Patents, Innovation, Intellectual Property, Complementarity, Concordance, Technological proximity, firm performance, firm growth

---

\*We thank the participants the Department of Economics Unibo internal seminar (June 2018), the European Commission Joint Research Center seminar (JRC-Seville, June 2018), the European Association for Evolutionary Political Economy (EAEPE) conference (September 2018) and the IP Statistics for Decision Makers (IPSDM) conference in Alicante (October 2018). We are grateful to Roberto Susanna (Infocamere press office) and to Bureau van Dijk for their technical assistance at various stage. We are also indebted to Hanna Hottenrott, Laura Magazzini, Arianna Martinelli, Daniele Moschella, Gabriele Pellegrino, Emanuele Pugliese, Enrico Santarelli, Antonio Vezzani and Nikolas J. Zolas for insightful comments. We gratefully acknowledge comments by two anonymous reviewers. All remaining errors are our own. This work received fundings by the Fondazione Cassa di Risparmio of Forlì, within the research project ‘Organizzazione d’impresa, innovazione e commercio internazionale’, grant ORGANIMPRE.

# 1 Introduction and related literature

This work investigates the effect of patents and trademarks on firm performance. To do this, we employ the virtual universe of Italian limited liability manufacturing firms to analyse the IP strategies of firms and we propose a novel measure to assess the concordance between the two instruments, when firms employ both.

The question of how innovation activity affects firms' performance has a long tradition in economics, with many recent contributions focusing on the analysis of *R&D* and patents, especially for large and medium firms. In particular, starting with Griliches (1981), many authors investigated the relationship between firm's market value and its *R&D* expenditures and number of patents, often focusing on large quoted companies (see e.g., Toivanen et al., 2002; Hall, 2000). Others, in addition to *R&D*, analyse the impact of product and process innovation based on survey data (see e.g., among the many others, Hall et al., 2009).

A recent wave of studies, starting with Mendonça et al. (2004), has suggested that, on top of patents, trademarks provide a useful proxy for firms' innovation activity. This is even more true for small and medium enterprises, SMEs henceforth, and for 'softer' types of (non-patentable, non-technological) innovation like service, marketing and organizational innovations as well as innovation activities closer to the market introduction stage (see e.g., Flikkema et al., 2015, 2014; Helmers and Rogers, 2010). Indeed, for instance, SMEs may not be in the position of affording other protection mechanisms, see for instance Block et al. (2015). Moreover, while patents have a finite duration, trademarks do not expire as long as all post registration maintenance documents are timely filed. A trademark can thus protect a product (or service) when a similarly aged patent would expire. Schautschick and Greenhalgh (2016) provide a detailed survey of the empirical works revealing significant correlations between innovation, patents and trademarks and testing for the suitability of trademarks as proxy for innovation.

In line with this recent strand of literature, we study how and to what extent, innovation activities, as proxied by patents and trademarks, affect firms' performance in terms of revenues and growth. We consider balance-sheet data matched with IP activity of the (virtual) universe of Italian manufacturing limited liability firms, independently of their size. This feature of the dataset allows us to contribute to the literature in several directions. First, we can provide a complete picture of patent and trademark activities of Italian firms and assess whether, for instance, there exists a size or sector-specific pattern. Second, we can dig into the effects of these innovation proxies, not only on large firms (as investigated, among others, by Castaldi and Dosso, 2018) but also on small and medium firms. Note that this is quite relevant as SMEs, which represent the vast majority of Italian firms, are more likely to trademark over patenting, relative to large firms. Moreover, we investigate whether there is some complementarity or substitutability among the two IP instruments. Finally, and to our knowledge completely novel, for firms reporting both patents and trademarks, we employ the available information to build a measure of concordance (or proximity) between the stocks of patents and trademarks owned by the same firm and we then investigate whether such concordance may have an impact on performance. This is a rather new domain, opened by the contemporaneous availability of firm level IP data and recent developments in Algorithmic Links with Probabilities (ALP) (see, among the others, Lybbert and Zolas, 2014). Although not perfect, the linkages established between patents and trademarks thanks to ALP enable to investigate a number of issues. In this work we consider the relative timing in the activities of patent-

ing and trademarking and the impact of the concordance measure on firm performance. In this respect, our prior, in line with the capability view of the firm, is that a higher coherence or overlap between the pool of knowledge, as proxied by patents, and the set of products sold, as proxied by trademarks, should be associated to better firm performance (see among the many others Teece et al., 1994; Markides and Williamson, 1996; Bryce and Winter, 2009).

Our findings reveal that both patents and trademarks do have a role in explaining firms performance, in terms of revenues and growth. In particular, firms' with both IP rights enjoy a higher premium than firms only owning patents or trademarks. Moreover, our findings suggest that the positive impact of having only trademarks is somewhat larger than the positive impact of owning patents alone. The data at our disposal do not allow to explore this result at greater detail, however we might conjecture that it is related to the very nature of trademarks as 'recognisable designations and symbols for goods and services' (Mendonça et al., 2004) hence playing a more crucial role in the process of marketing innovation *vis à vis* patents. Further, trademarks are also the most diffused instrument of IP protection for Italian firms. We have to notice however that when focusing on the degree of concordance between the stock of patents and trademarks, one faces the constraint, particularly binding for the Italian case, of restricting the sample to firms owning both IP instruments.

As said, following the seminal work of Mendonça et al. (2004), a number of contributions have focused on trademarks as an appropriate indicator of innovation output, in addition to the more standard proxy traditionally provided by patents. Due to data constraints, most of these works focus on large and medium firms that are publicly traded and find evidence of a positive effect of trademarks on firms' stock market value (see e.g., Greenhalgh and Rogers, 2006, 2012).

Interestingly, Llerena and Millot (2013) examine the separate and combined effects of patents and trademarks on the market value of a sample of French quoted companies. They develop a theoretical model and, based on supermodularity theory, they test whether the two types of IP are complementary or substitute. In line with their theoretical predictions, they find that results differ across sectors. In chemical and pharmaceutical industries, patents and trademarks tend to be complementary whereas in high-tech business sectors, where the crucial asset to be protected is the technology rather than the brand, they are found to be substitutes.

In general, there are few contributions encompassing also small and medium firms. An exception is Rogers et al. (2007) that cover UK small and medium firms over the period 2001-2004. They study the effect of the registration of a new trademark over the following years and find that, because of the investment needed to launch an innovation, the profits may initially decrease even when the new product is ultimately successful. In comparison with the other companies, trademark active firms are more concentrated in the lowest and highest quartiles of the profits distribution.

The rest of the paper is organized as follows. The description of the data and the definition of the variables of interest is reported in Section 2. We next present some non-parametric evidence based on univariate kernel density estimations and on the Fligner-Policello test of stochastic equality in Section 3. Section 4 illustrates our empirical strategy and discusses the empirical results. Section 5 concludes.

## 2 Data and variables of interest

The empirical analysis is based on AIDA (Analisi Informatizzata delle Aziende) and AMADEUS data, two datasets provided by Bureau van Dijk (BvD).

The AIDA dataset provides detailed information on Italian limited liability companies operating in both the manufacturing and services sectors which, in accordance to law, have to deposit their balance sheets to the Chamber of Commerce. The data provide financial and economic information for Italian firms, as well as a wide set of relevant indicators, including number of employees, incorporation year and sector of activity, among others.<sup>1</sup>

The AIDA dataset, due to its nature, covers virtually the universe of Italian limited liability firms independently of their size or age, thus representing an ideal set of data to study the dynamics of firms and industries. This is a noteworthy feature with respect to many previous works (see Fang et al. (2011), Chang et al. (2012), Dernis et al. (2015), Daiko et al. (2017) and Dosso and Vezzani (2017), among others) that, due to data constraints, could only focus on top corporate R&D investors or on firms operating in high-tech industries. The time period covered by our empirical analysis spans from 2006 to 2014 and we focus only on manufacturing firms.<sup>2</sup>

We integrate AIDA data with information on the stock of applied patents<sup>3</sup> and registered trademarks owned by firms in each relevant year. As for patents, AMADEUS provides some relevant information, including international patent classification (IPC), the application date and whether a patent has been granted or not, among others; while, for filed trademarks one can resort to NICE classification code, the filing date and information on their registration. In particular, information on the patents and trademarks classifications (IPC and NICE, respectively) are crucial in order to define a measure of economic concordance between the technological fields of the two measures of IP rights.

Before proceeding with some descriptive statistics on the dataset, we illustrate the strategy adopted to characterize firms' innovation activity and the degree of coherence between the domains of knowledge incorporated in the firms' stocks of patents and trademarks, respectively.

### 2.1 Patents, trademarks and a measure of IP instruments concordance

In order to obtain a suitable measure of firms' innovation activities as proxied by patents, we consider data on applied patents, irrespectively of the outcome of the application, that have been applied at the United States Patent and Trademark Office (USPTO), at

---

<sup>1</sup>We use two AIDA 'historic' disks in the release version of December 2015 and December 2016, respectively. For a detailed description of the procedures followed in building the dataset, refer to Grazzi et al. (2017).

<sup>2</sup>As suggested by the previous literature, see OHIM (2015); Schautschick and Greenhalgh (2016); Graham et al. (2018) among others, patenting activity is mainly observed within manufacturing sectors, while trademarks are generally used in the whole economy by firms independently from their sector of activity. Consistent evidences emerge using the AIDA-AMADEUS dataset. Consequently, we decided not to account for firms in service sectors because of the low number of firms with both IP instruments in these sectors. Indeed, only for firms with both patents and trademarks we can identify the measure of concordance.

<sup>3</sup>We have opted to focus on applied, rather than granted patents, to enlarge as much as possible the sample of observations, given the very small percentage of Italian firms with patents. Results on the smaller sample (granted patents) are also available upon request.

the European Patent Office (EPO), and/or at the Italian Patent and Trademark Office (IPTO).<sup>4</sup> The stock of patents for each firm, in each year, does not include patents applied more than 20 years before the year of interest. That is, if a patent was applied in 1991 by a firm, we include it in the patent count from 2006 to 2010, but not from 2011 onward. This choice allows us to forego patents which are too ‘old’ to adequately represent a valuable proxy of firms’ technological capabilities.<sup>5</sup> Note that the range of information on patents accessible through BvD is only a subset of what available through PATSTAT, see for instance De Rassenfosse et al. (2014). Also note that, as shown in Table 3, only a very small fraction of Italian firms, around 7%, holds patents, see also Malerba and Orsenigo (1999), Cefis and Orsenigo (2001) and the more recent Dosi et al. (2015). One known issue of the literature on patents is the matching of IP instruments to a given firm (owner). In this respect, note that we are employing patents and trademarks linked to firms as originally provided in BvD AMADEUS. An alternative attempt is provided, still on Italian firms, by Lotti and Marin (2013) on EPO applications which the authors match to a restricted sample of AIDA firms (the so called, BvD AIDA TOP). In the end, their effort results in 5485 patenting firms over the period 2000-2007. It is however not possible to directly compare the two final datasets. It is true that ours displays a higher number of firms with patents (8616 in 2006, see Table 3), but this is most likely due to employing the “full” AIDA version and, although to a lesser extent, to counting USPTO and IPTO patent applications.

With regard to trademarks, we focus on registered trademarks that have been filed at the United States Patent and Trademark Office (USPTO) or at the European Union Intellectual Property Office (EUIPO), formerly known as Office for Harmonization in the Internal Market (OHIM).<sup>6</sup> We identify the stock of registered trademarks for each firm, in each year, by considering trademarks applied before or in the year of interest and expiring after the year of analysis. Unfortunately, differently from patents, AMADEUS does not provide information for registered trademarks also at national level.<sup>7</sup>

We next define, at the firm level, a variable to measure the degree of concordance between the two classes of IP instruments, patents and trademarks. For this purpose, we need information on IPC codes associated to each patent and NICE code linked to each trademark.<sup>8</sup> In the AMADEUS dataset, a single patent can be associated with more

---

<sup>4</sup>Patents could be owned by more than one firm; in these cases we associated the patents to each owner, as suggested by the existing literature.

<sup>5</sup>Czarnitzki and Kraft (2004), Xu and Chiang (2005) and de Rassenfosse and Jaffe (2017), among others, have highlighted the importance to account for the decline in patents’ value during the life of patented inventions.

<sup>6</sup>As for patents, also trademarks can be jointly owned by more than one firm; in these cases we associated the trademarks to each owner, as suggested by the existing literature. Note however that the phenomenon of joint trademark ownership is much more limited with respect to patents.

<sup>7</sup>The lack of applications for trademarks at the national level is of course expected to reduce both the overall stock of trademarks in the country as well as the share of firms holding trademarks. It is however very difficult to find a reference to assess how large is the actual impact on our dataset. To the best of our knowledge, OHIM (2015) provides some guidance in terms of aggregate statistics even if it is at the EU level with no possibility to distinguish among countries. According to Table 8 (page 40) of the cited document, “38.1 per cent of all large companies and 8.6 per cent of all SMEs own trademarks”, either at national or EU. In our dataset, which includes USPTO and EUIPO trademarks, the corresponding percentages are respectively (see Table 4)  $57.9 \approx (679/1172)$  and  $12.4 \approx (3805+2653)/(44148+8016)$ . Hence, it would not appear that the lack of national patents greatly compromises the sample that we employ.

<sup>8</sup>The IPC is a classification for patents and utility models according to the different areas of technology to which they pertain. We use it at the 3-digit level. NICE is a 2-digit international classification of

than one IPC code, while for trademarks only a single NICE code is provided.

In order to build such concordance measure, we rely on a series of crosswalks linking different classifications of industries, products, IPC and NICE codes, for which we refer to the works of Lybbert and Zolas (2014), Zolas et al. (2017) and Goldschlag et al. (2016). As a preliminary step, we need to convert the IPC codes and the NICE codes to a common code. In particular, we associate the International Standard Industrial Classification (ISIC) to each IPC and to each NICE code, respectively, relying on the probabilistic algorithms developed by Lybbert and Zolas (2014) and Zolas et al. (2017). These algorithms provide a concordance of each IPC and NICE code to ISIC codes. Namely, for each 3-digit IPC code several 2-digit ISIC codes may be associated, each with a weight that identifies likelihood (or strength) of the linkage between that particular IPC code and that particular ISIC code. The sum of these weights is one. Similarly, each 2-digit NICE code is associated to several 2-digit ISIC codes, each with a weight that represents the probability of concordance between that particular NICE code and that particular ISIC code. The sum of these weights is then one.

We obtain the concordance measure between patents and trademarks following a three-step procedure. Note that this measure is provided for any firm with patents and trademarks and it might vary over time. *First*, for each firm ( $i$ ) and year ( $t$ ), we identify the sets of 3-digit IPC and 2-digit NICE codes associated to the stock of patents and trademarks, respectively, and we label  $N_L^{i,t}$  ( $N_K^{i,t}$ ) the number of IPC (NICE) codes associated to these sets.<sup>9</sup> In the *second* step, considering the ISIC codes and their probability weights,<sup>10</sup> we compute the overlapping coefficient,  $overlap_{L-K}$ , for each pair ( $L, K$ ) of IPC and NICE codes.<sup>11</sup> Formally:

$$overlap_{L-K} = \sum_{j=1}^{N_{LK}} \min\{p_L(isic_j), p_K(isic_j)\}, \quad (1)$$

where  $N_{LK}$  is the number of ISIC codes associated to each pair of 3-digit IPC and 2-digit NICE codes,  $p_L(isic_j)$  is the probability weight that identifies the likelihood of the linkage between IPC code  $L$  and ISIC code  $j$ , and  $p_K(isic_j)$  is the likelihood of the linkage between NICE code  $K$  and ISIC code  $j$ . The overlapping coefficient is higher than zero only when both the IPC and NICE codes are linked to at least one common ISIC code. Indeed, the probability weights in equation (1) are both positive only when an ISIC code is associated to both IPC and NICE codes. Contrarily, if an ISIC code is not linked to either IPC or NICE code, the probability weight is set equal to zero. In the *third* and final step, we compute, for each firm and year, the degree of concordance between (the stock of IPC and NICE codes associated to) the stock of patents and trademarks. In particular, as shown in equation (2), for each firm ( $i$ ) in each year ( $t$ ), we sum all the overlapping coefficients identified for each pair of IPC ( $L$ ) and NICE ( $K$ ) codes ( $overlap_{L-K}$ ). Moreover, we normalize our measure dividing it by the product between the number of elements in the

---

goods (codes from 1 to 34) and services (codes from 35 to 45) applied for the registration of trademarks that has been adopted with the Nice Agreement (1957).

<sup>9</sup>Clearly, we do not account for patents and trademarks without information on the IPC and NICE codes.

<sup>10</sup>For the methodology and weights refer to works of Lybbert and Zolas (2014), Zolas et al. (2017) and Goldschlag et al. (2016).

<sup>11</sup>The overlapping coefficient is a measure of agreement (or similarity) which refers to the area under two probability density functions simultaneously. Notice that the overlapping coefficient between an IPC and a NICE code is not firm or time specific.

set of 3-digit IPC codes ( $N_L^{i,t}$ ) and the number of elements in the set of 2-digit NICE codes ( $N_K^{i,t}$ ).

$$conc_{i,t} = \left( \sum_{L=1}^{N_L^{i,t}} \sum_{K=1}^{N_K^{i,t}} overlap_{L-K} \right) / (N_L^{i,t} * N_K^{i,t}) \quad (2)$$

The degree of concordance takes value zero if there are no common ISIC codes associated to the two sets of IPC and NICE codes, while strictly positive and higher values of  $conc_{i,t}$  suggest higher coherence between the stocks of patents and trademarks.

For illustrative purposes, consider the following example taken from our database for a firm (labeled 1) with four patents and nine trademarks in year 2006. As we can see from Table 1, the first patent (patent 1) has two IPC codes (A61 and H02), the second patent has one IPC code (C09) and both the third and fourth patents have one IPC code (A61). Thus, for firm 1 in year 2006 the set of 3-digit IPC codes includes the following three codes: A61, H02 and C09. Looking at the stock of trademarks, the first eight trademarks have the 2-digit NICE code 2, while the ninth trademark has the NICE code 35. Hence, for firm 1 in year 2006, the set of NICE codes includes the following two codes: 2 and 35. So that  $N_L^{1,2006} = 3$  and  $N_K^{1,2006} = 2$ . Based on the probabilistic algorithms by Lybbert and Zolas (2014) and Zolas et al. (2017), in Table 2 each IPC and NICE code is associated to the corresponding ISIC sectors (it can be more than one) with the relative probability weights. For example, we link the 3-digit IPC code A61 to the following ISIC codes 10, 20, 21, 32 and 36. For each ISIC code the algorithm provides a probability weight which identifies the likelihood of the linkage between the IPC code and each ISIC code. Similarly, for NICE code 2, we consider the following ISIC codes 2, 20, 25 and 41, each of them linked to a given probability weight.

In order to compute  $conc_{1,2006}$ , we first determine the overlapping coefficient for each of the six pairs of IPC and NICE codes (A61-2, A61-35, C09-2, C09-35, H02-2 and H02-35). For instance,  $overlap_{A61-2} = 0.0505$ , with  $N_{A61,2} = 8$ , that is the number of ISIC codes associated to the pair (A61,2).<sup>12</sup> Hence, the degree of concordance between the two IP instruments is given by:

$$conc_{1,2006} = (0.0505005 + 0 + 0.1203539 + 0 + 0 + 0) / (3 * 2) = 0.028475733. \quad (3)$$

(Table 1 and 2 around here)

## 2.2 Descriptive statistics

We next present some of the main highlights of the dataset that will be employed in the empirical analysis.

Table 3 reports the distribution of trademarks and patents for firms in the manufacturing sector. As expected on the basis of prior literature, the fraction of Italian firms owning IP instruments is very small and it displays different trends for patents and trademarks. Note indeed that the proportion of firms owning at least one registered

---

<sup>12</sup>A common 2-digit ISIC code is identified only for the pairs of IPC and NICE codes A61-2 and C09-2. In particular, both the IPC code A61 and the NICE code 2 are linked to the ISIC code 20; while, both the IPC code C09 and the NICE code 2 are linked to the ISIC code 20. Thus, the overlapping coefficients are different from zero only for these pairs on IPC and NICE codes (0.0505005 and 0.1203539 for pairs A61-2 and C09-2, respectively).

trademark increases during the period of investigation (from 4.062% of firms in 2006 to 6.857% in 2014); while it is not the same for the fraction of firms owning at least one patent (it is around 7% in each year).<sup>13</sup> Quite interestingly, in more recent years, the number of firms with trademarks exceeds the number of firms with patents. This points to the increasing importance of trademarks as instrument to protect IP, at least among Italian firms. However, also note, as shown by the last two columns of Table 3, that the total number of patents of Italian firms, even if decreasing in more recent years, is still bigger than the number of trademarks.<sup>14</sup> Thus we can conclude that, at least in Italy, ownership of patents, differently from trademarks, is concentrated in a narrower set of companies and that trademarks are becoming more diffused as instrument of IP protection. As suggested in Hall et al. (2013a), the concentration of patent ownership might be related to the fact that ‘(a) some firms do not automatically patent all of their patentable inventions, (b) some firms avoid the patent system altogether, either because of its cost or because patenting is perceived to yield no additional benefit, and (c) some innovations involve inventions that are not patentable.’ And, in addition, to the fact that there appears to be size threshold below which firms find it more difficult to afford the cost related to the functioning of a formal R&D laboratory. Moreover, the data reveal that around 7% of firms with patents are involved in co-ownership (*co-patents*) and the number of the co-owned patents has increased over the years (the share of patents with more than one owner rises from 2.694% in 2006 to 3.145% in 2014).<sup>15</sup>

(Table 3 around here)

Note that due to the small number of firms owning IP rights in general, when we restrict to the sub-sample of firms with *both* patents *and* trademarks, we are left only with around 2% of the firms making up the original dataset, column (5) of Table 3. The share of firms owning both patents and trademarks slightly increases over time, from 1.543% in 2006 to 2.190% in 2014. Focusing on the concordance between IP instruments, around 90% of firms reporting both patents and trademarks, column (6) of Table 3, displays a strictly positive value for the degree of concordance,  $conc_{i,t}$ , as defined in equation (2). Among these firms the average degree of concordance is around 0.18. The whole distribution is plotted in Figure 1 and it does not vary much over time because, at the firm level, the value of the concordance measure is quite stable over the years. In turn, this is due to the poor dynamics in patenting and trademarking, that is, very few firms are registering or applying for new trademarks or patents over the time span covered by the database; hence also the measure does not change much over time. Unfortunately, the co-occurrence of a very low number of firms with both IP and a poor underlying dynamics do not make the ‘Italian case’ the ideal testbed for the measure we are proposing here. Finally, note that the proposed concordance measure, at least for the sample under investigation, appears to be independent from the size of the firm, see Figure 2.

(Figures 1 and 2 around here)

---

<sup>13</sup>Such similar percentages of firms with patents and firms with trademarks might also be related to the absence of IPTO trademarks.

<sup>14</sup>A similar trend in the number of patents has been reported also by the Italian Observatory for patents, *Osservatorio Italiano Brevetti*, Osservatorio Italiano Brevetti (2014). Considering patents applied at the Italian Patent and Trademark Office (IPTO) from both Italian and foreign firms, the Observatory highlights a reduction in the number of applied patents starting from 2011.

<sup>15</sup>In the interest of space, we do not show descriptive statistics considering co-ownership of patents. They are available upon request.



In addition to the analysis of IP bundles of firms and of their internal coherence, it is also much relevant to investigate if any regularity emerges in terms of the timing of applications for patents *vis à vis* trademarks. In this respect, our prior would be that in a given process of innovation, moving from the original concept, to the technological development of the product and finally to market introduction, the patent application, being more technology related, pre-dates the trademark application, which is more related to bringing the innovation to the market (see also the contribution of Mendonça et al., 2004; Aaker, 2007; Helmers and Rogers, 2010; Llerena and Millot, 2013; Flikkema et al., 2014). In our dataset, as in most to date, we cannot directly observe: a) which IP is related to a given product or line of business and, as a consequence, b) the timing with which firms resort to patents and trademark along the innovation process. We can only indirectly verify the existence of such a sequential timing by exploiting positive occurrences of our concordance measure to identify innovation processes that we can expect to be related to the same product (or line of business). In particular, we consider, in each year and for each firm, the pairs of 3-digit IPC and 2-digit NICE codes associated to the stock of patents and trademarks, for which the overlapping coefficient is higher than zero, that is, when we are inclined to expect that the patent and the trademark are related to the same line of product. Moreover, for each pair of IPC and NICE codes, for each firm and year, we account for the least recent patent and trademark in the firm portfolio. Our descriptive evidence provides only mild support to our hypothesis on the relative timing: depending on the year of analysis, in 50% to 55% of such IPC-NICE pairs the patent pre-dates the trademark.<sup>16</sup>

Table 4 shows the distribution of patents and trademarks across firms according to their size in 2014. The size classes are defined conforming to the Eurostat definition.<sup>17</sup> We consider four size classes: ‘micro’ firms with less than 10 workers; ‘small’ with workers ranging between 10 and 49; ‘medium’ between 50 and 249 and ‘large’ with more than 250 workers.<sup>18</sup>

Consistently with the actual population of Italian firms, also in our dataset the highest share of firms is represented by micro-firms (around 58%). The share of small firms is around 35%, while large firms are only less than 1%. Due to the shape of the firm size distribution, in absolute numbers, the two categories of SMEs are those more represented in terms of IP’s holders, and represent the highest share of firms with concordant patents and trademarks (around 32% and 43% of firms with concordant IP rights, respectively). Note however that, in relative terms, while around 55% (645/1172) of large firms are patent holders, the relative share shrinks respectively to 30% for medium and to 8.6% for small firms.

(Table 4 around here)

Looking at the number of patents and trademarks we note that the largest number of trademarks falls in the medium class (around 35% of the total number of trademarks),

---

<sup>16</sup>Dinlersoz et al. (2018) perform a similar exercise on the relative timing for R&D, patents and trademark filings and report mild evidence in the opposite direction for US firms. Notice however that they do not employ any concordance measure.

<sup>17</sup>The classification is available at:

[http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Enterprise\\_size](http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Enterprise_size)

<sup>18</sup>Eurostat classification refers to the number of workers, thus including also the entrepreneur(s). AIDA dataset originally reports the number of employees and, as obvious, many SMEs report no employees. Hence, to make the two figures comparable, we approximate the number of workers as the number of employees plus one.

while the highest share of patents is owned by large firms (around 42% of the total). In this respect, the descriptive evidence for Italy confirms the findings from previous studies highlighting a higher propensity to patent for large firms and identify trademarks as the main IP instrument for small and medium firms (see Mendonça et al. (2004), Blind et al. (2006), Leiponen and Byma (2009) and Flikkema et al. (2014) among others).

Finally, Table 8 in the Appendix reports the distribution of patents and trademarks according to the sector of economic activity. Trademarks are mainly concentrated in food products (ATECO 10, with around 14% of total number of trademarks), manufacture of machinery and equipment n.e.c. (ATECO 28, with around 11% of total number of trademarks), chemical products and wearing apparel (ATECO 20 and 14, both with around 8% of total number of trademarks). Sectors reporting the highest number of patents are manufacture of machinery and equipment n.e.c. (ATECO 28, with around 29% of total number of patents), manufacture of computer, electronic and optical products (ATECO 26, with around 13% of total number of patents) and manufacture of fabricated metal products, except machinery and equipment (ATECO 25, with around 11% of total number of patents).

### 3 Non-parametric evidence

Before proceeding with more standard econometric analysis we report evidence from univariate kernel density estimations, which allows us to graphically compare the performance of different groups of firms.

We compare the empirical distributions of firms' performance across four groups of firms: without IP instruments, owning only trademarks, owning only patents and firms with both patents and trademarks. Throughout, we will refer to the firms of the first group as *non-innovative* and to the rest as *innovative* firms. We focus on (log of) total revenues and yearly growth rate of total revenues as measures of firms' performance. For the sake of completeness, we estimate kernel densities focusing on the first (2006) and last (2014) available years in the dataset, note however that there is not much intertemporal variation.

Graphically, we identify relevant differences between groups of firms in terms of total revenues. As reported in Figure 3, firms without any IP instrument underperform firms with patents and/or trademarks in terms of revenues. Among innovative firms, those with both patents and trademarks tend to show higher values in terms of total revenues; while firms owning only trademarks or only patents display similar distributions. On the contrary, these groups of firms do not seem to differ much in terms of growth rates, Figure 4. The only feature that is possible to appreciate from the graph is the higher concentration around the mean for firms with both patents and trademarks.

(Figure 3 around here)

(Figure 4 around here)

The descriptive evidence on Italian firms is in line with the existing empirical literature, which reports higher performance for firms with IP instruments, even employing different measures of firm performance. In particular, given limited data availability

on the universe, many contributions focus on market values of quoted companies or on specific sectors of the economy. Greenhalgh and Rogers (2012) and Sandner and Block (2011), among others, report a positive impact of both measures of IP on the markets' valuation of firms. Dosso and Vezzani (2017) and Llerena and Millot (2013) analyze and identify a positive influence of the combined use of patents and trademarks on firms' market value only for some industrial sectors. Graham and Somaya (2004), focusing on the software industry, also report evidence in favor of the complementarity hypothesis. Moreover, Zhou et al. (2016) identify a positive impact of the combined use of patents and trademarks on the start-up amount of venture capital funding. Other studies investigate the effect of patent and trademark activity on firms' survival and identify a positive correlation between these two measures of IP and a higher increase in the expected life span of firms (e.g., Jensen et al., 2008; Buddelmeyer et al., 2010; Helmers and Rogers, 2010; Wagner and Cockburn, 2010).

There is scant evidence on the relationship between IP instruments and firm growth. In particular, Guzman and Stern (2015) show that both patent and trademark activities positively affect growth of a sample of start-ups in the Silicon Valley; however, authors do not focus on the potential combined impact of both trademarks and patents. More recently, Castaldi and Dosso (2018) analyze the effects of the combined impact of patents and trademarks on firm growth. They find that patents (as well as firms' R&D stock) negatively affect firms' growth, although such negative impact is mitigated by the indirect positive effect of trademarks.

We now focus our attention on the restricted sample of firms with both patents and trademarks, and we divide these firms into two groups according to whether their degree of concordance,  $conc_{i,t}$  as defined in equation (2), is strictly positive (firms with concordant patents and trademarks) or nil. We graphically compare the distributions of these two groups. As exhibited in Figure 5, firms with concordant patents and trademarks mildly outperform those without concordant IP instruments in terms of total revenues. Contrarily, as shown in Figure 6, kernel density estimations do not show any difference in terms of growth for the two groups of firms.

(Figures 5 and 6 around here)

To confer statistical accuracy to the graphical analysis reported above, we perform non-parametric Fligner-Policello (FP henceforth) test of stochastic equality, proposed by Fligner and Policello (1981) to compare *innovative* and *non-innovative* firms. As FP test allows to compare two groups at a time, we gather together firms with any proxy of innovative activity (firms owning only trademarks, only patents and both). The test is defined as follows. Let  $F_I$  and  $F_{NI}$  be the distributions of the relevant variables of *innovative* and *non-innovative* firms, respectively. Denote with  $\mathbf{X}_I \sim F_I$  and  $\mathbf{X}_{NI} \sim F_{NI}$  the associated random variables, and with  $X_I$  and  $X_{NI}$  two respective realizations. Based on the FP test, the distribution  $F_I$  is said to stochastically dominate  $F_{NI}$  if  $Prob\{X_I > X_{NI}\} > 1/2$ .<sup>19</sup>

For each relevant variable (total revenues and growth) we compare, *innovative* with *non-innovative* firms in years 2006 and 2014, respectively. A positive sign of the FP statistic means that *innovative* firms have a higher likelihood to take on larger values of a given relevant variable (i.e. the distribution of *innovative* firms stochastically dominates

---

<sup>19</sup>The null hypothesis in the FP test implies equality of median among the compared distributions.

the distribution of *non-innovative* firms), and the opposite holds if the statistic takes a negative sign.<sup>20</sup>

FP statistics show that *innovative* firms perform better with respect to *non-innovative* firms in terms of total revenues, while they do not appear to differ in terms of growth, in both years considered. FP statistics are thus confirming the descriptive evidence of the plots above.

Focusing on firms with both patents and trademarks, we perform the FP test in order to compare performance between the two groups of firms with concordant IP instruments and firms without such concordance. Similarly to the graphical comparisons, even in this case, firms with concordant IP instruments result to be characterized by higher total revenues. However, according to the FP test, concordance between IP instruments does not imply a relevant difference between firms in terms of growth.

## 4 Parametric analysis of firms' performance

In this section we investigate the effects of patents and trademarks on firms' performance resorting to parametric analysis. The dependent variable,  $Y_{i,t}$ , is firms' performance proxied by, alternatively, firms' total revenues (the logarithmic transformation,  $\ln(\text{totrev}_{i,t})$ ) and yearly growth rate ( $\text{growth}_{i,t}$ ), measured as the logarithmic difference between firms total revenues in two consecutive years.

We relate  $Y_{i,t}$  to the innovation indicators and other controls through the following baseline specification:

$$Y_{i,t} = c + \beta_1 tm_{i,t-1} + \beta_2 pat_{i,t-1} + \beta_3 both_{i,t-1} + \beta_4 \ln(\text{workers}_{i,t-1}) + \beta_5 \ln(LP_{i,t-1}) + X'_{i,t} \alpha + u_{i,t} \quad (4)$$

where  $i$  denotes firms and  $t$  years. The explanatory variables that we are mostly interested in consist of three binary indicators  $tm_{i,t-1}$ ,  $pat_{i,t-1}$  and  $both_{i,t-1}$  equal to 1 if firms own, respectively, at least one registered trademark (but not patents), at least one (applied) patent (but not trademarks), both trademarks and patents.<sup>21</sup> These dummy variables are mutually exclusive so that we can better identify the different IP strategies of the firms; at the same time, we avoid problems of multicollinearity due to the high correlation between firms' trademarks and patents. Moreover, we include among independent variables firms' size, measured by the logarithmic transformation of number of workers ( $\ln(\text{workers}_{i,t-1})$ ) and the logarithmic transformation of labour productivity ( $\ln(LP_{i,t-1})$ ).<sup>22</sup> In order to reduce the potential endogeneity of our regressors, these variables are lagged one period with respect to the dependent variable. In the regression we account for a vector of controls,  $X'_{i,t}$ , which includes: a set of 2-digit ATECO industry dummy variables; three geographical area dummy variables which identify firms operating in the North, Centre or South of Italy, respectively; and nine year dummies (2006-2014). These binary variables allow to control for time invariant sectoral effects, for the omission of geographical specific time invariant characteristics which might bias our parameter estimates and, for the

<sup>20</sup>In the interest of space we do not show the FP tests, but they are available upon request.

<sup>21</sup>We build the three dummy variables based on the yearly number of patents and registered trademarks owned by firms. For more details on the construction of patents and trademarks stocks refer to Section 2.1.

<sup>22</sup>Labour productivity is computed as the ratio between added value and number of workers.

economic cycle and common macroeconomic factors, respectively.<sup>23</sup>

In order to assess whether the concordance between firms' stocks of patents and trademarks positively impact on firms' performance, in an extended specification, we also include the degree of concordance between IP rights,  $conc_{i,t-1}$ , as defined in equation (2), with a reporting lag of one year. Clearly, this specification only includes firms with both patents and trademarks. Hence, moving from the baseline specification to this extended one, the number of observations significantly decreases (from more than 600,000 to less than 18,000). In order to verify if results are driven by the lower number of observations, first we re-estimate the baseline specification on the sub-sample of firms with both patents and trademarks.<sup>24</sup>

We resort to Pooled OLS models for two reasons. First, given the rather poor dynamics in the patenting (trademarking) activities of Italian firms (i.e. firms are not registering or applying for many new trademarks or patents every year), for many companies the number of patents (trademarks) does not change over the period of observation, hence, for these firms, the effect of IP would get confounded with the firm fixed effect. Secondly, as shown by Angrist and Pischke (2008), estimating fully saturated dummy variables models with OLS is fully general, regardless of the distribution of the dependent variable.

Results, shown in Table 5, suggest that firms' innovative capacity is positively associated with their performance. In particular, coherently with the non-parametric results, estimates show that patents and trademarks exert a positive impact on firms' total revenues and growth. More in detail, the estimates report coefficients that are larger for the group of firms owning both IP instruments rather than either patents or trademarks.

(Table 5 around here)

In order to confer statistical precision to the results above, we perform one-sided t-tests to compare, on the one side, the magnitude of the coefficients for the group of firms owning both patents and trademarks and, on the other side, the magnitude of the coefficients of firms owning either only patents or trademarks. Tests show that owning both IP rights has a higher impact than owning only one of the two IP instruments on both firms' total revenues and growth. Notice that while the results of Table 5 largely confirms the non-parametric evidence, including the further controls of equation 4 is crucial to appreciate the growth premium of firms with both IP instruments.

Also note that it is not obvious to compare our results to the extant literature as, to our knowledge, we are the first to investigate the role of patents and trademarks on the entire population of firms, hence including also SMEs and micro ones. To the extent that such comparison is possible, our results are coherent with Zhou et al. (2016) who find that start-ups with both patents and trademarks obtain higher amount of venture capital funding than firms with only one of the two IP instruments. Moreover, our results do not drastically diverge from other studies which identify a combined effect of IP rights on

---

<sup>23</sup>Focusing on firms' growth as measure of performance, our regression model is in line with specifications suggested by the applied literature which empirically tests the validity of Gibrat's law (see Santarelli et al. (2006) for an exhaustive survey of empirical studies testing Gibrat's law). In this literature, firms' growth is modeled as a function of the initial size. Moreover, this baseline specification has been extended in order to account for several major control variables in level, such as productivity, export, external finance and R&D activity as well as others. We refer to, within the very vast literature that investigates the growth of firms, Gibrat (1931); Penrose (1959); Evans (1987b,a); Doms et al. (1995); Del Monte and Papagni (2003); García-Manjón and Romero-Merino (2012).

<sup>24</sup>Correlation tables for the variables included in our specifications are available upon request.

firms' performance which differ across sectors. For example, Llerena and Millot (2013) identify a positive combined effects of patents and trademarks on the market value only for French firms operating in chemical and pharmaceutical industries (they find a negative combined effect for firms operating in high-tech sectors), while Dosso and Vezzani (2017) show that having both patents and trademarks implies higher firms' market value only for the automobile sector.

As far as the comparison between the two instruments is concerned, our estimates suggest that firms owning only trademarks show higher performance than firms with only patents. Indeed, in all specifications, coefficients associated to the dummy variable indicating firms owning only trademarks are bigger than coefficients estimated for the dummy variable indicating firms with only patents (coefficients for  $tm_{i,t-1}$  and  $pat_{i,t-1}$  are 0.306 and 0.169 in column 1 and 0.0525 and 0.0358 in column 4, respectively). This result differs from Guzman and Stern (2015) and Castaldi and Dosso (2018) that focus on firms' growth as measure of performance. In particular, Guzman and Stern (2015) identify a higher effect of patents, compared to the effect of trademarks, on the start-up probability of growth. Conversely, focusing on top R&D investors, Castaldi and Dosso (2018) find only an indirect impact of trademarks on growth, with trademarks only mitigating the negative effect of patents on firm growth.

Focusing on the degree of concordance between the stock of patents and trademarks, as proxied by our proposed measure, parametric estimates on the entire sample do not suggest a significant relationship between the coherence of firms' IP rights and performance. Indeed, the degree of concordance never shows up as significant for the entire sample. We are cautious in interpreting this as a failure or irrelevance of such measure, but rather consider this as the consequence of the very small number of Italian firms for which it is possible to compute the measure, in fact, less than 2% of the complete sample. The same constraint was also affecting the analysis of the diversification patterns in knowledge (as proxied by patents) of Italian firms reported in Dosi et al. (2017). A certainly more appropriate context to assess its relevance would be provided by larger firms owning patents and trademarks, and also pursuing some - more or less deliberate - pattern of diversification in their technological knowledge.

In considering the other control variables of Table 5, we find that firms' size has a positive impact on firms' total revenues and a negative effect on growth.<sup>25</sup> Focusing on productivity, our findings suggest a positive and significant impact on firms' total revenues, regardless of the sample considered (whole sample in column 1, and sub-sample of firms with both IP rights in columns 2 and 3). When considering growth as dependent variable, productivity displays a negative and significant impact on firms' growth when considering the whole sample (column 4), whereas it turns out as not significant on the restricted sample (columns 4 and 5). The lack of the (expected) positive relation between productivity and growth is not much surprising, at least for the Italian case, see Bottazzi et al. (2005, 2008).

We next verify whether results on the proximity of IP instruments vary across sectors. To this purpose, we estimate the previous specifications on sectoral sub-samples. In particular, we consider the aggregation of the manufacturing sectors proposed by Eurostat. This classification, based on 2-digit NACE Rev. 2 (ATECO 2007 for Italy), groups sectors into high-technology, medium-high, medium-low and low-technology according to their

---

<sup>25</sup>The negative impact of initial firms' size on their growth is a common result in literature, also known as violation of Gibrat's law (see e.g., Mazzucato and Parris, 2015; Castaldi and Dosso, 2018; Grazzi and Moschella, 2018).

technological intensity (proxied by R&D expenditure/value added).<sup>26</sup> An alternative and much appropriate aggregation is the Pavitt taxonomy (Pavitt, 1984) that is based on the sources of technological knowledge, requirements of the users, and appropriability regimes. Unfortunately, in the AIDA-AMADEUS dataset sectors of economic activity are reported according to Nace Rev. 2 and the latter classification does not provide a fine-grained mapping of industries into Pavitt classes.<sup>27</sup> It is for this reason that we opt for the aggregation proposed by Eurostat. Results for these four groups are shown in Table 10 to Table 13 in the Appendix.

This further analysis suggests that the impact of IP on firms' total revenues and growth does not significantly differ across sectors. Differences across sectors instead arise with respect to the proxy of concordance between patents and trademarks. Indeed, focusing on the effect on firms' revenues, we find that owning concordant IP rights has a negative and significant impact for firms in the medium-high technology sectors (column 3 of Table 11). While these results might be mostly driven by the small number of observations, they are also coherent with a picture in which firms in the medium-high tech sector enjoy an advantage, in terms of sales, from having technological knowledge, as proxied by patents, that covers a different product spectrum than that embodied in trademarks. As for the whole sample, focusing on growth as measure of firms' performance, results do not identify any significant effect of the degree of coherence.

## 4.1 Robustness

In order to account for the potential endogeneity between labour productivity and the three binary indicators for the stocks of patents and trademarks ( $tm_{i,t-1}$ ,  $pat_{i,t-1}$  and  $both_{i,t-1}$ ), we present a robustness check where we estimate a structural model that describes the linkage between *R&D* activity, firms' innovation outputs, productivity and performance. We rely on the methodology proposed by Crepon et al. (1998) (CDM model henceforth) which allows to account for statistical biases related to selection and endogeneity. The CDM model is formalized in three steps: the first one comprises two equations which identify the factors underlying the decision and intensity of *R&D* expenditure, respectively; the second step takes into account the relationship between the *R&D* intensity and innovation output and the third step relates innovation output to labor productivity. By using predicted values of *R&D* and innovation output in the second and third steps respectively, the CDM model allows to account for simultaneity and reverse causality problems in the various steps. Moreover, we augmented the CDM model adding a fourth step, in which we analyze the relationship between (predicted) labor productivity and firms' performance (total revenues and growth, alternatively). A similar approach has been used by Antonietti and Cainelli (2011), that estimate a structural model of *R&D*, innovation output, productivity and export performance.

Other studies estimate different versions of the CDM model mainly using cross-sectional data. In particular we refer to Griffith et al. (2006) using data across four European countries (France, Germany, Spain and UK), Hall et al. (2013b), Marin and Lotti (2017) and Pellegrino and Piva (2014) for Italian firms, and Martin and Nguyen-Thi

<sup>26</sup>We show in Tables 8 and 9 in the Appendix how sectors are grouped in the Eurostat aggregation.

<sup>27</sup>Due to multiple correspondences, it is not possible to accurately convert 3-digit ATECO 2007 codes into 3-digit ATECO 2002 and, thus, it is hard to group 3-digit ATECO 2007 sectors into Pavitt industries. Some attempts have also been made in order to provide Pavitt taxonomy for industry classification Nace Rev. 2 (and ATECO 2007), refer for instance to Perani and Cirillo (2015) and Bogliacino and Pianta (2016).

(2015) that focus on firms in Luxembourg, among others (see Hall and Mairesse (2006) for a survey).<sup>28</sup> As standard in the literature, we also estimate the augmented CDM model with cross sectional data. Moreover, differently from other studies, we can avoid problems of simultaneity between the dependent and the explanatory variables by exploiting the availability of several years in our data and including one year lagged explanatory variables in each step. We then perform the robustness check for firms reporting relevant data for the period 2007-2011. In such a way, our data allow to account for the time lag between firms' *R&D* decisions, the resulting innovation output, its effect on productivity and the impact of productivity on firms' performance.

Our specification follows the basic model proposed by Crepon et al. (1998), with the exception for the fourth step that is not contemplated by the authors.

The first step includes two equations which account for the *R&D* decision and for its intensity. Following Griffith et al. (2006) and Pellegrino and Piva (2014) among others, we estimate the first step of the CDM model not only for innovative firms; thus, we use *R&D* intensity predicted values for all firms in the first step to proxy innovation input in the second step. This choice is related to the fact that all firms exert innovative effort, but not all firms report their *R&D* expenditure.

*R&D* intensity is measured as *R&D* expenditure per worker. In order to account for possible sample selection bias we estimate the first step using the Heckman two-step procedure (Heckman, 1976, 1979). Thus, in the first equation we estimate the firms' probability to invest in *R&D*, and, in the second equation, we estimate the determinants of *R&D* intensity. We account for the following selection equation, representing whether a firm performs (reports) *R&D* or not:

$$rd_{i,2008} = \begin{cases} 1, & \text{if } rd_{i,2008}^* = \beta_1 z_{i,2007} + u_{1,i} > 0 \\ 0, & \text{if } rd_{i,2008}^* = \beta_1 z_{i,2007} + u_{1,i} \leq 0 \end{cases} \quad (5)$$

where  $rd_{i,2008}$  is the observed dummy variable and it is equals to one if firm  $i$  reports positive *R&D* expenditure in year 2008.  $rd_{i,2008}^*$  is the corresponding latent variable such that firms invest in *R&D* if it is positive, and  $z_{i,2007}$  is a set of one year lagged variables explaining the *R&D* decision.

Conditional on performing (reporting) *R&D*, the research intensity equation is defined as follows:

$$rd\_int_{i,2008} = \begin{cases} rd\_int_{i,2008}^* = \beta_2 w_{i,2007} + u_{2,i}, & \text{if } rd_{i,2008} = 1 \\ 0, & \text{if } rd_{i,2008} = 0 \end{cases} \quad (6)$$

where  $rd\_int_{i,2008}$  is the observed *R&D* expenditure per employee (in natural logarithms) and  $w_{i,2007}$  is a set of one year lagged explanatory variables. We assume the joint normality of the disturbances ( $u_{1,i}$  and  $u_{2,i}$ ) in equations (5) and (6) (i.e., the generalized Tobit model assumption).

---

<sup>28</sup>Only few studies apply the CDM model using panel data. In particular, we refer to Marin (2014), that employs a panel of Italian manufacturing firms over the period 2000-2007, Heshmati and Kim (2011) using a panel of Korean *R&D* engaged firms from 1986 to 2002, and Chudnovsky et al. (2006) that use a panel of 718 firms interviewed in two consecutive innovation surveys carried out by INDEC (Argentina's National Statistical Institute) and covering the period 1992-2001. Moreover, Raymond et al. (2010), using data for firms in the Netherlands, from four waves of the Community Innovation Survey between 1994 and 2002, consider the innovation equation (the second step of the CMD model). The authors consider two equations, the first equation for firms' decision to innovate and the second equation for the innovation intensity.



Note that the explanatory variables in the two equations need not be the same; however, we do not have any a priori on the exclusion restriction to include in the first-stage selection equation and, as in Crepon et al. (1998), we consider the same set of explanatory variables ( $z_{i,2007} = w_{i,2007}$ ) in both equations and, as in their basic model, we include among regressors the number of workers and firms' market share, both of them expressed in logarithms. We compute firm  $i$  market share in 2007 as firm's total revenues over average 2 digit ATECO sector total revenues in the same year. Moreover we control for a set of 2 digit ATECO industry dummy variables and three geographical area dummy variables which identify firms operating in the North, Centre or South of Italy, respectively. Differently from Crepon et al. (1998), we do not have information which allow to measure firms' diversification.

In the second step the knowledge or innovation production function is estimated by applying Probit model as in Hall et al. (2013b) and Antonietti and Cainelli (2011).<sup>29</sup> Our dependent variable is a dummy variable equal to 1 if firm  $i$  owns any type of IP (only patents, only trademarks or both) in 2009 and 0 otherwise.

$$Prob(inno_{i,2009} = 1 | X_{i,2008}) = \Phi(\alpha_1 \widehat{rd\_int}_{i,2008} + \beta_3 X_{i,2008} + u_{3,i}) \quad (7)$$

where  $\Phi(\cdot)$  is the cumulative distribution function of a standard normal variable, with associated density  $\phi(\cdot)$ ,  $\widehat{rd\_int}_{i,2008}$  is the predicted value of the  $R\&D$  intensity from the first step, and  $X_{i,2008}$  is a set of explanatory variables including firm's size (measured as the natural logarithm of the number of employees in year 2008) and dummy variables for sector and geographical area in which firms operate. As suggested by Crepon et al. (1998), firms' market share only indirectly enters the innovation output equation through the  $R\&D$  intensity; while the inclusion of firms' size among regressors, in equation (7), allows to verify whether the effects of firms' size on innovation output is completely captured by the  $R\&D$  activity or not. The use of predicted values from the previous step as regressor generate invalid standard errors, for this reason we compute bootstrapped standard errors in all regressions in which we use predicted values as independent variables.

The third step consists in the estimation, applying OLS, of an augmented Cobb-Douglas production function with physical capital, workers and innovation output as main determinants:

$$\ln(LP_{i,2010}) = \alpha_2 \widehat{inno}_{i,2009} + \beta_4 r_{i,2009} + u_{4,i} \quad (8)$$

where  $\ln(LP_{i,2010})$  is the natural logarithm of value added per worker,  $\widehat{inno}_{i,2009}$  is the predicted value of the innovation output indicator from equation (7), in the second step, and  $r_{i,2009}$  is a set of one year lagged explanatory variables including total fixed assets per worker and the number of workers (both expressed in natural logarithm), dummy variables for sector and geographical area in which firms operate.

Finally, in the fourth step we estimate, by applying OLS, the impact of labor productivity on firms' performance  $Y_{i,2011}$ , measured alternatively, as firms' total revenues ( $\ln(totrev_{i,2011})$ ) and growth rate ( $growth_{i,2011}$ ), using the following model:

$$Y_{i,2011} = c + \alpha_3 \ln(\widehat{LP}_{i,2010}) + \beta_5 s_{i,2009} + u_{5,i} \quad (9)$$

---

<sup>29</sup>Hall et al. (2013b) define innovation as at least one of the process, product, and organizational innovations. In the version of the CDM model proposed by Antonietti and Cainelli (2011), the propensity to create a new product or a new process is considered as measure for innovation output.

where  $\ln(\widehat{LP}_{i,2010})$  is the predicted value of productivity from the third step and  $s_{i,2009}$  is a set of independent variables including workers (the natural logarithm), industry and geographical area dummy variables.

Tables 6 and 7 show CDM model estimates considering total revenues and growth as measure of firms' performance, respectively. According to the existing literature (i.e. Hall et al. (2013b) and Marin and Lotti (2017), among others), columns (1) and (2) (in both Tables 6 and 7) show that firms' size positively impact on *R&D* propensity, while *R&D* intensity decreases with size. As in Crepon et al. (1998), market share exerts a positive impact on both the decision to perform (report) *R&D* and on the size of the *R&D* investment. The likelihood-ratio tests ( $\chi^2(1)$ ), reported at the bottom of the two Tables, justify the inclusion of the Heckman selection equation in our models. This test computationally compares the joint likelihood of an independent Probit model for the selection equation and a regression model on the observed *R&D* intensity against the Heckman model likelihood. Coherently with Hall et al. (2013b) and Antonietti and Cainelli (2011), in columns (3), we find that the (predicted) probability to have any type of IP increases with *R&D* intensity and firm's size. Estimation of the production functions (columns (4)) reveals that, as expected, firms' productivity increases with their capital intensity and size. Moreover, we find that having any type of IP rises labor productivity. Finally, we verify the role of labor productivity on firms' performance. Focusing on total revenues as proxy for performance (see Tables 6 in column (5)) results are in line with our main estimates presented in Table 5 (column (1)). Indeed, we find that firms' size and labor productivity positively impact on firms' total revenues. Differently from the results shown in Table 5 (column (4)), considering the impact that patents and trademarks exert on growth, Table 7, column (5), shows that the impact of both size and productivity turn out to be positive.

(Tables 6 and 7 around here)

## 5 Conclusion

We have analyzed the relationship between innovation activities and firms' performance in terms of revenues and growth.

Our contribution is threefold. First, differently from most of empirical works, which focus on large and medium firms, on top corporate R&D investors or on firms operating in high-tech industries (see e.g., Greenhalgh and Rogers, 2006; Fang et al., 2011; Chang et al., 2012; Greenhalgh and Rogers, 2012; Dernis et al., 2015; Daiko et al., 2017; Dosso and Vezzani, 2017; Castaldi and Dosso, 2018), we have performed our analysis on IP taking into account the virtual universe of Italian limited liability manufacturing firms. Second, we have investigated whether there is a complementary or substitute relationship between the two instruments of IP. To the best of our knowledge, only few studies examine combined impacts of patents and trademarks on firms' performance (see e.g., Greenhalgh and Rogers, 2012; Sandner and Block, 2011; Dosso and Vezzani, 2017; Llerena and Millot, 2013; Jensen et al., 2008; Buddelmeyer et al., 2010; Helmers and Rogers, 2010; Wagner and Cockburn, 2010). Finally and completely novel, for firms owning both IP instruments, we have proposed a measure of concordance between a firm's stock of patents and trademarks, investigating the effect of this concordance on firms' performance.

Overall, our results indicate that IP rights exert a positive impact on firms' performance and, more importantly, identify complementarity between the use of patents and

trademarks. Interestingly, looking at the degree of concordance between IP rights, we find that the effect on firms' performance only marginally varies according to the sectoral technological intensity.

Although Italian firms do not represent the ideal testbed due to the very low share of firms with both patents and trademarks, we believe that the measure of concordance we propose can contribute to better understand how business firms engage in innovation activities and how this affects their performance. Establishing a linkage between patents and trademarks can shed some new light on the deliberate choice of the relative timing for the two activities. Do firms start with seeking protection for their technological innovation, or they rather decide to apply for a patent only when the introduction of a certain line of business, possibly protected by a trademark, has already proved to be successful? Further, the concordance measure or in other terms, the extent to which a given trademark (and the related product) is "backed up" by a patent, provides a more detailed perspective on the complementarity existing between IP instruments as it focuses on a specific product or line of business. Is the existence of patent-to-trademark linkage evidence of a "stronger" innovation, which is expected to generate higher profits?

We are of course aware of some limitations of our work. We remind the reader that for the construction of the measure we have extensively resorted to "Algorithmic Links with Probabilities (ALP)" which have themselves some limitations in terms of lack of precision. On a brighter side, also note that the public availability and use of such (and similar) ALP is also likely to generate some improvement over time. Second, and as apparent from this work even more relevant, the returns from employing the concordance measure largely depend on the characteristic of firms and in particular on their propensity to apply for patents and trademarks.

## References

- AAKER, D. (2007): "Innovation: Brand it or lose it," *California Management Review*, 50, 8–24.
- ANGRIST, J. D. AND J.-S. PISCHKE (2008): *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.
- ANTONIETTI, R. AND G. CAINELLI (2011): "The role of spatial agglomeration in a structural model of innovation, productivity and export: a firm-level analysis," *The Annals of Regional Science*, 46, 577–600.
- BLIND, K., J. EDLER, R. FRIETSCH, AND U. SCHMOCH (2006): "Motives to patent: Empirical evidence from Germany," *Research Policy*, 35, 655–672.
- BLOCK, J. H., C. O. FISCH, A. HAHN, AND P. G. SANDNER (2015): "Why do SMEs file trademarks? Insights from firms in innovative industries," *Research Policy*, 44, 1915–1930.
- BOGLIACINO, F. AND M. PIANTA (2016): "The Pavitt Taxonomy, revisited: patterns of innovation in manufacturing and services," *Economia Politica*, 33, 153–180.
- BOTTAZZI, G., M. GRAZZI, AND A. SECCHI (2005): "Input Output Scaling Relations in Italian Manufacturing Firms," *Physica A*, 355, 95–102.

- BOTTAZZI, G., A. SECCHI, AND F. TAMAGNI (2008): “Productivity, profitability and financial performance,” *Industrial and Corporate Change*, 17, 711–751.
- BRYCE, D. J. AND S. G. WINTER (2009): “A General Interindustry Relatedness Index,” *Management Science*, 55, 1570–1585.
- BUDELMEYER, H., P. H. JENSEN, AND E. WEBSTER (2010): “Innovation and the determinants of company survival,” *Oxford Economic Papers*, 62, 261–285.
- CASTALDI, C. AND M. DOSSO (2018): “From R&D to market: using trademarks to capture the market capability of top R&D investors,” Tech. rep., Joint Research Centre (Seville site).
- CEFIS, E. AND L. ORSENIGO (2001): “The persistence of innovative activities: A cross-countries and cross-sectors comparative analysis,” *Research Policy*, 30, 1139–1158.
- CHANG, K.-C., D.-Z. CHEN, AND M.-H. HUANG (2012): “The relationships between the patent performance and corporation performance,” *Journal of Informetrics*, 6, 131–139.
- CHUDNOVSKY, D., A. LÓPEZ, AND G. PUPATO (2006): “Innovation and productivity in developing countries: A study of Argentine manufacturing firms behavior (1992–2001),” *Research policy*, 35, 266–288.
- CREPON, B., E. DUGUET, AND J. MAIRESSE (1998): “Research, Innovation And Productivity: An Econometric Analysis At The Firm Level,” *Economics of Innovation and New Technology*, 7, 115–158.
- CZARNITZKI, D. AND K. KRAFT (2004): “Innovation indicators and corporate credit ratings: evidence from German firms,” *Economics Letters*, 82, 377–384.
- DAIKO, T., H. DERNIS, M. DOSSO, P. GKOTSIS, M. SQUICCIARINI, A. TUEBKE, AND A. VEZZANI (2017): “World Top R&D Investors: Industrial Property Strategies in the Digital Economy,” JRC Working Papers JRC107015, Joint Research Centre (Seville site).
- DE RASSENFOSSE, G., H. DERNIS, AND G. BOEDT (2014): “An introduction to the Patstat database with example queries,” *Australian Economic Review*, 47, 395–408.
- DE RASSENFOSSE, G. AND A. B. JAFFE (2017): “Econometric Evidence on the R&D Depreciation Rate,” NBER Working Papers 23072, National Bureau of Economic Research.
- DEL MONTE, A. AND E. PAPAGNI (2003): “R&D and the growth of firms: empirical analysis of a panel of Italian firms,” *Research policy*, 32, 1003–1014.
- DERNIS, H., M. DOSSO, F. HERVAS, V. MILLOT, M. SQUICCIARINI, AND A. VEZZANI (2015): “World Corporate Top R&D Investors: Innovation and IP bundles,” JRC Working Papers JRC94932, Joint Research Centre (Seville site).
- DINLERSOZ, E., N. GOLDSCHLAG, A. MYERS, AND N. ZOLAS (2018): “An Anatomy of U.S. Firms Seeking Trademark Registration,” in *Measuring and Accounting for Innovation in the 21st Century*, National Bureau of Economic Research, Inc, NBER Chapters.

- DOMS, M., T. DUNNE, AND M. J. ROBERTS (1995): “The role of technology use in the survival and growth of manufacturing plants,” *International Journal of Industrial Organization*, 13, 523–542.
- DOSI, G., M. GRAZZI, AND D. MOSCHELLA (2015): “Technology and costs in international competitiveness: from countries and sectors to firms,” *Research Policy*, 44, 1795–1814.
- (2017): “What do firms know? What do they produce? A new look at the relationship between patenting profiles and patterns of product diversification,” *Small Business Economics*, 48, 413–429.
- DOSSO, M. AND A. VEZZANI (2017): “Firm market valuation and intellectual property assets,” Tech. rep., Joint Research Centre (Seville site).
- EVANS, D. S. (1987a): “The Relationship between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries,” *Journal of Industrial Economics*, 35, 567–81.
- (1987b): “Tests of Alternative Theories of Firm Growth,” *The Journal of Political Economy*, 95, 657–674.
- FANG, E., R. W. PALMATIER, AND R. GREWAL (2011): “Effects of customer and innovation asset configuration strategies on firm performance,” *Journal of Marketing Research*, 48, 587–602.
- FLIGNER, M. AND G. POLICELLO (1981): “Robust rank procedures for the Behrens-Fisher problem,” *Journal of the American Statistical Association*, 76, 141–206.
- FLIKKEMA, M., C. CASTALDI, A.-P. DE MAN, AND M. SEIP (2015): “Explaining the trademark-innovation linkage: the role of patents and trademark filing strategies,” *Academy of Management Proceedings 2015*, 1.
- FLIKKEMA, M., A.-P. DE MAN, AND C. CASTALDI (2014): “Are Trademark Counts a Valid Indicator of Innovation? Results of an In-Depth Study of New Benelux Trademarks Filed by SMEs,” *Industry and Innovation*, 21, 310–331.
- GARCÍA-MANJÓN, J. V. AND M. E. ROMERO-MERINO (2012): “Research, development, and firm growth. Empirical evidence from European top R&D spending firms,” *Research Policy*, 41, 1084–1092.
- GIBRAT, R. (1931): *Les inégalités économiques*, Librairie du Recueil Sirey, Paris.
- GOLDSCHLAG, N., T. J. LYBBERT, AND N. J. ZOLAS (2016): “An Algorithmic Links with Probabilities Crosswalk for USPC and CPC Patent Classifications with an Application Towards Industrial Technology Composition,” Working Papers 16-15, Center for Economic Studies, U.S. Census Bureau.
- GRAHAM, S. AND D. SOMAYA (2004): “The Use of Patents, Copyrights, and Trademarks in Software: Evidence from Litigation,” in *Patents, Innovation and Economic Performance*, ed. by D. Guellec, OECD.

- GRAHAM, S. J., C. GRIM, T. ISLAM, A. C. MARCO, AND J. MIRANDA (2018): “Business dynamics of innovating firms: Linking U.S. patents with administrative data on workers and firms,” *Journal of Economics & Management Strategy*, 27, 372–402.
- GRAZZI, M. AND D. MOSCHELLA (2018): “Small, young, and exporters: New evidence on the determinants of firm growth,” *Journal of Evolutionary Economics*, 28, 125–152.
- GRAZZI, M., C. PICCARDO, AND C. VERGARI (2017): “Building a firm level dataset for the analyses of industrial dynamics and demography,” Quaderni - Working Papers DSE 2003, University of Bologna.
- GREENHALGH, C. AND M. ROGERS (2006): “Market Valuation of UK Intellectual Property: Manufacturing, Utility and Financial Services Firms,” in *The Management of Intellectual Property*, Edward Elgar Publishing, Chapters, chap. 7.
- (2012): “Trade Marks and Performance in Services and Manufacturing Firms: Evidence of Schumpeterian Competition Through Innovation,” *Australian Economic Review*, 45, 50–76.
- GRIFFITH, R., E. HUERGO, J. MAIRESSE, AND B. PETERS (2006): “Innovation and productivity across four European countries,” *Oxford review of economic policy*, 22, 483–498.
- GRILICHES, Z. (1981): “Market value, R&D, and patents,” *Economics Letters*, 7, 183–187.
- GUZMAN, J. AND S. STERN (2015): “Where is silicon valley?” *Science*, 347, 606–609.
- HALL, B. H. (2000): “Innovation and market value,” In: Barrell, R., Mason, G., O'Mahony, M. (Eds.), *Productivity, Innovation and Economic Performance*, 33, 13–33.
- HALL, B. H., C. HELMERS, M. ROGERS, AND V. SENA (2013a): “The importance (or not) of patents to UK firms,” *Oxford Economic Papers*, 65, 603–629.
- HALL, B. H., F. LOTTI, AND J. MAIRESSE (2009): “Innovation and productivity in SMEs: empirical evidence for Italy,” *Small Business Economics*, 33, 13–33.
- (2013b): “Evidence on the impact of R&D and ICT investments on innovation and productivity in Italian firms,” *Economics of Innovation and New Technology*, 22, 300–328.
- HALL, B. H. AND J. MAIRESSE (2006): “Empirical studies of innovation in the knowledge-driven economy,” *Economics of innovation and new technology*, 15, 289–299.
- HECKMAN, J. (1979): “Sample Selection Bias as a Specification Error,” *Econometrica*, 47, 153–61.
- HECKMAN, J. J. (1976): “The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models,” in *Annals of Economic and Social Measurement, Volume 5, number 4*, NBER, 475–492.

- HELMERS, C. AND M. ROGERS (2010): “Innovation and the survival of new firms in the UK,” *Review of Industrial Organization*, 36, 227–248.
- HESHMATI, A. AND H. KIM (2011): “The R&D and productivity relationship of Korean listed firms,” *Journal of Productivity Analysis*, 36, 125–142.
- JENSEN, P. H., E. WEBSTER, AND H. BUDELMEYER (2008): “Innovation, technological conditions and new firm survival,” *Economic Record*, 84, 434–448.
- LEIPONEN, A. AND J. BYMA (2009): “If you cannot block, you better run: Small firms, cooperative innovation, and appropriation strategies,” *Research Policy*, 38, 1478–1488.
- LLERENA, P. AND V. MILLOT (2013): “Are Trade Marks and Patents Complementary or Substitute Protections for Innovation,” Working Papers of BETA 2013-01, Bureau d’Economie Théorique et Appliquée, UDS, Strasbourg.
- LOTTI, F. AND G. MARIN (2013): “Matching of PATSTAT applications to AIDA firms: discussion of the methodology and results,” *Questioni di Economia e Finanza (Occasional Papers)* 166, Bank of Italy, Economic Research and International Relations Area.
- LYBBERT, T. J. AND N. J. ZOLAS (2014): “Getting patents and economic data to speak to each other: an algorithmic links with probabilities approach for joint analyses of patenting and economic activity,” *Research Policy*, 43, 530–542.
- MALERBA, F. AND L. ORSENIGO (1999): “Technological entry, exit and survival: an empirical analysis of patent data,” *Research Policy*, 28, 643–660.
- MARIN, G. (2014): “Do eco-innovations harm productivity growth through crowding out? Results of an extended CDM model for Italy,” *Research Policy*, 43, 301–317.
- MARIN, G. AND F. LOTTI (2017): “Productivity effects of eco-innovations using data on eco-patents,” *Industrial and Corporate Change*, 26, 125–148.
- MARKIDES, C. C. AND P. J. WILLIAMSON (1996): “Corporate Diversification And Organizational Structure: A Resource-Based View,” *Academy of Management Journal*, 39, 340–367.
- MARTIN, L. AND T. U. NGUYEN-THI (2015): “The Relationship between Innovation and Productivity based on R&D and ICT Use,” *Revue économique*, 66, 1105–1130.
- MAZZUCATO, M. AND S. PARRIS (2015): “High-growth firms in changing competitive environments: the US pharmaceutical industry (1963 to 2002),” *Small Business Economics*, 44, 145–170.
- MENDONÇA, S., T. S. PEREIRA, AND M. M. GODINHO (2004): “Trademarks as an indicator of innovation and industrial change,” *Research Policy*, 33, 1385–1404.
- OHIM (2015): “Intellectual property rights and firm performance in Europe: an economic analysis,” Tech. rep., Office for Harmonization in the Internal Market (OHIM).
- OSSERVATORIO ITALIANO BREVETTI (2014): “Analisi del brevetto italiano. Dati statistici e trends degli ultimi anni,” Technical report, OIB.

- PAVITT, K. (1984): “Sectoral patterns of technical change: towards a taxonomy and a theory,” *Research policy*, 13, 343–373.
- PELLEGRINO, G. AND M. PIVA (2014): “Do innovative inputs lead to different innovative outputs in mature and young firms?” Tech. Rep. dises1497, Università Cattolica del Sacro Cuore, Dipartimenti e Istituti di Scienze Economiche (DISCE).
- PENROSE, E. T. (1959): *The theory of the growth of the firm*, Oxford: Blackwell, 3rd ed.
- PERANI, G. AND V. CIRILLO (2015): “Matching industry classifications. A method for converting Nace Rev. 2 to Nace Rev. 1,” Working Paper WP-EMS 2012-02, University of Urbino.
- RAYMOND, W., P. MOHNEN, F. PALM, AND S. S. VAN DER LOEFF (2010): “Persistence of innovation in Dutch manufacturing: Is it spurious?” *The Review of Economics and Statistics*, 92, 495–504.
- ROGERS, M., C. HELMERS, AND C. GREENHALGH (2007): “An Analysis of the Characteristics of Small and Medium Enterprises That Use Intellectual Property,” Tech. rep., London: Report for the UK Intellectual Property Office.
- SANDNER, P. G. AND J. BLOCK (2011): “The market value of R&D, patents, and trademarks,” *Research Policy*, 40, 969–985.
- SANTARELLI, E., L. KLOMP, AND A. R. THURIK (2006): “Gibrats Law: An overview of the empirical literature,” in *Entrepreneurship, growth, and innovation*, Springer, 41–73.
- SCHAUTSCHICK, P. AND C. GREENHALGH (2016): “Empirical studies of trade marks—the existing economic literature,” *Economics of Innovation and New Technology*, 25, 358–390.
- TEECE, D. J., R. RUMELT, G. DOSI, AND S. WINTER (1994): “Understanding corporate coherence: Theory and evidence,” *Journal of Economic Behavior & Organization*, 23, 1–30.
- TOIVANEN, O., P. STONEMAN, AND D. BOSWORTH (2002): “Innovation and the Market Value of UK Firms, 1989-1995,” *Oxford Bulletin of Economics and Statistics*, 64, 39–61.
- WAGNER, S. AND I. COCKBURN (2010): “Patents and the survival of Internet-related IPOs,” *Research Policy*, 39, 214–228.
- XU, B. AND E. P. CHIANG (2005): “Trade, patents and international technology diffusion,” *The Journal of International Trade & Economic Development*, 14, 115–135.
- ZHOU, H., P. G. SANDNER, S. L. MARTINELLI, AND J. H. BLOCK (2016): “Patents, trademarks, and their complementarity in venture capital funding,” *Technovation*, 47, 14–22.
- ZOLAS, N., T. J. LYBBERT, AND P. BHATTACHARYYA (2017): “An Algorithmic Links with Probabilities Concordance for Trademarks with an Application Towards Bilateral IP Flows,” *The World Economy*, 40, 1184–1213.



Table 1: Number of elements in the set of 3-digit IPC and 2-digit NICE codes.

ID	PAT	IPC	ID	TM	NICE
1	1	A61	1	1	2
1	1	H02	1	2	2
1	2	C09	1	3	2
1	3	A61	1	4	2
1	4	A61	1	5	2
			1	6	2
			1	7	2
			1	8	2
			1	9	35

Table 2: Moving from IPC and NICE to ISIC codes.

IPC	ISIC	$P_{IPC}(ISIC)$	NICE	ISIC	$P_{NICE}(ISIC)$
A61	10	0.0361579	2	2	0.0491944
A61	20	0.0505005	2	20	0.1203539
A61	21	0.8616829	2	25	0.8065651
A61	32	0.0302412	2	41	0.0238867
A61	36	0.0214176	35	46	0.1524157
C09	20	0.9659607	35	63	0.0744019
C09	23	0.0340393	35	69	0.0240318
H02	24	0.0637551	35	70	0.0710176
H02	26	0.0888389	35	73	0.4864157
H02	27	0.8212442	35	78	0.0490504
H02	28	0.0261618	35	82	0.0307544
			35	90	0.0424586
			35	94	0.0694538

Table 3: Number of firms, trademarks and patents

Year	Firms	Firms with tm(%)*	Firms with pat (%)*	Firms with tm and pat (%)*	Firms with $conc_{i,t} > 0$ (%)**	Num of tm	Num of pat
2006	116507	4732 (4.062)	8616 (7.359)	1798 (1.543)	1623 (90.267)	13190	77312
2007	123007	5340 (4.341)	8884 (7.222)	2018 (1.641)	1822 (90.287)	15582	82797
2008	129334	6040 (4.670)	9100 (7.036)	2237 (1.730)	2013 (89.987)	18508	87536
2009	135091	6679 (4.944)	9261 (6.855)	2389 (1.768)	2147 (89.870)	21243	91482
2010	141070	7396 (5.243)	9506 (6.738)	2601 (1.844)	2347 (90.235)	24442	95635
2011	138641	8011 (5.778)	9591 (6.918)	2772 (1.999)	2499 (90.152)	27485	99510
2012	135299	8616 (6.368)	9419 (6.962)	2891 (2.137)	2606 (90.142)	30384	101552
2013	132731	9070 (6.833)	9062 (6.827)	2927 (2.205)	2644 (90.331)	32612	98872
2014	129253	8863 (6.857)	8608 (6.660)	2830 (2.190)	2556 (90.318)	31740	93229

Notes. We only consider firms operating in manufacturing sectors (we exclude firms operating in the following 2-digit ATECO 2007 code: 12 and 33).

\* In brackets, percentage of total firms.

\*\* In brackets, as a percentage of firms with trademarks and patents.

Table 4: Number of firms, trademarks and patents for micro, small, medium and large firms. Year 2014

Size	Firms(%)*	Firms with tm(%)*	Firms with pat (%)*	Firms with tm and pat (%)*	Firms with $conc_{i,t} > 0$ (%)*	Num of tm (%)**	Num of pat (%)**
0-9	73641 (57.996)	1668 (18.944)	1673 (19.556)	223 (7.911)	187 (7.345)	3076 (9.736)	5910 (6.366)
10-49	44148 (34.769)	3805 (43.214)	3805 (44.477)	933 (33.097)	805 (31.618)	9356 (29.612)	20141 (21.695)
50-249	8016 (6.313)	2653 (30.131)	2432 (28.428)	1208 (42.852)	1114 (43.755)	11246 (35.594)	27599 (29.728)
250 and more	1172 (0.923)	679 (7.712)	645 (7.539)	455 (16.140)	440 (17.282)	7917 (25.058)	39187 (42.211)

Notes. We only consider firms operating in manufacturing sectors (we exclude firms operating in the following 2-digit ATECO 2007 code: 12 and 33). The number of firms in this Table differ from the previous Table because 2276 firms do not have information on their size, measured in terms of workers, in 2014.

\* In brackets, percentage of total firms, total firms with trademarks, patents, trademarks and patents, concordant trademarks and patents, respectively.

\*\* In brackets, percentage of total number of trademarks and patents, respectively.

Table 5: POOLED OLS estimates, whole sample

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(totrev <sub>i,t</sub> )	ln(totrev <sub>i,t</sub> )	ln(totrev <sub>i,t</sub> )	growth <sub>i,t</sub>	growth <sub>i,t</sub>	growth <sub>i,t</sub>
ln(workers <sub>i,t-1</sub> )	0.898*** (0.000920)	0.965*** (0.00437)	0.965*** (0.00437)	-0.0183*** (0.000616)	-0.00521* (0.00310)	-0.00523* (0.00310)
ln(LP <sub>i,t-1</sub> )	0.742*** (0.00166)	0.782*** (0.0138)	0.782*** (0.0138)	-0.0538*** (0.00123)	-0.0110 (0.00960)	-0.0110 (0.00958)
tm <sub>i,t-1</sub>	0.306*** (0.00434)			0.0525*** (0.00253)		
pat <sub>i,t-1</sub>	0.169*** (0.00367)			0.0358*** (0.00234)		
both <sub>i,t-1</sub>	0.322*** (0.00480)			0.0759*** (0.00303)		
conc <sub>i,t-1</sub>			-0.0210 (0.0275)			-0.0153 (0.0185)
_cons	2.772*** (0.00873)	2.599*** (0.0797)	2.605*** (0.0789)	0.286*** (0.00584)	0.123** (0.0572)	0.127** (0.0552)
N	686252	17929	17929	684323	17917	17917
r <sup>2</sup>	0.781	0.877	0.877	0.0248	0.0444	0.0445
F	55579.6	2069.7	2008.3	357.6	28.30	27.45

Notes. Standard errors in parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

In each specification we include ATECO 2 digit sectors, geographical areas and years dummy variables.

Table 6: CDM MODEL, total revenues as firms performance measure

	(1)	(2)	(3)	(4)	(5)
	$rd_{i,2008}$	$rd\_int_{i,2008}$	$inno_{i,2009}$	$ln(LP_{i,2010})$	$ln(totrev_{i,2011})$
$ln(workers_{i,2007})$	0.258*** (0.0142)	-0.230*** (0.0512)			
$ln(mkt\_share_{i,2007})$	0.108*** (0.0119)	0.0844** (0.0400)			
$\_cons$	-1.684*** (0.0597)	0.0990 (0.296)			
$\widehat{rd\_int_{i,2008}}$			0.771*** (0.119)		
$ln(workers_{i,2008})$			0.548*** (0.0193)		
$\_cons$			-2.436*** (0.0463)		
$\widehat{inno_{i,2009}}$				0.146* (0.0754)	
$ln(cap\_emp_{i,2009})$				0.158*** (0.00355)	
$ln(workers_{i,2009})$				0.0532*** (0.00812)	
$\_cons$				2.856*** (0.0263)	
$\widehat{ln(LP_{i,2010})}$					1.394*** (0.0279)
$ln(workers_{i,2010})$					0.772*** (0.00403)
$\_cons$					0.563*** (0.104)
$\lambda$		0.599 (0.105)			
$\chi^2(1)$		33.71***			
N	43805	43805	43805	43805	43805
Censored		39390			
Uncensored		4415			
$PseudoR^2$			0.2132		
$R^2$				0.0900	0.6572

Notes. Coefficients and standard errors are reported.

Robust standard errors in parentheses in columns 1 and 2. Bootstrapped standard errors (50 repetitions) in parentheses in columns 3, 4 and 5.

In each specification we include ATECO 2 digit sectors and geographical areas dummy variables.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table 7: CDM MODEL, growth as firms performance measure

	(1)	(2)	(3)	(4)	(5)
	$rd_{i,2008}$	$rd\_int_{i,2008}$	$inno_{i,2009}$	$ln(LP_{i,2010})$	$growth_{i,2011}$
$ln(workers_{i,2007})$	0.258*** (0.0142)	-0.230*** (0.0512)			
$ln(mkt\_share_{i,2007})$	0.107*** (0.0119)	0.0842** (0.0400)			
$\_cons$	-1.683*** (0.0598)	0.0997 (0.296)			
$\widehat{rd\_int}_{i,2008}$			0.767*** (0.118)		
$ln(workers_{i,2008})$			0.547*** (0.0185)		
$\_cons$			-2.437*** (0.0534)		
$\widehat{inno}_{i,2009}$				0.158** (0.0764)	
$ln(cap\_emp_{i,2009})$				0.157*** (0.00359)	
$ln(workers_{i,2009})$				0.0505*** (0.00767)	
$\_cons$				2.868*** (0.0305)	
$\widehat{ln(LP)}_{i,2010}$					0.0456*** (0.0110)
$ln(workers_{i,2010})$					0.00986*** (0.00152)
$\_cons$					-0.275*** (0.0408)
$lambda$		0.598 (0.105)			
$\chi^2(1)$		33.70***			
N	43731	43731	43731	43731	43731
Censored		39316			
Uncensored		4415			
$PseudoR^2$			0.2128		
$R^2$				0.0893	0.0137

Notes. Coefficients and standard errors are reported.

Robust standard errors in parentheses in columns 1 and 2. Bootstrapped standard errors (50 repetitions) in parentheses in columns 3, 4 and 5.

In each specification we include ATECO 2 digit sectors and geographical areas dummy variables.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

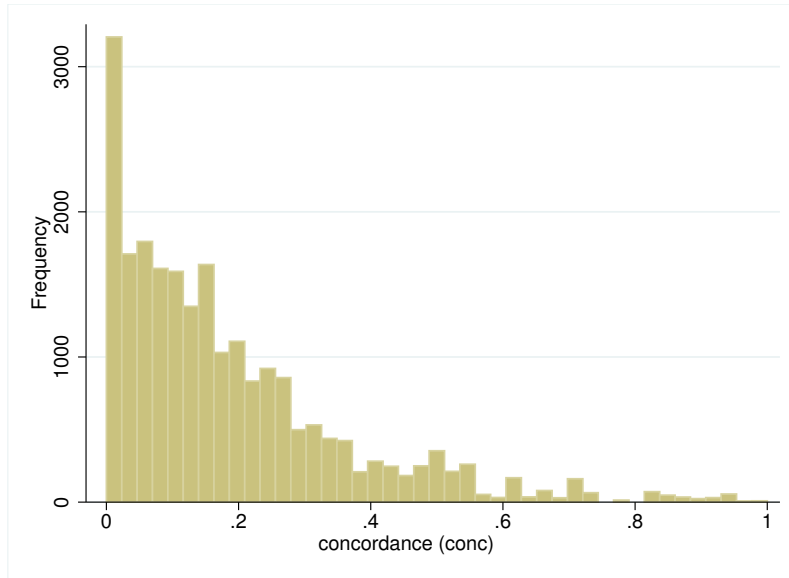


Figure 1: Distribution of the measure of IP concordance

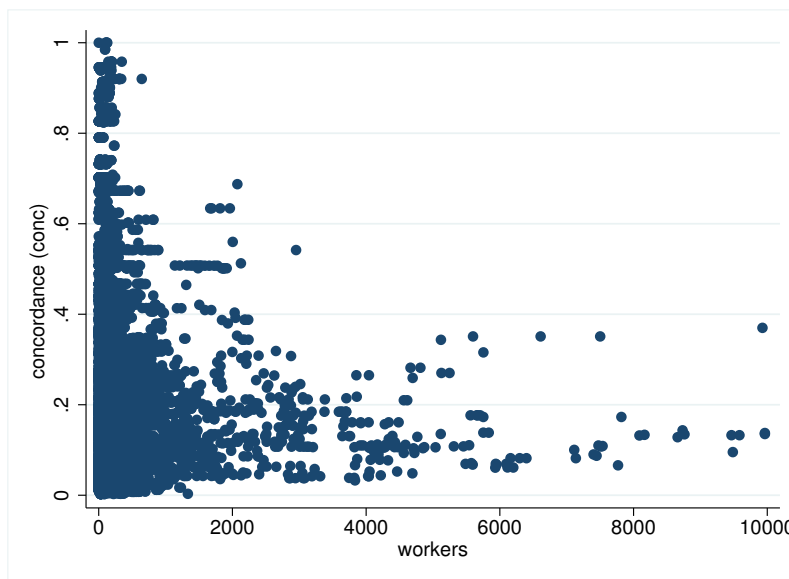


Figure 2: Firm size (number of workers) and concordance measure (stricly positive values)

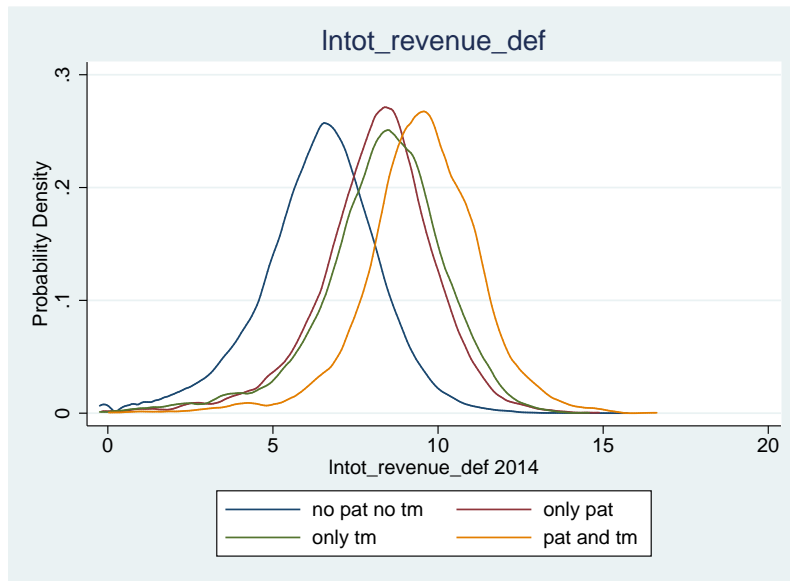


Figure 3: Empirical distribution of Total Revenues (in log) in 2014, *innovative vs non-innovative* firms

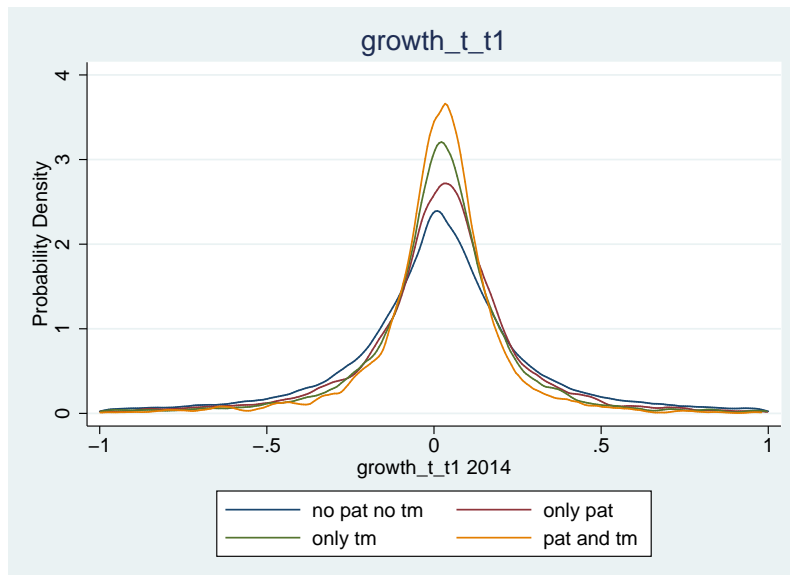


Figure 4: Empirical distribution of growth rates 2013-2014, *innovative vs non-innovative* firms

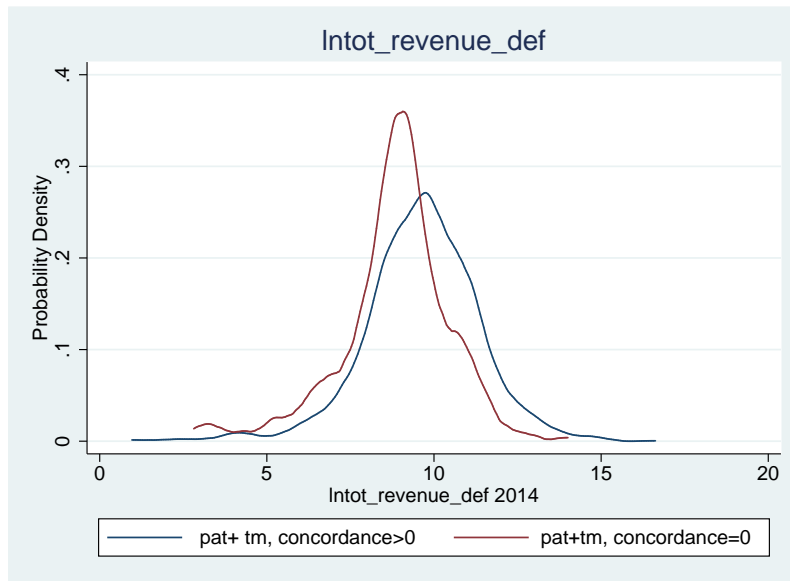


Figure 5: Empirical distribution of Total Revenues (in log) in 2014, concordant patents and trademarks vs non-concordant

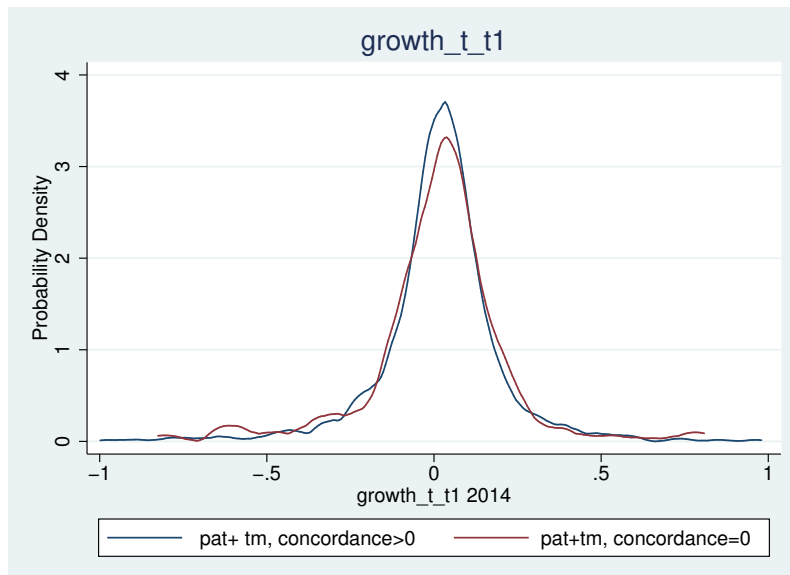


Figure 6: Empirical distribution of growth rates 2013-2014, concordant patents and trademarks vs non-concordant



## Appendix - Sectoral results

Table 8: Number of firms, trademarks and patents for 2 digit ATECO sectors. Year 2014

ATECO	Firms(%)*	Firms with tm(%)*	Firms with pat (%)*	Firms with tm and pat (%)*	Firms with concordance (%)*	Num of tm (%)**	Num of pat (%)**
10 <sup>LT</sup>	11560 (8.944)	954 (10.764)	210 (2.440)	97 (3.428)	79 (3.091)	4431 (13.960)	1505 (1.614)
11 <sup>LT</sup>	1754 (1.357)	335 (3.780)	29 (0.337)	15 (0.530)	10 (0.391)	2203 (6.941)	99 (0.106)
13 <sup>LT</sup>	5012 (3.878)	336 (3.791)	218 (2.533)	62 (2.191)	48 (1.878)	1087 (3.425)	1028 (1.103)
14 <sup>LT</sup>	7414 (5.736)	721 (8.135)	117 (1.359)	48 (1.696)	38 (1.487)	2545 (8.018)	684 (0.734)
15 <sup>LT</sup>	5262 (4.071)	467 (5.269)	156 (1.812)	72 (2.544)	60 (2.347)	1646 (5.186)	1081 (1.160)
16 <sup>LT</sup>	4529 (3.504)	102 (1.151)	117 (1.359)	34 (1.201)	32 (1.252)	240 (0.756)	441 (0.473)
17 <sup>LT</sup>	2229 (1.725)	105 (1.185)	147 (1.708)	32 (1.131)	27 (1.056)	528 (1.664)	896 (0.961)
18 <sup>LT</sup>	4671 (3.614)	67 (0.756)	75 (0.871)	10 (0.353)	9 (0.352)	108 (0.340)	315 (0.338)
19 <sup>MLT</sup>	289 (0.224)	17 (0.192)	16 (0.186)	7 (0.247)	7 (0.274)	74 (0.233)	46 (0.049)
20 <sup>MHT</sup>	3579 (2.769)	510 (5.754)	345 (4.008)	157 (5.548)	148 (5.790)	2777 (8.749)	3763 (4.036)
21 <sup>HT</sup>	586 (0.453)	146 (1.647)	164 (1.905)	89 (3.145)	88 (3.443)	1355 (4.269)	5252 (5.633)
22 <sup>MLT</sup>	5921 (4.581)	459 (5.179)	708 (8.225)	222 (7.845)	187 (7.316)	1514 (4.770)	6299 (6.756)
23 <sup>MLT</sup>	7430 (5.748)	282 (3.182)	292 (3.392)	86 (3.039)	81 (3.169)	914 (2.880)	1576 (1.690)
24 <sup>MLT</sup>	2039 (1.578)	106 (1.196)	163 (1.894)	52 (1.837)	48 (1.878)	257 (0.810)	1085 (1.164)
25 <sup>MLT</sup>	26951 (20.851)	799 (9.015)	1463 (16.996)	347 (12.261)	308 (12.050)	1888 (5.948)	10301 (11.049)
26 <sup>HT</sup>	4749 (3.674)	409 (4.615)	532 (6.180)	171 (6.042)	154 (6.025)	1228 (3.869)	12125 (13.006)
27 <sup>MHT</sup>	5576 (4.314)	478 (5.393)	573 (6.657)	208 (7.350)	189 (7.394)	1752 (5.520)	8211 (8.807)
28 <sup>MHT</sup>	15075 (11.663)	1376 (15.525)	2311 (26.847)	758 (26.784)	721 (28.208)	3557 (11.207)	27394 (29.384)
29 <sup>MHT</sup>	1687 (1.305)	172 (1.941)	200 (2.323)	70 (2.473)	64 (2.504)	619 (1.950)	3513 (3.768)
30 <sup>MHT</sup>	2081 (1.610)	148 (1.670)	114 (1.324)	43 (1.519)	38 (1.487)	616 (1.941)	2285 (2.451)
31 <sup>LT</sup>	5858 (4.532)	389 (4.389)	295 (3.4227)	103 (3.640)	93 (3.638)	945 (2.977)	1953 (2.095)
32 <sup>LT</sup>	5001 (3.869)	485 (5.472)	363 (4.217)	147 (5.194)	127 (4.969)	1456 (4.587)	3377 (3.622)

Notes. We only consider firms operating in manufacturing sectors (we exclude firms operating in the following 2-digit ATECO 2007 code: 12 and 33).

\* % is the share of total firms, total firms with trademarks, patents, trademarks and patents, concordant trademarks and patents, respectively.

\*\* % is the share of total number of trademarks and patents, respectively.

HT: high technology sectors; MHT: medium high technology sectors; MLT: medium low technology sectors; LT: low technology sectors. We refer to Eurostat aggregation of the manufacturing industry according to technological intensity.

Table 9: ATECO sectors

ATECO codes	Definitions
10 <sup>LT</sup>	Manufacture of food products
11 <sup>LT</sup>	Manufacture of beverages
13 <sup>LT</sup>	Manufacture of textiles
14 <sup>LT</sup>	Manufacture of wearing apparel
15 <sup>LT</sup>	Manufacture of leather and related products
16 <sup>LT</sup>	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
17 <sup>LT</sup>	Manufacture of paper and paper products
18 <sup>LT</sup>	Printing and reproduction of recorded media
19 <sup>MLT</sup>	Manufacture of coke and refined petroleum products
20 <sup>MHT</sup>	Manufacture of chemicals and chemical products
21 <sup>HT</sup>	Manufacture of basic pharmaceutical products and pharmaceutical preparations
22 <sup>MLT</sup>	Manufacture of rubber and plastic products
23 <sup>MLT</sup>	Manufacture of other non-metallic mineral products
24 <sup>MLT</sup>	Manufacture of basic metals
25 <sup>MLT</sup>	Manufacture of fabricated metal products, except machinery and equipment
26 <sup>HT</sup>	Manufacture of computer, electronic and optical products
27 <sup>MHT</sup>	Manufacture of electrical equipment
28 <sup>MHT</sup>	Manufacture of machinery and equipment n.e.c.
29 <sup>MHT</sup>	Manufacture of motor vehicles, trailers and semi-trailers
30 <sup>MHT</sup>	Manufacture of other transport equipment
31 <sup>LT</sup>	Manufacture of furniture
32 <sup>LT</sup>	Other manufacturing

*Notes.* We only consider firms operating in manufacturing sectors (we exclude firms in manufacture of tobacco products, ATECO 12, and repair and installation of machinery and equipment, ATECO 33, sectors).

*HT: high technology sectors; MHT: medium high technology sectors; MLT: medium low technology sectors; LT: low technology sectors.* We refer to Eurostat aggregation of the manufacturing industry according to technological intensity.

Table 10: POOLED OLS estimates, high technology sectors

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(totrev <sub>i,t</sub> )	ln(totrev <sub>i,t</sub> )	ln(totrev <sub>i,t</sub> )	growth <sub>i,t</sub>	growth <sub>i,t</sub>	growth <sub>i,t</sub>
ln(workers <sub>i,t-1</sub> )	0.908*** (0.00348)	0.946*** (0.0105)	0.945*** (0.0106)	0.000197 (0.00233)	-0.00619 (0.00791)	-0.00646 (0.00795)
ln(LP <sub>i,t-1</sub> )	0.705*** (0.00704)	0.731*** (0.0457)	0.729*** (0.0459)	-0.0429*** (0.00498)	-0.00895 (0.0173)	-0.00963 (0.0176)
tm <sub>i,t-1</sub>	0.264*** (0.0210)			0.0374** (0.0146)		
pat <sub>i,t-1</sub>	0.0971*** (0.0155)			0.0212** (0.0103)		
both <sub>i,t-1</sub>	0.361*** (0.0177)			0.0524*** (0.0112)		
conc <sub>i,t-1</sub>			0.117 (0.0823)			0.0419 (0.0568)
_cons	2.484*** (0.0338)	2.433*** (0.238)	2.426*** (0.237)	0.178*** (0.0230)	0.155 (0.136)	0.153 (0.134)
N	29798	1596	1596	29704	1595	1595
r <sup>2</sup>	0.831	0.904	0.904	0.0120	0.0119	0.0124
F	9456.4	923.9	860.6	18.85	1.805	1.699

Notes. Standard errors in parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

In each specification we include geographical areas and years dummy variables.

Table 11: POOLED OLS estimates, medium-high technology sectors

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(totrev <sub>i,t</sub> )	ln(totrev <sub>i,t</sub> )	ln(totrev <sub>i,t</sub> )	growth <sub>i,t</sub>	growth <sub>i,t</sub>	growth <sub>i,t</sub>
ln(workers <sub>i,t-1</sub> )	0.914*** (0.00181)	0.974*** (0.00666)	0.974*** (0.00666)	-0.0113*** (0.00129)	-0.00485 (0.00444)	-0.00484 (0.00444)
ln(LP <sub>i,t-1</sub> )	0.710*** (0.00354)	0.830*** (0.0182)	0.830*** (0.0182)	-0.0571*** (0.00273)	-0.00539 (0.0133)	-0.00538 (0.0133)
tm <sub>i,t-1</sub>	0.269*** (0.00853)			0.0526*** (0.00537)		
pat <sub>i,t-1</sub>	0.115*** (0.00620)			0.0316*** (0.00424)		
both <sub>i,t-1</sub>	0.275*** (0.00790)			0.0696*** (0.00522)		
conc <sub>i,t-1</sub>			-0.0720* (0.0403)			-0.0315 (0.0238)
_cons	2.484*** (0.0173)	1.936*** (0.0952)	1.951*** (0.0948)	0.260*** (0.0130)	0.0506 (0.0728)	0.0569 (0.0720)
N	154962	7731	7731	154495	7726	7726
r <sup>2</sup>	0.803	0.867	0.867	0.0273	0.0628	0.0630
F	35563.2	2025.6	1870.5	232.6	46.19	42.36

Notes. Standard errors in parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

In each specification we include geographical areas and years dummy variables.

Table 12: POOLED OLS estimates, medium-low technology sectors

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(totrev <sub>i,t</sub> )	ln(totrev <sub>i,t</sub> )	ln(totrev <sub>i,t</sub> )	growth <sub>i,t</sub>	growth <sub>i,t</sub>	growth <sub>i,t</sub>
ln(workers <sub>i,t-1</sub> )	0.919*** (0.00165)	1.004*** (0.00829)	1.004*** (0.00828)	-0.0230*** (0.00111)	-0.00562 (0.00607)	-0.00558 (0.00605)
ln(LP <sub>i,t-1</sub> )	0.743*** (0.00301)	0.859*** (0.0240)	0.859*** (0.0240)	-0.0621*** (0.00227)	-0.00218 (0.0211)	-0.00207 (0.0210)
tm <sub>i,t-1</sub>	0.311*** (0.00940)			0.0606*** (0.00528)		
pat <sub>i,t-1</sub>	0.210*** (0.00581)			0.0427*** (0.00344)		
both <sub>i,t-1</sub>	0.322*** (0.00898)			0.0713*** (0.00518)		
conc <sub>i,t-1</sub>			-0.0297 (0.0612)			-0.0139 (0.0334)
_cons	2.334*** (0.0143)	2.048*** (0.120)	2.052*** (0.121)	0.293*** (0.0104)	0.0705 (0.0965)	0.0721 (0.0972)
N	231639	4485	4485	231072	4481	4481
r <sup>2</sup>	0.772	0.872	0.872	0.0356	0.0531	0.0531
F	38535.1	1519.7	1406.5	442.4	24.47	22.47

Notes. Standard errors in parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

In each specification we include geographical areas and years dummy variables.

Table 13: POOLED OLS estimates, low technology sectors

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(totrev <sub>i,t</sub> )	ln(totrev <sub>i,t</sub> )	ln(totrev <sub>i,t</sub> )	growth <sub>i,t</sub>	growth <sub>i,t</sub>	growth <sub>i,t</sub>
ln(workers <sub>i,t-1</sub> )	0.893*** (0.00154)	0.957*** (0.00912)	0.958*** (0.00904)	-0.0196*** (0.000952)	-0.00166 (0.00644)	-0.00158 (0.00633)
ln(LP <sub>i,t-1</sub> )	0.780*** (0.00261)	0.749*** (0.0303)	0.750*** (0.0301)	-0.0451*** (0.00183)	-0.0135 (0.0218)	-0.0134 (0.0215)
tm <sub>i,t-1</sub>	0.406*** (0.00612)			0.0572*** (0.00334)		
pat <sub>i,t-1</sub>	0.166*** (0.00873)			0.0281*** (0.00494)		
both <sub>i,t-1</sub>	0.369*** (0.00981)			0.0784*** (0.00600)		
conc <sub>i,t-1</sub>			0.0518 (0.0506)			0.0120 (0.0365)
_cons	2.427*** (0.0124)	2.611*** (0.152)	2.602*** (0.150)	0.215*** (0.00815)	0.0363 (0.106)	0.0343 (0.102)
N	269853	4117	4117	269052	4115	4115
r <sup>2</sup>	0.749	0.868	0.868	0.0153	0.0212	0.0213
F	46006.9	1150.8	1089.9	196.5	11.32	10.37

Notes. Standard errors in parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

In each specification we include geographical areas and years dummy variables.