Crossing the Hurdle of Patenting. Effect of the First Patent on Productivity and Size of Italian Manufacturing SMEs

Giovanni Marin*

May 20, 2013

Abstract

The aim of this paper is to investigate the extent to which applying for the first patent affects firms' performance. It has been acknowledged that an hurdle exists for firms applying for their first patent. Patenting brings about explicit (application fee) and implicit costs (e.g. legal and technical support). Implicit costs tend to sharply decrease with experience in patenting, thus possibly giving great importance to the 'first patent'.

This paper investigates the effect of the first patent on measures of productivity and size for a large unbalanced panel of Italian manufacturing small and medium enterprises (SMEs). Results based on both a difference-in-differences approach and a propensity score matching approach show a generally positive and strong effect of applying for the first patent on productivity and firm size. Moreover, preliminary evidence shows that the effect is likely to occur since the first year of patenting but it is even bigger in the years after the first patent.

Keywords: patent, productivity, firm size, difference-in-differences, propensity score matching

JEL: O34, L25.

*CERIS-CNR, Via Bassini, 15, Milano, Italy, e-mail: g.marin@ceris.cnr.it.

1 Introduction: patent (propensity) and firms' performance

The aim of the paper is to assess the effect of 'crossing the hurdle' of patenting on firms' performance for a large panel of Italian small and medium enterprises (SMEs). The relationship between patenting behaviour of firms and their performance has been widely investigated (refer to Hall (2011) for a recent review of the literature on the relationship between innovation, including patents, and productivity). The role of patents for firms' performance is twofold. On the one hand, patents are a powerful indicator of innovative output (Griliches et al., 1988; Griliches, 1990) which, in turn, is likely to be the primary driver of firms' productivity and firms' growth. On the other hand, patenting is one alternative among many others available to firms to appropriate of the returns from innovations (Hall et al., 2012).

The relationship between patent applications and measures of firms' performance has been investigated from several perspectives. A first set of studies (Griliches, 1981; Pakes, 1985) based on U.S. data found a positive relationship between market value of the firm and number of patents. Crepon et al. (1998) use a structural model to describe the various steps of the innovation process at the firm level in a comprehensive way, from the drivers of investments in R&D to the role of R&D as a input in a knowledge production function and, finally, to the effect of knowledge on firm productivity. They apply this structural model (so-called CDM model) to a sample of French manufacturing firms and one of their indicators of innovation output is the count of patent applications. Their findings support the positive effect of successfully applying for patents on firms' performance. Looking at Italy, similar models based on patent applications as a measure of innovative output find a strong and positive effect of patent applications on firms' productivity (Marin, 2012; Lotti and Marin, 2013a).

While the general patent performance of firms has been found to be strongly related to firms' performance, Lotti and Schivardi (2005) highlight that the difference between patenting and non-patenting firms (i.e. patent propensity) is likely to be very important. In this light, crossing the hurdle of patenting for the first time makes subsequent patent applications 'easier'. This occurs because patenting brings about both explicit (application fee) and implicit costs (e.g. legal and technical support). Implicit costs tend to sharply decrease with experience in patenting, thus possibly giving great importance to the 'first patent'. This results in a quite persistent patenting behaviour for the small share of already patenting firms.

Despite the emphasis in highlighting the relevance of patent propensity 'in general' rather than actual intensity of patenting activity, just the article by Balasubramanian and Sivadasan (2011) explicitly deals with the issue of estimating the effect the first patent on firms' performance. Balasubra-

manian and Sivadasan (2011) investigate the effect of patenting on a large panel of US manufacturing firms by using US Economic Census data further extended to assignee information available in the NBER Patent Data. By means of a difference-in-differences approach they estimate the effect of patenting activity of manufacturing firms on a variety of measures of firms' performance (size, skill composition, productivity, capital intensity and number of products). In addition to the effect of general patenting activity, section VII of their article is devoted to a careful investigation of the effect of the first patent on firms' performance, complemented by a great variety of robustness checks. Balasubramanian and Sivadasan (2011) find a rather strong and robust effect of the first patent on firms' performance in terms of productivity, size and number of products, confirming the importance of crossing the hurdle of patenting.

The paper is organized as follows. Section 2 describes data sources and provides definitions for the variables of interest. Section 3 describe the two empirical approaches (difference-in-differences and propensity score matching) and the results. Section 4 outlines some possible extensions.

2 Data and variables

I use balance sheet information available in the AIDA database (the Italian counterpart of the AMADEUS database, by Bureau van Dijk), further merged with patent applications at the European Patent Office (EPO) as described in Lotti and Marin (2013b). The final dataset consists in an unbalanced panel of 34,023 Italian manufacturing SMEs for the period 2002-2007. The average number of observations for each firm is about 5.7. I use the Eurostat definition for SMEs (EUROSTAT)[1]. Two categories of indicators have been created: i) firm size; ii) firm productivity level. Firm size is measured in terms of real value added (VA), real sales and employee headcount (L). Firm productivity level is measured in terms of labour productivity (real value added per employee, VA/L) and in terms of Total Factor Productivity (TFP) estimated with the Levinsohn and Petrin (2003) (LP henceforth) structural model (TFP_LP) or with a simple production function with fixed effects[2] (TFP_FE). All monetary variables have been translated into real prices (euros in 2000) by means of sector-specific deflators. I excluded few outlying observations, identified as values three standard deviations above the third quartile or three standard deviations below the first quartile. Fi-

[1] SMEs are defined as 'as having less than 250 persons employed. They should also have an annual turnover of up to EUR 50 million, or a balance sheet total of no more than EUR 43 million' (Commission Recommendation of 6 May 2003).

[2] TFP has been estimated by means of sector-specific (2 digit Nace rev. 2) Cobb-Douglas production functions with real value added, labour input (employee headcount) and fixed capital stock (built with the perpetual inventory method). Results are available upon request.

nally, note that due to missing value for specific variables the sample of analysis changes depending on the indicator being evaluated.

The sample does not include all firms which applied for EPO patents before 2002. Out of 34,023 firms in the sample, 1,582 firms (4.65 percent) applied for their first EPO patent in the considered period. I define these firms as treated firms while a dummy variable for the year of first patent is defined as ‘First patent’.

Table 1 reports some descriptive statistics for the full sample while table 2 shows the distribution of observations by year and sector (Nace rev. 2).

3 Empirical approach and results

Table 3 reports simple OLS estimates (with year and sector fixed effects) in which, for each outcome variable, I compare the outcome of firms that are not going to apply for patents with the outcome of patenting firms before their first patent. These two groups of firms differ substantially, with firms that are going to apply for their first patent being much more productive and much bigger than other firms even before they apply for their first patent. This evidence suggests that a simple comparison between the outcome variables between patenting firms in the year of their first patent and non-patenting firms could be misleading due to pre-existing strong difference between the two groups. It is reasonable to assume that there is a process of self-selection into the category of patenting firms in which already innovative (here unobserved), more productive and bigger firms are substantially more likely to apply for their first patent than non-innovative, low-productivity and smaller firms. For that reason I employ two complementary methods to control for pre-treatment differences among firms: the difference-in-differences approach and the propensity score matching approach.

3.1 Difference-in-differences approach

I first adopt a difference-in-differences approach to identify the effect of applying for the first patent on firm size and productivity performance. The idea is to compare the change in the outcome variable when switching to the status of patenting firm with the change in the outcome of firms which did not experience the shift.

$$y_{it} = \mu_i + \beta * first\_patent + \tau_t + \epsilon_{it} \quad (1)$$

$$y_{it} = \mu_i + \beta * first\_patent + \phi_{jt} + \epsilon_{it} \quad (2)$$

Time dummies (τ_t) control for year-specific shocks while firm fixed effects (μ_i) average out firm-specific sample averages of the outcome variable. In

equation 2, I also add year-sector (Nace rev. 2 at 2-digits) specific fixed effects, thus allowing exogenous unobserved year-specific shocks to vary across sectors. In addition to these two baseline specifications, I estimate two similar models to investigate whether the effect of the first patent is persistent for the outcomes after the first patent (columns (3) c.t. - continuous treatment - and (4) c.t. in tables 4 and 5). Columns 5 in table 4 and 5 report preliminary results of a fixed effect model in which I include all patenting firms the year before their first patent and the year of their first patent together with a control sample selected with a propensity score matching algorithm (refer to the next section for additional details). Controls are matched to patenting firms based on their similarity in the propensity score estimated on pre-patenting features. For each patenting firm I selected 10 nearest neighbour matches. The estimated coefficient is computed by means of a fixed effect model that, when considering two period only, coincides with a first difference model. Observations for the control sample are weighted 0.1 each while observations for the patenting firms are weighted 1.

Finally, I also preliminary investigate the extent to which the effect changes future outcomes in different ways, depending on the lag between the first patent and the outcome variable. More specifically, I estimate the additional effects for the years after the third (included - columns 1, 3 and 5 in table 6) and the additional effects for the years after the fifth (included - columns 2, 4 and 6 in table 6).

Results

Tables 4 and 5 report the results for the various specifications of the difference-in-differences approach. Both productivity and size are generally positively affected by the decision of the firm to apply for its first patent at the EPO. The effect, however, changes in magnitude and significance depending on the specific measure and on the assumptions behind each specification. For all outcome variables except the logarithm of employees, the baseline difference-in-differences specification (column 1) shows a immediate, statistically strong and positive effect of applying for the first EPO patent. Productivity growth is between 2.91 percent (TFP_LP) and 3.65 percent (labour productivity) higher for firms in their first patenting year than for other firms. Value added and turnover grow, respectively, by 4.85% and 3.91% more in the first year of patenting. When controlling for sector-year specific shocks (column 2), the effect on productivity tends to shrink in magnitude and becomes statistically insignificant, while the effect on size is maintained, with the effect on labour being now greater and statistically significant.

While in specifications of columns 1 and 2 the patenting firms are removed from the dataset after the first year of patenting, in column 3 and 4 I maintain them in the dataset and I assume that the effect of patenting is

constant for all years after the first patent. In practice, the dummy 'First patent' now equals 1 for the year of the first patent and for all subsequent years. The effect of patenting for the first time on productivity turns out to be quite persistent, with the overall average differential growth in productivity relative to non-patenting firms being slightly higher than the immediate effect. This could be due to the fact that not only the effect of the first patent on productivity is persistent, but it tends to increase in the years after the first patent. When looking at firm size, however, the positive effect for value added and turnover and the insignificant effect on employment is confirmed in magnitude and statistical significance, suggesting that most of the effect of patenting occurs in the first year(s).

Finally, the last column of tables 4 and 5 report first-difference estimates of the effect of the first patent. The control sample is now selected by means of a propensity score algorithm based on pre-patenting observable characteristics of the firms. The positive effect of the first patent on productivity and firm size (except turnover, for which the effect turns out to be insignificant) is, on average, confirmed in sign, magnitude and statistical significance.

When I assumed the effect to be persistent and constant for the whole period after the first patent, I observe a generally greater effect than the one found for the first year, at least for measures of productivity. Table 6 deals with this issues by investigating whether a differential effect of the first patent on performance exists after t+2 and after t+4, with t being the year of the first patent. The effect for the first two years (the year of the first patent and the next one - Pat(t-T)) is very similar to the one found in columns 1 and 3 of tables 4 (measures of productivity) while it is substantially higher for measures of firm size (columns 1 and 3 of table 5). The delayed additional effect of the first patent on productivity is generally positive but statistically significant only for labour productivity. The overall effect on labour productivity after t+2 is on average 5.87 percent (3.51%+2.36%) the overall effect after t+4 is 8.54% (3.63%+1.21%+3.7%). When considering firm size, however, the positive effect found in the first two years of (Pat(t - T)) is partly compensated by negative effects in the following periods. These negative effects are small and statistically insignificant for value added and turnover while they are big and significant for labour, for which I observe a full compensation of the initial increase in labour. To conclude, this preliminary evidence shows that while the effect of entering the elite group of 'EPO applicants' has a persistent and increasing effect on productivity, the effect on size (especially when measured with employee headcount) tends to be only temporary.

3.2 Propensity score matching approach

An alternative way to measure the effect of the first patent on firms' performance would be to restrict the control group to those firms which are

similar to treated firms in the pre-treatment period. A way to deal with the identification of the control sample based on observables is to match each treated unit to a pool of control units which are as similar as possible to the treated ones. The propensity score matching approach consists in assigning (matching) to each treated unit a control unit (or a group of controls) characterized by the same or a similar probability of being treated and in computing the average treatment effect on treated as the difference in the outcome variable between treated and controls. For more detail on the propensity score matching approach and on the underlying assumptions refer to Rosenbaum and Rubin (1983) and Caliendo and Kopeinig (2008).

The propensity score is generally estimated as the predicted probabilities derived from a probit or logit model in which the probability of being treated is a function of observable characteristics of treated and control units. There is no causal concern when estimating the propensity score, meaning that the set of covariates in the probit or logit model will possibly include a full set of interactions and polynomials.

After estimating the propensity score and testing its properties, each treated unit should be matched to a control or set of control units. There are several ways of identifying the set of control units (refer to Caliendo and Kopeinig (2008) for more details). The simplest approach is to match each treated unit with its nearest control units in terms of estimated propensity score. Using the nearest neighbour only maximize the similarity between treated and control groups, thus reducing the expected bias, but results in imprecise estimates. On the contrary, using more than one nearest neighbour as controls, results in more precise estimates at the cost of a greater possible bias. A way to reduce the size of the bias in case of multiple neighbour matching is to limit the matched controls to those with a distance from the treated units in terms of propensity score below a certain threshold, called caliper. For all specifications of the propensity score matching approach, I used a caliper of 0.002. The propensity score matching approach is here implemented by using the user-written Stata module PSMATCH2 by Leuven and Sianesi (2003).

The explanatory variables in the probit model to estimate the propensity score are: sector, region and year dummies, a dummy variable for the firm reporting any R&D in its balance sheet and linear and quadratic terms of: the log of capital stock, the log of total assets, the log of leverage (total asset over equity), the log age, a logit transformation of the share of R&D in total assets. Moreover, to control for the pre-treatment differences in the outcome variable, I add the lag of the outcome variable as an additional explanatory variable in the probit model.

Results

The propensity score has been estimated separately for each of the outcome variables. The probability of applying for the first patent in the following year is, as expected, positively related with firm size (assets and capital), with the probability of performing R&D and with its intensity. Moreover, the lag of the outcome variable (both size and productivity indicators) positively affects the probability of applying for the first patent the following year, confirming the presence of a self-selection of more productive and bigger firms into the category of patenting firms. Due to the small relative number of treated units the explanatory power of the probit estimates is rather low, with a pseudo R squared always around 0.1. The balancing properties of the propensity score are on average satisfied, with no significant difference between treated and control sets in terms of observable characteristics. Looking at the distribution of the propensity score, the hypothesis of common support between treated and controls is always satisfied, with the support for treated being always in the range of the support for controls. For each specification I also report some measure on the improvement in the similarity between treated and controls when moving the unmatched sample to the matched sample. Finally, the distribution in terms of propensity score between treated and controls improves substantially when moving from all unmatched controls (left panel of 1 when considering labour productivity as outcome variable) to matched controls (right panel of figure 1). For sake of brevity, results of the probit models and additional tests concerning the balancing properties of the propensity score are not reported but remain available upon request.

Baseline results are reported in tables 7 and 8. First note that the sample size is smaller relative to the difference-in-difference results due to the need of using lagged variables in the estimation of the propensity score. I report results based, respectively, on five, ten, twenty and fifty nearest neighbours.

The average effect of applying for the first EPO patent on those firms which apply (average treatment effect on treated - ATT) is confirmed to be positive for all of the measures of productivity performance and firm size. However, while the magnitude of the estimated effects is in line with the one observed in the difference-in-differences approach, statistical significance is substantially weaker, with significant (at least 10% significance) effects for labour productivity, TFP_FE and value added. Note that both the magnitude and the significance depend quite substantially on the number of nearest neighbours used as control sample.

In addition to baseline results, I also report some preliminary results of the effect of the first patent on the outcome variable in the years after the first patent. While the matching is again performed on the basis of the characteristics of the firm in the year before the first patent, the difference in the outcome variable between treated and controls is then evaluated one

and two years after the first patent (table 9). For each outcome variable I report both the simple ATT (columns 1, 3 and 5) and the treatment effect computed as the difference-in-differences between t+1 (and t+2) and t-1 for treated and controls selected with the propensity score matching algorithm. In both cases, for each firm I select the 10 nearest neighbours with caliper of 0.002 and common support assumption. The delayed effects on productivity and firm size are, in most cases, statistically insignificant but in line with the estimated effects on outcome at time t. The only notable exception is value added, for which the delayed effect is always statistically significant and much higher that the one estimated at time t.

4 Further research

The results reported and discussed in the current paper are going to be further extended. First, I plan to investigate in more detail the timing of the effect of the first patent on firm's productivity. I will focus on the persistence and on the rate of decay of the effect. Second, I plan to investigate the patenting behaviour of firms after the first patent application. I will distinguish among different type of patenting firms based on the patenting intensity after the first patent and on the level of technological specialization or diversification.

References

Balasubramanian, N. and Sivadasan, J. 2011. What happens when firms patent? new evidence from U.S. economic census data. *The Review of Economics and Statistics* 93:126–146.

Caliendo, M. and Kopeinig, S. 2008. Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys* 22:31–72.

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

Griliches, Z. 1981. Market value, R&D, and patents. *Economics Letters* 7:183 – 187.

Griliches, Z. 1990. Patent statistics as economic indicators: A survey. *Journal of Economic Literature* 28:1661–1707.

Griliches, Z., Pakes, A., and Hall, B. H. 1988. The value of patents as indicators of inventive activity. NBER Working Papers 2083, National Bureau of Economic Research, Inc.

Hall, B. H. 2011. Innovation and productivity. NBER Working Papers 17178, National Bureau of Economic Research, Inc.

Hall, B. H., Helmers, C., Rogers, M., and Sena, V. 2012. The choice between formal and informal intellectual property: A literature review. NBER Working Papers 17983, National Bureau of Economic Research, Inc.

Leuven, E. and Sianesi, B. 2003. PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Statistical Software Components, Boston College Department of Economics.

Levinsohn, J. and Petrin, A. 2003. Estimating production functions using inputs to control for unobservables. *Review of Economic Studies* 70:317–341.

Lotti, F. and Marin, G. 2013a. Drivers and effects of eco-innovations using data on eco-patents. mimeo.

Lotti, F. and Marin, G. 2013b. Matching of PATSTAT applications to AIDA firms: discussion of the methodology and results. Questioni di Economia e Finanza forthcoming, Banca d'Italia.

Lotti, F. and Schivardi, F. 2005. Cross country differences in patent propensity: A firm-level investigation. *Giornale degli Economisti* 64:469–502.

Marin, G. 2012. Do eco-innovations harm productivity growth through crowding out? Results of an extended CDM model. EIC working paper series 3, IMT Institute for Advanced Studies Lucca.

Pakes, A. 1985. On patents, R&D, and the stock market rate of return. *Journal of Political Economy* 93:390–409.

Rosenbaum, P. and Rubin, D. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70:41–55.

Table 1: Descriptive statistics

Variable	N	Mean	Min	Median	Max	SD
log(VA/L)	144730	10.83	8.35	10.77	14.19	0.47
log(TFP_LP)	145978	9.62	3.33	9.63	16.02	0.75
log(TFP_FE)	145978	9.97	0.38	10.01	19.11	0.83
log(VA)	191650	14.05	4.74	13.98	20.16	0.99
log(Y)	192690	15.51	7.13	15.38	21.78	0.95
log(L)	147352	3.38	0	3.37	8.65	0.93

log(VA/L): logarithm of labour productivity (value added per employee); log(TFP_LP): logarithm of TFP obtained from sector-specific Cobb-Douglas production functions estimated with the Levinsohn and Petrin (2003) method; log(TFP_FE): logarithm of TFP obtained from sector-specific Cobb-Douglas production functions estimated with a fixed effect model; log(VA) is the log of real value added; log(Y) is the logarithm of real turnover; log(L) is the logarithm of employees headcount.

Table 2: Distribution of observations by sector (Nace rev. 2) and year

	2001	2002	2003	2004	2005	2006	2007	Total
CA	1,460	3,034	3,053	3,557	3,671	3,838	4,002	22,615
CB	1,653	3,591	3,572	4,096	4,254	4,385	4,529	26,080
CC	1,051	2,189	2,190	2,577	2,681	2,810	2,884	16,382
CD	63	110	108	126	127	128	130	792
CE	465	1,042	1,036	1,160	1,179	1,216	1,250	7,348
CF	68	160	156	182	188	199	206	1,159
CG	1,507	3,171	3,137	3,579	3,663	3,798	3,933	22,788
CH	2,199	5,343	5,256	6,082	6,289	6,555	6,771	38,495
CI	351	771	773	893	907	953	978	5,626
CJ	475	1,107	1,102	1,237	1,291	1,339	1,384	7,935
CK	1,403	3,170	3,155	3,572	3,657	3,779	3,892	22,628
CL	370	703	711	843	863	920	953	5,363
CM	1,069	2,191	2,195	2,554	2,664	2,763	2,819	16,255
Total	12,134	26,582	26,444	30,458	31,434	32,683	33,731	193,466

Table 3: Difference in outcomes before the first patent

	log(VA/L)	log(TFP_LP)	log(TFP_FE)	log(VA)	log(Y)	log(L)
Firms patenting in the future	0.0737***	0.143***	0.180***	0.492***	0.447***	0.419***
	(0.00634)	(0.00683)	(0.00738)	(0.0126)	(0.0121)	(0.0128)
F	528.6	6858.3	7393.8	253.0	295.6	200.9
N	141982	143213	143213	188384	189408	144575

All OLS regressions include year and sector dummies. Standard errors in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01. The sample include all non-patenting firms and patenting firms only in the years before their first patent (identified with the dummy variable 'Firms patenting in the future').

Table 4: Difference-in-differences: productivity measures

log(VA/L)	(1)	(2)	(3) c.t.	(4) c.t.	(5) nn=10
First patent	0.0365***	0.0163	0.0423***	0.0121	0.0298*
	(0.0138)	(0.0137)	(0.0129)	(0.0128)	(0.0158)
Year FE	Yes	Yes	Yes	Yes	Yes
Year/sect FE	-	Yes	-	Yes	-
N	93924	93924	97291	97291	1491
log(TFP_LP)	(1)	(2)	(3) c.t.	(4) c.t.	(5) nn=10
First patent	0.0291**	0.0137	0.0381***	0.0135	0.0401**
	(0.0135)	(0.0135)	(0.0138)	(0.0137)	(0.0179)
Year FE	Yes	Yes	Yes	Yes	Yes
Year/sect FE	-	Yes	-	Yes	-
N	95014	95014	98411	98411	1509
log(TFP_FE)	(1)	(2)	(3) c.t.	(4) c.t.	(5) nn=10
First patent	0.0304**	0.0174	0.0385***	0.0175	0.0346**
	(0.0131)	(0.0132)	(0.0137)	(0.0137)	(0.0174)
Year FE	Yes	Yes	Yes	Yes	Yes
Year/sect FE	-	Yes	-	Yes	-
N	95014	95014	98411	98411	1509

Standard errors clustered by firm in parentheses. * $p< 0.1$, ** $p< 0.05$, *** $p< 0.01$. In columns 1 and 2, patenting firms are dropped from the sample the year after patenting. In columns 3 and 4, the dummy variable 'First patent' equals 1 for the year of the first patent and for all subsequent years. The sample of column 5 includes all patenting firms and a selected control sample. The sample is selected by means of a propensity score matching procedure based on pre-patenting features. For each patenting firm, 10 control firms are selected, each of which is weighted 0.1.

Table 5: Difference-in-differences: size measures

log(VA)	(1)	(2)	(3) c.t.	(4) c.t.	(5) nn=10
First patent	0.0485***	0.0419***	0.0490***	0.0356**	0.0498***
	(0.0130)	(0.0130)	(0.0143)	(0.0143)	(0.0173)
Year FE	Yes	Yes	Yes	Yes	Yes
Year/sect FE	-	Yes	-	Yes	-
N	144244	144244	148776	148776	1987
log(Y)	(1)	(2)	(3) c.t.	(4) c.t.	(5) nn=10
First patent	0.0391***	0.0284**	0.0381***	0.0194	0.0197
	(0.0137)	(0.0137)	(0.0142)	(0.0142)	(0.0146)
Year FE	Yes	Yes	Yes	Yes	Yes
Year/sect FE	-	Yes	-	Yes	-
N	144939	144939	149499	149499	1997
log(L)	(1)	(2)	(3) c.t.	(4) c.t.	(5) nn=10
First patent	0.0217	0.0347**	0.0204	0.0340**	0.0168
	(0.0170)	(0.0170)	(0.0172)	(0.0172)	(0.0171)
Year FE	Yes	Yes	Yes	Yes	Yes
Year/sect FE	-	Yes	-	Yes	-
N	95640	95640	99060	99060	1515

Standard errors clustered by firm in parentheses. * $p< 0.1$, ** $p< 0.05$, *** $p< 0.01$. In columns 1 and 2, patenting firms are dropped from the sample the year after patenting. In columns 3 and 4, the dummy variable 'First patent' equals 1 for the year of the first patent and for all subsequent years. The sample of column 5 includes all patenting firms and a selected control sample. The sample is selected by means of a propensity score matching procedure based on pre-patenting features. For each patenting firm, 10 control firms are selected, each of which is weighted 0.1.

Table 6: Difference-in-differences: additional results

	(1)	(2)	(3)	(4)	(5)	(6)
	log(VA/L)	log(VA/L)	log(TFP_LP)	log(TFP_LP)	log(TFP_FE)	log(TFP_FE)
Pat(t - T)	0.0351***	0.0363***	0.0454***	0.0461***	0.0453***	0.0458***
	(0.0114)	(0.0113)	(0.0126)	(0.0125)	(0.0124)	(0.0124)
Pat(t+2 - T)	0.0236*	0.0121	0.0101	0.00306	0.00395	-0.000777
	(0.0125)	(0.0120)	(0.0127)	(0.0117)	(0.0125)	(0.0114)
Pat(t+4 - T)		0.0370**		0.0224		0.0151
		(0.0155)		(0.0173)		(0.0173)
F	1001.4	890.7	438.5	389.8	415.4	369.3
N	144730	144730	145978	145978	145978	145978
	log(VA)	log(VA)	log(Y)	log(Y)	log(L)	log(L)
Pat(t - T)	0.0689***	0.0684***	0.0541***	0.0535***	0.0325**	0.0317**
	(0.0148)	(0.0147)	(0.0144)	(0.0143)	(0.0148)	(0.0148)
Pat(t+2 - T)	-0.0152	-0.0114	-0.0140	-0.00850	-0.0332**	-0.0253**
	(0.0143)	(0.0122)	(0.0134)	(0.0116)	(0.0144)	(0.0128)
Pat(t+4 - T)		-0.0137		-0.0195		-0.0252
		(0.0204)		(0.0170)		(0.0181)
F	708.1	629.6	658.2	585.1	165.7	147.3
N	191650	191650	192690	192690	147352	147352

Standard errors clustered by firm in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Pat(t - T) is a dummy variable which equals 1 for the years of the first patent and for all subsequent years. Pat(t+2 - T) is a dummy varible which equals 1 for the third to the last year after the first patent. Pat(t+4 - T) is a dummy varible which equals 1 for the fifth to the last year after the first patent.

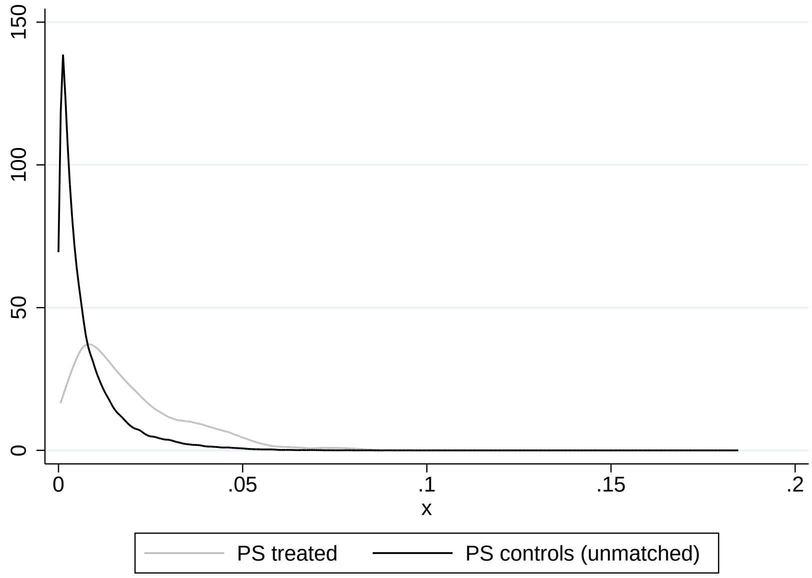
Table 7: Propensity score matching: productivity measures

log(VA/L)	nn=5	nn=10	nn=20	nn=50
ATT	0.0349	0.0298	0.0255	0.0322
s.e.	0.0171	0.0159	0.0156	0.0163
t-stat	2.039	1.877	1.631	1.975
p-value	0.0415	0.0605	0.103	0.0483
N	93924	93924	93924	93924
Unmatched				
Mean bias	14.4			
Chi2	736.78			
p-value	0			
Pseudo R2	0.085			
Matched				
Mean bias	1.2	0.9	0.8	0.6
Chi2	3.33	1.93	1.41	0.88
p-value	1	1	1	1
Pseudo R2	0.002	0.001	0.001	0
log(TFP_LP)	nn=5	nn=10	nn=20	nn=50
ATT	0.0387	0.0388	0.0293	0.0329
s.e.	0.0273	0.0261	0.0255	0.0251
t-stat	1.417	1.485	1.151	1.312
p-value	0.156	0.138	0.250	0.189
N	95014	95014	95014	95014
Unmatched				
Mean bias	14.8			
Chi2	741.86			
p-value	0			
Pseudo R2	0.084			
Matched				
Mean bias	1.5	1	0.8	0.6
Chi2	5.86	3.14	1.94	0.82
p-value	1	1	1	1
Pseudo R2	0.003	0.002	0.001	0
log(TFP_FE)	nn=5	nn=10	nn=20	nn=50
ATT	0.0569	0.0519	0.0431	0.0355
s.e.	0.0342	0.0328	0.0320	0.0315
t-stat	1.664	1.583	1.345	1.126
p-value	0.0961	0.113	0.179	0.260
N	95014	95014	95014	95014
Unmatched				
Mean bias	14.3			
Chi2	742.89			
p-value	0			
Pseudo R2	0.084			
Matched				
Mean bias	1.1	1.1	0.8	0.6
Chi2	3.81	2.59	1.33	0.78
p-value	1	1	1	1
Pseudo R2	0.002	0.001	0.001	0

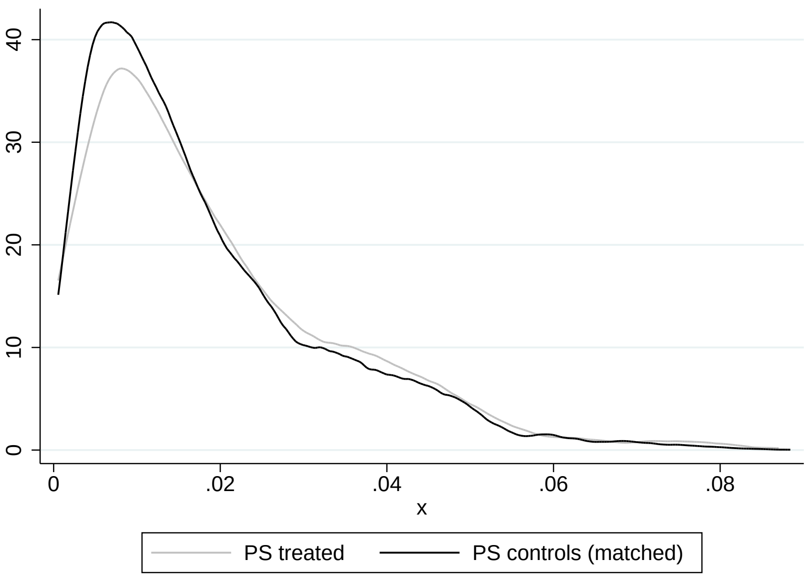
Table 8: Propensity score matching: size measures

log(VA)	nn=5	nn=10	nn=20	nn=50
ATT	0.0619	0.0558	0.0542	0.0523
s.e.	0.0339	0.0324	0.0316	0.0311
t-stat	1.825	1.725	1.716	1.681
p-value	0.0680	0.0845	0.0862	0.0927
N	144244	144244	144244	144244
Unmatched				
Mean bias		15.5		
Chi2		945.69		
p-value		0		
Pseudo R2		0.08		
Matched				
Mean bias	0.9	0.7	0.6	0.5
Chi2	3.35	1.76	0.96	0.7
p-value	1	1	1	1
Pseudo R2	0.001	0.001	0	0
log(Y)	nn=5	nn=10	nn=20	nn=50
ATT	0.0156	0.0218	0.0265	0.0278
s.e.	0.0339	0.0324	0.0316	0.0311
t-stat	0.459	0.673	0.838	0.893
p-value	0.646	0.501	0.402	0.372
N	144939	144939	144939	144939
Unmatched				
Mean bias		15.3		
Chi2		940.88		
p-value		0		
Pseudo R2		0.079		
Matched				
Mean bias	1	0.9	0.9	0.5
Chi2	3.8	2.58	2	0.78
p-value	1	1	1	1
Pseudo R2	0.001	0.001	0.001	0
log(L)	nn=5	nn=10	nn=20	nn=50
ATT	0.00663	0.0130	0.0150	0.0167
s.e.	0.0344	0.0329	0.0322	0.0317
t-stat	0.193	0.394	0.468	0.527
p-value	0.847	0.694	0.640	0.599
N	95640	95640	95640	95640
Unmatched				
Mean bias		15.5		
Chi2		750.31		
p-value		0		
Pseudo R2		0.085		
Matched				
Mean bias	1.3	1.1	0.7	0.5
Chi2	4.79	2.64	1.44	0.62
p-value	1	1	1	1
Pseudo R2	0.002	0.001	0.001	0

Figure 1: Distribution of the propensity score (outcome variable: ln(VA/L))



(a) Treated and unmatched controls



(b) Treated and matched controls (nearest neighbours (with caliper): 10)

Table 9: Propensity score matching: additional results

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome: t+1	log(VA/L)	log(VA/L)	log(TFP_LP)	log(TFP_LP)	log(TFP_FE)	log(TFP_FE)
First patent	0.0157 (0.0171)	0.0138 (0.0202)	0.0350 (0.0276)	0.0283 (0.0221)	0.0373 (0.0336)	0.0447** (0.0210)
N	81903	1379	82901	1399	82901	1399
Outcome: t+1	log(VA)	log(VA)	log(Y)	log(Y)	log(L)	log(L)
First patent	0.0525* (0.0310)	0.0499** (0.0204)	0.0191 (0.0312)	0.0112 (0.0182)	0.0131 (0.0328)	0.0154 (0.0225)
N	141885	1971	142808	1981	83542	1401
Outcome: t+2	log(VA/L)	log(VA/L)	log(TFP_LP)	log(TFP_LP)	log(TFP_FE)	log(TFP_FE)
First patent	0.0234 (0.0172)	0.0257 (0.0234)	0.0352 (0.0280)	0.0404 (0.0253)	0.0430 (0.0340)	0.0463* (0.0251)
N	81396	1373	82439	1395	82439	1395
Outcome: t+2	log(VA)	log(VA)	log(Y)	log(Y)	log(L)	log(L)
First patent	0.0768** (0.0317)	0.0771*** (0.0243)	0.0477 (0.0319)	0.0351 (0.0220)	0.0245 (0.0324)	0.0300 (0.0229)
N	141142	1967	142383	1977	83274	1403

Standard errors in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01. Columns 1, 3 and 5 report results of the propensity score matching method while columns 2, 4 and 6 report results of first-difference ($Y_{t+s} - Y_{t-1}$, with $s = 1, 2$ and t being the year of the first patent) estimates.