# Regional innovation system (in)efficiency and its determinants: an empirical evidence from Italian regions

Barra, Cristian and Zotti, Roberto

University of Salerno - Department of Economics and Statistics

1 October 2015

# Regional innovation system (in)efficiency and its determinants: an empirical evidence from Italian regions

**Cristian Barra***

**Roberto Zotti***

## Abstract

This paper investigates the regional innovation system (RIS) efficiency, and its determinants, in Italy through a Stochastic Frontier Analysis and using the concept of a knowledge production function. The contribution of universities', private and public sectors' resources devoted to research and development (R&D), in generating innovation, has been examined, as well as the impact of several exogenous environmental variables on RIS efficiency. The empirical findings suggest the importance of R&D investments taking place in the universities and in the private sector, which benefit the most to regional innovation activities; labour market and industries' characteristics are found to have an important role on RIS efficiency.

**Keywords:** Regional innovation system, Technical efficiency. Knowledge production function

**JEL-Codes:** O31; C14; C67; R12

*Corresponding authors: Department of Economics and Statistics, University of Salerno, Via Giovanni Paolo II, 132 - 84084 - Fisciano (SA) - Italy
E-mail: cbarra@unisa.it; rzotti@unisa.it

## 1. Introduction

Research and development (R&D) activities, likewise patents, could be considered as new ideas and pieces of knowledge that may turn into innovation when commercially exploited (Schumpeter, 1934; 1942); thus, innovation may be seen as the ability to use such knowledge to generate, develop and improve new products, processes and services. Generally speaking, there are different forms of innovation: (i) the introduction of new ways of doing things (Porter, 1990); (ii) the ability to use resources to create value (Drucker, 1993); (iii) the commercial exploitation of an idea (UK Department of Trade and Industry Innovation, 1994) and (iv) the output of a research and development process (Tidd, et al. 1997). However, regardless of the meaning of innovation adopted in the literature, it is still debated by the researchers why a system should innovate and what are the related benefits of maintaining innovation activities. Indeed, many are the factors involved in the innovation process such as economic (growth, competitiveness, internationalization), social (human capital development, employment and entrepreneurship), business (improvement of performance, value creation, competitive advantage) and scientific (development and enhancement of the knowledge). It is, instead, more clear that innovation is fundamental to economic growth of a region which, as a consequence, may increase technological capital by innovating; knowledge and technological progress are, indeed, among the main engines of economic dynamics in most endogenous growth models (Romer, 1990). In other words, advantages of regions, in terms of innovation outputs, could be also related to the ability of regional firms to develop their innovation (Krugman, 1991; Maskell and Malberg, 1999). Thus, it is especially important to find out what components of an R&D system are most decisive as engines of innovation and what are the factors determining systems' innovatory capacities. See Capello and Lenzi (2014) for more details on the role played by knowledge and innovation as drivers of regional economic growth and on their spatial impact and McCann and Simonen (2005) for the role played by geography in the promotion of innovation.

An important empirical approach to analyse the process of innovation creation is the knowledge production function (KPF), originally formalized by Griliches (1979) and Pakes and Griliches (1984), showing that knowledge is mainly generated through R&D activities carried out by firms, universities and other research institutions (see Acs et al. 2002). Both knowledge creation and regional innovation through research and technology transfer represent relevant channels. Promoting enterprise, business development and growth, all activities linked to the possibility of busting a more entrepreneurial culture and a more favourable business environment, also have to be considered. Empirical evidence from firm surveys (Mansfield, 1995; 1997; Cohen et al. 2002; Veugelers and Cassiman, 2005) confirms the importance of university research for corporate innovation performances. Knowledge transfers from academia have been investigated through licensing (again Shane, 2002), academic spin-off activities (Shane, 2002) and citation to academic patents

(Henderson et al. 1998). See Maietta (2015), on the channels through which university–firm R&D collaboration impacts upon firm product and process innovations, and Caniëls and van den Bosch (2010), on the role of higher education institutions (HEIs) in building regional innovation systems. In the literature, KPF has been implemented at regional level (see among others Crescenzi, 2005; Rodrıguez-Pose and Crescenzi, 2008; Sterlacchini, 2008; Marrocu, Paci, and Usai, 2013) showing evidence of the key role of knowledge inputs (i.e. R&D expenditures or employees) in generating knowledge outputs (i.e. patents).

The purpose of the paper is main-fold. Firstly, we investigate the production of knowledge of a regional innovation system (RIS), by estimating a RIS technical efficiency based on the concept of a knowledge production function as suggested by Griliches (1979) and Jaffe (1989) in the Italian regional context; in order to study the relationship between inputs and outputs of the innovation process, we use a Stochastic Frontier Analysis (SFA), which has been widely used to study technical efficiency in various settings since its introduction by Aigner et al. (1977), and Meeusen and van den Broeck (1977). We specifically rely on highly disaggregated proxies for measuring the inputs to the innovation process such as the expenditure and the staff employed in R&D activities in the public sector, in the universities and in the private sector (see Section 3.1. for more details on the production set). This allows us to better analyse the factors that have a direct impact on innovation outputs (as measured by patent applications). The capacity of generating local knowledge, and of turning knowledge into growth, has long been identified with the presence of territorial conditions in the area. Therefore, secondly, this paper directly investigates whether RIS efficiency is influenced by some exogenous characteristics of the regional environment - i.e. labour market and industries' characteristics - (see Sections 2 and 3.2 for more details on the way these variables are included into an SFA single stage approach); indeed, failing to model the exogenous factors leads to bias estimation of the technical efficiency scores (e.g. Caudill and Ford 1993; Caudill et al. 1995; Hadri 1999; Wang 2003). More specifically, we look at the effect of variables like a measure of urbanization such as the density of the population, a control for the financial market and some indicators of the labour market structure and of the industries' characteristics such as the rate of unemployment, a control for employment in services and in industry sectors, the involvement of firms in export activities and the capacity of firms of doing product and process innovation. In other words, we explore whether the environmental channel can explain regional differences in term of diffusion of knowledge and innovation.

To anticipate the results, we show evidence of the importance of R&D investments taking place in the universities and in the private sector, which benefit the most to regional innovation activities; indeed, both expenditures and staff employed in R&D activities have a positive and statistically significant effects on the innovation process. The findings show that regions in the Central-North area (North-Western, North-Eastern and Central) outperform the Southern area. Furthermore, the exogenous environmental variables such as labour market and industries' characteristics are found to have an important role on RIS efficiency. Statistical significance of both inputs variables and efficiency scores' determinants is not majorly affected by clustering the production function at regional level. The paper is organized as follows. Section 2 introduces the methodology used to estimate RIS efficiency. Section 3 describes the data, the production set and the specification of the models implemented in the analysis. The empirical evidence is described in Section 4, while Section 5 provides a sensitive analysis. Finally Section 6 concludes.

## 2. Measuring the Regional Innovation System Efficiency

Following Fritsch and Slavtchev (2011), we measure RIS efficiency through the concept of technical efficiency as introduced by Farrell (1957). In other words, a given unit is technically efficient if it is able to produce the possible maximum output from a given amount of input. A KPF [1], based on a Cobb-Douglas production function formulation (see Griliches, 1979 and Jaffe, 1989), is estimated, in order to analyse the relationship between inputs and outputs of the innovation process, which is essential for assessing RIS technical efficiency.

The problem of assessing economic performances of a given unit under analysis (i.e. regions), is also exacerbated by inefficiency in production; then, when modeling production and cost functions, it must be kept in mind that a given unit is likely to produce using their inputs in a sub-optimal way. An available approach for incorporating inefficiency into the estimation of production is the method named Stochastic Frontier Analysis (SFA), proposed by Aigner et al. (1977) and Meeusen and Van den Broeck (1977); econometrically, the method assumes that the error term is composed by two components with different distributions (see Kumbhakar and Lovell (2000) for analytical details on stochastic frontier analysis). The first component, regarding the "inefficiency", is asymmetrically distributed (typically as a semi-normal), while the second component, concerning the "error", is distributed as a white noise. In this way, it is necessary to assume that both components are uncorrelated (independent) to avoid distortions in the estimates. This approach is particularly suitable considering our context, as one of the main advantages of SFA is that statistical inference can be drawn, obtaining information on the determinants of inefficiency.

Formally, taking the logarithm version, the KPF is described as follows:

$$y_{it} = \alpha_i + f(\boldsymbol{x}_{it}, \boldsymbol{\beta})\exp\{v_{it} - u_{it}\} \tag{1}$$

where $a$ is the constant, $y_{it}$ is the output of region $i$ at time $t$; $\boldsymbol{x}_{it}$ is a vector of input quantities of region $i$ at time $t$ (see Section 3.1 for more details on the input-output framework used in the production set); $\boldsymbol{\beta}$ is a vector of unknown

[1] This is based on the assumption that R&D activities are the main source of inventions and innovation.

parameters; $f(\boldsymbol{x}_{it}, \boldsymbol{\beta})$ is the production function or conventional regression model; $v_{it}$ is a vector of random variables related to the idiosyncratic or stochastic error term of region $i$ at time $t$ assumed to be independently and identically distributed (i.i.d.) as $N(0, \sigma_v^2)$ and independent of the $u_{it}$, while $u_{it}$ is a vector of non-negative random variables measuring the inefficiency term of region $i$ at time $t$ assumed to be independently but not identically distributed. They are obtained from the truncation to zero of the distribution $N(m_{it}, \sigma_u^2)$, where $m_{it} = \mu + z_{it}\delta$, $\mu$ denoting the location parameter and $z_{it}$ a vector of determinants of (technical) RIS inefficiency of region $i$ at time $t$ (see Section 3.2. for more details on the variables used in $z$), and $\delta$ is a vector of unknown coefficients. The linear homogeneity of degree 1, where the sum of the coefficients associated to any input is assumed to be 1, is imposed.

All coefficients of parameters in equation (1) and technical efficiency are estimated through a maximum likelihood estimator (MLE) using the STATA 13 software. Following Kalirajan and Shand (1999), we estimate the technical efficiency assuming that output elasticity associated to any input (i.e. $\boldsymbol{\beta}$) is identical for all Italian regions ($i$=1,.., 20). In other words, the produced output may fall systematically below the maximum, not because of lower output elasticities of the factors of production, but because of a lower level of the function.

## 3. Data, production set and model specification

### *3.1 The production set*

The empirical analysis is based on data collected from the Italian National Institute of Statistics (ISTAT) covering a 10 years time-span (from 2000 to 2009). The production technology is specified with six inputs (both regarding the R&D expenditures and the number of R&D employees in different sectors) and one output (number of disclosed regional patent applications in the years 2000-2009). See Table 1 below for more details on the model specification. More specifically, the first set of inputs consists in the amount of R&D expenditures in the public sector (RD_EXP_PUBL), in the higher education institutions (RD_EXP_HEI) and in the private sector (RD_EXP_PRIV)[2]. See Bottazzi and Peri (2003), Fritsch and Franke (2004), Buesa et al. (2010) and Tavassoli and Carbonara (2014), for the use of similar innovation inputs. The second set of inputs, instead, consists in the number of R&D employees in the public sector (RD_EMPL_PUBL), in the higher education institutions (RD_EMPL_HEI) and in the private sector (RD_EMPL_PRIV).[3] See Fritsch and Slavtchev (2011) and Buesa et al. (2010) for the use of such kind of innovation inputs. As underlined by Buesa et al. (2010), the

[2] The number of R&D expenditures is expressed in thousands of euros. We decide to separate out R&D expenditures in the universities from those in the public sector in order to further analyse the contribution of the higher education institutions to the innovation activities.
[3] The number of R&D employees refers to the researchers, technical employees and any other operator in R&D activities, respectively in the public sector, in the universities and in the private sector. It is expressed as full time equivalent units. We decide to separate out R&D employees in the universities from those in the public sector in order to further analyse the contribution of the higher education institutions to the innovation activities.

choice of inputs is based on the conclusion that "innovatory outputs depend in the first place on the effort made in allocating resources, regardless of whether the latter is measured via expenditures or staff employed in R&D". Therefore, the two set of inputs are alternatively used in the knowledge production function in order to explore potential differences due to the way R&D investments are measured in the literature (i.e. R&D expenditures or employees).

[Table 1 around here]

Moving to the output side, we use a standard measure for innovation activities such as the number of disclosed regional patent applications registered at the European Patent Office (PAT) in the years 2000–2009. Although there are some limitations regarding the use of the number of patents as a measure of innovation output[4], they can be used as a good approximation of innovative ideas (see Bottazzi and Peri, 2003; Buesa et al. 2010; Fritsch and Slavtchev, 2011; Tavassoli and Carbonara 2014). We follow Fritsch and Slavtchev (2011), assuming that a certain amount of time is required before R&D activities will result in a patent; indeed, several months (usually from 12 to 18) are needed such that patents applications are published[5]. Therefore, a time lag between innovation inputs and the output of at least one or two years should be assumed[6]. It is true, however, that the time lag between R&D inputs and patent applications also depends on the reliability of the data. Indeed, different solutions have been exploited in the literature; Acs et al. (2002) report that innovation records result from inventions made 4 years previously, Fritsch and Slavtchev (2007) used a time lag of three years between patent applications and innovation input, Fischer and Varga (2003) used a two-year lag, Ronde and Hussler (2005) estimate regional innovation performances linking the R&D efforts to the number of patents registered one, two and three years later. Fritsch and Slavtchev (2011) reduced the time lag to a period of one year. In order to meet both data constraints and the main literature and also in order to take into account the time required to transform competences into concrete innovation as well as innovation into patents, following Fischer and Varga (2003), we assume a time lag of two years between innovation inputs and outputs.

All inputs and the output variables are in log-levels so that overall the positive skewness of variables is reduced.

When looking at the descriptive statistics (see Tables 2 below for more details), it is interesting to notice that, considering the main four geographical areas of the country, and taking into account the inputs, the Southern area shows the lowest

[4] First, patents are granted for an invention, but that invention is not necessarily transformed into an innovation, i.e., a new product or production technology. Second, patents are for products rather than for processes. Third, because there are other ways besides patenting to appropriate the returns of successful R&D activities, the number of patents might underestimate actual innovation output (see, among others, Cohen et al. 2000 and Cohen et al. 2002 on this points).

[5] This corresponds to the amount of time the patent office needs to verify whether an application fulfills the basic preconditions for being granted a patent and to complete the patent documents (Greif and Schmiedl, 2002).

[6] Assuming such a time lag also helps to avoid potential problems of endogeneity between R&D inputs and outputs.

number of both the amount of R&D expenditures and the number of employees in the R&D activities. Moreover, considering the performances in term of R&D outputs (i.e. patents) by geographical areas, again the North-Central areas outperform the Southern area. See Figures 1 and 2 below for a graphical representation of the inputs and the output used in the production function.

[Table 2 around here]

[Figures 1 and 2  around here]

### *3.2 Determinants of RIS (in)efficiency*

In a stochastic frontier model with heterogeneity, failure to model the exogenous factors leads to a biased estimation of the production frontier model and of the level of technical inefficiency, hence leading to poor policy conclusions (e.g. Caudill and Ford 1993; Caudill et al. 1995; Hadri 1999; Wang 2003). Indeed, differences in the economic environment might have an important impact upon RIS inefficiency and various control variables could be used to model this impact. These variables are considered exogenous in the sense that they influence the production process but are not themselves either inputs or outputs. They, in fact, influence the efficiency with which inputs are turned into outputs. Allowing inefficiency to depend on regional environmental characteristics enables researchers to examine the determinants of inefficiency, and to suggest policy interventions to improve efficiency. In other words, the basic assumption of the model is that the environmental factors influence the degree of technical inefficiency and then the innovation production function must include environmental variables which directly influence the inefficiency term. In the specific framework of SFA sometimes a two-stage estimation approach is used, where the first stage involves the specification and estimation of a stochastic frontier and the prediction of the technical efficiency scores of the units and the second one the specification of a regression model where the technical efficiency is regressed on explanatory factors relevant to the analysis; this approach will, however, lead to inconsistent estimates (Kumbhakar and Lovell, 2000); therefore, we apply, instead, a single stage approach (see, for example Battese and Coelli, 1995) where environmental factors are assumed to directly affect technical inefficiency.

In order to adequately measure the effects of some exogenous characteristics on innovation output (i.e. patents), we include in the inefficiency component[7] (see Section 2 for more analytical details) the following explanatory variables: population density (PD); financial volume (FV); unemployment rate (UR); employment in services (SERV) and in industry (IND)

[7] For a similar approach, see Fiordelisi et al. (2011) and Destefanis et al. (2014).

sectors, export (EP), and product and process innovation (INNO_PROD_PROC). More specifically population density (PD), measured as the number of inhabitants in the region by squared kilometres, aims to measure both the effects of urbanization economies and the unobserved region-specific effects. High population density should boost innovation activities as it provides opportunity for intensive contacts and cooperation (for a similar view, see Feldman, 2000 and Fritsch, 2000). Therefore a negative sign is expected for this variable on RIS inefficiency. In order to control for the lending market effects and then for the size of financial market, we use Financial Volume (FV), measured as the ratio between aggregate credits and value added. Recently this variable has been also used as a proxy for financial development (Hasan et al. 2009 and Destefanis et al. 2014). We assume a positive relationship between finance and innovation since the higher is the amount of loans provided by financial intermediaries in a certain area, the higher is the probability that the area can innovate. A negative sign on innovation inefficiency is then expected to be found on this variable. The unemployment rate (UR), measured as the number of people actively looking for a job as a percentage of the labour force, is intended to capture labour market effects. A positive sign on innovation inefficiency is expected to be found on unemployment. Take into account regional differences in the industry structure is crucial since patenting propensity differs across industries; therefore, in order to control for the impact of regional specialization in certain industries, following Bottazzi and Peri (2013), we use two variables such as the percentages of employment in services (SERV) and in industry (IND) sectors; specifically, SERV and IND are measured, respectively, as the number of employees in services and industry sectors, over the number of total employees in each region. We also use a variables indicating whether the firms are involved in export (EP) activities; specifically, EP is measured, as the values of exports as percentage of Gross Domestic Product. A negative sign is expected for these three variables on RIS inefficiency. In order to control for the firm innovative activity, a dummy variable indicating whether the firm has introduced product or process innovation (INNO_PROD_PROC) has been included[8]; a negative association between this variable and innovation inefficiency is expected to be found. Finally, time trend control for exogenous effects. See tables 3 and 4 for the definition, the expected sign and the correlation of these variables.

[Tables 3 and 4 around here]

[8] Both product and process innovation are analysed by Fitjar and Rodríguez-Pose (2012), González-Pernía et al. (2014), Nieto and Santamaria (2007) and Robin and Schubert (2013).

## 4. Empirical evidence

**4.1.** *Efficiency scores*

The estimated parameters of a KPF based on Cobb-Douglas specification are presented in Table 5 below. Results are showed when the amount of R&D expenditures are used as inputs (see Table 1 for the specification of the models) and the number of disclosed regional patents application are used as output. In order to take into account that a certain time is required before R&D activities will turn into a patent, a time lag between innovation inputs and the output of two years is assumed. We pay particular attention to both the assumption behind the production function used in the analysis and to inference issues; therefore, we report two estimates for the technical choices and two estimates for the standard errors. Table 5, Columns 1 and 2, reports estimates taking into account the linear homogeneity of degree 1 in inputs, whereas Table 5, Columns 3 and 4, reports estimates relaxing such imposition. Columns 1 and 3 report standard errors robust to heteroscedasticity, whereas Columns 2 and 4 report standard errors clustered at regional and year level. Cluster-adjusted standard errors correct for the possible correlation in innovative performances in the same regions over time. The asymptotic approximation relevant for clustered standard errors relies on a large number of clusters (see Donald and Lang, 2007). We have 160 clusters (8 years* 20 regions) which should be enough to deal with this issue. First of all, the null hypothesis that there is no heteroschedasticity in the error term has been tested and rejected, at 1% significance level, using a Likelihood Ratio Test (LR), giving credit to the use of some exogenous variables, according to which the inefficiency is allowed to change. In other words, the validity of the heteroschedastic assumption has been confirmed, leading to the significance of the inefficiency term. See Figures 3 and 4 for a graphical representation and for boxplots of the RIS efficiency scores.

[Table 5 around here]

[Figures 3 and 4 around here]

The coefficients show that all inputs variables have a positive and statistically significant effects on the innovation output, in all the specifications. When looking at the (average) technical efficiency scores by geographical areas (see Table 6 below), the estimates reveal that the Central-North area (North-Western, North-Eastern and Central) outperform the Southern area. Taking for instance estimates in Table 6, Column 1, the highest estimated gap efficiency scores exist between the Southern and the North-Western areas, in the order of 60%. Therefore, the average efficiency of the North-Western area is estimated around 75% - in other words, the output expected can be expanded by 25% using the same amount of inputs. Instead, the Southern area is around 15%, thus their inputs can be used more efficiently for producing almost three/fourth more outputs.

Estimates are quite similar when the linear homogeneity of degree 1 in inputs has and has not been imposed, but slightly higher in the former case.

[Table 6 around here]

*4.2. (In)efficiency score determinants*

When considering the exogenous factors included in the analysis, our findings show that the variables used to control for the different environment have an important role in describing the inefficiency term (Table 5). Population density (PD) has a significant and negative effect on RIS inefficiency, indicating that higher level of inhabitants in the region by squared kilometres is associated with higher levels of region's efficiency. This confirms the presence of urbanization economies already found in Fritsch and Slavtchev (2011), where the authors suggest that "densely populated regions provide a variety of opportunities for interaction and rich supplies of inputs, as well as a comprehensive physical and institutional infrastructure is advantageous for innovation activities". The control for the labour market (UR) seems to have an important role, too. Indeed, a positive and statistically significant coefficient has been found on the unemployment rate variable, meaning that the higher is number of people actively looking for a job as a percentage of the labour force (the higher is the chance of having more workers being involved in innovation activities) the lower is the inefficiency of innovation activities. Regional specialization in certain sectors seem to have relevance on the efficiency of the innovation processes, according to the negative sign of the percentages of employment in services (SERV) and in industry (IND) sectors, suggesting that RIS performances are positively affected by the share of employees in services and industry sectors, (which are supposed to have highly level of patenting). Product or process innovation (INNO_PROD_PROC) has a significant and positive effect on RIS inefficiency, meaning that those activities are a significant contributor to the firms' patenting behavior; from a statistical viewpoint, this result implies that patent statistics are biased towards product and process innovations, and therefore that the measures of innovation performance based on patent counts underestimate innovation output in the lines of business where both product and process innovation is important as compared to firms which are not engaged in these activities. A negative and statistically significant coefficient has been found on the export variable (EP), meaning that innovation activities, in regions where firms have high values of exports, are more efficient. Results also show a negative but not always statistically significant effect of the financial development proxy (FV). The weak relationship between finance and innovation could be explained by the fact that not all the innovation inputs turn into patents through the loans provided by the financial intermediaries; moreover, the financial recession period that took place around 2007 could have also triggered the credit rationing, leading to a reduction of the investments for innovation activities. Finally, according to the negative and significant coefficient (i.e. lower inefficiency) of the dummy variable for location in the North area (NORTH-WESTERN and NORTH-EASTERN)[9], innovation activities in regions located in the western and eastern part of the country are more efficient than those in South, suggesting that there are still considerable differences in the efficiency of the innovative process in the two parts of the country.

## 5. Sensitivity analysis: Does a different measure of innovation inputs affect the estimates?

In order to take into account the possible evidence of variation in the regional system efficiency and to examine whether an alternative measure of innovative inputs affects the analysis, we use the number of R&D employees (see Fritsch and Slavtchev, 2011, and Buesa et al. 2010 on the use of such inputs) instead of the amount of R&D expenditures (see Table 1 for more details on the production set). More specifically, we again disentangle the contribution to the regional innovative system, by public research institutions, private and public sector. Indeed, the set of inputs consists in the number of R&D employees in the public sector (RD_EMPL_PUBL), in the higher education institutions (RD_EMPL_HEI) and in the private sector (RD_EMPL_PRIV). The innovative output measure still consists in the number of disclosed regional patents application, used as output (again, in order to take into account that a certain time is required before R&D activities will turn into a patent, a time lag between innovation inputs and the output of two years is assumed). We report again two estimates for the technical choices and two estimates for the standard errors. Table 7, Columns 1 and 2, reports estimates taking into account the linear homogeneity of degree 1 in inputs, whereas Table 7, Columns 3 and 4, reports estimates relaxing such imposition. Columns 1 and 3, report standard errors robust to heteroscedasticity, whereas Columns 2 and 4, report standard errors clustered at regional and year level. The Likelihood Ratio Test (LR), still confirms the validity of the heteroschedastic assumption has been confirmed, leading to the significance of the inefficiency term.

[Table 7 around here]

Results still show that the coefficients of all inputs variables have a positive and statistically significant effects on the innovation output, in all the specifications, except for the number of R&D employees in the public sector (RD_EMPL_PUBL). This means that when the R&D employees are used as innovative input, the empirical findings suggest the importance of R&D investments taking place in the universities and in the private sector, which benefit the most to regional innovation activities. When looking at the (average) technical efficiency scores by geographical areas (see Table 8 below), the estimates confirm the presence of some geographical effects (by macro-areas) with Central-North regions

[9] The use of such variables allows us to control for potential geographical differences and therefore for the heterogeneity within the country.

(North-Western, North-Eastern and Central) outperforming those in the Southern area. When considering the exogenous factors included in the analysis, the findings confirm that the variables used to control for the different environment have an important role in describing the inefficiency term (Table 7). Population density (PD), the control for the labour market (UR), the export variable (EP) and being involved in product and process innovation (INNO_PROD_PROC) confirm the statistical significance and the sign expected. Regional specialization in certain sectors (SERV and IND variables) seem to have relevance on the efficiency of the innovation processes only when the linear homogeneity of degree 1 in inputs has not been assumed (see Table 7, Columns 3 and 4); evidence of a stronger relationship between finance and innovation has also been found according to the negative and always statistically significant effect of the financial development proxy (FV).

[Table 8 around here]

## 6. Summary and conclusions

In this paper, we investigate the regional innovation system efficiency in the Italian context, by estimating a measure of efficiency based on the knowledge production function concept. More specifically, a Stochastic Frontier Analysis, in order to analyse the relationship between inputs and outputs of the innovation process, has been applied. This parametric approach is particularly suitable considering the context analysed, as one of its advantage is that statistical inference can be drawn; indeed, obtaining information on the determinants of inefficiency and consequently on the estimated parameters, may attract the interest of regulators and decision makers towards the adoption of improving policies regarding the production of knowledge within a region leading to innovation activities and patents applications. The contribution of private and public sector resources devoted to research and development, in generating innovation, has been considered, as well as the impact of several exogenous environmental variables. Taking into account the measures of inputs in the innovative process, we disentangle the contribution (both considering the amount of R&D expenditures and the number of R&D employees) to the regional innovative system output (number of disclosed regional patents application) by public research institutions, private and public sector. Several exogenous variables such as labour market and industries' characteristics are used in order to examine whether the economic environment has an impact upon RIS inefficiency. The coefficients show that the input variables, almost in all specifications, have a positive and statistically significant effect on the regional innovation system efficiency; more specifically, this relationship is particularly evident for the R&D expenditure. While, when the staff employed in R&D activities has been used, the contribution of the private sector and of the universities is more evident. This evidence suggests that it is particularly the contribution of higher education institutions' and private firms' research activities to increase regional innovation efficiency. Findings also show that regions in the Central-North area (North-

Western, North-Eastern and Central) outperform the Southern area with the highest estimated gap efficiency scores existing between the Southern area and the North-Eastern area. A number of factors were found having a positive impact on RIS efficiency. Population density has a positive effect on innovation performances meaning that R&D activities are more productive in area more urbanized; RIS performances are found to be also influenced by the labour market and firm characteristics; indeed, innovation performances seem to be positively influenced by the rate of employment and by the presence of firms with high values of exports; product or process innovation is a significant contributor to the firms' patenting behavior. RIS performance is positively affected by the share of employees in services and industry sectors.

The empirical evidence provided calls into question possible limitations, some important policy implications as well as important issues to be further analysed in some future research. Indeed, a potential concern of our analysis regards the limited sample and the possibility of drawing robust conclusions with a max 160 observations. It has to be said, however, that although focusing on regional data in one country may bring to life some problems regarding the number of observations, it also reduces the heterogeneity, counting on a higher level of cultural, political and economic homogeneity country-wise. It could have been optimal to use more disaggregated data, such as at province level; unfortunately, we cannot investigate the innovation system at such territorial level due to the lack of information (more specifically, we cannot disentangle, for provinces or municipalities, the single contribution of HEIs, private and public sectors' investments in R&D activities – in term of expenditures and number of employees – on the innovative output. Keeping this discussion in mind, we believe that some lessons can be learned from this analysis. Firstly, the gap in efficiency among the macro-areas of the country requires some explanation, which can be useful for defining consistent policies that can improve the innovation productivity of the overall system; we claim that maintaining State-level policies can be detrimental for overall efficiency, and instead special interventions for regions in the South should be designed. Moreover, a policy that aims at improving RIS efficiency should be able to identify the most efficient channels through which knowledge transfer and innovation activities could be stimulated. The findings provide a clue towards the expansion of the importance of R&D investments taking place in the universities and in the private sector, which benefit the most to regional innovation activities. Further research is needed, using more disaggregated data, in order to disentangle the policy implications at province level.

## References


Acs, Z. J., Anselin, L., and Varga, A. (2002). Patents and innovation counts as measures of regional production of knowledge. Research Policy, 31(7), 1069–1083.

Aigner, D., Lovell, K. and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. Journal of Econometrics, 6, 21–37.

Battese, G. E. and Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. Empirical Economics, 20, 325–332.

Bottazzi, L. and Peri, G. (2003). Innovation and spillovers in regions: Evidence from European patent data. European Economic Review, 47, 687 – 710

Buesa, M., Heijsa, J. and Baumert, T. (2010). The determinants of regional innovation in Europe: A combined factorial and regression knowledge production function approach. Research Policy, 39, 722–735.

Caniëls, M. C. J. and van den Bosch, H. (2010). The role of Higher Education Institutions in building regional innovation systems. Papers in Regional Science, 90 (2), 271-286.

Capello, R. and Lenzi, C. (2014). Spatial heterogeneity in knowledge, innovation, and economic growth nexus: conceptual reflections and empirical evidence. Journal of Regional Science, 54, (2), 186–214.

Caudill, S.B., and Ford, J.M. (1993). Biases in Frontier Estimation Due to Heteroskedasticity. Economics Letters, 41, 17-20.

Caudill, S.B., Ford, J.M., and Gropper, D.M. (1995). Frontier estimation and firm specific inefficiency measures in the presence of heteroskedasticity. Journal of Business and Economic Statistics, 13, 105-111.

Cohen, W.M., Nelson, R.R. and Walsh, J.P. (2000). Protecting their intellectual assets: appropriability conditions and why U.S. manufacturing firms patent (or not). Working Paper No. 7552, National Bureau of Economic Research (NBER), Cambridge, MA.

Cohen, W.M., Nelson, R.R. and Walsh, J. P. (2002). Links and impacts: the influence of public research on industrial R&D, Management Science, 48, 1–23.

Crescenzi, R. (2005). Innovation and regional growth in the enlarged europe: the role of local innovative capabilities, peripherality, and education. Growth and Change, 36 (4), 471–507.

Destefanis, S., Barra, C. and Lubrano-Lavadera, G. (2014). Financial development and economic growth: evidence from highly disaggregated Italian data. Applied Financial Economics, 24 (24), 1605-1615.

Department of Trade and Industry (1994). Innovation – Your Move. London.

Donald, S. G., and Lang, K. (2007), Inference with Differences-in-Differences and Other Panel Data. Review of Economics and Statistics, 89 (2), 221-233.

Drucker, P. F. (1993). Post-Capital Society, Harper and Collins, New York.

Farrel, M.J. (1957). The measurement of productive efficiency. Journal of Royal Statistical Society, 120, 253-290.

Feldman, M.P. (2000). Location and innovation: the new economic geography of innovation, spillovers, and agglomeration, in Clark, G.L., Feldman, M.P. and Gertler, M.S. (Eds) Oxford Handbook of Economic Geography, 373–394. Oxford University Press, Oxford.

Fiordelisi, F., Marques-Ibanez, D. and Molyneux, P. (2011). Efficiency and risk in European banking. Journal of Banking and Finance, 35, 1315-1326.

Fischer, M.M. and Varga A. (2003). Spatial knowledge spillovers and university research: evidence from Austria, Annals of Regional Science, 37, 303–322.

Fitjar, R.D. and Rodríguez-Pose, A. (2012).Firm collaboration and modes of innovation in Norway. Research Policy, 42, 128-138.

Fritsch M., and Franke G. (2004) Innovation, regional knowledge spillovers and R&D cooperation. Research Policy, 33 (2), 245-255.

Fritsch, M. (2000) Interregional differences in R&D activities – an empirical investigation. European Planning Studies, 8, 409–427.

Fritsch, M., and Slavtchev, V. (2007). Universities and innovation in space. Industry and Innovation, 14(2), 201–218.

Fritsch, M., and Slavtchev, V. (2011). Determinants of the Efficiency of Regional Innovation Systems. Regional Studies, 45 (7), 905–918.

González-Pernía, J., Parrilli, M.D., and Peña-Legazkue, I. (2015). STI-DUI learning modes, firm university collaboration and innovation. The Journal of Technology Transfer, 40, (3), 475-492.

Greif S. and Schmiedl D. (2002). Patentatlas Deutschland. Deutsches Patent- und Markenamt, Munich.

Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. Bell Journal of Economics, 10, 92–116.

Hadri, K. (1999). estimation of a doubly heteroscedastic stochastic frontier cost function. Journal of Business and Economic Statistics, 17, 359-363.

Hasan, I., Koetter, M. and Wedow, M (2009). Regional growth and finance in Europe: is there a quality effect of bank efficiency?. Journal of Banking and Finance, 33, 1446-1453.

Henderson, R., Jaffe, A. and Trajtenberg, M. (1998). Universities as a source of commercial technology: a detailed analysis of university patenting, 1965-1988. The Review of Economics and Statistics, 80 (1), 119-127.

Jaffe, A. B. (1989). Real effects of academic research. American Economic Review, 79, 957–970.

Kalirajan, K.P. and Shand, R.T. (1999). Frontier production functions and technical efficiency measures. Journal of Economic Surveys, 13, 149–172.
Krugman, P. (1991). Geography and Trade. MIT Press, Cambridge, MA.
Kumbhakar, S.C. and Lovell, C.A.K. (2000). Stochastic Frontier Analysis. New York, Cambridge University Press.
Maietta, O. (2015). Determinants of university–firm R&D collaboration and its impact on innovation: a perspective from a low-tech industry. Research Policy, 44, 1341-1359.
Mansfield, E. (1995). Academic research underlying industrial innovations: sources, characteristics and financing. The Review of Economics and Statistics, 55-56.
Mansfield, E. (1997). Links between academic research and industrial innovations, in: David, P. & E. Steinmueller (Eds.), A production tension: university-industry collaboration in the era of knowledge-based economic development (Stanford University Press, Palo Alto).
Marrocu, E., Paci, R. and Usai, S. (2013). Productivity Growth in the Old and New Europe: The Role of Agglomeration Externalities. Journal of Regional Science, 53(3), 418–442.
Maskell, P. and Malmberg, A. (1999). The competitiveness of firms and regions: ubiquitification and the importance of localized learning. European Urban and Regional Studies, 6, 9-25.
McCann, P. and Simonen, J. (2005). Innovation, knowledge spillovers and local labour markets. Papers in Regional Science, 84 (3), 465-485.
Meeusen, W., and Van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. International Economic Review, 18, 435-444.
Nieto, M. J., and Santamaria, L. (2007). The importance of diverse collaborative networks for the novelty of product innovation. Technovation, 27, 367-377.
Pakes, A. and Griliches, Z. (1984). Patents and R & D at the firm level: a first look. In R & D, Patents, and Productivity, edited by Zvi Griliches. Chicago: Univ. Chicago Press.
Porter, M.E. (1990). The Competitive Advantage of Nations, Free Press, New York, NY.
Robin, S. and Shubert, T. (2013). Cooperation with public research institutions and success in innovation: evidence from France and Germany. Research Policy, 42, 149-166.
Rodrıguez-Pose, A. and Crescenzi, R. (2008). Research and Development, Spillovers, Innovation Systems, and the Genesis of Regional Growth in Europe. Regional Studies, 42 (1), 51–67.
Romer, P.M. (1990). Endogenous technological change. Journal of Political Economy 98 (5), 71–102.
Ronde, P. and Hussler C. (2005). Innovation in regions: what does really matter. Research Policy 34, 1150–1172.
Schumpeter, J.A.. (1934). The Theory of Economic Development. London: Oxford University Press.
Schumpeter, J.A. (1942). Capitalism, Socialism and Democracy. New York: Harper and Row.
Shane, S., (2002). Selling university technology: patterns from MIT. Management Science, 48 (1), 122–137.
Sterlacchini, A. (2008). R&D, Higher Education and Regional Growth: Uneven Linkages among European Regions. Research Policy, 37, 1096–1107.
Tavassoli , S. and Carbonara, N. (2014). The role of knowledge variety and intensity for regional Innovation. Small Business Economics, 43, 493–509.
Tidd, J., Bessant, J., Pavitt, K. (1997). Managing Innovation: Integrating Technological, Market and Organizational Change. Wiley, Chichester.
Veugelers, R. and Cassiman, B. (2005). R&D Cooperation between Firms and Universities, some empirical evidence from Belgian Manufacturing. International Journal of Industrial Organisation, 23, 5-6, 355-379.
Wang, H.J. (2003). A Stochastic Frontier Analysis of Financing Constraints on Investment: The Case of Financial Liberalization in Taiwan. Journal of Business and Economic Statistics, 21, 406-419.

## *Tables and Figures*

***Table n. 1 - Specification of inputs, outputs and exogenous factors in SFA models***

| | *Model A* | *Model B* |
|---|---|---|
| ***Inputs*** | *RD_EXP_PUBL; RD_EXP_HEI; RD_EXP_PRIV* | *RD_EMP_PUBL; RD_EMP_HEI; RD_EMP_PRIV* |
| ***Outputs*** | *PAT* | *PAT* |
| ***Explaining the inefficiency E(U)*** | *PD; FV; UR; SERV; IND; EP; INNO_PROD_PROC* | *PD; FV; UR; SERV; IND; EP; INNO_PROD_PROC* |

*Notes:*
*RD_EXP_PUBL: R&D expenditures in the public sector*
*RD_EXP_HEI: R&D expenditures in higher education institutions*
*RD_EXP_PRIV: R&D expenditures in the private sector*
*RD_EMP_PUBL: Number of R&D employees in the public sector*
*RD_EMP_HEI: Number of R&D employees in higher education institutions*
*RD_EMP_PRIV: Number of R&D employees in the private sector*
*PAT: Disclosed regional patent applications*
*PD: Population density*
*FV: Financial volume*
*UR: Unemployment rate*
*SERV: Employment in the services sector*
*IND: Employment in the industry sector*
*EP: Export*
*INNO_PROC_PROD: Product and process innovation*

***Table 2 – Descriptive statistics by macro areas***

| ***Variables*** | ***North-Western*** | ***North-Eastern*** | ***Central*** | ***Southern*** | ***Total*** |
|---|---|---|---|---|---|
| ***Production function parameters*** | | | | | |
| *PAT (in log)* | *558.91* | *391.281* | *164.309* | *37.445* | *236.33* |
| *RD_EXP_PUBL* | *104731.6* | *101565.3* | *355971.9* | *52628.6* | *134911* |
| *RD_EXP_HEI* | *309261.9* | *261839.5* | *358602.6* | *207158.1* | *269075.7* |
| *RD_EXP_PRIV* | *115262.0* | *465575.2* | *343497.3* | *109142.5* | *430998.5* |
| *RD_EMP_PUBL* | *1404.18* | *1232.819* | *4435.216* | *749.596* | *1730.625* |
| *RD_EMP_HEI* | *3927.057* | *3394.253* | *4610.809* | *2798.816* | *3509.617* |
| *RD_EMP_PRIV* | *10543.41* | *5533.737* | *3260.847* | *1107.137* | *4271.612* |
| ***Explaining the inefficiency - E(U)*** | | | | | |
| *PD* | *234.689* | *167.430* | *180.103* | *162.0001* | *180.806* |
| *FV* | *1.055* | *0.629* | *1.005* | *0.3292* | *0.669* |
| *UR* | *4.676* | *3.665* | *5.906* | *12.479* | *7.822* |
| *SERV* | *0.6755* | *0.6162* | *0.6643* | *0.6699* | *0.6592* |
| *IND* | *0.2950* | *0.3394* | *0.3032* | *0.2548* | *0.2894* |
| *EP* | *30.96* | *22.509* | *29.512* | *34.448* | *30.316* |
| *INNO_PROD_PROC* | *0.2999* | *0.3454* | *0.2770* | *0.2184* | *0.2718* |

*Notes*: Patents (PAT) represent the output. The first set of inputs consists in the amount of R&D expenditures in the public sector (RD_EXP_PUBL), in the higher education institutions (RD_EXP_HEI) and in the private sector (RD_EXP_PRIV). The second set of inputs, instead, consists in the number of R&D employees in the public sector (RD_EMPL_PUBL), in the higher education institutions (RD_EMPL_HEI) and in the private sector (RD_EMPL_PRIV). PD: Population density (measured as the number inhabitants in the region by squared kilometer); FV: Financial volume (measured as the ratio between aggregate credits and value added); UR: Unemployment rate (measured as the number of people actively looking for a job as a percentage of the labour force); SERV: employment in the services sector (measured as the number of employees in the services sector over the total number of employees); IND: employment in the industry sector (measured as the number of employees in the industry sector over the total number of employees); EP: export (measured as the values of exports as a percentage of the Gross Domestic Product); INNO_PROD_PROC: product and process innovation (measured as the number of firms which introduces product and/or process innovation over the total number of firms). All monetary aggregates are in thousands of deflated 2005 Euros.

***Table 3 – Definition of the variables and expected sign***

| ***Symbol*** | ***Description*** | ***Expected sign*** |
|---|---|---|
| ***Production function parameter*** | | |
| *RD_EXP_PUBL* | *R&D expenditures in the public sector* | + |
| *RD_EXP_HEI* | *R&D expenditures in higher education institutions* | + |
| *RD_EXP_PRIV* | *R&D expenditures in the private sector* | + |
| *RD_EMP_PUBL* | *Number of R&D employees in the public sector* | + |
| *RD_EMP_HEI* | *Number of R&D employees in higher education institutions* | + |
| *RD_EMP_PRIV* | *Number of R&D employees in the private sector* | + |
| ***Explaining the inefficiency - E(U)*** | | |
| *PD* | *Population density* | - |
| *FV* | *Financial volume* | - |
| *UR* | *Unemployment rate* | + |
| *SERV* | *Employment in the services sector* | - |
| *IND* | *Employment in the industry sector* | - |
| *EP* | *Export* | - |
| *INNO_PROD_PROC* | *Product and process innovation* | - |

*Notes*: The first set of inputs consists in the amount of R&D expenditures in the public sector (RD_EXP_PUBL), in the higher education institutions (RD_EXP_HEI) and in the private sector (RD_EXP_PRIV). The second set of inputs, instead, consists in the number of R&D employees in the public sector (RD_EMPL_PUBL), in the higher education institutions (RD_EMPL_HEI) and in the private sector (RD_EMPL_PRIV). PD: Population density (measured as the number inhabitants in the region by squared kilometer); FV: Financial volume (measured as the ratio between aggregate credits and value added); UR: Unemployment rate (measured as the number of people actively looking for a job as a percentage of the labour force); SERV: employment in the services sector (measured as the number of employees in the services sector over the total number of employees); IND: employment in the industry sector (measured as the number of employees in the industry sector over the total number of employees); EP: export (measured as the values of exports as a percentage of the Gross Domestic Product); INNO_PROD_PROC: product and process innovation (measured as the number of firms which introduces product and/or process innovation over the total number of firms).

***Table 4 – Correlation between variables***

| | | ***1*** | ***2*** | ***3*** | ***4*** | ***5*** | ***6*** | ***7*** | ***8*** | ***9*** | ***10*** | ***11*** | ***12*** | ***13*** | ***14*** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *1* | *RIS* | *1.000* | | | | | | | | | | | | | | |
| *2* | *PAT* | *0.8038* | *1.000* | | | | | | | | | | | | | |
| *3* | *RD_EXP_PUBL* | *-0.0514* | *0.1508* | *1.000* | | | | | | | | | | | | |
| *4* | *RD_EXP_HEI* | *0.3179* | *0.6387* | *0.5753* | *1.000* | | | | | | | | | | | |
| *5* | *RD_EXP_PRIV* | *0.6916* | *0.9184* | *0.2978* | *0.7002* | *1.000* | | | | | | | | | | |
| *6* | *RD_EMPL_PUBL* | *-0.0503* | *0.1530* | *0.9903* | *0.5979* | *0.3157* | *1.0000* | | | | | | | | | |
| *7* | *RD_EMPL_HEI* | *0.2965* | *0.6238* | *0.5665* | *0.9798* | *0.6927* | *0.5905* | *1.000* | | | | | | | | |
| *8* | *RD_EMPL_PRIV* | *0.7368* | *0.9270* | *0.2537* | *0.6857* | *0.9844* | *0.2694* | *0,6857* | *1.000* | | | | | | | |
| *9* | *PD* | *0.2397* | *0.5183* | *0.4206* | *0.7621* | *0.5837* | *0.4296* | *0.7545* | *0.5675* | *1.000* | | | | | | |
| *10* | *FV* | *0.5510* | *0.5441* | *-0.1584* | *0.1806* | *0.4654* | *-0.1519* | *0.2031* | *0.4956* | *0.1738* | *1.000* | | | | | |
| *11* | *UR* | *-0.7079* | *-0.4514* | *-0.1584* | *-0.0415* | *-0.3527* | *-0.0371* | *-0.0112* | *-0.3816* | *0.0508* | *-0.4145* | *1.000* | | | | |
| *12* | *SERV* | *-0.5242* | *-0.4083* | *0.4471* | *0.1041* | *-0.2258* | *0.4518* | *0.1056* | *-0.2879* | *0.1759* | *-0.5201* | *0.3634* | *1.000* | | | |
| *13* | *IND* | *0.6872* | *0.5515* | *-0.2740* | *0.0612* | *0.3961* | *-0.2798* | *0.0511* | *0.4540* | *0.0086* | *0.6237* | *-0.5953* | *-0.9198* | *1.000* | | |
| *14* | *EP* | *-0.1893* | *0.0217* | *0.4952* | *0.2087* | *0.1906* | *0.4973* | *0.2026* | *0.1412* | *0.2763* | *-0.2838* | *0.2025* | *0.2245* | *-0.1996* | *1.000* | |
| *15* | *INNO_PROC_PROD* | *0.7524* | *0.6230* | *0.0498* | *0.2666* | *0.5452* | *0.0473* | *0.2611* | *0.5882* | *0.1499* | *0.3548* | *-0.6512* | *-0.4862* | *0.6396* | *-0.1016* | *1.000* |

*Notes*: Regional Innovation System (RIS) efficiency denotes the technical efficiency calculated using a knowledge production function (KPF). Patents (PAT) represent the output. The first set of inputs consists in the amount of R&D expenditures in the public sector (RD_EXP_PUBL), in the higher education institutions (RD_EXP_HEI) and in the private sector (RD_EXP_PRIV). The second set of inputs, instead, consists in the number of R&D employees in the public sector (RD_EMPL_PUBL), in the higher education institutions (RD_EMPL_HEI) and in the private sector (RD_EMPL_PRIV). PD: Population density (measured as the number inhabitants in the region by squared kilometer); FV: Financial volume (measured as the ratio between aggregate credits and value added); UR: Unemployment rate (measured as the number of people actively looking for a job as a percentage of the labour force); SERV: employment in the services sector (measured as the number of employees in the services sector over the total number of employees); IND: employment in the industry sector (measured as the number of employees in the industry sector over the total number of employees); EP: export (measured as the values of exports as a percentage of the Gross Domestic Product); INNO_PROD_PROC: process and product innovation (measured as the number of firms which introduces product and/or process innovation over the total number of firms).

***Table 5 - Estimates for the knowledge production function and for the inefficiency components according to the stochastic frontier approach - Mean values***

| *Variables* | *Model A* | | | |
|---|---|---|---|---|
| | ***(1)*** | ***(2)*** | ***(3)*** | ***(4)*** |
| ***Production function parameters*** | | | | |
| *RD_EXP_PUBL (in log)* | *0.239**** | *0.239*** | *0.157**** | *0.157** |
| | *(0.050)* | *(0.106)* | *(0.046)* | *(0.089)* |
| *RD_EXP_HEI (in log)* | *0.422**** | *0.422*** | *0.537**** | *0.537**** |
| | *(0.077)* | *(0.203)* | *(0.055)* | *(0.073)* |
| *RD_EXP_PRIV (in log)* | *0.338**** | *0.338** | *0.170*** | *0.170* |
| | *(0.078)* | *(0.183)* | *(0.073)* | *(0.107)* |
| ***Explaining the inefficiency*** | | | | |
| *PD (in log)* | *-0.458**** | *-0.458**** | *-0.525**** | *-0.525**** |
| | *(0.079)* | *(0.129)* | *(0.076)* | *(0.165)* |
| *FV (in log)* | *-0.196*** | *-0.196* | *-0.060* | *-0.060* |
| | *(0.086)* | *(0.171)* | *(0.078)* | *(0.152)* |
| *UR (in log)* | *0.957**** | *0.957**** | *1.089**** | *1.089**** |
| | *(0.215)* | *(0.263)* | *(0.186)* | *(0.231)* |
| *SERV* | *-2.395* | *-2.395* | *-4.170** | *-4.170** |
| | *(2.472)* | *(2.219)* | *(2.269)* | *(2.327)* |
| *IND* | *-4.620*** | *-4.620** | *-6.260**** | *-6.260*** |
| | *(2.530)* | *(2.796)* | *(2.246)* | *(2.500)* |
| *EP (in log)* | *-0.372**** | *-0.372**** | *-0.322**** | *-0.322**** |
| | *(0.092)* | *(0.137)* | *(0.080)* | *(0.092)* |
| *INNO_PROD_PROC (in log)* | *-0.484*** | *-0.484** | *-0.463*** | *-0.463** |
| | *(0.251)* | *(0.276)* | *(0.228)* | *(0.267)* |
| *NORTH-WESTERN* | *-1.283**** | *-1.283*** | *-1.406**** | *-1.406*** |
| | *(0.305)* | *(0.540)* | *(0.294)* | *(0.742)* |
| *NORTH-EASTERN* | *-0.622*** | *-0.622* | *-0.614*** | *-0.614* |
| | *(0.340)* | *(0.423)* | *(0.270)* | *(0.444)* |
| *CENTRAL* | *-0.037* | *-0.037* | *-0.069* | *-0.069* |
| | *(0.210)* | *(0.291)* | *(0.191)* | *(0.314)* |
| *Log-likelihood* | *-40.9265* | *-40.9265* | *-24.5773* | *-24.5773* |
| *LR test for null inefficiency component* ($p > \chi^2$) | *207.84* | - | *239.63* | - |
| | *(0.0000)* | | *(0.0000)* | |
| *Wald statistic* | *40.72* | *396.15* | *1170.87* | *12486.87* |
| $\sigma_u$ | *0.352**** | *0.352**** | *0.338**** | *0.338**** |
| | *(0.037)* | *(0.104)* | *(0.027)* | *(0.074)* |
| $\sigma_v$ | *0.137**** | *0.137*** | *0.084**** | *0.084* |
| | *(0.033)* | *(0.061)* | *(0.031)* | *(0.098)* |
| $\lambda$ | *2.570**** | *2.570**** | *4.006**** | *4.006**** |
| | *(0.065)* | *(0.157)* | *(0.052)* | *(0.159)* |
| *N* | *160* | *160* | *160* | *160* |
| *Time dummies* | *Yes* | *Yes* | *Yes* | *Yes* |
| *Time trend in E(U)* | *Yes* | *Yes* | *Yes* | *Yes* |

Notes: Model A, Column 1, with the imposition of the linear homogeneity of degree 1 in inputs; Model A, Column 2, with the imposition of the linear homogeneity of degree 1 in inputs and standard errors clustered at region and year level; Model A, Column 3, without the imposition of the linear homogeneity of degree 1 in inputs; Model A, Column 4, without the imposition of the linear homogeneity of degree 1 in inputs and standard errors clustered at region and year level. The set of inputs consists in the amount of R&D expenditures in the public sector (RD_EXP_PUBL), in the higher education institutions (RD_EXP_HEI) and in the private sector (RD_EXP_PRIV). PD: Population density (measured as the number inhabitants in the region by squared kilometer); FV: Financial volume (measured as the ratio between aggregate credits and value added); UR: Unemployment rate (measured as the number of people actively looking for a job as a percentage of the labour force); SERV: employment in the services sector (measured as the number of employees in the services sector over the total number of employees); IND: employment in the industry sector (measured as the number of employees in the industry sector over the total number of employees); EP: export (measured as the values of exports as a percentage of the Gross Domestic Product); INNO_PROD_PROC: product and process innovation (measured as the number of firms which introduces product and/or process innovation over the total number of firms). All models consider time dummies in the frontier and in the inefficiency component. Southern area is our benchmark group. Standard deviation in brackets. The LR test evaluates the restricted and unrestricted models with and without the exogenous factors in the inefficiency term (the null hypothesis that there is no heteroschedasticity in the error term is rejected, at 1% significance level in all the models).

***Table 6 – Technical efficiency by macro areas and by regions according to the stochastic frontier approach***

| | Model A | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** |
| | ***Mean*** | ***Std. Dev.*** | ***Min*** | ***Max*** | ***Mean*** | ***Std. Dev.*** | ***Min*** | ***Max*** |
| *Geographical areas* | | | | | | | | |
| *Central* | *0.3661* | *0.2168* | *0.0969* | *0.8287* | *0.2878* | *0.1405* | *0.1289* | *0.5698* |
| *Nord-Eastern* | *0.6997* | *0.1989* | *0.3093* | *0.9455* | *0.6540* | *0.2229* | *0.2771* | *0.9472* |
| *North-Western* | *0.7446* | *0.2616* | *0.2434* | *0.9559* | *0.6967* | *0.2793* | *0.2259* | *0.9718* |
| *Southern* | *0.1379* | *0.0624* | *0.0209* | *0.3050* | *0.1015* | *0.0501* | *0.0096* | *0.2297* |
| *Regions* | | | | | | | | |
| *Abruzzo* | *0.2591* | *0.0327* | *0.1973* | *0.3050* | *0.2038* | *0.0212* | *0.1623* | *0.2297* |
| *Basilicata* | *0.0967* | *0.0303* | *0.0400* | *0.1274* | *0.0664* | *0.0222* | *0.0242* | *0.0879* |
| *Calabria* | *0.1628* | *0.0446* | *0.1055* | *0.2550* | *0.0958* | *0.0239* | *0.0637* | *0.1385* |
| *Campania* | *0.1104* | *0.0240* | *0.0745* | *0.1501* | *0.0995* | *0.0238* | *0.0628* | *0.1387* |
| *Emilia.Romagna* | *0.8348* | *0.0762* | *0.6728* | *0.9117* | *0.8166* | *0.0745* | *0.6793* | *0.9246* |
| *Friuli.Venezia Giulia* | *0.5663* | *0.0983* | *0.4172* | *0.7330* | *0.4600* | *0.0731* | *0.3340* | *0.5574* |
| *Lazio* | *0.1074* | *0.0115* | *0.0969* | *0.1322* | *0.1406* | *0.0148* | *0.1289* | *0.1718* |
| *Liguria* | *0.3511* | *0.0715* | *0.2434* | *0.4235* | *0.3262* | *0.0649* | *0.2259* | *0.4015* |
| *Lombardia* | *0.9389* | *0.0144* | *0.9069* | *0.9520* | *0.9632* | *0.0111* | *0.0390* | *0.9718* |
| *Marche* | *0.6641* | *0.1302* | *0.4634* | *0.8287* | *0.4561* | *0.0842* | *0.3450* | *0.5698* |
| *Molise* | *0.0822* | *0.0598* | *0.0209* | *0.1885* | *0.0409* | *0.0296* | *0.0096* | *0.0928* |
| *Piemonte* | *0.8848* | *0.0439* | *0.7823* | *0.9197* | *0.8648* | *0.0733* | *0.6966* | *0.9326* |
| *Puglia* | *0.1447* | *0.0233* | *0.1127* | *0.1803* | *0.1154* | *0.0174* | *0.0933* | *0.1404* |
| *Sardegna* | *0.1031* | *0.0282* | *0.0607* | *0.1410* | *0.0710* | *0.0197* | *0.0402* | *0.0974* |
| *Sicilia* | *0.1306* | *0.0326* | *0.0810* | *0.1844* | *0.1041* | *0.0259* | *0.0633* | *0.1469* |
| *Toscana* | *0.4020* | *0.0523* | *0.3514* | *0.4848* | *0.3706* | *0.0433* | *0.3271* | *0.4434* |
| *Trentino Alto Adig* | *0.4810* | *0.0971* | *0.3093* | *0.5900* | *0.4350* | *0.0911* | *0.2771* | *0.5453* |
| *Umbria* | *0.2911* | *0.0512* | *0.2162* | *0.3883* | *0.1839* | *0.0318* | *0.1362* | *0.2383* |
| *Valle d'Aosta* | *0.8231* | *0.2001* | *0.4201* | *0.9559* | *0.6112* | *0.2271* | *0.2487* | *0.8897* |
| *Veneto* | *0.9168* | *0.0251* | *0.8638* | *0.9455* | *0.9043* | *0.0360* | *0.8335* | *0.9472* |

Notes: Model A, Columns 1, 2, 3 and 4, with the imposition of the linear homogeneity of degree 1 in inputs; Model A, Columns 5, 6, 7 and 8, without the imposition of the linear homogeneity of degree 1 in inputs.

***Table 7 - Estimates for the knowledge production function and for the inefficiency components according to the stochastic frontier approach - Mean values, 2000-2009 period***

| *Variables* | *Model B* | | | |
|---|---|---|---|---|
| | *(1)* | *(2)* | *(3)* | *(4)* |
| ***Production function parameters*** | | | | |
| *RD_EMPL_PUBL (in log)* | *0.031* | *0.031* | *0.099* | *0.099* |
| | *(0.079)* | *(0.212)* | *(0.080)* | *(0.176)* |
| *RD_EMPL_HEI (in log)* | *0.537**** | *0.537**** | *0.363**** | *0.363**** |
| | *(0.068)* | *(0.164)* | *(0.097)* | *(0.135)* |
| *RD_EMPL_PRIV (in log)* | *0.430**** | *0.430* | *0.293**** | *0.293*** |
| | *(0.098)* | *(0.339)* | *(0.067)* | *(0.152)* |
| ***Explaining the inefficiency*** | | | | |
| *PD (in log)* | *-0.222**** | *-0.222* | *-0.413**** | *-0.413**** |
| | *(0.086)* | *(0.249)* | *(0.064)* | *(0.122)* |
| *FV (in log)* | *-0.251**** | *-0.251* | *-0.225**** | *-0.225*** |
| | *(0.097)* | *(0.211)* | *(0.073)* | *(0.098)* |
| *UR (in log)* | *1.270**** | *1.270**** | *0.881**** | *0.881**** |
| | *(0.267)* | *(0.411)* | *(0.194)* | *(0.272)* |
| *SERV* | *1.093* | *1.093* | *-4.601*** | *-4.601** |
| | *(2.600)* | *(3.361)* | *(2.268)* | *(2.411)* |
| *IND* | *4.080* | *4.080* | *-4.719*** | *-4.719** |
| | *(3.565)* | *(9.580)* | *(2.221)* | *(2.563)* |
| *EP (in log)* | *-0.226**** | *-0.226* | *-0.184*** | *-0.184* |
| | *(0.089)* | *(0.143)* | *(0.083)* | *(0.124)* |
| *INNO_PROD_PROC (in log)* | *-0.497*** | *-0.497**** | *-0.406*** | *-0.406** |
| | *(0.245)* | *(0.182)* | *(0.223)* | *(0.233)* |
| *NORTH-WESTERN* | *-0.239* | *-0.239* | *-0.722**** | *-0.722* |
| | *(0.421)* | *(1.270)* | *(0.271)* | *(0.567)* |
| *NORTH-EASTERN* | *-0.239* | *-0.239* | *-0.543*** | *-0.543* |
| | *(0.421)* | *(0.397)* | *(0.266)* | *(0.417)* |
| *CENTRAL* | *-0.529* | *-0.529* | *-0.030* | *-0.030* |
| | *(0.329)* | *(0.208)* | *(0.189)* | *(0.286)* |
| *Log-likelihood* | *-49.9566* | *-49.9566* | *-19.1042* | *-19.1042* |
| *LR test for null inefficiency component* ($p > \chi^2$) | *187.95* | - | *243.82* | - |
| | *(0.0000* | | *(0.0000)* | |
| *Wald statistic* | *47.17* | *141.79* | *1.34e+09* | *1.62* |
| $\sigma_u$ | *0.294**** | *0.294**** | *0.348**** | *0.348**** |
| | *(0.050)* | *(0.103)* | *(0.022)* | *(0.068)* |
| $\sigma_v$ | *0.227**** | *0.227**** | *8.35e-09* | *8.35e-09* |
| | *(0.040)* | *(0.053)* | *(3.34e-06)* | *(8.08e-09)* |
| $\lambda$ | *1.290**** | *1.290**** | *4.17e+07**** | *4.17e+07**** |
| | *(0.084)* | *(0.133)* | *(0.022)* | *(0.068)* |
| *N* | *160* | *160* | *160* | *160* |
| *Time dummies* | *Yes* | *Yes* | *Yes* | *Yes* |
| *Time trend in E(U)* | *Yes* | *Yes* | *Yes* | *Yes* |

Notes: Model B, Column 1, with the imposition of the linear homogeneity of degree 1 in inputs; Model B, Column 2, with the imposition of the linear homogeneity of degree 1 in inputs and standard errors clustered at region and year level; Model B, Column 3, without the imposition of the linear homogeneity of degree 1 in inputs; Model B, Column 4, without the imposition of the linear homogeneity of degree 1 in inputs and standard errors clustered at region and year level. The set of inputs consists in the number of R&D employees in the public sector (RD_EMPL_PUBL), in the higher education institutions (RD_EMPL_HEI) and in the private sector (RD_EMPL_PRIV). PD: Population density (measured as the number inhabitants in the region by squared kilometer); FV: Financial volume (measured as the ratio between aggregate credits and value added); UR: Unemployment rate (measured as the number of people actively looking for a job as a percentage of the labour force); SERV: employment in the services sector (measured as the number of employees in the services sector over the total number of employees); IND: employment in the industry sector (measured as the number of employees in the industry sector over the total number of employees); EP: export (measured as the values of exports as a percentage of the Gross Domestic Product); INNO_PROD_PROC: product and process innovation (measured as the number of firms which introduces product and/or process innovation over the total number of firms). All models consider time dummies in the frontier and in the inefficiency component. Southern area is our benchmark group. Standard deviation in brackets. The LR test evaluates the restricted and unrestricted models with and without the exogenous factors in the inefficiency term (the null hypothesis that there is no heteroschedasticity in the error term is rejected, at 1% significance level in all the models).

***Table 8 – Technical efficiency by macro areas and by regions according to the stochastic frontier approach***

| | Model B | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** |
| | ***Mean*** | ***Std. Dev.*** | ***Min*** | ***Max*** | ***Mean*** | ***Std. Dev.*** | ***Min*** | ***Max*** |
| *Geographical areas* | | | | | | | | |
| *Central* | *0.4373* | *0.1806* | *0.1699* | *0.7403* | *0.3264* | *0.1407* | *0.1500* | *0.6333* |
| *Nord-Eastern* | *0.8252* | *0.1032* | *0.5806* | *0.9494* | *0.6555* | *0.2529* | *0.3087* | *0.9999* |
| *North-Western* | *0.7036* | *0.2424* | *0.3083* | *0.9360* | *0.5831* | *0.2840* | *0.1682* | *0.9999* |
| *Southern* | *0.1978* | *0.0714* | *0.0527* | *0.4054* | *0.1224* | *0.0528* | *0.0100* | *0.2341* |
| *Regions* | | | | | | | | |
| *Abruzzo* | *0.2983* | *0.0329* | *0.2511* | *0.3407* | *0.1994* | *0.0242* | *0.1692* | *0.2341* |
| *Basilicata* | *0.1563* | *0.0393* | *0.0925* | *0.2076* | *0.0694* | *0.0284* | *0.0262* | *0.1006* |
| *Calabria* | *0.2763* | *0.0734* | *0.1718* | *0.4054* | *0.1366* | *0.0436* | *0.0810* | *0.2004* |
| *Campania* | *0.1329* | *0.0135* | *0.1042* | *0.1443* | *0.1197* | *0.0264* | *0.0700* | *0.1499* |
| *Emilia.Romagna* | *0.8748* | *0.0509* | *0.7992* | *0.9494* | *0.8323* | *0.0794* | *0.7012* | *0.9248* |
| *Friuli.Venezia Giulia* | *0.6676* | *0.0656* | *0.5806* | *0.7598* | *0.4380* | *0.0715* | *0.3348* | *0.5263* |
| *Lazio* | *0.2005* | *0.0132* | *0.1699* | *0.2114* | *0.1747* | *0.0188* | *0.1500* | *0.1944* |
| *Liguria* | *0.4166* | *0.0683* | *0.3083* | *0.5187* | *0.2886* | *0.0591* | *0.2181* | *0.3563* |
| *Lombardia* | *0.8717* | *0.0327* | *0.8161* | *0.9206* | *0.9714* | *0.0544* | *0.8607* | *0.0999* |
| *Marche* | *0.6735* | *0.0748* | *0.5355* | *0.7403* | *0.4871* | *0.0879* | *0.3611* | *0.6333* |
| *Molise* | *0.1389* | *0.0677* | *0.0527* | *0.2386* | *0.0480* | *0.0363* | *0.0100* | *0.1114* |
| *Piemonte* | *0.7044* | *0.0562* | *0.5093* | *0.7566* | *0.6482* | *0.0602* | *0.5419* | *0.7331* |
| *Puglia* | *0.2037* | *0.0220* | *0.1649* | *0.2341* | *0.1517* | *0.0298* | *0.0982* | *0.1982* |
| *Sardegna* | *0.1977* | *0.0449* | *0.1334* | *0.2812* | *0.1060* | *0.0309* | *0.0586* | *0.1477* |
| *Sicilia* | *0.1634* | *0.0280* | *0.1266* | *0.2026* | *0.1298* | *0.0347* | *0.0746* | *0.1762* |
| *Toscana* | *0.4802* | *0.0638* | *0.4021* | *0.5868* | *0.4154* | *0.0476* | *0.3335* | *0.4742* |
| *Trentino Alto Adig* | *0.8599* | *0.0295* | *0.8245* | *0.9094* | *0.3993* | *0.0603* | *0.3087* | *0.4903* |
| *Umbria* | *0.3948* | *0.0529* | *0.3175* | *0.4784* | *0.2285* | *0.0399* | *0.1676* | *0.2792* |
| *Valle d'Aosta* | *0.8611* | *0.1356* | *0.5863* | *0.9360* | *0.3711* | *0.1141* | *0.1682* | *0.4999* |
| *Veneto* | *0.8984* | *0.0267* | *0.8458* | *0.9309* | *0.9523* | *0.0531* | *0.8424* | *0.9999* |

Notes: Model B, Columns 1, 2, 3 and 4, with the imposition of the linear homogeneity of degree 1 in inputs; Model B, Columns 5, 6, 7 and 8, without the imposition of the linear homogeneity of degree 1 in inputs.

*Figure 1 – Inputs and outputs used in the production function over 2000–2009 time-span, by regions*

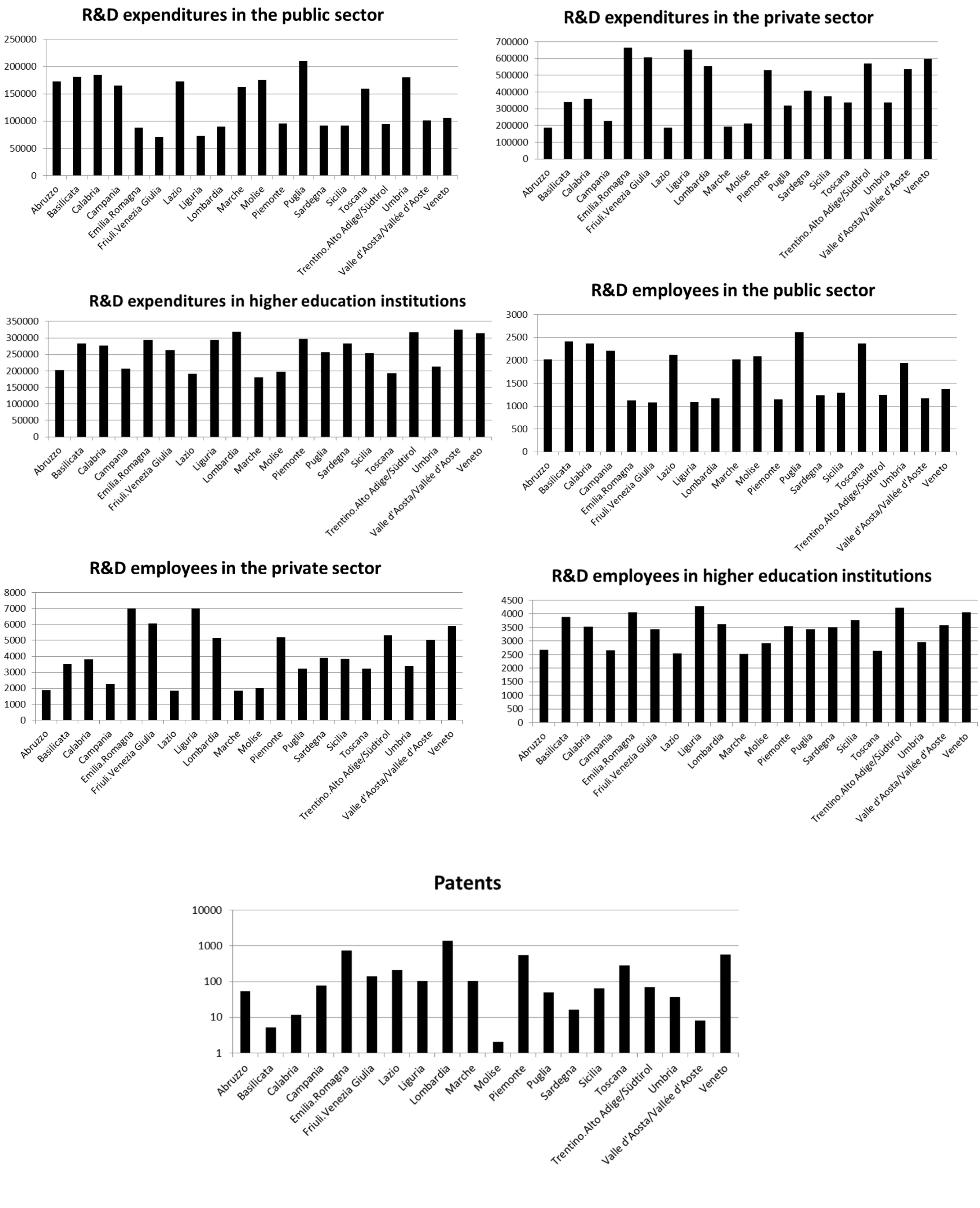


***Figure 2 – Inputs used in the production function, by sectors, by regions and by tertiles***

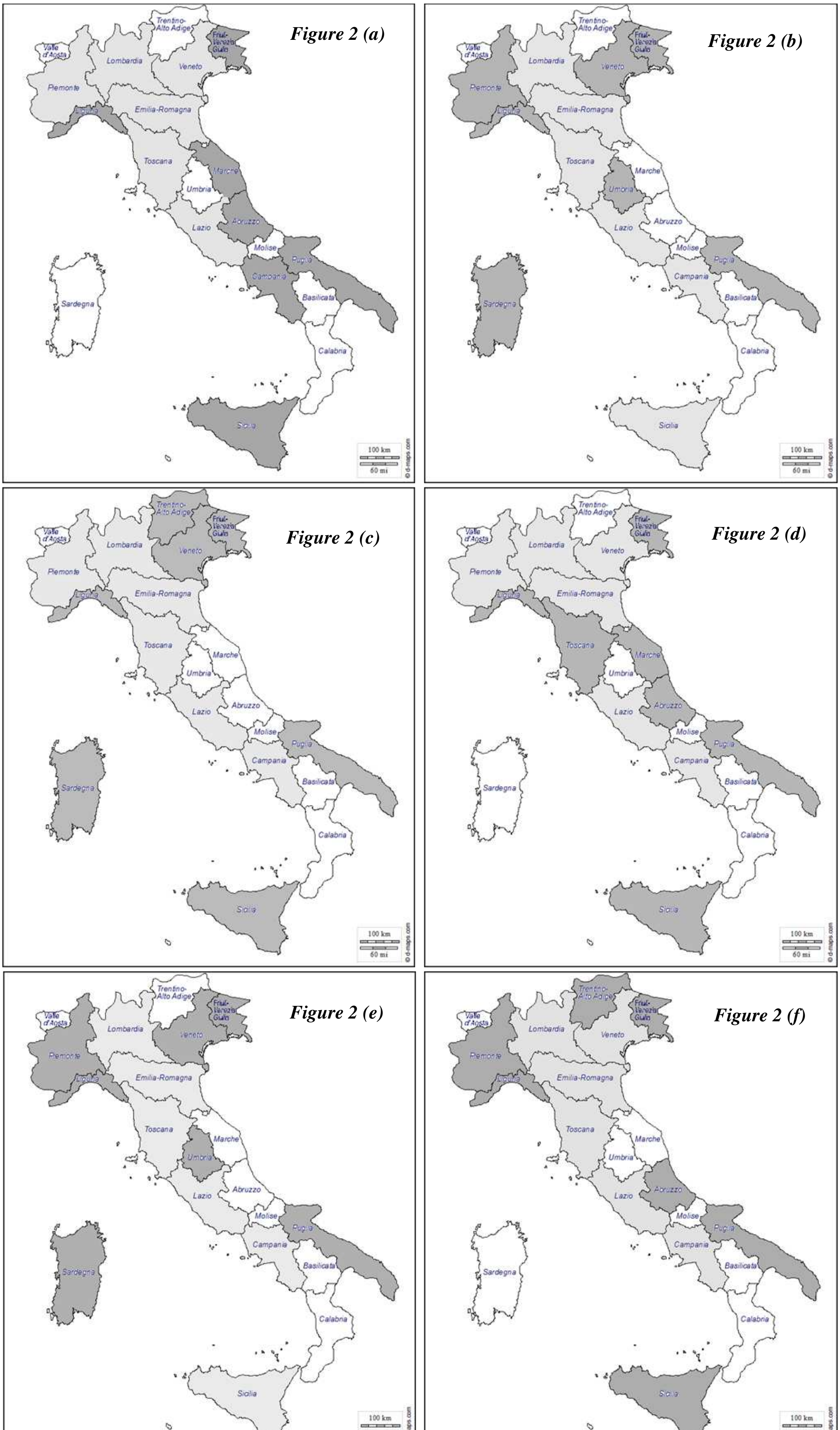



*Figure 2(a)-2(b)-2(c) shows the number of R&D expenditures, respectively, in private sector, higher education institutions and public sector.*
*Figure 2(d)-2(e)-2(f) shows the number of R&D employees, respectively, in private sector, higher education institutions and public sector.*
*In white, regions within the first tertile; in dark grey, regions within the second tertile; in light grey, regions within the third quartile.*

*Figure 3 – Patents and RIS efficiency scores - by regions and by tertiles*

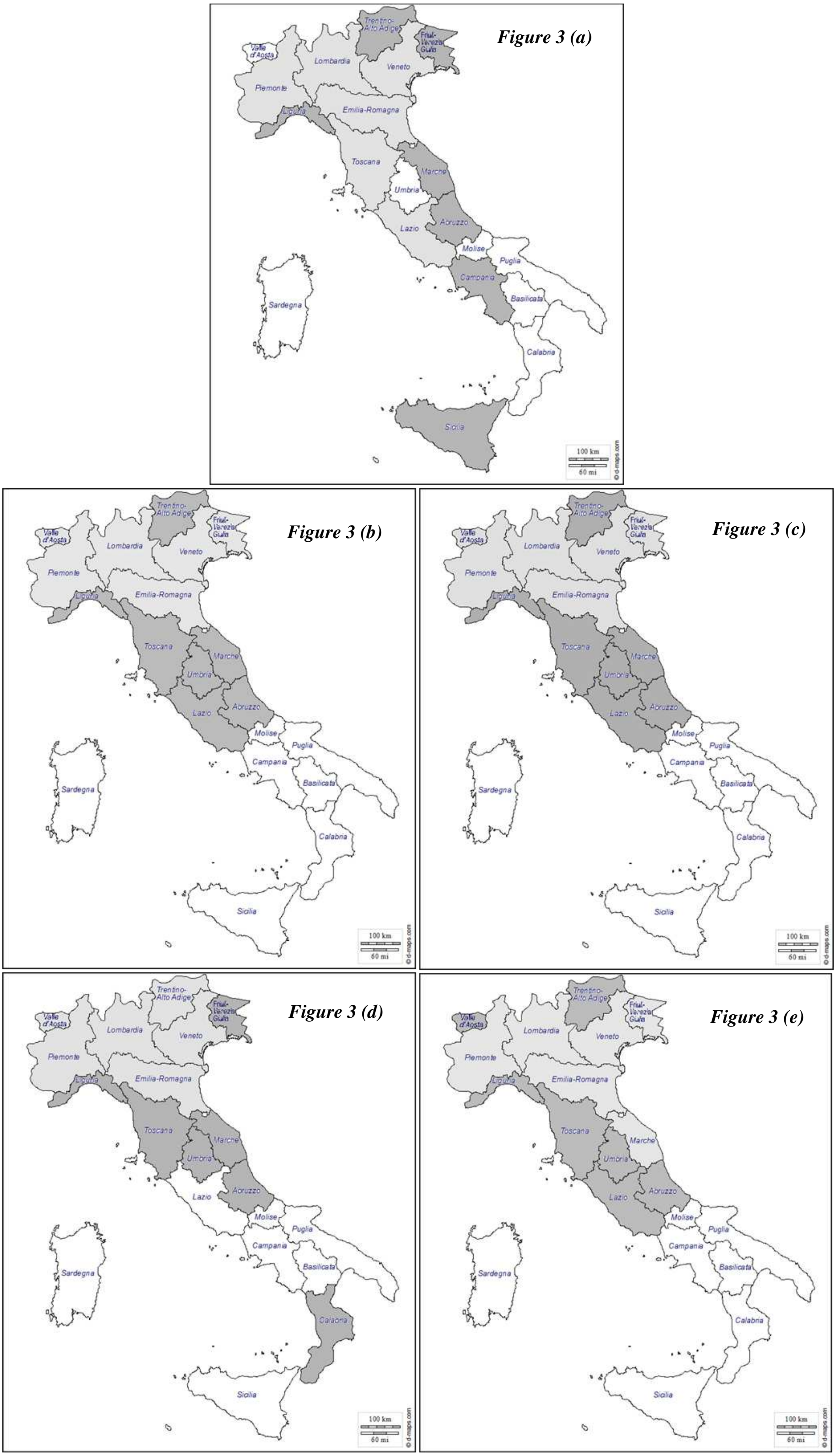
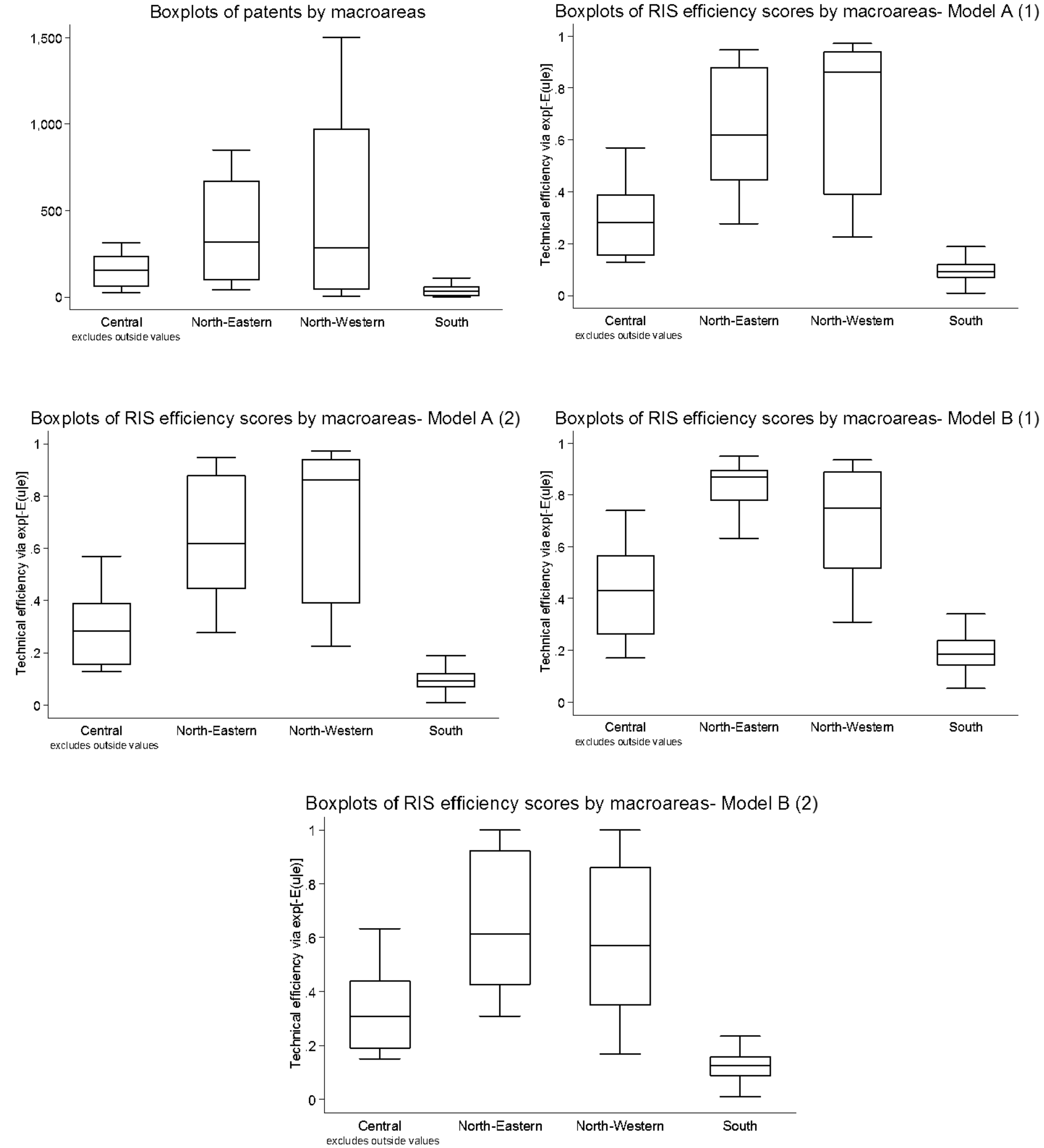





*Figure 3(a) shows the number of patents*
*Figure 3(b)-3(c) shows RIS efficiency scores when R&D expenditures are used as innovative inputs (see Model A in Table 1 and Column 1 in Table 6)*
*Figure 3(d)-3(e) shows RIS efficiency scores when R&D employees are used as innovative inputs (see Model B in Table 1 and Column 1 in Table 8)*
*In white, regions within the first tertile; in dark grey, regions within the second tertile; in light grey, regions within the third quartile.*

*Figure 4 – Patents and RIS efficiency scores, by macroareas*




*Model A (1) refers to RIS efficiency scores when R&D expenditures are used as innovative inputs with the imposition of the linear homogeneity of degree 1 in inputs (see Model A in Table 1 and Column 1 in Table 6);*
*Model A (2) refers to RIS efficiency scores when R&D expenditures are used as innovative inputs without the imposition of the linear homogeneity of degree 1 in inputs (see Model A in Table 1 and Column 5 in Table 6);*
*Model B (1) refers to RIS efficiency scores when R&D employees are used as innovative inputs with the imposition of the linear homogeneity of degree 1 in inputs (see Model B in Table 1 and Column 1 in Table 8);*
*Model B (2) refers to RIS efficiency scores when R&D employees are used as innovative inputs without the imposition of the linear homogeneity of degree 1 in inputs (see Model B in Table 1 and Column 5 in Table 8).*