Doctoral Thesis

University of Trento

CIFREM
Interdepartmental Centre for Research Training
in Economics and Management

Doctoral School in Economics and Management

Essays on Productivity and Efficiency Analysis

a dissertation

submitted to the doctoral school of economics and management

in partial fulfillment of the requirements

for the Doctoral degree

(Ph.D.)

in Economics and Management

Fabio Pieri

November 2010

Advisors

Advisor:
Prof. Enrico Zaninotto
University of Trento

Co-Advisor:
Prof. Davide Castellani
University of Perugia

Referee Committee

Prof. Christopher L. Gilbert
University of Trento

Prof. Antonio Alvarez
University of Oviedo (Spain)

Doctoral Committee

Prof. Massimo Riccaboni
University of Trento

Dr. Ornella Tarola
"Sapienza", University of Rome

Prof. Howard Gospel
King's College London (UK)

Acknowledgments

At the end of this long journey, I really should write another manuscript, explaining how I got here, the people I met and the things I discovered, mostly about myself. Actually, the only thing I would like to emphasize here (and in my mind) is the large group of people who have guided and helped me over these five years.

Thanks first go to my two supervisors. Enrico Zaninotto helped and encouraged me since the beginning of the Ph.D. program, never imposing his own views but leaving me free to discover and deepen my research interests: it was not easy at the beginning, but now I understand and appreciate his approach. Davide Castellani 'bet' on me without being insured about my 'success', and I will always be grateful to him for that: the two years spent working with him, I learned many things not only from a scientific point of view. I would also like to thank Christopher L. Gilbert and Antonio Alvarez, as members of the Referee Committee, for their insightful and valuable comments on the first version of the text: by working on the revisions they suggested, I was able to improve the thesis considerably. Alessandro Petretto pushed me in the direction of the Ph.D. program in Trento, and I thank him for this.

I would like to express my sincere gratitude to all the professors and colleagues I met during the conferences and seminars I attended: they spent time reading and discussing with me various issues raised in this thesis, and I am indebted to all of them. Here is a (non-exhaustive) list: Ornella Wanda Maietta, Alan Wall, William H. Greene, Hung-Jen Wang, Oleg Badunenko, William Horrace, Stefano Comino, Marco Corsino, Enrico Tundis, Sandro Trento, Roberto Gabriele, Marco Zamarian, Loris Gaio, Stefano Schiavo, Chiara Tomasi, Paola Villa, Cristiano Perugini, Francesco Venturini, Paolo Polinori, Fabrizio Pompei, David Aristei, Gianfranco Di Vaio, Tiziana Laureti, Angelo Zago, Camilla Mastromarco, Kumaraswamy Velupillai, Stefano Zambelli, Secondo Rolfo, Concetta Castiglione, Gabriel Walton and Jill Giraldi. Thoughts also go to my colleagues at CIFREM, with whom I spent difficult but also stimulating days, especially during the first year of the doctorate at Trento: Dominique, Matteo, Fabio T., Riccarda, Sridhar, Binh, Nicole, Francesca M., Francesca F., Giulia, Stefania, Gianna, Sofia and my very good friend Mariana ("in the end ... everything will turn out all right!").

Special thanks go to my parents, my sister Silvia, and my grandparents: they put up with me during bad times and rejoiced during good ones, all of them always close to me. I also want to thank a large group of friends (just because they are friends): Thomas & Francy, Umbe, Paolino, Laura, Alessio & Cristina (& Andrea!), la Fedora, Andrea, Federico, Marco, Tommy, Chiara, la Gabri, David & Francesca, Francesco ("... dottorandi ... !!!"), Beppe, Mattia, Nicola, Betta, Toni, Domenico, Chiara, Simone, Rossana and Stefania. I thank them because they were with me in various circumstances while I was working on this thesis and each one of them also contributed to it, because they contributed to my personal growth. Surely I am forgetting someone: sorry for that! Yet more thanks go to all the people who said "No!", putting a spoke in my wheel: they allowed me to develop and have more self-confidence, sometimes changing my ideas and beliefs. And they made me remember that all of us must be the greatest 'cheerleaders' of ourselves!

This thesis is dedicated to all the people who care about me, and all the people I care about.

“The Navy Diver is not a fighting man, he is a salvage expert. If it is lost underwater, he finds it. If it’s sunk, he brings it up. If it’s in the way, he moves it. If he’s lucky, he will die young, 200 feet beneath the waves, for that is the closest he’ll ever get to being a hero.”
(from the movie *Men of Honor*; USA, 2000)

Per aspera sic itur ad astra
(Latin phrase)

“E dove dunque vogliamo arrivare? Al di là del mare? Dove ci trascina questa possente avidità, che è più forte di qualsiasi altro desiderio? Perché proprio in quella direzione, laggiù dove sono fino ad oggi tramontati tutti i soli dell’umanità?”
(*Aurora*, Friedrich Wilhelm Nietzsche)

“Pieno di Dio non temo l’avvenire, perché qualunque cosa accada non sarà mai più grande di questa mia Anima”
(Peruvian prayer)

Contents

List of Tables

List of Figures

Chapter 1

Productivity and efficiency measurement

1.1 Introduction

This thesis is made up of four chapters on productivity and efficiency analysis. The first chapter is a critical review of theoretical and empirical literatures related to this broad field, which has attracted a considerable amount of economic research in the last years; the other three chapters are original contributions in different directions. Chapter 2 consists of an extensive Monte Carlo exercise on the misspecification of the inefficiency distribution in stochastic frontier models, Chapter 3 investigates, both theoretically and empirically, the relationship between vertical integration and firm efficiency in the Italian machine tool industry, and, finally, Chapter 4 sheds light on the effect of both inward and outward foreign direct investments on regional productivity growth in Europe. Although each chapter has its own independence, two features characterize the entire thesis: the detection of large differences in production performance both at the micro and aggregate level, and the attempt to relate these differences to other aspects of the production units, starting from economic theory.

The remainder of this chapter is organized as follows: Section 1.2 introduces concepts of productivity and efficiency and Section 1.3 overviews different approaches to the theory of production in economics; Section 1.4 introduces a framework of analysis which deals with productivity and efficiency in a coherent way. A list of the available methodologies in order to measure productivity and efficiency is presented in Section 1.5, with a particular attention devoted the strengths and the weakness of each of them. This digression will be particularly useful for understanding the taken choices regarding methodologies which have been employed in the applied works of the thesis. An investigation into the determinants of productivity and efficiency, both at the micro level and at the macro level, is presented in section 1.6. Concluding remarks and links to other chapters are presented in Section 1.7.

1.2 Productivity and efficiency in economics

Productivity and *efficiency* are two economic concepts frequently used both in scientific articles and in the popular press. They deal with the economic performance of the production unit under observation (either a firm or an organization, an industry or a country). They refer to the production process which the producer accomplishes, transforming a set of inputs, either in form or in location[1], into a set of 'useful' outputs (Greene, 2008, p.97) Productivity and efficiency are frequently used as overlapping terms in order to indicate the performance of a production unit; however, they are two related but separate concepts. In the recent survey by Fried, Lovell, and Schmidt (2008), productivity is defined as the ratio of the outputs of a production process to its inputs, while efficiency refers to the comparison between observed and 'optimal' outputs and observed and 'optimal' inputs. Productivity is a residual: difference in productivity among producers in the same time period, or variations in a given period of time (productivity growth) can be defined as the unexplained part of the variation of output after having taken the variation in inputs into account. Efficiency is a residual too, but it also requires the existence of a benchmark (a best practice) in order to be put into operation. Overall efficiency, usually called economic efficiency, has a technical and an allocative component. The technical component refers to the ability to avoid waste, either by producing as much output as input usage allows (output orientation) or by using as little input as required by technology and the output production (input orientation). The allocative component refers to the ability to combine inputs and/or outputs in optimal proportion in light of prevailing prices. Thus if technical efficiency only pertains to the adherence to the own production plan and does not require any assumption on the producer behaviour, economic efficiency needs an *a priori* on the economic objective of the producer and information on relevant prices.

The aim of this chapter is threefold. The first one is to provide a common framework of analysis for the two concepts. The second one is to summarize the different methodologies available to the researcher who wish to perform efficiency and productivity analyses. Finally, the relevant literature regarding the determinants of productivity and efficiency will be overviewed.

1.3 An historic overview of the theory of production

The concepts of productivity and efficiency are grounded in the broad theory of production in economics. Theories of production can be mostly grouped into the marginalist approach and the classical approach which focuses on the *surplus*. A

[1] Production of services mostly consists of rearranging or redistributing information and resources, which is to say, moving resources rather than transforming them.

rather different approach is that by Georgescu Roegen (1966) who developed a model of production based on stocks and fluxes, which devotes attention to the length of time of the production process; the Romanian economist underlined the possibility that a given production process can be realized with different forms of activations, each of them being related to specific problems regarding to the efficient utilization of the basic elements of the production. Under-utilizations of the elements of production have to be avoided in order to pursue efficiency. The classical approach has been reinvigorated by the work of Sraffa (1960) who investigated the production as a circular phenomenon, grounding his model on the theories developed by classical economists like Marx[2].

In the neoclassical approach, important contributions to the theory of production have been provided by Walras (1874) and Pareto (1927) even if the school of Losanna was much more focused on issues regarding the general equilibrium analysis more than on problems relating to single production units. Following the Paretian path, other scholars have provided important contributions as Samuelson (1956) and Frisch (1965); Shepard (1953) and McFadden (1966) contributed to the implementation of the duality theory. Another important model for the theory of production is the activity analysis model, which was mainly developed by Koopmans (1951) and Debreu (1951): they developed a model in which the problem of choice among alternative possibilities is related to a problem of optimal utilization of the available resources (further comments on activity analysis are provided in Section 1.4). The neoclassical approach studies the production process with analytical tools, and the process is basically viewed as a vector of $\mathbf{z}$ elements, in which the positive elements $\mathbf{y}$ measure the outputs of the process, while the negative ones $\mathbf{x}$ are the inputs. The optimal process is selected as the process which guarantees the maximum net profit, or a given output level at the minimum cost. Prices are given both for outputs, $\mathbf{p}$ and for inputs $\mathbf{w}$, and the net profit is obtained via the expression $\mathbf{py} - \mathbf{wx}$. The tool that is used for the representation of the production possibilities is the production function, which identifies for each vector of inputs, the maximum attainable output level. The net profit maximization behaviour allows, given the price vectors in the market (price-taker hypothesis), the producer to select the amount of quantity of inputs to use, and the amount of output which is going to be produced. In this way, it is possible to detect the optimal production vector for each producer, and each producer is able to reach productive efficiency. In the following Sections it will be explained that several modern methods to measure efficiency and productivity turn from the requirement that all production units reach the full productive efficiency (via mechanics of the market).

[2] See Degasperi and Fredholm (2010) for the development and application of a method of productivity accounting based on production prices, which draws on the scheme developed by Leontief and Sraffa.

1.4 The modern productivity and efficiency analysis: a unified framework

Drawing on Van Biesebroeck (2007) and Del Gatto, Di Liberto, and Petraglia (2010), a unified framework of analysis is presented in this section in order to get a straighter treatment of the previously introduced concepts. At this point it is useful to introduce a tool which is used for representing production processes in economics: the production function[3].

The function

$$Y_{i,t} = A_{i,t} F\left(\mathbf{X}_{i,t}\right), \tag{1.1}$$

relates the output Y of a production unit i (either a firm, an industry or a country) in a given period of time t to the vector of inputs employed by the production unit, $\mathbf{X}_{i,t}$. The function $F(\cdot)$ represents the body of knowledge available to the producer and $A_{i,t}$ is the index of productivity. In this case, the index of productivity is an index of multifactor productivity or *total factor productivity*, while in the case in which just one input would be considered (i.e., $X_{i,t}$ is a scalar), $A_{i,t}$ would be an index of partial productivity[4]. Formally,

$$TFP_{i,t} \equiv A_{i,t} = \frac{Y_{i,t}}{F\left(\mathbf{X}_{i,t}\right)}, \tag{1.2}$$

the TFP index results to be the ratio of produced output to total inputs employed. This approach dates back to Abramovitz (1956) and Solow (1957), and was started as a tool for the analysis of productivity of countries using aggregate data. The framework can be used to evaluate either variations in productivity among producers in the same time period, or variations in a period of time (productivity growth). Comparisons of productivity among producers who share the same body of knowledge, $F(\cdot)$, can be measured with the ratio

$$\frac{TFP_{i,t}}{\overline{TFP_t}} = \frac{A_{i,t}}{\overline{A_t}}, \tag{1.3}$$

where $\overline{A_t}$ is the average productivity of all the producers in the sample. Productivity growth for the same unit between two periods of time t and $t+1$ can be written as

$$\frac{TFP_{i,t+1}}{TFP_{i,t}} = \frac{A_{i,t+1}}{A_{i,t}}. \tag{1.4}$$

It is relevant to note that in this framework, the observed output is equal to the potential level of production, i.e. the frontier output, at each moment in time. In other

[3]In section 1.5, different methods will be introduced which aim at measuring productivity and efficiency: one of the features by which those methods can be categorized is the need to specify or estimate a specific functional form.

[4]The most widely used measure of partial productivity is *labour productivity*.

words, there is no room for technical inefficiency, and A captures only technological change.

Allowing for the presence of technical inefficiency in production processes, Equation 1.2 becomes

$$Y_{i,t} \leq A_{i,t} F\left(\mathbf{X}_{i,t}\right), \tag{1.5}$$

where the observed level of production, $Y_{i,t}$, does not necessarily turn out to be equal to the potential output.

At this point, it is necessary to introduce the output-oriented measure of technical efficiency. The formal definition of technical efficiency is due to Koopmans (1951) in the framework of activity analysis; Farrell (1957) operationalized the concept, both referring to the work by Koopmans than to the Debreu (1951)'s 'coefficient of resource utilization'[5]

If only a single output is produced, an output-oriented measure of technical efficiency is given by the function

$$TE_o\left(\mathbf{X}_{i,t}, Y_{i,t}\right) = \left[max\left\{\phi : \phi Y_{i,t} \leq A_{i,t} F\left(\mathbf{X}_{i,t}\right)\right\}\right]^{-1}. \tag{1.6}$$

Rearranging Equation 1.6, it follows that

$$Y_{i,t} = TE_o\left(\mathbf{X}_{i,t}, Y_{i,t}\right) \cdot A_{i,t} F\left(\mathbf{X}_{i,t}\right), \tag{1.7}$$

where $TE_o\left(\mathbf{X}_{i,t}, Y_{i,t}\right) \leq 1$.

Equation 1.7 indicates that if the framework allows for technical inefficiency, maximum potential output $A_{i,t} F\left(\mathbf{X}_{i,t}\right)$ will be equal to the observed output $Y_{i,t}$, corrected for the output-oriented technical efficiency 'score', which is equal to 1 just for fully efficient firm (thus going back to Equation 1.1).

A comparison of relative TFP among producers i and j in the same time period t can be obtained by modifying Equation 1.3 in the following way:

$$\frac{\frac{Y_{i,t}}{F(\mathbf{X}_{i,t})}}{\frac{Y_{j,t}}{F(\mathbf{X}_{j,t})}} = \frac{TE_o\left(\mathbf{X}_{i,t}, Y_{i,t}\right) \widetilde{A}_{i,t}}{TE_o\left(\mathbf{X}_{j,t}, Y_{j,t}\right) \widetilde{A}_{j,t}}, \tag{1.8}$$

where $\widetilde{A}_{i,t}$ and $\widetilde{A}_{j,t}$ are, respectively, $\frac{A_{i,t}}{A_t}$ and $\frac{A_{j,t}}{A_t}$.

Going back to productivity comparison during two periods of time (productivity growth), we can re-write Equation 1.3 now accounting for technical inefficiency in the

[5]An historical treatment of the definition and the implementation of the technical efficiency concept is far from the main objective of this chapter and the thesis as a whole. However, Farrell wrote that even if his analysis was "largely inspired by activity analysis [...] no reference is made to this in the exposition. The professional economist can easily draw the necessary parallels for himself as indeed, he can note the similarity of the measure of 'technical efficiency' and Debreu's 'coefficient of resource utilization' (Debreu, 1951)(p. 11)". The coefficient of resource utilization was for sure a reference point for Farrell, even if the work by Debreu basically dealt with economic systems, and welfare economics, thus not just production.

production process. It becomes

$$\frac{TFP_{i,t+1}}{TFP_{i,t}} = \frac{A_{i,t+1}}{A_{i,t}} \frac{TE_o\left(\mathbf{X}_{i,t+1}, Y_{i,t+1}\right)}{TE_o\left(\mathbf{X}_{i,t}, Y_{i,t}\right)}, \tag{1.9}$$

TFP growth is here decomposed into two parts: technological change (the first ratio on the right hand side) and change in technical efficiency (the second ratio on the right hand side). This measure of productivity growth will be equivalent to the one in Equation 1.3 only in the absence of inefficiency, i.e. only if TFP change is explained solely in terms of technological change. If the researcher aims at separating the contribution due to technological change from the contribution due to efficiency change, the 'augmented' Equation 1.7 should be preferred to 1.1

Two different remarks should be made at this point. On the one hand, from a theoretical point of view, while conventional economic theory can justify the presence of variations in productivity due to differences in technology, differences in the scale of production and in the operating environment, heterogeneity in efficiency levels (i.e. the observation of not fully efficient units, which stay below the frontier) does not fit easily with conventional microeconomic theory. Nonetheless empirical analyses and real-life cases do not rule out the presence of inefficiency (at least observed, if not actual), and some motivations have been addressed in the literature, which will be discussed in Section 1.6 . On the other hand, for purposes of empirical measurement, Van Biesebroeck (2007) claims that the distinction between the two concepts is —to some extent— 'definitional', because firms which are observed as being inefficient are only those firms that are just behind the most productive one(s) in frameworks which assume all of them to be technically efficient.

Next section introduces a taxonomy of methods for productivity and efficiency analysis.

1.5 Methods for measuring productivity and efficiency

The task of measuring productivity or efficiency in a fair way is not an easy task. The researcher interested in productivity analysis faces a batch of methods which can be classified according to the assumptions they lead to regarding production process, the behavior taken by the unit under analysis, and the data required by each of them.

The objective of this section is to cover most of the available methodologies for productivity and efficiency estimation and to outline the relevant pros and cons of each of them[6]. Three different criteria have been chosen here in order to classify the most used methodologies: *frontier* versus *non-frontier*, *parametric* versus *non-parametric* (and *semi-parametric*), and *stochastic* versus *deterministic*. Table 1.1

[6] A complete and formal introduction to modern methods for efficiency and productivity analysis is provided by Coelli, Rao, O'Donnell, and Battese (2005).

summarizes the methods.

Table 1.1: Methods for measuring productivity and efficiency

	Deterministic		Stochastic		
	Parametric	Non-parametric	Parametric	Semi-parametric	Non-parametric
Frontier	*L/Q programming*; *COLS & MOLS*	*DEA & FDH*	*Stochastic frontiers*		*Stochastic non-parametric frontiers*
Non-frontier	*Growth accounting*	*Index numbers*	*Growth regressions*	*IV & Proxy variables*	

Though all these methods can be used to measure productivity, only frontier methods account for technical inefficiency in the production process: thus, the researcher must first decide whether or not to choose a method that takes technical inefficiency into account. I do not provide a formal description of each method in Table 1.1, rather I will focus on the main features of each class of them. The interested reader is cross-referred to Del Gatto, Di Liberto, and Petraglia (2010).

Macro versus micro. Some of these methods have been developed and employed in the macroeconomic literature like *growth accounting* (Abramovitz, 1956; Solow, 1957) and *growth regression* (Mankiw, Romer, and Weil, 1992; Islam, 1995). Growth accounting has been used to estimate TFP both at the country level and at the sectoral level. It is probably the most popular method to measure productivity growth at the aggregate level. Thus it is worth to spend some more words on it. Taking logs and derivatives with respect to time Equation 1.1 becomes:

$$\frac{\dot{y}}{y} = \frac{\dot{a}}{a} + \sum_{n=1}^{N} \beta_n \frac{\dot{x}_n}{x_n}, \tag{1.10}$$

where $(\dot{a}/a)$ is the TFP growth rate and β_n are the inputs social marginal products. Thus, knowing the growth rates of factors of production and their social marginal products, the TFP growth rate can be calculated as a residual (the Solow residual):

$$SR = \frac{\dot{a}}{a} = \frac{\dot{y}}{y} - \sum_{n=1}^{N} \beta_n \frac{\dot{x}_n}{x_n}. \tag{1.11}$$

As it can be easily seen by this equation, the rate of change of TFP represent the change in national income that is not explained by changes in the level of inputs used. On the other hand, in the growth regression approach TFP is not estimated as a residual, and technology (disembodied productivity) evolves exogenously, i.e. the growth rate of the technology frontier is constant: this approach tries to answer to the question of whether TFP convergence is taking place, and under what conditions.

Others methods have been prevalently used in the microeconomic (i.e. industrial economics) literature like the *proxy variables* approach (Olley and Pakes, 1996; Levinsohn and Petrin, 2003). The proxy variable approach deal explicitly with the 'simultaneity' problem (i.e. the endogenous decision of inputs by the firm which observes its TFP). In order to cope with this issue, it explicitly makes two assumptions: (i) it assumes a two-factors production function, and (ii) it hypothesizes that firm productivity evolves according to a first-order Markow process. The taking logs in Equation 1.1 and adding a noise term $e_{i,t}$ it follows that:

$$y_{i,t} = \beta_k k_{i,t} + \beta_l l_{i,t} + a_{i,t} + e_{i,t}, \tag{1.12}$$

where $a_{i,t} = E\left[a_{i,t}|a_{i,t-1}\right] + u_{i,t}$, where $u_{i,t}$ denotes innovation in $a_{i,t}$. The approach is based on several (quite restrictive) hypothesis: the proxy variable —investment in the specification by Olley and Pakes (1996)— is a strictly monotonic function of the unobservable $a_{i,t}$,

$$i_{i,t} = i(k_{i,t}, a_{i,t}); \tag{1.13}$$

moreover, investment and capital are orthogonal and both decided at time $t-1$, while labour is chosen at time t when firm productivity is observed. The i function can be inverted, thus giving a proxy for the TFP which can be included in Equation 1.12, and it follows:

$$y_{i,t} = \beta_l l_{i,t} + \Phi_{i,t}(i_{i,t}, k_{i,t}) + e_{i,t}; \tag{1.14}$$

the regressors are no longer correlated with the error and the parameter of labour can be estimated in the 'first stage', while the parameter of capital has to be estimated in a 'second stage' through non-linear least squares.

Finally, other methods again have been applied in both group of studies even in different proportions as all the frontier methods (both deterministic and stochastic, parametric and non-parametric) and *index numbers* (Caves, Christensen, and Diewert, 1982a,b).

Parametric versus non-parametric. Another choice has to be made considering the assumptions the method makes on the production function. All parametric methods need a specification of the functional form $F(\cdot)$ in Equation 1.1 which is common to all producers in the sample[7]. When in the sample under analysis a high degree of technological heterogeneity is at work, it is difficult to impose a common functional form and the analysis can bring to misleading results on estimates concerning parameters and productivity.

Non-parametric methods has an appealing feature in this sense: they do not require the calculation or estimation of the production function parameters. Index

[7]Here we skip the growing literature on models of production which allow for different parameters in the production function(s) adopted in the same sample under analysis; however, basic references for random coefficients models are Mairesse and Griliches (1990), Klette (1999).

numbers (e.g. *Malmquist productivity index*) rely on a theoretically motivated (perfect competition in inputs and output, optimizing behaviour by firms, constant return to scale and the absence of measurement errors) aggregation method for inputs and outputs without estimating any production function, while deterministic frontier approaches (Data Envelopment Analysis-DEA and Free Disposal Hull-FDH) build the upper bound of the production possibility set 'passing through' the outermost observations (viewed as inputs/outputs combinations), thus not making any claim on the production function.

Stochastic versus deterministic. Finally, one has to choose between methods which account for measurement errors in variables and sources of noise in the model, and methods which are fully deterministic. In the second case, outliers and measurement errors in the data can bring to unreliable measurements of productivity (as in the case of index numbers) or efficiency (even if more advanced statistical techniques proposed by Simar and Wilson, 1998, 1999, try to cope with this limitation). Deterministic frontier models can be distinguished between methods that 'parameterize' the technology and non-parametric methods. The former class of methods are relatively uncommon in nowadays applications. Linear and quadratic programming (proposed by Aigner and Chu, 1968), corrected ordinary least squares (COLS, early proposed by Winsten, 1957; Gabrielsen, 1975), and modified ordinary least squares (MOLS, developed by Afriat, 1972; Richmond, 1974) have been largely substituted by non-parametric methods in the estimation of deterministic frontiers. DEA and FDH are the two most popular non-parametric frontier methods in recent applications. DEA follows directly from the work by Farrell (1957), while Deprins, Simar, and Tulkens (1984) developed the FDH relaxing the assumption on convexity of the production possibility set.

Among the methods which account for measurement errors in variables, and variations in productivity due to factors which are not under the control of the firm (i.e. bad weather, significant machines breakage), it is important to remind the *stochastic frontiers*. These models are composed error models in which technical inefficiency is separated away from noise, assuming a specific functional form for both components. Starting from Equation 1.1, the stochastic frontier model can be written as

$$y_{it} = a + \beta_k k_{i,t} + \beta_l l_{i,t} + v_{it} - u_{it}, \tag{1.15}$$

where $v_{i,t}$ accounts for noise in the model and u_{it} captures technical inefficiency, i.e. output distance from the frontier function. The $v_{i,t}$ component is normally distributed, while u_{it} is usually assumed to follow a one-sided distribution, either half-normal, exponential or truncated normal. Both terms are assumed to be distributed independently from each other and from the inputs. The parameters of the production frontier are usually estimated via maximum likelihood, while distances to it (u_{it}, inefficiency) can be estimated via the the Jondrow, Lovell, Materov, and Schmidt (1982) estimator or the estimator proposed by Battese and Coelli (1988).

Since the contemporaneous introduction by Aigner, Lovell, and Schmidt (1977) and by Meeusen and van den Broeck (1977) this field of analysis has experienced incredible advancements, and different directions of research are currently investigated, for instance:

- The separation of technological heterogeneity from inefficiency, and the possibility to allow for different technologies in the sample under analysis: this target has been addressed in different ways up to now, and it is still under debate: see Greene (2005), for example, who has proposed the *true-fixed effects* and *true-random effects* models in order to separate firm-specific heterogeneity from time-variant inefficiency, or O'Donnell, Rao, and Battese (2008) who has proposed a *meta-frontier* approach to account for observable heterogeneity in technology parameters. Huang (2004) has estimated a random coefficient stochastic frontier model.

- The tentative to cope with the endogeneity of inputs, the so called 'simultaneity' problem, which has been a rather neglected issue in this field of efficiency analysis until this time. See, for instance, the work by Guan, Kumbhakar, Myers, and Lansink (2009) who investigate the excess capital capacity in a sample of Dutch cash crop farms, taking into account the endogeneity of inputs.

- The effort of expanding stochastic frontier models to 'environments' which were considered not favorable to them, i.e. samples in which inefficiency could not be detected by the conventional estimators. See, for example the work by Carree (2002), on samples of data with positive skewness in the overall residuals.

- The decomposition of the aggregate productivity growth, both at the country level and at the regional level. Some examples can be reminded: Kumbhakar and Wang (2005) have decomposed the Malmquist index for 82 countries using a stochastic frontier framework, while Alvarez (2007) have used a stochastic frontier methodology for decomposing regional productivity growth for Spanish regions.

Strengths and weakness. Van Biesebroeck (2007, 2008) explores strengths and weaknesses of some of the above methods, trying to suggest when each methodology is expected to be particularly appropriate[8]. Using both real and simulated data,

[8]More precisely, he compare results from index numbers, DEA, stochastic frontier estimators by Cornwell, Schmidt, and Sickels (1990) and Battese and Coelli (1992), instrumental variables (GMM) estimators and proxy variable approach by Olley and Pakes (1996).

the author reaches some conclusions which are interesting starting points for further analysis and debate.

- If one is only interested in measuring (either estimating or calculating) the residual, the chosen method is not very important: the residual is similar among different methods, and this is even more evident in the comparison of productivity growth rates (see Equation 1.4). Results from non-parametric methods are surprisingly well in line with those obtained by parametric ones.

- Non-parametric techniques (either frontier or not) work well when high technological heterogeneity is at work: eligible cases are those samples with pool of firms coming from very different industries, at different stages in their lifecycle, or operating in countries characterized by different stages of development.

- The author lists also some distinctive features of the parametric methods: stochastic frontiers work well when productivity differences are constant over time and observations share the same technology[9]; *instrumental variables* methods (IV) cope well with the problem of 'simultaneity' of productivity and input choices and heterogeneity in technology, and the same holds for semi-parametric methods (Olley and Pakes, 1996; Levinsohn and Petrin, 2003).

Summing up, the availability of different methods makes the researcher able to cope with different issues which arise in productivity and efficiency analysis. Nonetheless, methods seem —with a reasonable degree of approximation— to bring to similar results, and this is even more evident for estimates of productivity growth rates than for estimates of productivity levels.

1.6 Determinants of productivity and efficiency: modeling the unobservable

The purpose of productivity and efficiency analysis is (most of the times) not only the computation of 'scores' of performance, but also the characterization and the analysis of the causes of observed performance. This is true both for studies which seek to understand the causes of productivity variations among firms, organizations and other single agents, and for studies which aim at finding the drivers of aggregate productivity growth and (more recently) the determinants of productivity differentials throughout regions and countries. Thus, the measurement of economic performance goes hand in hand with the analysis of the causes of its variations among production units, mainly because an improper measurement of the first is more likely to bring to

[9]However, it is important to stress that conclusion regarding stochastic frontiers are driven by the particular type of estimators considered by Van Biesebroeck. Other estimators can take into account time-varying productivity differences and heterogeneity in technology.

unreliable results regarding the second. The aim of this Section is to itemize the most frequently studied determinants of economic performance both at the micro and the macro level, also providing the main results obtained by empirical studies.

Most of the considered determinants have been alternatively assumed to be determinants of productivity or determinants of efficiency in empirical works. Does this common practice have a theoretical underpinning? It is useful to go back to the two remarks at the end of Section 1.4 and to keep the discussion on two separate levels.

From a theoretical point of view, the observation of differences in productivity levels and in productivity growth rates is due to differences in factors relating to technology, scale of production and externalities. Inefficiency is, instead, not contemplated in conventional economic theory (see Section 1.3), in which first order and second order optimizing conditions are satisfied (see also Fried, Lovell, and Schmidt, 2008, p.5), but in a long-standing debate, several scholars have brought explanations in favor of the effective detection of economic inefficiency. Borrowing from Kumbhakar and Lovell (2000, Introduction), Fried, Lovell, and Schmidt (2008, Section 1.2) and Greene (2008, Section 2.1.2), it is possible to list the main contribution in the history of economic thought to the treatment of economic inefficiency.

Knight (1933) claimed that if it would be possible to include all outputs and all inputs (in quantities) in the transformation function of the producer, since 'nothing can be created nor destroyed', all producers would achieve the same unitary productivity (and efficiency) evaluation. However, economists are more interested in the ratio of 'useful' outputs to inputs, where usefulness is mainly represented by weights incorporating market prices. This thing raises the problem on how to deal with productivity and efficiency when not enough outputs or inputs are taken into consideration. Stigler (1976), reacting to the works by Leibenstein (1966, 1976), considered as a lack in an incomplete model what Leibenstein actually named 'X-inefficiency', i.e. a series of motivations which deal with agency problems, incomplete contracts and inadequate motivation. If the researcher fails to incorporate all relevant variables, and if she/he is not able to specify the right economic objectives and the right constraints faced by the production unit, these failures will be detected by her/him as 'inefficiency'. However, from a practical standpoint this situation will be the case almost always, thus bringing to measured (if not effective) inefficiency. Possible sources of measured inefficiency which are linked to the notion of 'X-inefficiency' provided by Leibenstein are the 'bounded rationality' of managers, who (because of a limited information processing ability) engage in a 'satisificing' behaviour (Simon, 1955), and the transaction costs economizing behaviour of firms, which was investigated by Williamson (1975). Transaction costs economics have 'enlarged' the perspective on the conventional (production) costs minimizing behaviour of the firm to the unconsidered costs of the market.

However, from an empirical point of view, all frontier methods allow for a framework in which just some units operate on the technological frontier, while the rest is observed below it. Thus, these methods basically differ from non-frontier methods in the way they model the unobservable, TFP. The two frameworks are not incompatible at all, as it has been showed in Section 1.4. The outcome of all frontier methods can be seen as a rescaled TFP score with respect to the most productive firm(s), in the case of cross-section data. The parallel is even more immediate in the case of panel data, because TFP growth can be decomposed into movements of the frontier (technological change), movement along the frontier (either inputs deepening or contraction, i.e. scale changes) and movements toward the frontier (efficiency change, or *catching-up*). From a certain point of view, frontier methods allow for a finer decomposition and specification of the residual which can be seen as the results of three (and not just two!) components: this characteristic of frontier methods can be appealing especially for studies at the macro level[10].

Summing up the above paragraphs, if the distinction between determinants of productivity and determinants of efficiency is meaningful from a theoretical standpoint, it is less tenable in applied works. Frontier and non-frontier methods either come to two measures which capture the same *unobservable* in a different scale (in the case of a cross-section data), or result in a different breaking down of the same (in the case of panel data). There is no general theory which brings to the choice of relating a part of the output variation to the change of a set of inputs included in the specification (the $\mathbf{X}_{it}$ vector in Equation 1.7), or to the variation of the component of the residual which should capture 'inefficiency' and which measures the distance of the unit to the observed technological frontier (Lovell, 1993). Moreover, modern frontier methods allow for the inclusion of determinants of inefficiency together with conventional inputs in the framework, thus leading to an augmented model which include in the inefficiency specification a vector of determinants, $\mathbf{Z}_{it}$[11]. It is mainly up to the researcher to ground the modeling of the residual on economic meaningful hypotheses, given the constraints on available data.

In view of this, the proposed taxonomy will make no distinction between the determinants of productivity and the determinants of efficiency (even because the empirical literature has almost ever made any distinction), and the main criterion that has been used regards the level of analysis: the productivity of single agents (firms, organizations) or the aggregate productivity. In fact, despite that there is a common basis regarding the set of determinants of productivity and efficiency, the micro and macro literatures have followed two paths of evolution, both in terms of theoretical models, and with reference to empirical tools (mainly due to different data

[10]Few more examples can be reminded, together with those listed in the paragraph regarding the stochastic frontier models: Kumar and Russell (2002) have decomposed labour productivity growth into relative contributions of technological change, *catching-up* and capital deepening, while Henderson and Russell (2005) have investigated the role of human capital in enhancing TFP growth.

[11]See Chapter 7 in Kumbhakar and Lovell (2000) for a detailed explanation of the inclusion of exogenous factors in stochastic frontier models.

availability).

1.6.1 The literature on productivity at the micro level

The literature of the determinants of productivity at the firm and plant level is wide. However, notwithstanding the large amount of evidences about the role of relevant factors (other than labour and capital) in explaining a significant part of productivity heterogeneity among firms —both in levels and in the growth rates—, productivity still remains a measure of our ignorance (Griliches and Mairesse, 1983; Bartelsman and Doms, 2000). Several factors have been investigated as determinants of productivity differentials among firms.

- *The regulatory environment.* The effect of regulation policy on firm productivity is not easy to be estimated. In fact, regulation affects decisions firms make today, but also the future market structure, by altering incentives for innovating, investing, entering in the market and the possibility for gaining market shares. Alchian and Kessel (1962) characterized regulated industries as market situations in which firms are either limited in their pursuit of efficiency or threatened by antitrust action, which can be also a limitation for efficiency. Olley and Pakes (1996) have studied successive stages of deregulation in the U.S. Telecommunications Equipment Industry, and they have found that considerable resource reallocation followed deregulation. Deregulation affected productivity of the industry in two different ways: first it changed choices of producers with respect to their innovative activity, the adopted inputs and production volumes, and second it exerted a crowding-out effect on less efficient plants. Pozzana and Zaninotto (1989) study the effect of the market structure on productive efficiency in a sample of firms in the Italian retail industry.

- *The role of management and different types of ownership.* Choices of technology, inputs, and production are made by management and, thus, better managers may make better choices. Two lines of research have been developed regarding the role of management and the type of ownership with respect to firms' productivity. The first one deals with the effect of mergers on productivity growth. Lichtenberg (1992) and McGuckin and Nguyen (1995), exploring the issue in a large panel of U.S. manufacturing plants, found that establishments which faced ownership change also enjoyed above-average productivity growth for several years after a change: this could be due to a reduction in corporate overhead and a reduction in auxiliary offices. The second one deals with differences in performance of private and State-owned enterprises. Alchian (1965) backed the inferior efficiency pursued by managers of the public sector enterprises, due to the looser control exerted by owners with respect to owners of private enterprises; Pestieau and Tulkens (1993) analyzed the difference in technical efficiency between private and State-owned enterprises, while Bottasso and Sembenelli (2004) provided an interesting analysis of differences in technical efficiency in a representative sample of Italian manufacturing enterprises, finding no difference in efficiency between private firms and affiliates to national groups, while State-owned enterprises show the lowest levels of efficiency.

- *Technology and the human capital.* Physical and human capital provide two sources of productivity differentials among firms. Nelson (1981) emphasized the importance of understanding the way in which technology is generated and distributed through firms, and many empirical studies have documented the correlation between some measure of technology and productivity at the micro level (see Dunne, 1994; Lichtenberg, 1996, among others), unfortunately suffering of a possible 'reverse causality' explanation which goes from productivity to the adoption of more advanced technologies in the organization of the firm. Interestingly enough technology has been found to be strictly related to labour quality in the study by Doms, Dunne, and Troske (1997), in which the presence of workers with skills above of the average was found to be related to the adoption of advanced technology.

- *Firm international exposure.* The literature on the relationship between firm productivity and the export status (exporter versus non-exporter firms) has increased since the nineties. Since the early works by Bernard and Jensen (1995, 1999) on U.S. exporters, and by Roberts and Tybout (1997) and Clerides, Lach, and Tybout (1998) on a sample of developing economies, an open debate started on the direction of the relationship found between the exporting activity and firm productivity. The hypothesis of self-selection claims for an auto-selection operated by more productive firms to the export activity: these firms can exploit their comparative advantage thus being more suited to overcome obstacles related to the exporting activity; on the other hand, firms engaged in export activities could learn new technologies in the host country, thus improving their efficiency (the so called learning effect). While the former hypothesis has found a robust support in empirical works, the latter has generated contradictory results. However, a group of studies using econometric techniques able to control for the 'endogenous' exporting choice have supported the evidence of a learning effect: Aw, Chung, and Roberts (2000) provided evidence for Korea and Van Biesebroeck (2003) did the same for Sub-Saharan manufacturing plants. Castellani (2002) and Serti and Tomasi (2008) have provided econometric evidence supporting the hypothesis that export behavior cause learning effects in different representative samples of Italian manufacturing firms. Another strand of the literature has pointed out that firms engaging in foreign direct investments show higher level of productivity than domestic firms and simple exporters, first because they need to overcome the cost of doing business abroad (Helpman, Melitz, and Yeaple, 2004), but also because investing abroad they may be able to access foreign knowledge and reaping the benefit of higher economies of scale (Cantwell, 1995; Fosfuri and Motta, 1999).

- *Firm structure decision.* As Syverson (2010) has underlined, the organizational structure of the firm can be related to its productivity level. In particular the control over vertical links of production seems a strategic choice which brings to different performances: more integrated structure can have a better control over the production chain, both allowing for an easier movement of physical and intermediates inputs along the chain and for a sharing of human capital and management skills among different phases and activities; however, disintegrated structure —which have become more and more common in the world in recent years— may focus on their core competences, leaving unproductive phases to the 'outside' and reaching an higher flexibility.

1.6.2 The literature on productivity at the macro level

In 1957 Robert Solow came out from his analysis on U.S. productivity growth with a large portion of change in aggregate output not explained by a growth in conventional inputs, i.e. labor and capital: this unexplained part, which was attributed to technological progress, is nowadays called the total factor productivity. The empirical literature on aggregate productivity growth have tried to relate the residual to particular drivers.

Later developments of the neoclassical model were provided by Solow (1960) and Salter (1960): Salter, in a work based on his Ph.D. thesis, developed a vintage model of capital in which technical progress takes place only if there is investment. Works by Griliches (1960, 1963a), Denison (1962) and Jorgenson and Griliches (1967) have tried to 'whittling away the residual' (Stone, 1980) investigating several factors which could have explained productivity variation along a given period of time. In a (not so) recent article overviewing the historical evolution of the analysis of the residual, Griliches (1994) listed the more investigated factors and the still not studied (from his point of view) factors which would have been deserved more attention in the recent future.

Borrowing from his work, it is possible to list the following factors as determinants of the aggregate productivity growth:

- *improvement in labour quality and capital* (frequently not taken into account);
- *formal and informal R&D investments by individuals, firms, governments;*
- *unmeasured contributions by science and other spillovers.*

In this paper, Griliches tried to motivate how this framework of analysis resulted not to be satisfactory, especially after the events which characterized the seventies and the eighties: beginning in 1974 (or perhaps already in 1968) productivity growth slowed down significantly in the United States and abroad[12], and this fact was at odds with the above framework. The author also raised some possible explanations of this 'failure'. The first one was the poor attention devoted to a thoughtful use of aggregate data on R&D and the output or research (i.e. patents) that were available in that period[13]. The second one, was the possibility that the framework of analysis was rather incomplete: Griliches claimed that the framework did not take into account several important sources of aggregate productivity growth which could be the objectives of fruitful improvements in the existing framework: externalities, heterogeneous expectations, the rise of new products and technologies, X-inefficiency, changes in political and regulatory environment. The above framework of analysis falls under the name of *growth accounting*.

Another strand of the literature has focused on the *catching-up* phenomenon (Gerschenkron, 1962; Abramovitz, 1979), hypothesizing that productivity growth should be expected to be negatively correlated with the level of productivity. Countries behind the world innovation frontier, it is argued, can grow faster by copying technologies already developed in technologically more advanced economies. This literature emphasized the importance of *investments in physical and human capital*, *social and institutional factors* and *technological congruence* —as possible constraint— for the outcome of the *catching-up* process (Fagerberg, 1995, p.10). Overall, the *catching-up* literature has brought to a clear conclusion: a simple model with one independent variable is not sufficient to explain differences in growth and we should look for additional variables to be included in the model.

A third framework of analysis of aggregate productivity is that which falls under the name of the *Shumpeterian perspective*. In this framework, both innovation and imitation (catching-up to the frontier) are important for productivity growth. It is not possible to surpass the technological leaders without passing innovative activity to them as well (Pavitt and Soete, 1982). Thus the Schumpeterian framework allows for both divergence and convergence. Fagerberg (1991) tested (in a sample of developed and newly industrialized countries) a model which included three variables as explanatory for productivity growth: foreign-produced knowledge, growth in national innovative activities, and effort (proxied by investments). The results showed that in

[12] It is interesting to see that after the slowing down of the aggregate productivity, the attention increased for studies which explicitly taken into account possible technical inefficiency at the more disaggregated level (industry, firms or even plants): see Caves and Barton (1990), among others, for a remarkable example of application of stochastic frontier models using U.S. Census data on individual manufacturing establishments for a large number of industries. The authors tried also to examine a large set of possible factors explaining variation in efficiency at the establishment level.

[13] The major message was that available data were often misinterpreted, because of inadequate attention to how they were produced.

order to *catch-up* with the developed countries, semi-industrialized countries have to increase their national technological activities.

Finally, a fourth framework of analysis goes under the name of the *new growth theory*: in this framework there basically two different views on the relationship between technology and productivity growth:

- Lucas (1988), Romer (1990) and others have developed models in which growth in new knowledge is analyzed as a by-product (externality) of other economic activities (investments in physical and human capital).

- Grossman and Helpman (1991) and Aghion and Howitt (1992) introduced models in which innovation take place, because innovating firms can appropriate (for) a period of time of the advantage/rent due to it. However, innovation is also characterized by technological spillovers that facilitate subsequent innovation projects. Thus, it is the dual public-private character of the innovation process that allows growth to go on in these models.

A typical result is that the rate of growth is proportional to the amount of resources devoted to innovation.

From an empirical point of view, many studies have recently followed the new growth theory and the catching-up debate. The variables taken into account in these studies may be divided into three groups: (i) GDP per capita, as a proxy for the scope for catching-up; (ii) Variables reflecting attempts to affect the gap, such as investment, education and resources devoted to - or output from - innovation activities; (iii) other variables of a 'structural' or political nature assumed to affect growth (such as the degree of openness to trade, country size, share of public sector in GDP, population growth as suggested by Fagerberg, 1995).

Even more recently, the attention on productivity heterogeneity (both in levels and in growth rates) has also regarded sub-national levels of aggregation, like regions and municipalities. One explanation for this attention is the possibility that these units of observations (nearer to firms and micro organizations) give for the analysis of externalities, spillover effects, and social factors which were indicated by Griliches as those factors not yet taken into account by the *growth accounting* tradition. Moreover, it is probably more meaningful to control for heterogeneity in the quality of conventional inputs (especially the quality of the labour force, human capital and the quality of physical capital) at regional or local levels, given the well known variation in these characteristics which make the more advanced regions of a country far enough from the last ones. The literature on regional performance has investigated the role of agglomeration economies, human capital, infrastructures, and the industrial composition in explaining productivity differences among regions and driving productivity growth.

1.7 Concluding remarks and links to the other chapters

This chapter focuses on the analysis of performance of production units, either firms or organizations, regions or countries. Two basic concepts have been analyzed, productivity and efficiency: they have often been used as interchangeable, but it has been stressed that they are not overlapping. Productivity equals the ratio of the outputs that a unit produces to the inputs that it uses. Efficiency deals with the placement of the unit with respect to the production (or technology) frontier. If the unit can rise the production of an output, without having to increase any input, or diminishing in the use of an input without having to reduce any of its outputs, the unit can improve its degree of technical efficiency, because it is not on the frontier. Allocative efficiency deals with the optimal combination of inputs, given market prices. The researcher interested in productivity and efficiency analysis has at his disposal a large set of methodologies from which she/he can choose paying attention to the characteristics of the phenomenon which has to be analyzed and to the constraints imposed by available data. Strengths and weakness of each method have been detailed, as for the hypotheses which each method needs in order to get reliable measures of productivity and efficiency. In the following chapters, different research questions have been addressed, making use of some of the introduced methods for productivity and efficiency analysis. For a matter of coherence, I introduce the basic motivations for the adoption of different methods below, leaving the discussion on the results I have obtained in each study to the relative chapter and to the conclusions of the thesis.

In Chapter 2, which is a methodological work, I have used Monte Carlo simulations in order to perform a set of experiments in the framework of stochastic frontier models. In this chapter, I have investigated the consequences of a misspecification of the inefficiency distribution on both inefficiency scores levels and ranking. Scholars of the field have questioned whether the assumption on the specific distribution for the inefficiency term is relevant and may actually drive the results of the analysis: a common practice is to compare the results obtained by estimating differing —in the specification of the inefficiency distribution— stochastic frontier models from the same sample of production units; previous evidence indicates general concordance among set of estimated inefficiency scores. However, an extensive exercise on this issue is still lacking in the literature. The use of Monte Carlo simulations allows me to design appropriate data generating processes (DGP): the performance of the most frequently used models - normal-half-normal, normal-exponential and normal-truncated normal - are analyzed to estimate the efficiency scores, both when distribution has been correctly specified and when it has not. Overall, the news for practitioners are encouraging. If inefficiency ranking is the main concern of the analysis, the three most frequently estimated models give the same result, so that the specification of the inefficiency distribution does not matter.

Chapters 3 and 4 are applied works on real data. In Chapter 3 I have investigated the relationship between vertical integration and firm efficiency in the Italian Machine Tool industry. The control of vertical links of production, i.e. the decision about which phases of production to keep inside to the firm (vertical integration) and which ones to leave to the 'outside' (outsourcing) is certainly related to the firm productive performance and, even if it has been analyzed in previous works, those have not converged to clear-cut results. In order to come up with a testable hypothesis, I have first set up a theoretical model (in line with previous models on productivity heterogeneity and organizational choices, as the one proposed by Antras and Helpman, 2004): in this model more efficient firms decide to produce as vertically integrated, bearing higher (organizational) fixed costs while less efficient firms choose to outsource part of production process buying an intermediate input from other firms, thus reducing fixed costs but bearing higher marginal costs. This result is confirmed by a stochastic frontier analysis on a sample of more than 500 machine tool producers. The heteroskedastic frontier model allows me to jointly estimate the parameters of the production function and the coefficients of the variables related to inefficiency, in particular a measure for vertical integration. The empirical analysis shows that vertical integrated firms present a lower variance (and lower mean) of the inefficiency distribution, after having controlled for firm size, type of ownership, agglomeration economies and the economic cycle. Thus, vertical integrated firms are, *ceteris paribus* more efficient than disintegrated firms.

In Chapter 4, exploiting an original and extensive dataset on foreign direct investments (FDIs), I have investigated the relationship between FDIs and labour productivity growth in a large set of NUTS2 regions in almost all countries of the Enlarged Europe (EU-27). The results of the econometric analysis support that both inward and outward FDIs have positive effects on productivity growth at the regional level, after controlling for a relevant set of regional characteristics, such as human capital, technology capital and the industry mix: in particular, inward foreign investments have a positive effect on regional productivity only above a certain threshold level, while outward investments have a positive effects up to a certain threshold, which is however very high in our sample. This is an interesting result, given the increasing role of regions in the European context and the relevance –in terms of GDP– of inward and outward FDIs in the European Union. The econometric analysis has provided –to my knowledge for the first time– a robust evidence of positive effects. This is an original contribution to the international economics literature in several dimensions: previous studies with a regional perspective have focused on comparisons within single countries and have addressed only the role of 'inward' investments as a driver of increasing local performance. Moreover, those few studies which have attempted to assess the specific role of outward investments on productivity have taken a country perspective, almost neglecting the sub-national level of analysis. These results have been showed to be robust to different specifications of the econometric model, like the inclusion of regional characteristics (in levels and growth), the diversity in technological regimes between regions belonging to the EU-15 and regions belonging to the EU-12, and spatial dependence in labour productivity across European regions.

Chapter 2

Misspecification of the Inefficiency Distribution in Stochastic Frontier Models: a Monte Carlo analysis

2.1 Introduction

Stochastic production frontier models are used in productivity analysis to measure the performance of firms, industries, regions and countries in terms of observed distances to the productive 'best practice'. The frontier production function is an empirical model based on the theoretical premise that a production function represents an ideal, i.e. the maximum output attainable with a given set of inputs. Estimation of the frontier parameters is usually performed by maximum likelihood methods, but this requires specific distributional assumptions regarding the components of the overall error term. In particular, a specific distributional form is needed for the technical inefficiency component and the relevant literature has mostly specified it as being half-normal, exponential or truncated normal. Scholars of the field have questioned whether the assumption on the specific distribution of the inefficiency term is relevant and can actually drive the results of the analysis: a common practice is to compare the results obtained by estimating differing —in the specification of the inefficiency distribution— stochastic frontier models from the same sample of production units; previous evidence indicates general concordance among set of estimated inefficiency scores. However, an extensive exercise on this issue is still lacking in the literature.

The aim of this paper is to fill this gap, by assessing the performance of stochastic frontier models in correctly estimating true inefficiency values both when the inefficiency distribution is correctly specified and when it is not: in order to do this, we make use of a set of Monte Carlo experiments. The main advantage of using Monte Carlo simulations is complete control over the data generating process: in other words, once the true inefficiency distribution is known, the researcher can monitor the behaviour of the stochastic frontier estimators. This can never happen with real data. Along this lines, we explored eighteen combinations of true and assumed distributions, covering the largest set of misspecifications of the inefficiency term which has never been considered until now.

The first stable result which emerges from all experiments is that for each of the six inefficiency distributions considered, the three estimated models —i.e., normal-half-normal, normal-exponential and normal-truncated normal—, reproduce the inefficiency ranking with the same precision. This is also true for both correctly specified models (those which assume the correct inefficiency distribution) and for misspecified ones. This is a useful piece of evidence for practitioners, because if the ranking of inefficiency is the main object of the analysis, the three most frequently employed models give the same results. Conversely, if one is interested in the inefficiency value *per se*, it is important to specify the correct distribution of the inefficiency term: for each of the estimated models the average difference between true inefficiency scores and estimated ones is lower when the model is correctly specified than when it is misspecified. Lastly, comparing results from a given stochastic frontier model in estimating inefficiency values, once they have been generated from different distributions with the same variance, we may conclude that there are 'qualitative' differences among groups of inefficiency distributions, but the consequences on the estimated inefficiency scores are only marginal.

The rest of the paper is organized as follows: Section 2.2 explains in mode detail the reasons for this simulation study, and relates it to previous Monte Carlo analyses of stochastic frontier models; Section 2.3 briefly summarizes the characteristics and properties of the three most frequently used empirical models in stochastic frontier analysis and the characteristics of the technical efficiency estimator; Section 2.4 describes the design of the experiments and the simulation protocol; Section 2.5 presents the results; and Section 2.6 discusses the experimental results and suggests some steps for further research. A Data Appendix concludes the paper.

2.2 Motivation: is the choice of the inefficiency distribution relevant?

Stochastic production frontier models (SFMs) were originally proposed by Aigner, Lovell, and Schmidt (1977) and Meeusen and van den Broeck (1977). In SFMs, deviations from the frontier are attributed to two factors: motivations which are not under the control of the production unit (for instance, bad luck; faulty machinery and breakdowns; adverse weather in agricultural production) enter the generic term called *noise*, together with measurement errors in output, whereas factors which stem from non-optimal use of technology are captured by *technical inefficiency* term[1]. The noise and the technical inefficiency terms sum to the composed error of a SFM which, for a cross-section of production units, may be formulated as follows:

$$Y_i = f(\mathbf{X_i}, \boldsymbol{\beta}) \cdot exp\left\{\epsilon_i\right\} \tag{2.1}$$

where:

$$\epsilon_i = v_i - u_i. \tag{2.2}$$

Y_i indicates the observed level of output, $\mathbf{X_i}$ is the vector of inputs used in the production process, $\boldsymbol{\beta}$ is the vector of unknown technological parameters to be estimated, v_i is the error component which refers to noise, and u_i is a non-negative random term which accounts for technical inefficiency.

Taking logs on both sides of Equation 2.1, the model may be rewritten as:

$$y_i = f\left(\mathbf{x_i}, \boldsymbol{\beta}\right) + v_i - u_i, \tag{2.3}$$

which is the linear form of the model usually employed in empirical applications. Estimation of the model in the Equation 2.3 is performed, in most cases, by maximum likelihood methods (ML) in order to have consistent and asymptotically efficient estimators of the frontier parameters. Estimation of the inefficiency scores are recovered in a second step by means of the estimator developed by Jondrow, Lovell, Materov, and Schmidt (1982), which is based on the information on u_i contained in the overall residual. In order to implement the ML estimation, some assumptions for the components of the error term ϵ are required; the usual being:

1. $v_i \sim iid\ N(0, \sigma_v^2)$;
2. u_i is a non-negative random term which follows a one-sided distribution;
3. v_i and u_i are distributed independently of each other;
4. v_i and u_i are distributed independently of the regressors.

The first assumption is conventional in econometric models and the third assumption is established in almost all works using SFMs[2]. Thus, assumptions 1 and 3 seem innocuous, the other two are worthy of further comments.

A violation of the fourth assumption regards the possibility that one (or more) inputs comes to be endogenous in SFMs and, consequently, the estimators of the frontier parameters are neither unbiased nor consistent. However, in this paper we

[1]In this framework, technical inefficiency may be measured either as output-oriented or input-oriented: in the first case, used in this paper, it refers to the expansion of output which can be obtained, given the number of inputs and the degree of technology which is available to the production unit.

[2]Greene (2008, p.135) suggests using the copula method to specify models in which inefficiency is correlated with noise.

do not examine this issue[3], but focus on the second assumption, i.e. the correct specification of the one-sided distribution.

The second assumption has traditionally been operationalized by assuming u_i to be distributed as half-normal, exponential or truncated normal [4]. The preference for these distributions is grounded on two main motivations (Kumbhakar and Lovell, 2000, p.74):

- Tractability: the distribution of the sum of v_i and u_i is relatively easy to be derived under assumptions 1 and 2, and with one of the three above distributions. Aigner, Lovell, and Schmidt (1977) derived the distribution of the composed error for the normal-half-normal case and both Aigner, Lovell, and Schmidt (1977) and Meeusen and van den Broeck (1977) did so for the normal-exponential model; Stevenson (1980) introduced the normal-truncated normal formulation.

- Economic proposition: the economic model underlying the three most frequently assumed distributions is the one in which most producers show low values of technical inefficiency, as higher values of inefficiency become increasingly less likely[5].

Since the true inefficiency distribution is unknown in empirical applications, researchers have wondered about the importance of the choice of the distribution for u_i. Some have checked the robustness of results using alternative distributions[6], and their studies report high rank correlation coefficients between pairs of inefficiency scores estimated by assuming different distributions for u_i. Some authors, using Monte Carlo experiments, have explored the consequences of the misspecification of the distribution assumed for u_i. Ruggiero (1999) analyzed the performance of the misspecified exponential distribution relative to a correctly specified half-normal, and

[3]Biased estimates of frontier parameters may led to incorrect estimation of the distance to it and thus undermine attempts to estimate firm-specific inefficiency scores correctly. The endogeneity of inputs has been widely debated in the traditional econometric literature regarding the estimation of average production functions and total factor productivity (see Blundell and Bond, 2000; Olley and Pakes, 1996; Levinsohn and Petrin, 2003, among others), but it seems to have been quite neglected in the theoretical and empirical literature regarding SFMs. Gong and Sickles (1992) used Monte Carlo simulations to compare the performance of deterministic and stochastic frontier models in correctly measuring technical inefficiency when there is a correlation between inputs and level of inefficiency. In a recent paper, Guan, Kumbhakar, Myers, and Lansink (2009) attempted to solve this issue, by holding on a two stage procedure which implement a vector of instruments in the first stage in order to cope with the endogeneity of inputs.

[4]Stevenson (1980) and Greene (1990) also proposed the gamma distribution, although the greatly increased complexity of the resulting formulation has somewhat inhibited its application (see Greene, 2008, pp.124-126, for an explanation and computations regarding this model).

[5]Incidentally, this assumption seems fairly plausible in a competitive market structure, but this might not be the case in other types of markets: for instance, we can imagine an industry in which the structure is not that of *perfect competition* for several reasons, like a low number of producers or differentiated products.

[6]For example, Greene (1990) estimated a stochastic cost frontier for 123 U.S. electric utilities using all the three one-sided densities and reported sample mean inefficiencies; Wang (2003) studied the effect of financial constraints on the investment efficiency of a panel of Taiwanese manufacturing firms, assuming alternatively that u_i was distributed as truncated normal or exponential.

found that in most of cases, the incorrectly assumed exponential outperformed the correctly specified half-normal in estimating the rank of the true inefficiency values. More recently, Jensen (2005) used Monte Carlo simulations in order to explore the consequences of misspecification in a broader set of cases. His main results may be summarized as follows. In cases in which the inefficiency distribution was correctly specified (i.e., the assumed distribution overlapped the true one), the truncated normal distribution outperformed the half-normal and exponential in reproducing the ranking of inefficiency scores; however, the half-normal specification was the one which reached the minimum (in absolute values) difference between true and estimated scores in the sample, followed by the truncated normal and the exponential. Another result, which was not throughly explored by the author, was that, for each true distribution, both the model with the correctly specified distribution and the misspecified one obtained the same average rank correlation between the true inefficiency scores and the estimated ones. Unfortunately, this result was limited to two cases: one in which the inefficiency scores were generated from an exponential distribution and were estimated by assuming the distribution to be either exponential or half-normal; and another in which the true scores followed a truncated normal and they were estimated by assuming either a truncated normal or a half-normal. Several combinations between true and estimated distributions are missing in Jensen's experiments. In addition, neither Ruggiero (1999) nor Jensen (2005) offer plausible reasons for these results. Thus, a more broadly exercise on the importance of the assumed distribution for accurate estimation of u_i (in both levels and ranking) would enrich this interesting albeit evidence.

The aim of this paper is to assess the performance of the three most frequently assumed one-sided distributions —half-normal, exponential and truncated normal— in correctly estimating the true inefficiency scores via a set of Monte Carlo experiments, both when they coincide with the true inefficiency distribution and when they do not. The set of possible data generating processes (DGP) was increased by including three more 'unusual' densities, uniform, log-normal and Weibull. From an economic point of view, each of these distributions may have a peculiar meaning[7]; however, that is not the concern of this paper: they may simply be viewed as a further challenge for testing the ability of 'traditional' distributions (i.e. half-normal, exponential and truncated normal) in correctly reproducing true inefficiency scores. Overall, three (assumed distributions) $\times$ six (true distributions) possible combinations are explored in our Monte Carlo experiments, in order to cover the largest set of possible misspecifications of the inefficiency distribution previously examined.

[7]For example, log-normal and Weibull distributions may stand for an environment in which there are extensive asymmetries in the performance of production units.

2.3 Stochastic frontier models

The following Sections briefly describe the theory on stochastic production frontiers, relevant to this simulation study.

2.3.1 Modeling a production frontier

This Section introduces the three empirical models which have been estimated in the Monte Carlo simulations. The SFMs, introduced by Equation 2.3, aim at identifying the 'best practice' in the sample of production units. The production function can be specified allowing for different levels of flexibility: however, the Cobb-Douglas and the translogarithmic are the most employed functional forms in empirical applications. Estimation of the vector of frontier's parameters, $\boldsymbol{\beta}$, is performed by ML methods. Starting from Equation 2.3 and from assumptions on error components, it is straightforward to write the log-likelihood function for the three most frequently used models.

1. If u_i is half-normal
$$u_i \sim N^+(0, \sigma_u^2), \tag{2.4}$$
the log-likelihood function can be written as:
$$l(\boldsymbol{\beta}, \sigma, \lambda) = constant - I\ln(\sigma) + \sum_{i=1}^{I} \left\{ \ln\Phi\left[-\frac{\epsilon_i\lambda}{\sigma}\right] - \frac{1}{2}\left[\frac{\epsilon_i}{\sigma}\right]^2 \right\}, \tag{2.5}$$
where I is the number of observations, $\epsilon_i = y_i - \boldsymbol{\beta}'\mathbf{x_i}$ indicates the overall error term, Φ is the cumulative distribution of a standard normal, and $\lambda = \frac{\sigma_u}{\sigma_v}$ and $\sigma^2 = \sigma_u^2 + \sigma_v^2$ are the two variance parameters of the SFMs. This likelihood function refers to the normal-half-normal (NHN) model.

2. If u_i is distributed as exponential:
$$u_i \sim Exp(\eta), \tag{2.6}$$
where $\eta = \sigma_u$ is the scale parameter, the log-likelihood may be written as
$$l(\boldsymbol{\beta}, \sigma_v, \sigma_u) = \sum_{i=1}^{I} \left[-\ln\sigma_u + \frac{1}{2}\left(\frac{\sigma_v}{\sigma_u}\right)^2 + \ln\Phi\left(\frac{-(\epsilon_i + \sigma_v^2/\sigma_u)}{\sigma_v}\right) + \frac{\epsilon_i}{\sigma_u} \right]. \tag{2.7}$$
Equation 2.7 refers to the normal-exponential (NEX) model.

3. Lastly, if u_i is truncated-normal:
$$u_i \sim N^+\left(\mu, \sigma_u^2\right), \tag{2.8}$$

where μ and σ_u^2 are the mean and the variance of the pre-truncated normal, respectively, the log-likelihood may be written as

$$l(\boldsymbol{\beta}, \sigma, \lambda, \mu) = constant - I\left[\ln \sigma + \ln \Phi(\mu/\sigma_u)\right] + \\ + \sum_{i=1}^{I}\left[-\frac{1}{2}\left(\frac{\epsilon_i + \mu}{\sigma}\right)^2 + \ln \Phi\left(\frac{\mu}{\sigma\lambda} - \frac{\epsilon_i \lambda}{\sigma}\right)\right]. \quad (2.9)$$

Both errors and variance parameters are defined as in the NHN model, and this likelihood function refers to the normal-truncated-normal model (NTN)[8].

Variance parameters deserve further attention: λ is a useful parameterization of the contribution of technical inefficiency to the overall error term, and as $\lambda \to 0$ (either $\sigma_v^2 \to \infty$ or $\sigma_u^2 \to 0$), the idiosyncratic component dominates the inefficiency term, while as $\lambda \to \infty$ (either $\sigma_v^2 \to 0$ or $\sigma_u^2 \to \infty$), the inefficiency term dominates the noise term and the stochastic frontier converges to a 'deterministic' one. However, λ is not equal to the ratio of variances of the error components in the NHN and NTN model, but it is in the NEX model. Parameter σ^2 relates to the variability of the overall error term, but again it is not equal to the sum of variances of the error components in the NHN and NTN models (see Greene, 2008, p.118). The explanation is simple: while σ_v^2 is always equal to the variance of the noise term, σ_u^2 equals the variance of the inefficiency term only in the case of exponential distribution, but does not in the case of half-normal or truncated-normal distributions. A useful example is given in Data Appendix 2.7.1.

2.3.2 The estimator of technical inefficiency

Although inefficiency scores, u_i, are not directly recoverable because they are part of the overall error term, they can be estimated via the conditional mean function proposed by Jondrow, Lovell, Materov, and Schmidt (1982)[9]. Its general formulation is given by (see Greene, 2008, p.177):

$$E\left(u_i|\epsilon_i\right) = \frac{\int_0^\infty u_i f_u\left(u_i\right) f_v\left(\epsilon_i + u_i\right) du_i}{\int_0^\infty f_u\left(u_i\right) f_v\left(\epsilon_i + u_i\right) du_i}. \quad (2.10)$$

For the NHN, NEX and NTN models Equation 2.10 has a closed form which is reported by several authors. For the NHN model, the conditional mean function

[8] The NTN model was introduced by Stevenson (1980), who argued that the zero mean of the pre-truncated distribution (as in the half-normal case) was an unnecessary restriction. This assumption has been relaxed by the present author, truncating a normal random variable at zero with possibly non-zero mean.

[9] Other estimators of the inefficiency scores (less used in empirical applications) are the conditional mode function and the estimator proposed by Battese and Coelli (1988), $E(TE_i|\epsilon_i) = E(exp(-u_i)|\epsilon_i)$, which is also built on information contained in the overall error term.

takes the following form

$$E\left(u_i|\epsilon_i\right) = \frac{\sigma_u\sigma_v}{\sigma}\left[\frac{\phi(\frac{\epsilon_i\lambda}{\sigma})}{1-\Phi(\frac{\epsilon_i\lambda}{\sigma})} - \frac{\epsilon_i\lambda}{\sigma}\right], \tag{2.11}$$

where ϕ and Φ are, respectively the probability density function and the the cumulative function of a standard normal distribution, ϵ_i are overall errors and $\sigma_u, \sigma_v, \sigma, \lambda$ are previously defined variance parameters. The Jondrow formula for the NEX model is

$$E\left(u_i|\epsilon_i\right) = \sigma_v\left[\frac{\phi\left(\frac{\epsilon_i}{\sigma_v}+\frac{1}{\lambda}\right)}{1-\Phi\left(\frac{\epsilon_i}{\sigma_v}+\frac{1}{\lambda}\right)} - \left(\frac{\epsilon_i}{\sigma_v}+\frac{1}{\lambda}\right)\right]. \tag{2.12}$$

In the NTN model, the conditional mean function is:

$$E\left(u_i|\epsilon_i\right) = \frac{\sigma_u\sigma_v}{\sigma}\left[\frac{\phi(\frac{\epsilon_i\lambda}{\sigma}-\frac{\mu}{\sigma\lambda})}{1-\Phi(\frac{\epsilon_i\lambda}{\sigma}-\frac{\mu}{\sigma\lambda})} - \left(\frac{\epsilon_i\lambda}{\sigma}-\frac{\mu}{\sigma\lambda}\right)\right]. \tag{2.13}$$

The empirical counterpart of function $E(u_i|\epsilon_i)$ is point estimate $E(u_i|e_i)$, where $e_i = y_i - \hat{\boldsymbol{\beta}}'\mathbf{x_i}$ is the residual for the ith unit and $\sigma_u, \sigma_v, \lambda, \sigma, \mu$ are replaced by their estimates in the sample: $\hat{\sigma_u}, \hat{\sigma_v}, \hat{\lambda}, \hat{\sigma}, \hat{\mu}$[10].

Some properties and characteristics of Jondrow's formula need to be discussed. First, the conditional distribution of $(u_i|\epsilon_i)$ is a normal distribution truncated at zero, when unconditional u_i is distributed as half-normal, exponential or truncated normal[11]. Second, whatever the unconditional distribution of u_i, the conditional mean function is a non-negative and strictly decreasing function in ϵ_i (Jondrow, Lovell, Materov, and Schmidt, 1982; Wang and Schmidt, 2009). Third, the empirical estimator $E(u_i|e_i)$ is neither an unbiased nor a consistent estimator of u_i: it does not estimate u_i unbiasedly in the sense that, in repeated sampling, the mean of a set of observations on $E(u_i|e_i)$ would equal u_i[12]. Nevertheless, $E(u_i|e_i)$ is a consistent but not unbiased estimator of $E(u_i|\epsilon_i)$, because —according to ML estimates— it converges to the true conditional mean function (see Greene, 2008, p.178). Taking the analysis a step further, Wang and Schmidt (2009) observe that the distributions of u_i and of $E(u_i|\epsilon_i)$ are different, in the sense that $E(u_i|\epsilon_i)$ is a shrinkage of u_i towards its mean: on average, it will overestimate u_i when it is small and underestimate it when it is large. The above authors show that the distribution of $E(u_i|\epsilon_i)$ collapses on the distribution of u_i as $\sigma_v^2 \to 0$ if the value of σ_u^2 is fixed (i.e., $\lambda \to \infty$), but it collapses on point $E(u_i)$ as $\sigma_v^2 \to \infty$ if the value of σ_u^2 is fixed (i.e., $\lambda \to 0$). This is why

[10] As Wang and Schmidt (2009) claim, $E(u_i|\epsilon_i)$ and $E(u_i|e_i)$ are different because of the contribution of the estimation error in $\boldsymbol{\beta}$; however, the authors stress that the intrinsic randomness in $E(u_i|\epsilon_i)$, being a function of ϵ_i, counterbalances the randomness due to the estimation error in β.

[11] See Theorems 1 and 2 in Jondrow, Lovell, Materov, and Schmidt (1982).

[12] In fact $E(u_i|\epsilon_i)$ is an unbiased estimator of u_i in Theil's sense, i.e. zero expected prediction error $E(E(u_i|\epsilon_i)) = E(u_i)$ (Waldman, 1984, p.355).

they suggest not comparing the distribution of inefficiency estimates, $E(u_i|e_i)$ with the true distribution of inefficiency values u_i, but with the theoretical distribution of $E(u_i|\epsilon_i)$.

Thus empirical estimator $E(u_i|e_i)$ is only an 'indirect estimator' of u_i, because it is based on e_i (the overall residual), whereas it is a point estimate of $E(u_i|\epsilon_i)$. However, it represents also the 'best' information which can be recovered on the inefficiency level of a single production unit in a stochastic frontier framework, and the computed measure in the majority of empirical works. This is why, from the perspective of applied researchers, we compare $E(u_i|e_i)$ scores with true inefficiency values u_i[13].

2.4 Simulation protocol

2.4.1 Data generating process

In order to examine the performance of the three stochastic frontier models, NHN, NEX and NTN in estimating true inefficiency values, both when the models include the true inefficiency distribution, and when they do not (misspecification), we set up an experimental environment in which we mainly focused on inefficiency distribution. Starting from Equation 2.3, all units are assumed to produce following a Cobb-Douglas functional form with two inputs. Inputs x_1 and x_2 were generated from uniform distributions on the interval (5,15), independent of each other and of random components: they were generated once and taken fixed in repeated samples, in order to limit unnecessary randomness in the data. Technological parameters, $\beta_0 = .7$, $\beta_1 = .4$ and $\beta_2 = .6$ were used in all experiments. The dependent variable was generated after each realization (sample) of the error components as:

$$y_i = 0.7 + 0.4 \cdot x_1 + 0.6 \cdot x_2 + v_i - u_i. \tag{2.14}$$

This Cobb-Douglas specification has constant returns to scale. Since the inputs were generated as independent (i.e. not correlated), adding or removing another orthogonal regressor would not change the property of the ML estimator of the frontier parameters (see Olson, Schmidt, and Waldman, 1980, pp.76-78); the choice to use two inputs was only a question of managing a familiar type of Cobb-Douglas form. Passing on to how the stochastic terms of the model were generated, we performed experiments in two different settings, depending on how the variance parameters of the stochastic frontier model were fixed in order to generate the error components.

In the first setting, we monitored the performance of the three SFMs by correctly

[13]We also compared $E(u_i|e_i)$ scores with $E(u_i|\epsilon_i)$ values, but do not report results here: it is sufficient to note that the $E(u_i|e_i)$ scores are much more in 'line' with the $E(u_i|\epsilon_i)$ than with the u_i values. However, it is beyond doubt that the $E(u_i|e_i)$ scores are also generally taken as estimates of u_i, the estimation of which is the main concern of any kind of efficiency analysis.

estimating the inefficiency scores, taking σ^2 as fixed and making λ assume different values. This setting is similar to that used by Ruggiero (1999) and Jensen (2005) in their experiments, except for the fact that they did not keep σ^2 fixed. However, it is more convenient to make only one parameter vary at a time, in order to have better control of the underlying process[14]. In this first setting, the samples of the stochastic terms were generated in each replication in the following way:

- $v_i \sim N(0, \sigma_v^2)$ in all experiments;
- samples of u_i were generated from one of the following six different distributions
  - half-normal: $u_i \sim N^+(0, \sigma_u^2)$;
  - exponential: $u_i \sim Exp(\eta)$, where $\eta = \sigma_u$;
  - truncated normal: $u_i \sim N^+(\mu, \sigma_u^2)$;
  - uniform: $u_i \sim U(0, b)$, where b is the upper bound of the support of the distribution;
  - log-normal: $u_i \sim \log N(0, \sigma_u^2)$, where σ_u is the standard deviation of the natural log of the variable: the underlying normal distribution;
  - Weibull: $u_i \sim Weibull(k, \eta)$, where k is the shape parameter and $\eta = \sigma_u$, is the scale parameter of the distribution [15].
- $\sigma^2 = \sigma_u^2 + \sigma_v^2$=.5 in all experiments;
- $\lambda = \frac{\sigma_u}{\sigma_v}$ is the most important parameter in this exercise: its variation allows us to examine the performance of the three estimated models in estimating the true inefficiency values in samples with differing error terms. Six values were considered:

  $$\lambda = .5, 1, 5, 10, 20.$$

  The relative contribution of inefficiency to the overall error terms raises as λ increases;
- consequently, given the values of σ^2 and λ, σ_u and σ_v take the couples of values in Table 2.1.
- For the truncated normal distribution, the lower bound was set at zero ($a = 0$), and $\mu = (1.5) \cdot \sigma_u$ in order to obtain a shape of the truncated normal distribution which was sufficiently different from that of the nested half-normal.

[14] There is another reason for keeping σ^2 fixed while moving λ: as Coelli (1995, p.254) suggests, if the experiment is conducted for a particular point inthe parameter space $\left(\beta, \sigma^2, \lambda, \boldsymbol{X}\right)$, and then repeated for the same point the only alteration being a doubling of the value assumed for σ^2, then this would have the effect of multiplying each random error by $\sqrt{2}$. Thus, it is possible, without loss of generality, to assume $\sigma^2 = .5$ in all our experiments.

[15] Note that if k=1, the Weibull distributions collapses to an exponential distribution when the scale parameter is equal to η.

Table 2.1: Variance parameters - Setting 1

λ	σ_u	σ_v
0.500	0.316	0.632
1.000	0.500	0.500
5.000	0.693	0.138
10.000	0.703	0.070
20.000	0.706	0.035

- As the standard deviation of the uniform distribution depends on the interval over which the distribution is defined, (a, b), the lower bound (a) was set at zero, and the upper bound (b) was chosen in order to obtain a standard deviation which equals the above values for σ_u;

- The shape parameter of the Weibull distribution, k, was set constant in all experiments and is equal to .75 in order to obtain a rather different shape from that of the nested exponential; the scale parameter η is equal to σ_u.

A characteristic of this setting is that it does not permit us to compare samples with the same variances of the error terms. Going back to the observation at the end of Section 2.3.1 and taking —for purposes of explanation— half-normal and exponential distributions, samples generated from two DGPs which are equal in all relevant parameters ($\boldsymbol{\beta'}, \lambda, \sigma^2, \sigma_u, \sigma_v$), but which are different in the inefficiency distribution, will present error components with different variances. In particular, the sample in which the inefficiency term follows a half-normal distribution will be generated from a population with a value of $Var(u)$ which is almost one-third of its counterpart in the exponential population[16]; consequently, the two samples will also differ in the variance of overall error term $Var(\epsilon)$.

In order to see whether differences in variances could drive the simulation results, we also ran the experiments in another setting, keeping the variance of overall error term, $Var(\epsilon)$, fixed and making (the square root of) the ratio of the variances $\lambda^* = \sqrt{\frac{Var(u)}{Var(v)}}$ move. In this setting it is possible to monitor the performance of the three SFMs in correctly estimating the inefficiency scores in samples which are equal in error variances but not in the 'shape' of the inefficiency distribution[17]. In this case, the stochastic terms were generated as follows:

- $v_i \sim N(0, \sigma_v^2)$ in all experiments, where $Var(v) \equiv \sigma_v^2$;

- samples of u_i were generated following one of the six different distributions introduced above, characterized in each experiment by a particular value of $Var(u)$;

[16] $Var(u) = \sigma_u^2 \left[\frac{\pi-2}{\pi}\right]$ in the half-normal case and $Var(u) = \sigma_u^2$ in the exponential one

[17] This setting appears to be more in line with applications of SFMs to real data: in fact, applied researchers estimate models with different specifications of the inefficiency distribution, but the variance of the overall error term is given in the sample of units.

- $Var(\epsilon) = Var(u) + Var(v)$=.5 in all experiments;
- Five values of $\lambda^* \equiv \sqrt{\frac{Var(u)}{Var(v)}}$ were considered:

$$\lambda^* = .5, 1, 5, 10, 20.$$

  Given the monotonic relationship between λ^* and λ, also in this setting the relative contribution of the inefficiency term to the overall error raises as λ^* increases;
- The values of $Var(u)$ and $Var(v)$ considered in this setting and the corresponding values of σ_u for each distribution are listed in Table 2.2.
- Also in this setting, the lower bound of the truncated normal was set equal at zero ($a = 0$), and $\mu = (1.5) \cdot \sigma_u$ due to the explanations given above;
- As in the first setting, the upper bound (b) of the uniform distribution was chosen in order to obtain the above values for $Var(u) \equiv \sigma_u$;
- Also in this setting, the shape parameter of the Weibull distribution, k, was set constant in all experiments and is equal to .75; the scale parameter, η is equal to σ_u;

All simulations were performed in the Stata 10.1 environment. Both u_i and v_i were generated in each replication with the Stata pseudo-random number generator and a common random 'seed' (101010). In all experiments 1000 samples (i.e., the number of replications) were generated; as this work is not mainly concerned with small sample properties of the ML estimators of the frontier parameters or of the Jondrow formula, we focused on samples of 1000 units[18].

2.4.2 Estimated models

The three frontier models, NHN, NEX and NTN were estimated: their task was to estimate the true inefficiency scores as well as possible, both when the inefficiency distribution has been correctly specified and when it has not. In the estimated models, the form of the production function was correctly specified (Cobb-Douglas with the proper number of inputs, x_1 and x_2): this was done in order to isolate the effect of the sole misspecification of the inefficiency distribution[19].

Stochastic frontier models were estimated by means of the command `frontier`. In each experiment, maximization of the log-likelihood function was performed by

[18] Nonetheless, the same experiments were also performed with samples of 100 units and the results are in line with those discussed here: they are available from the author upon request.

[19] Admittedly, the correct specification of the production function is another important problem in SFMs and, more in general, in all parametric methods for measuring efficiency and productivity. However, some papers have already monitored undesirable consequences of this kind of misspecification: an extensive exercise on this issue was carried out by Giannakas, Tran, and Tzouvelekas (2003).

Table 2.2: Variance parameters - Setting 2

	λ	σ_u	σ_v	$\sigma^2 = \sigma_u^2 + \sigma_v^2$	$\lambda^* = \sqrt{\frac{Var(u)}{Var(v)}}$	$Var(u)$	$Var(v)$	$Var(\epsilon) = Var(u) + Var(v)$
Half-normal	0.833	0.527	0.632	0.677	0.5	0.100	0.400	0.5
	1.666	0.833	0.500	0.944	1	0.250	0.250	0.5
	8.333	1.150	0.138	1.355	5	0.480	0.020	0.5
	16.666	1.173	0.070	1.38	10	0.495	0.005	0.5
	33.437	1.177	0.035	1.387	20	0.498	0.002	0.5
Exponential	0.5	0.316	0.632	0.500	0.5	0.100	0.400	0.5
	1	0.500	0.500	0.500	1	0.250	0.250	0.5
	5	0.693	0.138	0.500	5	0.480	0.020	0.5
	10	0.703	0.070	0.500	10	0.495	0.005	0.5
	20	0.705	0.035	0.500	20	0.498	0.002	0.5
Truncated normal	0.570	0.360	0.632	0.529	0.5	0.100	0.400	0.5
	1.140	0.570	0.500	0.575	1	0.250	0.250	0.5
	5.700	0.790	0.138	0.643	5	0.480	0.020	0.5
	11.380	0.800	0.070	0.645	10	0.495	0.005	0.5
	23.140	0.810	0.035	0.657	20	0.498	0.002	0.5
Uniform	0.5	0.316	0.632	0.500	0.5	0.100	0.400	0.5
	1	0.500	0.500	0.500	1	0.250	0.250	0.5
	5	0.693	0.138	0.500	5	0.480	0.020	0.5
	10	0.703	0.070	0.500	10	0.495	0.005	0.5
	20	0.705	0.035	0.500	20	0.498	0.002	0.5
Log-normal	0.47	0.296	0.632	0.488	0.5	0.100	0.400	0.5
	0.86	0.433	0.500	0.438	1	0.250	0.250	0.5
	3.97	0.551	0.138	0.323	5	0.480	0.020	0.5
	7.91	0.556	0.070	0.315	10	0.495	0.005	0.5
	15.94	0.558	0.035	0.313	20	0.498	0.002	0.5
Weibull	0.312	0.197	0.632	0.439	0.5	0.100	0.400	0.5
	0.624	0.311	0.500	0.347	1	0.250	0.250	0.5
	3.120	0.432	0.138	0.206	5	0.480	0.020	0.5
	6.240	0.438	0.070	0.197	10	0.495	0.005	0.5
	12.470	0.440	0.035	0.195	20	0.498	0.002	0.5

iterating the numerical procedure up to 300 times at most (otherwise declaring 'non-convergence'), and by switching between the Newton-Raphson (`NR`) and Broyden-Fletcher-Goldfarb-Shanno (`BFGS`) method every 50 iterations up to convergence of the maximization procedure[20]. After estimation of the frontier parameters via ML, the mean of the conditional distribution of u_i given overall residual e_i was computed in order to estimate inefficiency scores, picking up for each of the three estimated models the proper Equation, as 2.11, 2.12, 2.13.

To asses the performance of the three estimated models in correctly estimating the true inefficiency scores, two measures were computed in each replication:

- the Spearman rank correlation coefficient between true and estimated inefficiency scores[21]
- the average absolute difference between true and estimated inefficiency scores; this measure may be defined as:

$$\overline{diff} = \sum_i \frac{|u_i - E(u_i|e_i)|}{n},$$

  where u_i are true inefficiency values, $E(u_i|e_i)$ are estimated inefficiency scores and n is the number of units in the sample.

We also estimated and reported technology parameter estimates ($\hat{\boldsymbol{\beta}}$) and variance parameters estimates ($\hat{\lambda}, \hat{\sigma^2}$), as they are included in the Jondrow formula and, thus, if biased, may have influenced the estimation of the inefficiency scores.

To sum up, the Monte Carlo experiments conducted in the first setting features three treatments (estimated models), six different true distributions (DGP) for u_i, and five values for λ, with a total of $3 \times 6 \times 5 = 90$ possible combinations. The second setting adds 90 more combinations, i.e. 180 scenarios. To our knowledge, this is the largest exercise on possible misspecifications of the inefficiency distribution which have been examined in the stochastic frontier literature until now. The detailed Stata code is available from the author upon request.

2.5 Results

2.5.1 Usual distributions - Setting 1

We start by commenting on the results obtained in the first Setting and in those cases in which u_i were generated following a half-normal, exponential or truncated

[20] Switching between techniques is a good way of finding the maximum of a difficult-to-maximize function. See Gould, Pitblado, and Sribney (2006), pp. 16-20 for an introduction to the maximization methods employed.

[21] The well-known coefficient which assesses how well the relationship between two variables can be described with a monotonic function, can be written as $\rho = 1 - \frac{6\sum d_i^2}{n(n-1)}$; in our case, d_i =rank(u_i)−rank($E(u_i|e_i)$) and n is the number of units;

normal. First, we look at the estimates of the technology parameters of the three estimated models of Table 2.3: the three main columns list the estimated model, and the rows list the distributions from which the u_i values were generated. Thus, the three blocks on the diagonal list the results for correctly specified models, and the off-diagonal blocks list the six cases of misspecification of distribution. The number of successful replications (converged maximization procedures) is listed in the far right column. The estimates of the intercept are unbiased only in some cases: in particular,

Table 2.3: Technology parameter estimates - Setting 1

		Estimated models											
		NHN				NEX				NTN			
DGP	λ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	s.reps	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	s.reps	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	s.reps
Half-normal	0.5	0.666	0.400	0.600	1000	0.579	0.400	0.600	992	0.917	0.400	0.600	1000
	1	0.679	0.400	0.600	1000	0.522	0.400	0.600	1000	0.870	0.400	0.600	999
	5	0.700	0.400	0.600	998	0.569	0.400	0.600	1000	0.702	0.400	0.600	998
	10	0.700	0.400	0.600	1000	0.612	0.400	0.600	1000	0.700	0.400	0.600	1000
	20	0.700	0.400	0.600	999	0.645	0.400	0.600	1000	0.700	0.400	0.600	999
Exponential	0.5	0.876	0.400	0.600	1000	0.685	0.400	0.600	999	0.878	0.400	0.600	1000
	1	0.928	0.400	0.600	963	0.696	0.400	0.600	1000	0.730	0.400	0.600	963
	5	0.799	0.400	0.600	1000	0.698	0.400	0.600	983	0.702	0.400	0.600	970
	10	0.754	0.400	0.600	999	0.698	0.400	0.600	1000	0.702	0.400	0.600	646
	20	0.728	0.400	0.600	999	0.698	0.400	0.600	1000	0.700	0.400	0.600	594
Truncated normal	0.5	0.414	0.400	0.600	1000	0.320	0.400	0.600	993	0.725	0.400	0.600	998
	1	0.282	0.400	0.600	1000	0.104	0.400	0.600	999	0.735	0.400	0.600	1000
	5	0.227	0.400	0.600	1000	-0.069	0.400	0.600	1000	0.710	0.400	0.600	997
	10	0.234	0.400	0.600	1000	-0.075	0.400	0.600	1000	0.707	0.400	0.600	972
	20	0.238	0.400	0.600	1000	-0.075	0.400	0.600	1000	0.700	0.400	0.600	952
True value		0.700	0.400	0.600		0.700	0.400	0.600		0.700	0.400	0.600	

they are almost unbiased in the blocks on the diagonal (correctly specified models), but are not in cases of misspecification of the inefficiency distribution. The correctly specified NHN and NEX models underestimates the intercept for $\lambda = .5$, but correctly estimate it for higher values of the parameter; conversely, the correctly specified NTN model overestimates the intercept for $\lambda < 5$. Intercept biases which were also found by Coelli (1995) and Olson, Schmidt, and Waldman (1980), are due to identification problems between that parameter and the mean of the inefficiency component. One interesting result is that, for medium and high values of λ, the NTN model can correctly estimate the frontier intercept, even when the inefficiency distribution is misspecified and the true u_i were generated as half-normal or exponential. This is not so for the NHN and NEX models, if the assumed distribution does not coincide with the true one. The superiority of the NTN model in estimating the intercept in cases of misspecification may be explained by the fact that the truncated normal is a more 'flexible' distribution which has two parameters and not just one like the half-normal or exponential. It can thus better adapt to the inefficiency distribution when it is generated from simpler (e.g., exponential), or even nested (e.g., half-normal) distributions[22].

[22]However, this has a cost: the NTN model is well known to prevent convergence of iterations quite frequently, because the log-likelihood is ill-behaved when μ is unrestricted (see Greene, 2008, p.130). This

Overall, the role of λ is evident; Figure 2.1 shows kernel density estimates of the intercept in the three correctly specified cases: the higher the value of λ, the more concentrated are the kernels around the true values. Instead, the estimates of input coefficients (output elasticities), $\widehat{\beta_1}$ and $\widehat{\beta_2}$, are unbiased in both correctly specified frontier models and misspecified ones: this result was expected, in view of the properties of the ML estimators and of the fact that x_1 and x_2 were generated as independent of the error terms and of each other[23]. Data Appendix 2.7.4 provides kernel density estimates for $\widehat{\beta_1}$ and $\widehat{\beta_2}$.

Moving on to variance parameters, the patterns of estimates are similar to those of technology parameters, listed in Tables 2.4 and 2.5. In particular, the correctly

Table 2.4: λ parameter estimates - Setting 1

		Estimated models					
		NHN		NEX		NTN	
DGP	λ	$\hat{\lambda}$	extreme v.	$\hat{\lambda}$	extreme v.	$\hat{\lambda}$	extreme v.
Half-normal	0.5	0.450	0	0.202	0	1.081	83
	1	0.959	0	0.410	0	1.863	76
	5	5.156	1	2.075	0	5.247	4
	10	10.590	0	3.941	0	10.733	1
	20	22.393	20	7.606	0	22.535	17
Exponential	0.5	1.073	0	0.487	0	4.351426	17
	1	2.183	0	1.009	0	11.70022	0
	5	11.203	24	5.151	0	75.789	2
	10	23.565	61	10.475	0	92.588	4
	20	53.985	196	21.939	11	171.250	21
Truncated normal	0.5	0.461	0	0.206	0	1.299	74
	1	0.871	0	0.355	0	1.646	114
	5	2.133	0	0.702	0	5.789	53
	10	2.360	0	0.738	0	12.239	156
	20	2.437	0	0.749	0	22.188	403

Table 2.5: σ^2 parameter estimates (true value=.5) - Setting 1

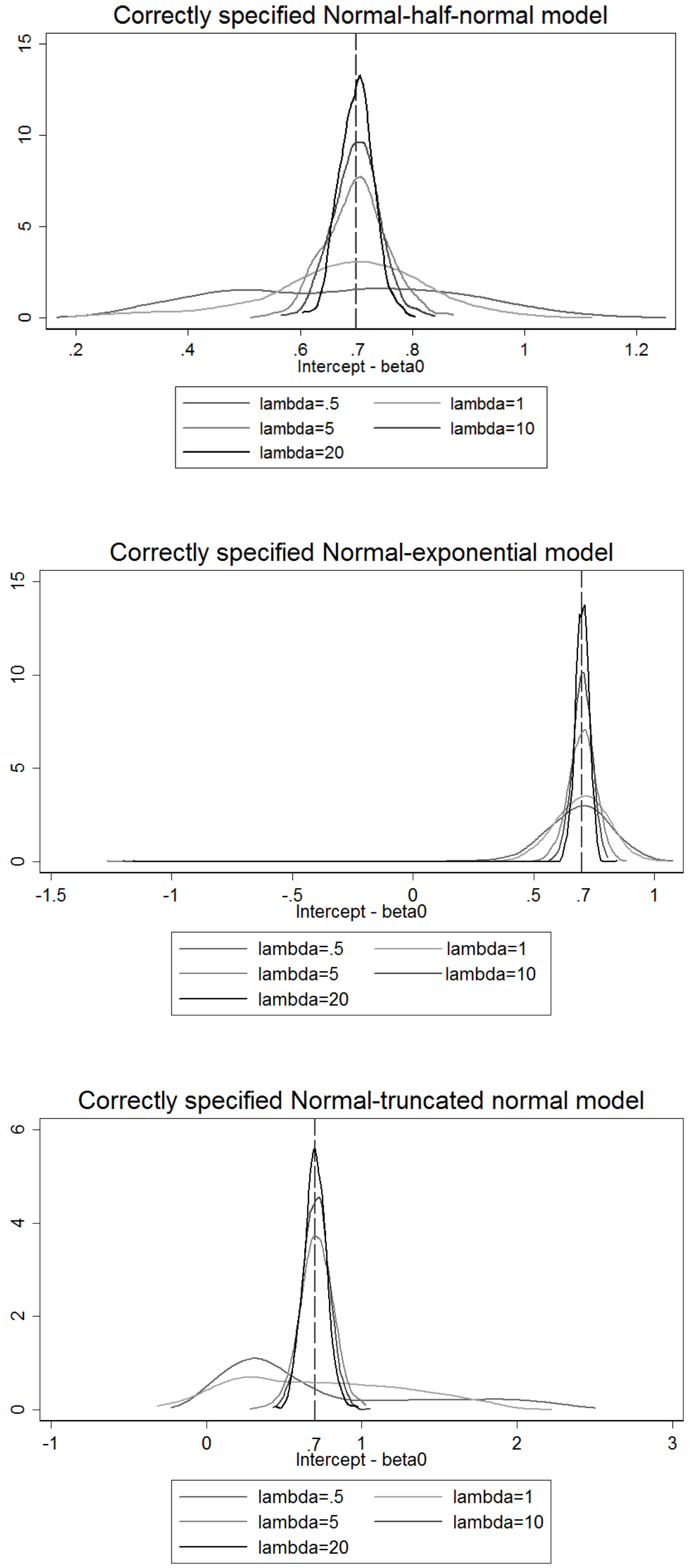
		Estimated models					
		NHN		NEX		NTN	
DGP	λ	$\hat{\sigma}^2$	extreme v.	$\hat{\sigma}^2$	extreme v.	$\hat{\sigma}^2$	extreme v.
Half-normal	0.5	0.511	0	0.435	0	0.501	152
	1	0.493	0	0.341	0	0.592	157
	5	0.499	0	0.221	0	0.510	4
	10	0.499	0	0.240	0	0.507	3
	20	0.498	0	0.263	0	0.504	2
Exponential	0.5	0.755	0	0.497	0	9.940	41
	1	1.032	0	0.498	0	51.637	0
	5	1.128	0	0.499	0	154.023	0
	10	1.073	0	0.499	0	41.851	24
	20	1.038	0	0.498	0	30.833	15
Truncated normal	0.5	0.558	0	0.474	0	0.537	170
	1	0.615	0	0.442	0	0.539	142
	5	0.819	0	0.394	0	0.501	0
	10	0.847	0	0.392	0	0.500	0
	20	0.854	0	0.390	0	0.501	0

specified models (main diagonal blocks) estimate parameter λ quite well and show

is confirmed by the relatively low number of successful replications reported in the far right column of Table 2.3.

[23] Of course, the result is independent of the value of λ.

Figure 2.1: Kernel densities of $\widehat{\beta}_0$: correctly specified models - Setting 1



unbiased estimates of parameter σ^2. Nevertheless, for $\lambda = 20$, parameter λ has extreme values[24] in some replications, and their numbers are listed next to each column: the NTN model suffers worst from this problem. For values lower than or equal to 1, the negative bias of parameter λ is in line with the evidence reported by Coelli (1995). In correctly specified models, parameter σ^2 shows a positive bias for small values of λ and a negative or no bias at all for large values of λ, as already documented by Olson, Schmidt, and Waldman (1980). Overall, the NTN model presents a large number of extreme values in variance parameter estimates, both when it has been correctly specified and when u_i comes from a half-normal or an exponential: as stated above, the higher flexibility of the two-parameters distribution involves the cost of a greater computational difficulty, which may result in extreme values in other parameter estimates.

After having examined the estimates of technology and variance parameters, we can now focus on the performance of the three models in estimating true inefficiency values. The main problem here is the consequence of inefficiency distribution misspecification on the correct estimation of the values and the ranking of u_i. Table 2.6 lists the Spearman rank correlation coefficients and absolute differences between true and estimated inefficiency values: both measures are averaged across all replications. As above, results related to the correctly specified models are listed in the main diagonal blocks and, cases of misspecification on the other blocks.

Table 2.6: Rank correlations and absolute differences between u_i and $E(u_i|e_i)$ - Setting 1

		Estimated models					
		NHN		NEX		NTN	
DGP	λ	$\widehat{\rho}$	$\overline{diff}$	$\widehat{\rho}$	$\overline{diff}$	$\widehat{\rho}$	$\overline{diff}$
Half-normal	0.5	0.268	0.208	0.268	0.185	0.268	0.444
	1	0.482	0.221	0.482	0.243	0.482	0.373
	5	0.931	0.101	0.931	0.156	0.931	0.103
	10	0.980	0.055	0.980	0.098	0.980	0.055
	20	0.994	0.029	0.994	0.059	0.994	0.029
Exponential	0.5	0.372	0.291	0.372	0.210	0.372	0.334
	1	0.597	0.331	0.598	0.246	0.597	0.254
	5	0.946	0.135	0.946	0.101	0.946	0.102
	10	0.982	0.073	0.983	0.055	0.983	0.055
	20	0.994	0.039	0.995	0.029	0.994	0.030
Truncated normal	0.5	0.390	0.344	0.390	0.396	0.390	0.551
	1	0.647	0.460	0.647	0.608	0.647	0.511
	5	0.972	0.488	0.972	0.779	0.973	0.121
	10	0.991	0.479	0.991	0.785	0.993	0.065
	20	0.996	0.475	0.996	0.785	0.998	0.036

A first result, in line with previous works, is that the rank correlation coefficients between true inefficiency scores and estimated ones rises as λ increases. This result is robust for all three estimated models both correctly specified and misspecified. This means that, as the inefficiency term dominates the noise term in the sample, SFMs are better able to reproduce the correct ranking of inefficiency. The explanation is that,

[24] Taking the interquartile range of the distribution of $\widehat{\lambda}$, $IQR = Q3 - Q1$, we define a value $\widetilde{\lambda}$ as 'extreme' if $\widetilde{\lambda} < Q1 - 3IQR$ or $\widetilde{\lambda} > Q3 + 3IQR$.

as λ increases, σ_v^2 becomes relatively smaller than σ_u^2 and $\epsilon_i \to -u_i$, so we effectively estimate something that is closer to u_i and asymptotically $E(u_i|e_i) = u_i$. Instead, as λ becomes smaller, i.e., σ_v^2 becomes larger with respect to σ_u^2, ϵ_i contains less useful information about u_i, which may consequently be worse estimated. The same result holds good for the differences between true inefficiency scores and estimated ones: the difference decreases as λ increases, thus indicating that all the models can better reproduce the inefficiency values as the one-sided term dominates the symmetric error component. This fact again holds good both for correctly specified and misspecified models: two exceptions regard the misspecification of the NHN and NEX models. If the true inefficiency distribution is truncated normal, the two models estimate worse inefficiency scores as λ becomes larger[25].

A second result, which is the most interesting, is that for each of the true inefficiency distributions (rows), the three estimated models (columns) reach the same rank correlation values, thus indicating a perfect agreement in the reproduced ranking[26]. The degree of similarity is astonishing, and, because the choice of the column is the only thing that is under the researcher's control, results show that the choice does not matter from the point of view of the resulting ranking of inefficiency scores. This result strengthens two quite common claims — only partially supported until now— in literature on stochastic frontier models: the general concordance in inefficiency ranking among different estimated models (Kumbhakar and Lovell, 2000) and the general preference for the simplest inefficiency distributions, basically half-normal and exponential, over most flexible but also computationally burdensome truncated normal and gamma distributions. The explanation rests on the fact that the Jondrow formula is a non-linear transformation of the residuals. We refer readers to the Data Appendix 2.7.2 for some algebra on this result. Practically, the only difference between the vectors of residuals in the three estimated models is given by the difference between the true value of the intercept (.7 in our case) and its estimate, which is a linear factor that shifts the whole distribution of e_i. Now, taking the first row of Table 2.6, in the case of the correctly specified model:

$$\beta_0 \cong \hat{\beta}_{0(hn,NHN)} \to e_{(hn,NHN)} \cong v_{(hn)} - u_{(hn)},$$

where hn is the true inefficiency distribution from which the u_i values were generated and NHN is the estimated model. The misspecified models k =NEX,NTN contain a shift-factor, given by the quantity $\left(\beta_0 - \hat{\beta}_{0(hn,k)}\right)$ which may be positive or negative.

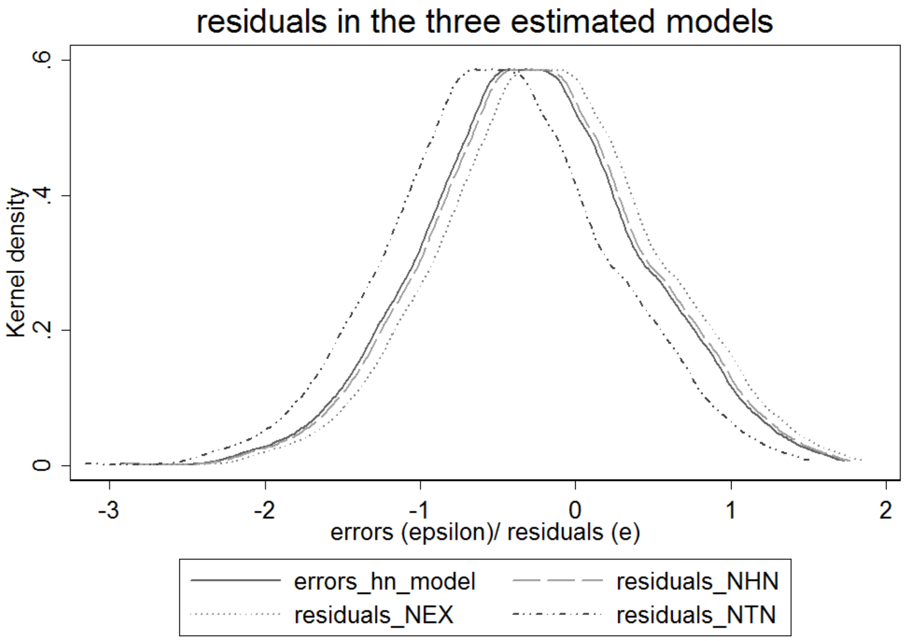
[25]This unexpected behaviour may be explained by the way in which the truncated normal has been generated: in particular, given that the mean of the pre-truncated distribution has been generated as $\mu = (1.5) \cdot \sigma_u$ and given that as λ increases also σ_u increases, higher values of λ are associated to truncated normal distributions with the central tendency further away from zero. The half-normal and exponential distribution seem to not adapt well to such distributions, especially the second one.

[26]This is in line with some results provided by Jensen (2005), although not emphasized and explored by the author.

If this difference is positive (i.e., $\hat{\beta}_{0(hn,k)}$ has a downward bias), the distribution of the residuals is shifted to the right with respect to the underlying errors; if it is negative (i.e. $\hat{\beta}_{0(hn,k)}$ has an upward bias), the residuals are shifted to the left of the error distribution[27]. However, this factor is constant across observations, and ranking of the residuals is not affected by it.

Figure 2.2 illustrates this fact. For explanatory purpose, we assume that $\lambda = .5$ and compare the distribution of errors in which u_i were generated from a half-normal, the distribution of the residuals of the three estimated models, NHN, NEX and NTN, takes the averages of the estimated parameters listed in the first row of Tables 2.3, 2.4 and 2.5 as the estimates in a given sample. The distribution of residuals which is closer to the distribution of errors is clearly that of the correctly specified model (NHN), whereas the other two distributions of e_i are shifted (with respect to ϵ_i) by $\left(\beta_0 - \hat{\beta}_{0(hn,k)}\right)$: the direction of the shift is in line with our expectations. The whole

Figure 2.2: Kernel densities of residuals

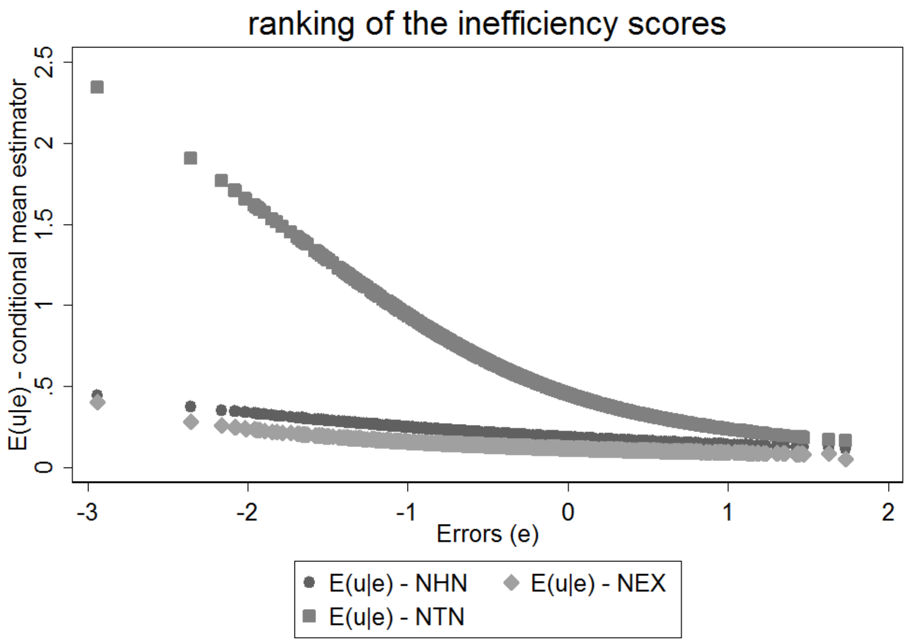


distribution of residuals in each of the three estimated models is shifted either left or right of the error distribution, but the ranking is unaffected. As the Jondrow estimator is a decreasing and monotonic function of residuals, whatever the true distribution of u_i, the three models (with their respective Jondrow's formulas) predict the same ranking of the inefficiency scores, as shown in Figure 2.3. Incidentally, the estimates of the variance parameters do not seem to affect the inefficiency ranking, even if they appear in the Jondrow formulas[28].

[27] Of course, the shape of the distribution of the residuals, e_i, does not coincide with that of the errors, ϵ; the difference is given by the estimation error in vector β.

[28] Table 2.4 shows that for each DGP and for $\lambda = .5$, $\hat{\lambda}$ assumes different values depending on the

Figure 2.3: Conditional mean estimators



A third result is that if we are interested in inefficiency values, the choice of the correct distribution does matter. For each of the estimated models the average difference between u_i and $E(u_i|e_i)$ is lower in the correctly specified model than in the two misspecified ones. Interestingly enough, when the true inefficiency distribution is half-normal or truncated normal, the NTN model can reproduce the inefficiency scores better than the NEX model. This is in line with the fact that the half-normal distribution is nested into the truncated normal.

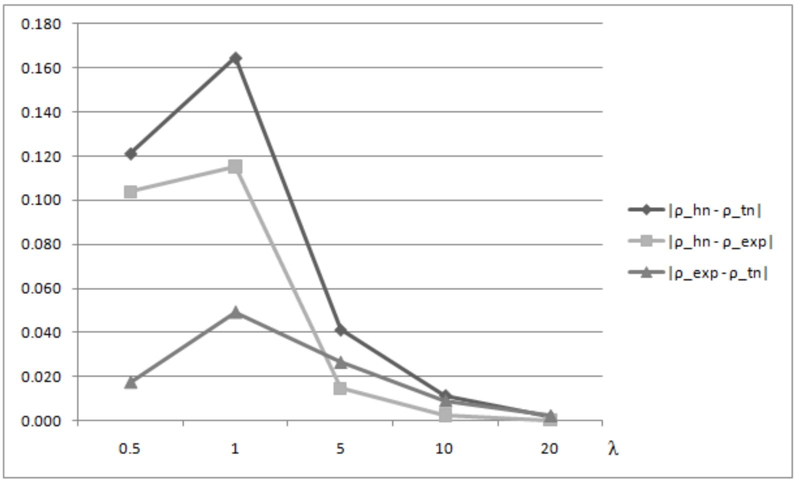
2.5.2 Usual distributions - Setting 2

Table 2.6 shows that the same estimated model has different rank correlation coefficients depending on the true inefficiency distribution. As an example, let us take the first main column of Table 2.6, which refers to the NHN model: when the true inefficiency values are generated as half-normal and $\lambda = .5$, the rank correlation is (.268), but when they are generated as exponential the rank correlation is (.372). Lastly, when u_i are generated as truncated normal, the rank correlation is (.390). These differences are especially evident for values of λ lower than 5; for values greater than or equal to 5, the rank correlation coefficients are almost equal and this is true for any of the three estimated models. Figure 2.4 shows the absolute value of the difference between rank correlation coefficients obtained by estimating the same model (NHN), if the inefficiency values are generated from a half-normal, exponential or truncated

estimated model, but this does not affect the results on ranking at all. However, there are fewer differences in the estimates of σ^2 among the three models.

normal distribution. Each rank correlation coefficient was compared with the other two. As λ increases, differences between pairs of rank correlation coefficients decrease, and the true inefficiency distribution does not seem to matter much for the estimation of the inefficiency ranking. Nonetheless, differences are not negligible for 'low' values of λ. Starting from analogous observations, Ruggiero (1999) and Jensen (2005) sug-

Figure 2.4: Differences in rank correlation coefficients: NHN model - Setting 1



gested that researchers should think in terms of preferring some models rather than others. However, it is necessary to understand what drives these differences before embarking on any kind of suggestion. The first plausible motivation relates is that in each row and for each value of λ in Table 2.6, samples of u_i were generated from populations with the same value of σ_u, but different variances. Recalling Section 2.4.1, it is interesting to see whether differences in the variances of the error components give rise to this result. The expectation is that samples generated from different DGPs which are different only in the 'shape' of the inefficiency distribution should be more easily comparable and produce more 'aligned' results.

In order to test this hypothesis, we re-ran the experiments keeping $Var(\epsilon) = Var(u) + Var(v)$ fixed and equal to (.5) and moving $\lambda^* = \sqrt{\frac{Var(u)}{Var(v)}}$. In order to save space, we directly report the performance measures of the three estimated models in the new setting in Table 2.7, relegating the technology and variance parameter estimates to Data Appendix 2.7.3. Looking at the first main column of Table 2.7, for cases in which $\lambda^* = .5$, when the true inefficiency scores are generated following a half-normal, the rank correlation between true and estimated inefficiency scores is (.419), and when the u_i are generated as truncated normal, the rank correlation is (.434). The two are really closer to each other with respect to the same case in Setting 1. Conversely, when the inefficiency values are generated as exponential the rank correlation is (.372), showing a underlying distribution which is (even though it was generated with the same variance) is 'qualitatively' different from the other two probability density functions. Thus, once $Var(u)$ has been fixed, the half-normal

Table 2.7: Rank correlations and absolute differences between u_i and $E(u_i|e_i)$ - Setting 2

		Estimated models					
		NHN		NEX		NTN	
DGP	λ^*	ρ	$\bar{diff}$	ρ	$\bar{diff}$	ρ	$\bar{diff}$
Half-normal	0.5	0.419	0.261	0.419	0.273	0.419	0.489
	1	0.669	0.276	0.669	0.337	0.669	0.315
	5	0.971	0.107	0.971	0.185	0.971	0.113
	10	0.992	0.057	0.992	0.113	0.992	0.058
	20	0.998	0.030	0.998	0.069	0.998	0.031
Exponential	0.5	0.372	0.291	0.372	0.210	0.372	0.334
	1	0.597	0.331	0.598	0.246	0.597	0.254
	5	0.946	0.135	0.946	0.101	0.946	0.102
	10	0.982	0.073	0.983	0.055	0.983	0.055
	20	0.994	0.039	0.995	0.029	0.994	0.030
Truncated normal	0.5	0.434	0.385	0.434	0.453	0.434	0.587
	1	0.695	0.504	0.695	0.686	0.695	0.501
	5	0.978	0.551	0.978	0.887	0.979	0.123
	10	0.993	0.543	0.993	0.893	0.993	0.067
	20	0.997	0.545	0.997	0.902	0.998	0.038

and truncated normal lead to almost identical rank correlation coefficients. Instead, as Greene (2008, p.120) has emphasized, the exponential implies tighter clustering of the inefficiency values near zero, explaining the different rank correlation coefficients obtained. Although researchers have no 'control' over the true inefficiency distribution, we believe that this result will reassure them regarding the decision on how to model inefficiency.

2.5.3 Unusual distributions - Setting 1

In order to further check the performance of SFMs in estimating inefficiency values accurately, in another set of experiments we generated true inefficiency values from three unusual distributions: log-normal, Weibull and uniform distributions. We estimated the three models, NHN, NEX and NTN, questioning their flexibility in adapting to inefficiency distributions which are quite different from those on which they are usually built. Nonetheless, the results are in line with those related to the usual distributions. We do not report here the technology and variance parameter estimates in order to save space, as their behaviour with respect to λ is in line with the cases shown in Tables 2.3 2.4 2.5, but directly comment on the performance of the three models in estimating inefficiency scores[29]. Table 2.8 lists the average rank correlation coefficient and the average absolute difference between true and estimated inefficiencies. It is important to bear in mind that all the blocks in this matrix are types of misspecification of the inefficiency distribution: there are no correctly specified models here. However, the main result of our analysis is stable: whatever

[29]Tables of the technology and variance parameters are available from the author upon request.

the true inefficiency distribution (row), each of the three estimated models reaches almost the same rank correlation coefficient; in other words, even when the true u_i were generated following 'unusual' distributions, each model performs equally in correctly reproducing the ranking of inefficiency.

Table 2.8: Rank correlations and absolute differences between u_i and $E(u_i|e_i)$ - Setting 1

		Estimated models					
		NHN		NEX		NTN	
DGP	λ	$\hat{\rho}$	$diff$	$\hat{\rho}$	$diff$	$\hat{\rho}$	$diff$
Log-normal	0.5	0.437	0.351	0.437	0.463	0.437	0.485
	1	0.691	0.301	0.691	0.412	0.691	0.396
	5	0.971	0.133	0.971	0.242	0.971	0.241
	10	0.991	0.104	0.992	0.227	0.992	0.224
	20	0.997	0.098	0.998	0.222	0.997	0.218
Weibull	0.5	0.460	0.477	0.460	0.307	0.460	0.311
	1	0.662	0.577	0.665	0.320	0.657	1.059
	5	0.943	0.216	0.945	0.132	0.944	0.136
	10	0.978	0.120	0.980	0.075	0.980	0.077
	20	0.992	0.062	0.993	0.042	0.992	0.045
Uniform	0.5	0.445	0.421	0.445	0.460	0.445	0.614
	1	0.714	0.742	0.714	0.807	0.714	0.768
	5	0.980	1.136	0.980	1.174	0.980	0.689
	10	0.993	1.157	0.993	1.194	0.993	0.658
	20	0.997	1.163	0.997	1.198	0.997	0.646

2.5.4 Unusual distributions - Setting 2

As we have stressed previously, in order to have more comparable data generating processes samples must be generated from inefficiency distributions which are equal in terms of variance but different in terms of the shape of the distribution. Table 2.9 lists the rank correlation coefficients and average absolute differences between true inefficiency values and estimated ones, once u_i are generated from the three unusual distributions considered above, keeping $Var(\epsilon)$ fixed and making λ^* move, and thus considering the same variance of the inefficiency distributions for each value of λ^*.

Table 2.9: Rank correlations and absolute differences between u_i and $E(u_i|e_i)$ - Setting 2

		Estimated models					
		NHN		NEX		NTN	
DGP	λ	$\hat{\rho}$	$\widehat{diff}$	$\hat{\rho}$	$\widehat{diff}$	$\hat{\rho}$	$\widehat{diff}$
Log-normal	0.5	0.414	0.361	0.414	0.461	0.414	0.510
	1	0.641	0.299	0.641	0.431	0.641	0.392
	5	0.960	0.184	0.960	0.332	0.959	0.307
	10	0.988	0.178	0.989	0.318	0.988	0.304
	20	0.996	0.177	0.997	0.314	0.996	0.299
Weibull	0.5	0.328	0.373	0.328	0.233	0.328	0.290
	1	0.526	0.408	0.526	0.269	0.526	0.276
	5	0.897	0.188	0.897	0.120	0.897	0.120
	10	0.959	0.108	0.960	0.069	0.959	0.070
	20	0.985	0.060	0.986	0.039	0.981	0.045
Uniform	0.5	0.445	0.421	0.445	0.460	0.445	0.614
	1	0.714	0.742	0.714	0.807	0.714	0.768
	5	0.980	1.136	0.980	1.174	0.980	0.689
	10	0.993	1.157	0.993	1.194	0.993	0.658
	20	0.997	1.163	0.997	1.198	0.997	0.646

Interestingly enough, the values of the ranking correlation coefficients obtained by the estimated models when the inefficiency scores follow a log-normal distribution, are in line with the coefficients obtained when u_i are either generated as half-normal or truncated normal (see Table 2.7). Instead, the rank correlation coefficients obtained by the three models when the inefficiency values follow a Weibull distribution is in line with the coefficients when the inefficiency distribution is exponential.

Overall, the six distributions are qualitatively different: although they share the same variance, the shape of the distribution seems to —at least for low values of λ— to influence the performance of each of the three estimated models in reproducing inefficiency score rankings. Nested distributions (like exponential with Weibull, or half-normal with truncated normal) produce similar results.

2.6 Concluding remarks and steps for further research

In stochastic frontier analysis, the technical inefficiency term is usually assumed to be half-normal, exponential or truncated normal. Researchers using frontier models have questioned whether the specific form of the inefficiency distribution, u_i, is important for their conclusions, and whether it can actually give rise to the results: a common robustness check compares the results by estimating stochastic frontier models which are different for the assumed inefficiency distribution. Although previous evidence indicates general concordance among the sets of estimated inefficiency scores in a non-negligible number of applied works, a systematic evidence corroborated by detailed explanations is still lacking in the literature[30]. In this paper we assessed the performance of stochastic frontier models in correctly estimating true inefficiency scores, both when the inefficiency term has been correctly specified and when it has not. In order to monitor the behavior of models in a fully controlled environment we used of a set of Monte Carlo experiments, which allowed us to analyze in more depth the performance of the three most frequently employed models, normal-half-normal, normal-exponential and normal-truncated normal, as the true inefficiency distribution is known.

Overall, the news for practitioners are encouraging. If inefficiency ranking is the main concern of the analysis, the three most frequently estimated models give the same result, so that the specification of the inefficiency distribution does not matter. This is consistent with the analytical results of Ondrich and Ruggiero (2001), who demonstrated that for any stochastic frontier model[31], the Jondrow estimator is a monotonic decreasing function of residuals and that the rank correlation between the

[30] Greene (2008, p.180) argues that the question does not have an analytical answer.

[31] The authors show that the rank correlation between the Jondrow estimates of technical inefficiency and the maximum likelihood composed error is one in models in which v_i follows a log-concave distribution. Consequently, deterministic models based on ML estimation achieve the same ranking as stochastic frontier models.

two is equal to one. However, they take it as a 'negative' result which questions the need to use a stochastic frontier approach when a deterministic one could simply be applied, without having to separate noise from inefficiency. Instead, we think that there is also a positive message: if we are interested in measuring inefficiency, and try to separate it from noise, stochastic frontier models must be used, especially in samples with a mixture of noise and inefficiency in the data (see the evidence provided by Waldman, 1984, pp.357-358). In addition, researchers can be confident that the choice of the model to be estimated, in terms of the assumed inefficiency distribution, does not really matter as regard the resulting inefficiency scores ranking. This result also indicates that the analysis should be started by assuming simple distributions (e.g., half-normal or exponential) if the main concern is to estimate inefficiency ranking[32]. It is important to note that the results are robust to various types of misspecifications. Although the 'true' inefficiency distribution is quite different from the usually assumed ones, the ranking yielded by the three estimated models is equal. Conversely, if inefficiency values *per se* are the focus of interest, specification of the correct distribution does matter. Unfortunately the true inefficiency distribution is never known in applied works and this second result cannot be exploited further for practical suggestions on which model should be preferred.

From a methodological point of view, this paper examined the role of variance of error components. In previous experiments on the misspecification of the inefficiency distribution, like those of Ruggiero (1999) and Jensen (2005), the results 'suffered' from the fact that the authors compared inefficiency distributions with the same value of σ_u, but different values of $Var(u)$. In this paper we performed all experiments in two different settings: the first was similar to that of previous studies for purposes of comparison, and the second, in order to check the robustness of the results, kept the variance of the overall error term fixed and move the (square root of) the ratio of variances, $\sqrt{\frac{Var(u)}{Var(v)}}$. The second setting allowed us to compare distributions of inefficiency which are equal in terms of variance but different as regards the 'shape' of the distribution. Although the main result of the paper is also stable in the second setting, and the three estimated models show the same rank correlation coefficients for each of the true inefficiency distribution, the results also reveal that the six distributions examined are qualitatively different: the shape of the true inefficiency distribution seems —at least for low values of λ— to influence the performance of each of the three estimated models in reproducing inefficiency score ranking. Nested distributions also seem to yield similar results.

A further development of this study could be to examine misspecification of noise term, v_i, on the correct estimation of the inefficiency scores: this is a type of misspecification which has been almost completely neglected in previous works and which

[32]This is also in line with suggestions provided from other scholars in the field such as Ritter and Simar (1997) and Koop (2001).

was briefly considered only by Jensen (2005). It would also be interesting to explore the consequences of the correct estimation of the inefficiency scores of a neglected correlation between the two random terms, u_i and v_i, which are always assumed to be uncorrelated.

2.7 Data Appendix

2.7.1 Variance of error components

This Section provides an example to clarify the relationship between parameter σ_u and the standard deviation in each of the six true inefficiency distributions examined in this simulation study. On one hand, for half-normal and truncated normal distributions, parameter σ_u is the standard deviation of the pre-truncated distribution; in the case of log-normal distribution, σ_u is the standard deviation of the natural logarithm of the variable. On the other hand, for exponential and Weibull distributions, σ_u is the scale parameter, and for uniform distribution it is equal to its standard deviation. For each of the six distributions, the variance may be written as a function of σ_u. For half-normal distribution:

$$Var(u) = \sigma_u^2 \left[\frac{\pi - 2}{\pi} \right]; \tag{2.15}$$

for exponential distribution:

$$Var(u) = \sigma_u^2; \tag{2.16}$$

in the truncated normal case, with the lower truncation point at $a = 0$:

$$Var(u) = \sigma_u^2 \left[1 + \frac{\frac{-\mu}{\sigma_u}\phi\left(\frac{-\mu}{\sigma_u}\right) - \frac{b-\mu}{\sigma_u}\phi\left(\frac{b-\mu}{\sigma_u}\right)}{\Phi\left(\frac{b-\mu}{\sigma_u}\right) - \Phi\left(\frac{-\mu}{\sigma_u}\right)} - \frac{\phi\left(\frac{-\mu}{\sigma_u}\right) - \phi\left(\frac{b-\mu}{\sigma_u}\right)}{\Phi\left(\frac{b-\mu}{\sigma_u}\right) - \Phi\left(\frac{-\mu}{\sigma_u}\right)} \right]; \tag{2.17}$$

for uniform distribution:

$$Var(u) = \sigma_u^2; \tag{2.18}$$

in the log-normal case:

$$Var(u) = \left(e^{\sigma_u^2} - 1\right) \cdot e^{2\mu + \sigma_u^2}, \tag{2.19}$$

where μ is the mean of the natural logarithm of the variable;

for the Weibull distribution,

$$Var(u) = \sigma_u^2 \left[\cdot\Gamma(1 + \frac{2}{k}) - \Gamma^2(1 + \frac{1}{k}) \right], \tag{2.20}$$

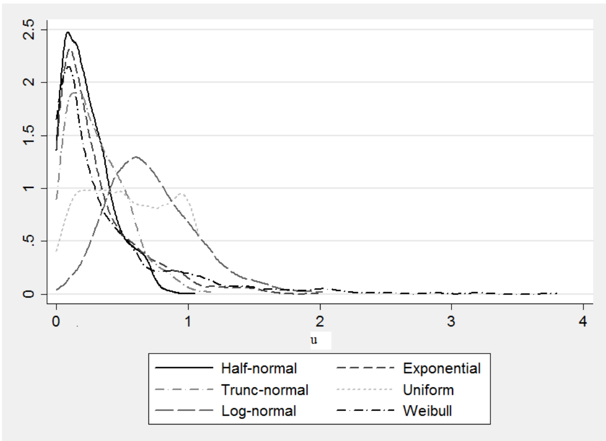
where k is the shape parameter of the Weibull distribution.

Looking at Table 2.10 and Figure 2.5, we see that the six true distributions — all generated with the same value of σ_u =0.316— have different standard deviations and variance values. The Weibull distribution is the one with the largest standard deviation, followed by log-normal, uniform and exponential distributions. For the same value of σ_u, and thus λ, the half-normal and truncated normal distributions have the lowest standard deviations (and variances).

Table 2.10: Standard deviations for true inefficiency distributions; $\sigma_u = 0.316$

Inefficiency distribution	St. deviation
Half-normal	0.183
Exponential	0.312
Truncated normal	0.222
Uniform	0.315
Log-normal	0.334
Weibull	0.481

Figure 2.5: Kernel densities of six true inefficiency distributions for $\sigma_u = 0.316$



This simple example clarifies why the first Setting cannot compare inefficiency distributions which differ only in 'shape'. More comparable samples require distributions with equal variance, and that is what we did in the second setting.

2.7.2 Some algebra on results relating to rank correlation coefficients

This Section provides some algebra to explain why the three estimated models reach the same rank correlation coefficient whatever the true inefficiency distribution is (see Table 2.6). For each of the true inefficiency distribution (DGP) l, each estimated model k and each value of λ the vector of the residuals is:

$$e_{(l,k)} = y_{(l)} - \boldsymbol{\hat{\beta}'_{(l,k)}}\mathbf{x_i}^{33}. \tag{2.21}$$

In the first row of Table 2.6 in which l=(half-normal) and k=(NHN,NEX,NTN), the residuals may be computed respectively as:

$$e_{(hn,NHN)} = y_{(hn)} - \widehat{\beta'}_{(hn,NHN)}\mathbf{x_i}; \tag{2.22}$$

$$e_{(hn,NEX)} = y_{(hn)} - \widehat{\beta'}_{(hn,NEX)}\mathbf{x_i}; \tag{2.23}$$

$$e_{(hn,NTN)} = y_{(hn)} - \widehat{\beta'}_{(hn,NTN)}\mathbf{x_i}. \tag{2.24}$$

Now, from Table 2.3 it is clear that $\hat{\beta}_{1(hn,k)} = \hat{\beta}_{1(hn)}$ and $\hat{\beta}_{2(hn,k)} = \hat{\beta}_{2(hn)}$ $\forall k$, i.e., for all three estimated models the estimates of the technology parameters are asymptotically equal. The only difference lies in the estimate of the intercept, which is unbiasedly estimated only in the correctly specified model (NHN). Thus, Equations 2.22, 2.23 and 2.24 can be rewritten as

$$e_{(hn,NHN)} = y_{(hn)} - (\hat{\beta}_{0(hn,NHN)} + c); \tag{2.25}$$

$$e_{(hn,NEX)} = y_{(hn)} - (\hat{\beta}_{0(hn,NEX)} + c); \tag{2.26}$$

$$e_{(hn,NTN)} = y_{(hn)} - (\hat{\beta}_{0(hn,NTN)} + c); \tag{2.27}$$

where $c = \hat{\beta}_{1(hn)}x_1 + \hat{\beta}_{2(hn)}x_2$ for all estimated models. In other words, the three vectors of residuals show a linear factor c, which shifts their distribution in the same direction and by the same 'amount'. In addition, $y_{(hn)} = \beta_0+\beta_1x_1+\beta_2x_2+v_{(hn)}-u_{(hn)}$, and $\beta_1x_1+\beta_2x_2$ is a constant across the estimated models, we can replace endogenous variable $y_{(hn)}$ with $\beta_0 + b + v_{(hn)} - u_{(hn)}$ in all the above equations, thus obtaining:

$$e_{(hn,NHN)} = \left(\beta_0 - \hat{\beta}_{0(hn,NHN)}\right) + (b-c) + v_{(hn)} - u_{(hn)}; \tag{2.28}$$

$$e_{(hn,NEX)} = \left(\beta_0 - \hat{\beta}_{0(hn,NEX)}\right) + (b-c) + v_{(hn)} - u_{(hn)}; \tag{2.29}$$

$$e_{(hn,NTN)} = \left(\beta_0 - \hat{\beta}_{0(hn,NTN)}\right) + (b-c) + v_{(hn)} - u_{(hn)}. \tag{2.30}$$

[33] Remember that $\mathbf{x_i}$ is fixed $\forall l, k$ and that $y_{(l,k)} = y_{(l)}$ $\forall k$, i.e. the data generating process is 'fixed' along a given row of Table 2.3.

As $(b-c)$ goes to zero asymptotically, the only difference between the vectors of residuals in the three estimated models is given by the difference between the true value of the intercept (.7 in our case) and its estimate, which is a linear factor shifting the whole distribution of e_i. Now, in the case of the correctly specified model

$$\beta_0 \cong \hat{\beta}_{0(hn,NHN)} \rightarrow e_{(hn,NHN)} \cong v_{(hn)} - u_{(hn)},$$

whereas in the misspecified models k =NEX,NTN there is a shift-factor, given by amount $\left(\beta_0 - \hat{\beta}_{0(hn,k)}\right)$, which may be positive or negative.

2.7.3 Technology and variance parameters

Estimates of technology parameters —in Setting 2— are listed in Table 2.11: the results are very similar to those observed for the corresponding parameters in Setting 1. This holds good both for the observed patterns in the bias of the intercept in the misspecified cases, and for the unbiasedness of output elasticity estimates, $\widehat{\beta}_1$ and $\widehat{\beta}_2$. In line with Setting 1, the NTN model is better able to estimate the intercept of the frontier in the misspecified cases than the NHN and NEX models. As in Setting 1, the

Table 2.11: Technology parameter estimates - Setting 2

			Estimated models								
			NHN			NEX			NTN		
DGP	λ^*	λ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
Half-normal	0.5	0.833	0.658	0.400	0.600	0.499	0.400	0.600	0.935	0.400	0.600
	1	1.666	0.700	0.400	0.600	0.447	0.400	0.600	0.747	0.400	0.600
	5	8.333	0.701	0.400	0.600	0.537	0.400	0.600	0.708	0.400	0.600
	10	16.666	0.700	0.400	0.600	0.596	0.400	0.600	0.701	0.400	0.600
	20	33.437	0.701	0.400	0.600	0.635	0.400	0.600	0.701	0.400	0.600
Exponential	0.5	0.500	0.876	0.400	0.600	0.685	0.400	0.600	0.878	0.400	0.600
	1	1.000	0.928	0.400	0.600	0.696	0.400	0.600	0.730	0.400	0.600
	5	5.000	0.799	0.400	0.600	0.698	0.400	0.600	0.702	0.400	0.600
	10	10.000	0.754	0.400	0.600	0.698	0.400	0.600	0.702	0.400	0.600
	20	20.000	0.728	0.400	0.600	0.698	0.400	0.600	0.700	0.400	0.600
Truncated normal	0.5	0.569	0.370	0.400	0.600	0.261	0.400	0.600	0.718	0.400	0.600
	1	1.140	0.237	0.400	0.600	0.027	0.400	0.600	0.738	0.400	0.600
	5	5.700	0.165	0.400	0.600	-0.175	0.400	0.600	0.711	0.400	0.600
	10	11.380	0.172	0.400	0.600	-0.181	0.400	0.600	0.709	0.400	0.600
	20	23.140	0.169	0.400	0.600	-0.191	0.400	0.600	0.699	0.400	0.600
True value			0.700	0.400	0.600	0.700	0.400	0.600	0.700	0.400	0.600

three correctly specified models (main diagonal blocks) estimate quite well parameter λ and the NTN model presents the highest number of extreme values[34] with respect to the other two models. The patterns of the estimates of parameter σ^2 are also in line with Setting 1.

[34] Taking the interquartile range of the distribution of $\widehat{\lambda}$, $IQR = Q3 - Q1$, we define a value $\widetilde{\lambda}$ as 'extreme' if $\widetilde{\lambda} < Q1 - 3IQR$ or $\widetilde{\lambda} > Q3 + 3IQR$.

Table 2.12: λ parameter estimates - Setting 2

			Estimated models					
			NHN		NEX		NTN	
DGP	λ^*	λ	$\hat{\lambda}$	extreme v.	$\hat{\lambda}$	extreme v.	$\hat{\lambda}$	extreme v.
Half-normal	0.5	0.833	0.764	0	0.329	0	1.785	73
	1	1.666	1.679	0	0.712	0	1.892	52
	5	8.333	8.721	3	3.316	0	8.847	32
	10	16.666	18.236	8	6.373	0	18.426	16
	20	33.437	40.568	68	12.823	0	40.897	62
Exponential	0.5	0.500	1.073	0	0.487	0	4.351	17
	1	1.000	2.183	0	1.009	0	11.700	0
	5	5.000	11.203	24	5.151	0	75.789	2
	10	10.000	23.565	61	10.475	0	92.588	4
	20	20.000	53.985	196	21.939	11	171.250	21
Truncated normal	0.5	0.569	0.506	0	0.221	0	1.467	74
	1	1.140	0.983	0	0.393	0	1.644	120
	5	5.700	2.192	0	0.712	0	6.596	67
	10	11.380	2.382	0	0.741	0	13.935	217
	20	23.140	2.445	0	0.750	0	24.068	450

Table 2.13: σ^2 parameter estimates - Setting 2

				Estimated models					
				NHN		NEX		NTN	
DGP	λ^*	$Var(\epsilon)$	σ^2	$\widehat{\sigma^2}$	extreme v.	$\widehat{\sigma^2}$	extreme v.	$\widehat{\sigma^2}$	extreme v.
Half-normal	0.5	0.500	0.677	0.667	0	0.500	0	0.767	166
	1	0.500	0.944	0.942	0	0.506	0	1.018	81
	5	0.500	1.355	1.354	0	0.633	0	1.356	2
	10	0.500	1.380	1.378	0	0.712	0	1.394	2
	20	0.500	1.387	1.387	0	0.772	0	1.398	1
Exponential	0.5	0.500	0.500	0.755	0	0.497	0	9.940	41
	1	0.500	0.500	1.032	0	0.498	0	51.637	0
	5	0.500	0.500	1.128	0	0.499	0	154.023	0
	10	0.500	0.500	1.073	0	0.499	0	41.851	24
	20	0.500	0.500	1.038	0	0.498	0	30.833	15
Truncated normal	0.5	0.500	0.529	0.596	0	0.497	0	0.583	159
	1	0.500	0.575	0.730	0	0.500	0	0.612	115
	5	0.500	0.643	1.065	0	0.507	0	0.646	0
	10	0.500	0.645	1.097	0	0.506	0	0.647	0
	20	0.500	0.657	1.127	0	0.515	0	0.662	0

2.7.4 Kernel densities of technology and variance parameters

This Section describes the kernel densities of the output elasticity estimates for different values of λ in Setting 1: the higher λ, the more concentrated the kernels around the true values. This is reasonable, given that, as λ increases the variance of the overall error term, $Var(\epsilon)$ decreases (given that σ^2 is taken to be constant). This makes the empirical variance of the output elasticity estimates decrease, as Olson, Schmidt, and Waldman (1980) have suggested in their work (see p.71). This result is also in line with the fact that the more asymmetric the distribution becomes, the better the ML estimator of the frontier —which specifically takes the asymmetry of the distribution of the disturbance into account— performs.

Figure 2.6: Kernel density estimates of $\widehat{\beta}_1$: correctly specified models - Setting 1

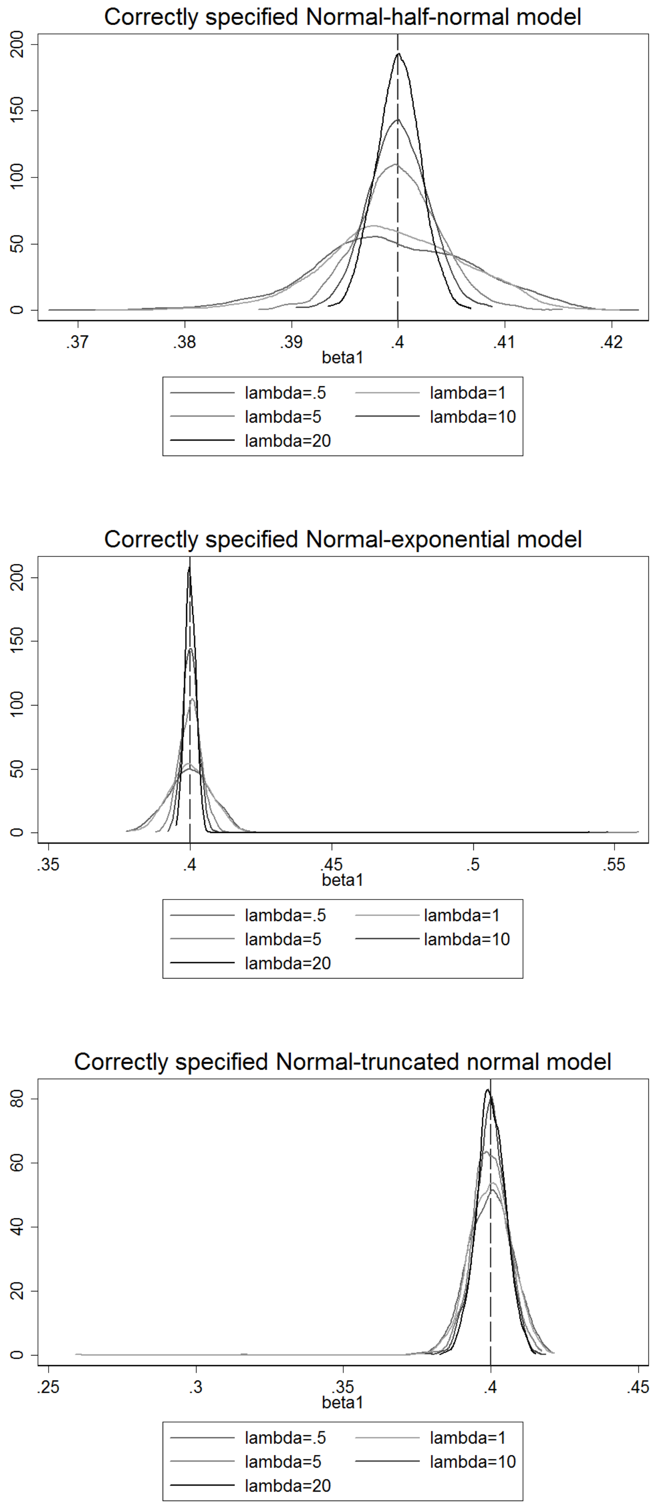
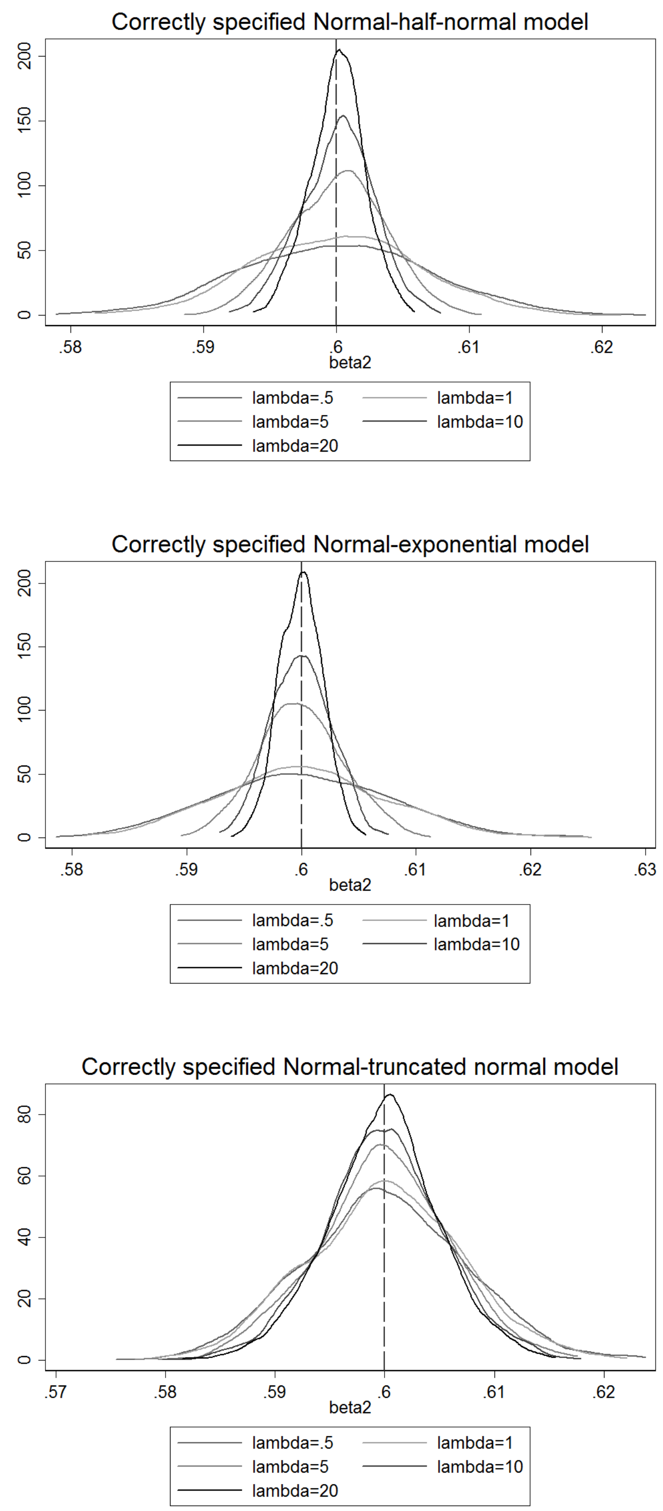


Figure 2.7: Kernel density estimates of $\widehat{\beta}_2$: correctly specified models - Setting 1



Chapter 3

Vertical Integration and Efficiency: an Application to the Italian Machine Tool Industry[1]

3.1 Introduction

Empirical studies on productivity and efficiency at the micro level have found large heterogeneity across firms or plants even within narrowly defined industries (see Bartelsman and Doms, 2000; Bartelsman, Scarpetta, and Schivardi, 2005; Fried, Lovell, and Schmidt, 2008; Dosi, Grazzi, Tomasi, and Zeli, 2010, among others). Differences in performance among production units have been mainly attributed to variations in management skills, human capital, R&D and technological capital, product innovation, firm's international exposure (exports and foreign direct investments), together with factors which are external to the firm, like technological spillovers, the intra-industry degree of competition and the regulatory environment (see Syverson, 2010, for an extensive review on the topic).

The control of vertical links of production, i.e. the decision about which phases of production to keep inside to the firm (vertical integration) and which ones to leave to the 'outside', is another factor which is related to the firm productive performance: vertically integrated structures can be either justified by the search for an optimal provision of specific physical inputs in a production process[2], or by a better supervision over each phase of production (which stands for a better use of management as in Hortaçsu and Syverson, 2009); moreover, a backward integration may allow to avoid a double marginalization in the market for inputs or may be a channel to rise up the costs of competitors, buying the main part of essential input of a backward

[1]This Chapter draws on a joint work with Enrico Zaninotto (University of Trento).

[2]This motivation has been mainly studied in the transaction costs and property rights literature (Williamson, 1971; Grossman and Hart, 1986). See Lafontaine and Slade (2007) for an up-to-date survey on this field of analysis.

market[3]. Different degrees of vertical integration are observable in all kinds of industries and across different countries and, in the last decades, a tendency toward a disintegration of the production processes has been extensively documented by researchers (see Feenstra, 1998; Grossman and Helpman, 2005, among others) and the popular press. This phenomenon has generically fallen under the name 'outsourcing', and it has been justified by different motivations ranging from the need for focusing on 'core competences' to the raise of information technologies, which have lowered transaction costs typical of fragmented organizations (Hitt, 1999; Baldwin, 2006).

Given the relevance of the phenomenon, the relationship between the vertical organization of production and productive efficiency has generated an amount of empirical research in the last years, but results are still not unambiguous. The wide collection of cases presented by Berger (2006) illustrates vividly how firms can follow different outsourcing strategies while getting similar profitability. Heshmati (2003) offers a wide survey of studies on the relationship between outsourcing and productive efficiency, with particular reference to service outsourcing, from which not clear-cut patterns emerge. A similar result of wide heterogeneity of outsourcing choices, and not clear patterns of its effects on productivity emerges from the more recent survey proposed by Olsen (2006). More recently, some slight evidence in favor of a negative impact of disintegration on productivity have been proposed, as in the study on German manufacturing firms by Broedner, Kinkel, and Lay (2009), or by Federico (2010) whose study of Italian manufacturing firms finds evidence of a productivity ordering where vertical integration is chosen by the most productive firms while outsourcing is chosen by the least productive firms.

In this paper we study the relationship between firm efficiency and vertical integration in a representative sample of Italian machine tool (MT) builders. Given the debated relationship, in order to come up with an empirical testable hypothesis we have set-up a theoretical model (largely inspired by Antras and Helpman, 2004; Syverson, 2004) of entry and competition within an industry in which firms can choose the vertical organization of production, i.e. to be vertically integrated or not. The main prediction of the models is that the most efficient firms self-select in being vertically integrated while less efficient firms prefer a disintegrated structure and they both coexist in the market in equilibrium. The coexistence of different organizational choices is made possible because firms trade off organizational fixed costs, which are higher in a vertical integrated structure, with marginal costs of production which are higher in a disintegrated structure.

In the second part of the paper, drawing on this result, we have empirically tested the relationship between efficiency and vertical integration. The Italian MT industry seems a natural candidate for this exercise: in fact, this industry is characterized

[3]Of course, these are just few motivations supporting the vertical integration choice: see Perry (1989) for an extensive discussion on this issue.

by the coexistence of different types of organizational forms (see Rolfo, 1998) and large heterogeneity in productive efficiency. A stochastic frontier framework has been adopted in order to estimate the relationship between firm efficiency and the level of vertical integration. Using an novel panel dataset including around 500 MT builders, our empirical findings show that vertically integrated firms delineate the frontier technology, thus confirming the theoretical prediction.

Overall, this work's main contributions regard a better understanding of the functioning of those industries —as the MT industry— which are characterized by differences in the productive performance among firms and wide heterogeneity in organizational choices, and this has been done both setting up a proper theoretical framework and detailed empirical analysis. From a methodological point of view, the use of a stochastic frontier framework allows us to jointly estimate the parameters of the production function, the level of efficiency and the correlation between firm efficiency and the degree of vertical integration: this can be considered as an improvement to previous studies on the topic, in which productive efficiency scores (total factor productivity) have been usually regressed on the covariate in a second step of the econometric analysis, raising several econometric problems related to the 2-step estimation[4].

The paper is structured as follows: in Section 3.2 we give a general overview of the industry under analysis; Section 3.3 illustrates the theoretical model from which the main hypothesis is derived; Section 3.4 presents the choices adopted for the empirical evaluation; Section 3.5 presents the data and Section 3.6 shows the results of our empirical analysis. Section 3.7 discusses some issues and suggests steps for further research.

3.2 Industry overview

The MT industry gathers all the producers of metal working machines (and components), which are capital goods that are used for manufacturing final goods in other industries. The main user of machine tools is the broader mechanical engineering industry (which uses around 40% of the produced machines); the automotive industry and models and dies industry are two other important clients. The three main productions of the MT industry are (i) the *forming machines* (such as presses, sheet metal deformation machines, shearing machines), the (ii) *cutting machines* (such as machining centers, turning machines-lathes, grinding machines) and the (iii) *non conventional machines* (such as machines for marking and cutting with laser); other types of machines are marginal and can be grouped in a residual class of *other machines* (which comprehends mechanical arms, measuring-control machines, and heat treatment machines). As Rolfo (1998) underlines the industry is characterized by

[4]See Hortaçsu and Syverson (2009) and Federico (2010) among others.

a low rate of product diversification, where the vast majority of firms have not expanded their traditional production to other types of machines as time has passed:: instead, they have focused on shaping the machine characteristics to the consumer needs. Almost all types of products are characterized by the existence of niches in which the ability to solve customers' specific problem is fundamental. The role of customization has especially been important for small enterprises, which have developed a particular ability in interpreting and matching the customer demand (Wengel and Shapira, 2004). The industry is also characterized by relatively low barriers to entry, because new firms can be set up with relatively small capital and little technological know-how.

Taking an aggregate perspective, the MT industry is very representative of Italian competitiveness in the broader mechanical engineering sector (Rolfo and Calabrese, 2006): in 2007, Italy was in the third place for export value and fourth for value of production, making it one of the world leaders for production of MT[5]. Table 3.1 provides an overview of the value of production trends since 1998, and Table 3.2 provides country rankings for exports value: after Japan, Germany and (more recently) China, Italy is among the leaders.

Table 3.1: Value of production by country - trend

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Japan	8018	7074	9564	8470	5712	6189	7504	9382	9634	9406
Germany	6822	7167	7559	8640	7427	6818	7206	7876	8075	9282
China	1690	1747	2445	2928	2487	2635	3280	4100	5653	7360
Italy	3258	3519	4163	4240	4007	3678	3735	3912	4554	5330
South Korea	436	808	1851	1521	1653	1792	1985	2320	3300	3319
Taiwan	1419	1432	2056	1825	1879	1874	2321	2737	3058	3193
U.S.	4216	3980	4534	3670	2570	2129	2554	2788	2937	2610
Switzerland	1753	1905	1965	2319	1930	1664	1878	2120	2363	2543
Spain	844	910	929	990	915	820	822	904	979	1048
France	703	363	517	500	405	418	574	692	762	845

Source: Ucimu,*Industry Report*, 2007; Millions of euro

Table 3.2: Exports value by country - ranking

	2007
Germany	6686
Japan	6501
Italy	2968
Taiwan	2485
Switzerland	2215
South Korea	1312
U.S.	1210
China	1167
United Kingdom	672

Source: Ucimu,*Industry Report*, 2007; Millions of euro

[5]For a detailed report on the evolution of the industry in terms of value of production, exports and imports see the industry reports by Ucimu (2007a,b).

The Italian MT industry is characterized by the coexistence of a small group of large firms, which are able to compete both in domestic and in foreign markets, and a large tier of smaller firms, ranging from highly specialized machine (or components) makers to firms that provide buffer capacity and help larger firms to level out their plant utilization (see Rolfo, 1998). According to a survey conducted by Ucimu (the Italian Machine Tools, Robots and Automation Manufacturers Association) in 2006, 71% of MT manufacturers invoiced less than €12.5 millions, and 75.8% had less than 100 employees. On the other hand, firms with more than 100 employees produced 67.8% of the overall value of production and accounted for 69.7% of the overall exports value. Moreover, turnover per employee ranged from €127,000 for smaller firms, to €143,300 for larger companies. Le largest percentage of MT facilities is in the North of Italy, also because the majority of clients is located there: Lombardy (the region of Milan) accounts for 46% of the production units. The explanation for the existence of a large bunch of small firms has to be searched, among the other things, in the Italian regulatory environment, which has made easier for small firms to reduce employment and report fiscal accounts, thus conferring to these firms an innate flexibility advantage, which however has decreased with the raise of international competition and the introduction of several technological innovations (as flexible automation) that have counterbalanced the advantage of smaller firms. Despite the high fragmentation among smaller and larger firms and their geographical agglomeration in just few regions, the structure of Italy's MT industry has experienced a transformation from the typical 'industrial district' to networks of firms where the physical proximity is not essential anymore and the leader of the network is the main actor (both in terms of exchange of resources and in developing new technologies, as documented by Wengel and Shapira, 2004).

The vertical structure of the Italian MT industry took different configurations since the 1950s (see Rolfo, 1998, 2000). At that time the most important mechanical engineering firms produced their own MT in-house (from foundry to finished products) thus the prevailing model was that of vertically integrated firms. The 1960s saw, a significant increase in internal demand which stimulated the growth of an independent MT industry and the 1970s were characterized by the 'small firm model', and a consequent vertical dis-integration of firms: electronic and computer components tended to be outsourced. Although there have been slight changes over time, this low level of vertical integration has tended to dominate for the majority of Italian MT firms[6]. Presently, MT builders basically 'leave to the outside' the manufacture of some components (as electronics), but there is not a clear path between firm size and vertical integration strategy as it has been documented by Wengel and Shapira (2004) in a small but significant sample of around 200 firms: on the one hand small

[6]Italian manufacturing firms have traditionally showed lower levels of vertical integration than their counterparts in other European countries e.g. Germany and the UK (see Arrighetti, 1999).

firms show an higher frequency of in-house mechanical components production, while on the other hand larger firms are more oriented to keep in-house the electronic assembling and the software planning. Overall, almost all firms undertake designing, mechanical assembling and testing in-house, which appear as the core competences. Again, this general evidence confirms the tendency of the Italian machine tool firms in producing customer-specific interfaces.

The vertical position of the firm along the production chain, therefore, is a key dimension in this industry, which has consequences both for firms' productive efficiency, and also for the control of the knowledge and innovation processes (Poledrini, 2008).

3.3 Theory: firm efficiency and vertical integration

The model which follows is mainly inspired by the works of Antras and Helpman (2004) and Syverson (2004).

Preferences and demand The industry under analysis is modeled as a continuum of final good producers of measure N. Each producer makes a distinct variety (indexed by i) of the industry's products-machines. Following Melitz and Ottaviano (2008), the representative consumer has preferences over these varieties given by the following quadratic utility function:

$$U = q_0 + \alpha \int_i q_i di - \frac{1}{2}\gamma \int_i q_i^2 di - \frac{1}{2}\eta \left(\int_i q_i di \right)^2, \tag{3.1}$$

where q_0 is the quantity of a numeraire good, q_i is the quantity of good i consumed and $Q = \int_i q_i di$ is the total consumption over all varieties. α and η are the indicators of the substitution patterns between the differentiated varieties and the numeraire, while γ index the product differentiation between the varieties. If $\gamma = 0$ only the consumption level over all the varieties matter, because varieties are perfect substitutes.

The inverse demand function for each variety is thus:

$$p_i = \alpha - \gamma q_i - \eta Q. \tag{3.2}$$

Equation 3.2 can be inverted in order to get the linear market demand system for these varieties:

$$q_i = \frac{\alpha}{\eta N + \gamma} - \frac{1}{\gamma} p_i + \frac{\eta N}{\eta N + \gamma} \frac{1}{\gamma} \bar{p}, \tag{3.3}$$

where N is the measure of producers, p_i is the price of good i and $\bar{p}$ is the average price among industry producers. The price bound, p_i^{max}, at which the demand for variety i goes to zero, can be obtained as:

$$p_i^{max} = \frac{\gamma\alpha + \eta N}{(\eta N + \gamma)}.$$

The price bound results to be an increasing function of the γ parameter (a higher product differentiation leads to an higher upper bound in terms of feasible price for variety i), a decreasing function of the measure of consumed varieties N, and an increasing function of the average price of the varieties $\bar{p}$.

Production and firm behaviour Each variety of machines needs two inputs to be produced. Capital, K_i , which is available to the machine-tool maker internally and which has a unit cost equal to w_K and an intermediate input, M_i, which can be either produced by the machine tool maker or acquired from the outside. In the first case, the intermediate input has a unit cost equal to w_{Mv} (where v stands for for vertical integration) and the producer is vertically integrated, while in the second case, the price of the intermediate input is equal to w_{Mo} (where o stands for firms engaging in the outsourcing strategy or simply disintegrated firms) and the producer is disintegrated.

- **Assumption 1**: $w_{Mv} < w_{Mo}$

This assumption does not seem to be restrictive, given that the internally produced input is evaluated at its marginal cost, while if it is acquired in the market and this is not perfectly competitive, that can bring to a price which is higher than the marginal cost (due to double marginalization). Moreover, this is a pretty realistic assumption for the Italian MT industry: in fact, due to the highly differentiated nature of final products, the market of components is in turn differentiated.

On the other hand, a vertically integrated firms face higher organizational fixed costs:

- **Assumption 2**: $f_v > f_o$

This assumption, which relates to the additional managerial tasks which are needed in order to supervise the production of the intermediate input is in line with the theoretical literature on productivity heterogeneity and different organizational forms (Antras and Helpman, 2004; Grossman and Helpman, 2005). Moreover, given the complexity of some phases of the production of a machine tool, as it has been explained in Section 3.2, it is reasonable to think that an expansion along the vertical production chain would imply higher organizational costs.

Production of each variety i is modeled as a Cobb-Douglas function, which is characterized by constant return to scale (CRS), for purpose of simplicity[7].

$$q_i = \left(K_i^{\beta} M_i^{1-\beta}\right) e^{-U}, \tag{3.4}$$

where $0 < \beta < 1$, and U is a firm-specific random term which is extracted from a known nonnegative distribution ($G(U)$, $U > 0$). U reflects the firm-specific level of technical inefficiency, i.e. a factor which shifts the firm away from the technology frontier (production function). In this framework, the production function or technological frontier is reached by the most efficient firms only, i.e. those with $U = 0$, while all the other firms are below it. We derive the total and the marginal cost function of the firm producing q_i, given the vector of input prices. In equilibrium, the optimal level of inputs solves the following system of equations:

$$\begin{cases} q_i & = \left(K_i^{\beta} M_i^{1-\beta}\right) e^{-U} \\ \frac{MP_M}{MP_K} & = \frac{w_{Ml}}{w_K}, \end{cases}$$

where $l = \{v, o\}$. We can compute the marginal productivity of input M as

$$MP_M = \frac{\partial q_i}{\partial M_i} = \left[K_i^{\beta}\,(1-\beta)\,M_i^{(1-\beta)-1}\right] e^{-U},$$

and the marginal productivity of input K as

$$MP_H = \frac{\partial q_i}{\partial K_i} = \left[\beta K_i^{\beta-1} M_i^{(1-\beta)}\right] e^{-U}.$$

Thus, the marginal rate of technical substitution is

$$MRTS_{K,M} = \frac{MP_M}{MP_K} = \left(\frac{1-\beta}{\beta}\right)\frac{K_i}{M_i}.$$

The second equation of the system 3.3 can be re-arranged in order to obtain

$$K_i = \left(\frac{w_{Ml}}{w_K}\right) M_i \left(\frac{\beta}{1-\beta}\right), \tag{3.5}$$

[7] The main result of the theoretical analysis do not change if there are more than one inputs available internally to the firm, or the technology is characterized by non-constant returns to scale.

which can be substituted in the production function, in order to obtain the conditional demand (optimal quantity) of input M_i^* [8]:

$$M_i^* = q_i \left(e^U\right) \left(\frac{w_K}{w_{Ml}}\right)^{\beta} \left(\frac{1-\beta}{\beta}\right)^{\beta}. \tag{3.6}$$

Now, we can substitute the conditional demand of M_i^* into Equation 3.5 in order to obtain the conditional demand of input K_i^*[9]:

$$K_i^* = q_i \left(e^U\right) \left(\frac{w_K}{w_{Ml}}\right)^{\beta-1} \left(\frac{1-\beta}{\beta}\right)^{\beta-1}. \tag{3.7}$$

The total cost function, TC_i, can be written as:

$$TC_{il} = q_i \left(e^U\right) \left(\frac{w_K}{w_{Ml}}\right)^{\beta} \left(\frac{1-\beta}{\beta}\right)^{\beta} \cdot w_{Ml} + q_i \left(e^U\right) \left(\frac{w_K}{w_{Ml}}\right)^{\beta-1} \left(\frac{1-\beta}{\beta}\right)^{\beta-1} \cdot w_K. \tag{3.8}$$

The marginal cost function can be easily derived, as:

$$\frac{\partial TC_{il}}{\partial q_i} = c_{il} = \left(e^U\right) \left(\frac{w_K}{w_{Ml}}\right)^{\beta} \left(\frac{1-\beta}{\beta}\right)^{\beta} \cdot w_{Ml} + \left(e^U\right) \left(\frac{w_K}{w_{Ml}}\right)^{\beta-1} \left(\frac{1-\beta}{\beta}\right)^{\beta-1} \cdot w_K. \tag{3.9}$$

The marginal cost is idiosyncratic to each MT producer, and it is a function of the technical inefficiency term and the relative price. In particular from Equation 3.9 it follows that, *ceteris paribus*:

- $\partial c_{il}/\partial U > 0$, firms which present higher level of inefficiency show higher marginal costs;

Holding on Equation 3.3, the profit function of the producer of ith variety can be written as:

$$\pi_{il} = \left(\frac{\alpha}{\eta N + \gamma} - \frac{1}{\gamma} p_i + \frac{\eta N}{\eta N + \gamma} \frac{1}{\gamma} \bar{p}\right) \cdot (p_i - c_{il}) - f_l, \tag{3.10}$$

where f_l are the organizational fixed costs, which are different between vertical integrated and disintegrated firms.

Equilibrium The MT industry is modeled as a Bertrand-Nash model with differentiated products (see Mas-Colell, Whinston, and Green, 1995, p.395-400): this seems reasonable, given the industry characteristics which have been introduced in Section 3.2. Each producer sells its product on the market at the price which maximizes

[8] As can been easily verified $\partial M_i^*/\partial U > 0$, i.e. an increase in the use of the input is positively related to an increase in technical inefficiency , given the level of q_i; moreover, $\partial M_i^*/\partial w_K > 0$ and $\partial M_i^*/\partial w_{Ml} < 0$ indicate the substitution between inputs.

[9] The same considerations on technical inefficiency and the relative price apply to this input too.

its profits (see Syverson, 2004, p.537). The optimal price can be found solving the following condition:

$$\frac{\partial \pi_{il}}{\partial p_i} = -\frac{1}{\gamma}\left(p_i - c_{il}\right) + \left(\frac{\alpha}{\eta N + \gamma} - \frac{1}{\gamma}p_i + \frac{\eta N}{\eta N + \gamma}\frac{1}{\gamma}\bar{p}\right) = 0. \qquad (3.11)$$

Solving for p_i, we get:

$$p_i^* = \frac{\alpha\gamma}{2\left(\eta N + \gamma\right)} + \frac{\eta N}{2\left(\eta N + \gamma\right)}\bar{p} + \frac{c_i l}{2}, \qquad (3.12)$$

which can be substituted into Equation 3.3, in order to obtain the quantity sold by the producer of variety i at the optimal price:

$$q_i^* = \frac{\alpha}{2\left(\eta N + \gamma\right)} + \frac{\eta N}{2\gamma\left(\eta N + \gamma\right)}\bar{p} - \frac{c_{il}}{2\gamma}. \qquad (3.13)$$

The maximized profits formula can thus be written using Equations 3.12 and 3.13:

$$\pi_{il}^* = q_i^* \cdot \left(p_i^* - c_{il}\right) - f_l = \left(\frac{\alpha}{2\left(\eta N + \gamma\right)} + \frac{\eta N}{2\gamma\left(\eta N + \gamma\right)}\bar{p} - \frac{c_{il}}{2\gamma}\right) \cdot$$
$$\left(\frac{\alpha\gamma}{2\eta N + \gamma} + \frac{\eta N}{2\eta N + \gamma}\bar{p} + \frac{c_i l}{2} - c_{il}\right) - f_l \qquad (3.14)$$

$$\pi_{il}^* = \frac{1}{4\gamma}\left(\frac{\alpha\gamma}{\eta N + \gamma} + \frac{\eta N}{\eta N + \gamma}\bar{p} - c_{il}\right)^2 - f_l. \qquad (3.15)$$

A sunk cost needs to be paid before entering in the market, f_E[10]. After doing that, the producer can observe its actual inefficiency level, U, which determines a firm-specific marginal cost; thus, firms choose either to start the production, earning the corresponding profits or to exit the market. In the first case they can also face the decision on how to organize the production, i.e. to be vertically integrated or not. In the other case, the marginal cost results to be above a given threshold and that is due to an inefficiency shock above a given upper bound. In order to assess the existence of firms with different levels in inefficiency and different organizational form in equilibrium, we need to study the maximized profit function in relationship with the inefficiency term U.

It is possible to set $k^* = \frac{1}{4\gamma}\frac{\alpha\gamma}{\eta N+\gamma} + \frac{\eta N}{\eta N+\gamma}\bar{p}$, and substituting Equation 3.9 into Equation 3.15 we get:

$$\pi_{il}^* = \frac{1}{4\gamma}\left[k^* - \left(\left(e^U\right)\left(\frac{w_K}{w_{Ml}}\right)^{\beta}\left(\frac{1-\beta}{\beta}\right)^{\beta} \cdot w_{Ml} + \left(e^U\right)\left(\frac{w_K}{w_{Ml}}\right)^{\beta-1}\left(\frac{1-\beta}{\beta}\right)^{\beta-1} \cdot w_K\right)\right]^2 - f_l \qquad (3.16)$$

First, it is possible to verify that the maximized profit function is decreasing in U;

[10] Which of course do not appear in Equation 3.15 of the operating profits.

in other words, higher levels of inefficiency imply lower profits*ceteris paribus*:

$$\frac{\partial \pi_{il}^*}{\partial U} = \frac{1}{4\gamma} \cdot (2) \cdot \left(k^* - \left(\left(e^U\right)\left(\frac{w_K}{w_{Ml}}\right)^{\beta}\left(\frac{1-\beta}{\beta}\right)^{\beta} \cdot w_{Ml} + \left(e^U\right)\left(\frac{w_K}{w_{Ml}}\right)^{\beta-1}\left(\frac{1-\beta}{\beta}\right)^{\beta-1} \cdot w_K\right)\right) \cdot$$
$$\cdot (-1) \cdot \left[\left(e^U\right)\left(\frac{w_K}{w_{Ml}}\right)^{\beta}\left(\frac{1-\beta}{\beta}\right)^{\beta} \cdot w_{Ml} + \left(e^U\right)\left(\frac{w_K}{w_{Ml}}\right)^{\beta-1}\left(\frac{1-\beta}{\beta}\right)^{\beta-1} \cdot w_K\right] < 0 \quad (3.17)$$

Given that the first two terms are always positive, the third one needs to be positive for all the firms operating in the industry, and the last one (equal to the marginal cost) is always positive, the multiplicative constant (-1) makes profits in Equation 3.17 to be a negative function of inefficiency

From Equation 3.16 it is possible to see that there is an upper-bound level of inefficiency at which profits go to zero, and firms do not have any incentive to produce in the market. This level of inefficiency can be computed solving Equation 3.16, for $\pi_{il}^* = 0$.

$$\pi_{il}^* = \frac{1}{4\gamma}\left(k^* - c_{il}\right)^2 - f_l = 0$$
$$\left(k^* - c_{il}\right)^2 = f_l 4\gamma$$
$$k^* - c_{il} = 2\sqrt{f_l\gamma}$$
$$k^* - 2\sqrt{f_l\gamma} = \left(e^U\right)\left(\frac{w_K}{w_{Ml}}\right)^{\beta}\left(\frac{1-\beta}{\beta}\right)^{\beta} \cdot w_{Ml} + \left(e^U\right)\left(\frac{w_K}{w_{Ml}}\right)^{\beta-1}\left(\frac{1-\beta}{\beta}\right)^{\beta-1} \cdot w_K$$
$$e^U = \frac{\left(k^* - 2\sqrt{f_l\gamma}\right)}{\left[\left(\frac{w_K}{w_{Ml}}\right)^{\beta}\left(\frac{1-\beta}{\beta}\right)^{\beta} \cdot w_{Ml} + \left(\frac{w_K}{w_{Ml}}\right)^{\beta-1}\left(\frac{1-\beta}{\beta}\right)^{\beta-1} \cdot w_K\right]}$$
$$\overline{U} = \ln\left[\frac{\left(k^* - 2\sqrt{f_l\gamma}\right)}{\left[\left(\frac{w_K}{w_{Ml}}\right)^{\beta}\left(\frac{1-\beta}{\beta}\right)^{\beta} \cdot w_{Ml} + \left(\frac{w_K}{w_{Ml}}\right)^{\beta-1}\left(\frac{1-\beta}{\beta}\right)^{\beta-1} \cdot w_K\right]}\right]$$

It follows that:

- $\frac{\partial \overline{U}}{\partial f_l} < 0$, and
- $\frac{\partial \overline{U}}{\partial w_{Ml}} < 0$.

Thus, all else equal, higher fixed organizational costs and variable costs result in lower $\overline{U}$, which is the highest level of inefficiency that firms in the market can bear in order to have non-negative operating profits.

In equilibrium, the free entry condition pins down the value of $\overline{U}$: in fact, it must set the net expected profits of entry into the industry, π^e, equal to zero:

$$\pi^e = \int_0^{\overline{U}} \left[\frac{1}{4\gamma}\left(k^* - c_{il}\right)^2 - f_l\right] \cdot G(U)dU - f_E = 0; \quad (3.18)$$

this condition ensures that all producers make non-negative profits and that entry occurs until the net expected value of taking an inefficient draw is 0. When model's parameters change $(\alpha, \eta, \gamma, f_l, w_M l)$, $\overline{U}$ changes to maintain the equilibrium.

Conditional to the entry equilibrium, vertically integrated firms will face a different upper bound of inefficiency from that experienced by disintegrated firms. For purpose of simplicity, let us assume $\beta = 1/2$ and compute the two upper bounds. Vertically integrated firms face an upper bound $\overline{U_v}$,

$$\overline{U_v} = \ln\left[\frac{\left(k^* - 2\sqrt{f_v \gamma}\right)}{2\left(w_K w_{Mv}\right)^{\frac{1}{2}}}\right], \tag{3.19}$$

while firms which acquire the intermediate input from the outside face the upper bound $\overline{U_o}$,

$$\overline{U_o} = \ln\left[\frac{\left(k^* - 2\sqrt{f_o \gamma}\right)}{2\left(w_K w_{Mo}\right)^{\frac{1}{2}}}\right]. \tag{3.20}$$

It is interesting to derive the conditions under which $\overline{U_o}$ is higher, equal or lower than $\overline{U_v}$ in terms of fixed and variable costs. In this way it is possible to infer how firms with different levels of inefficiency select different vertical organizational configurations.

Case 1 - $\overline{U_o} > \overline{U_v}$. The inefficiency thresholds can be rewritten as:

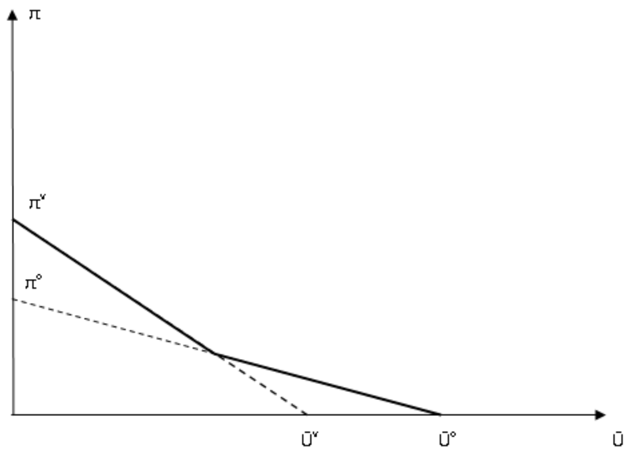
$$\ln\left[\frac{\left(k^* - 2\sqrt{f_o \gamma}\right)}{2\left(w_K w_{Mo}\right)^{\frac{1}{2}}}\right] > \ln\left[\frac{\left(k^* - 2\sqrt{f_v \gamma}\right)}{2\left(w_K w_{Mv}\right)^{\frac{1}{2}}}\right]$$

$$\frac{\left(k^* - 2\sqrt{f_o \gamma}\right)}{\left(w_{Mo}\right)^{1/2}} > \frac{\left(k^* - 2\sqrt{f_v \gamma}\right)}{\left(w_{Mv}\right)^{1/2}}$$

$$\frac{k^* - 2\sqrt{f_o \gamma}}{k^* - 2\sqrt{f_v \gamma}} > \frac{\left(w_{Mo}\right)^{1/2}}{\left(w_{Mv}\right)^{1/2}}.$$

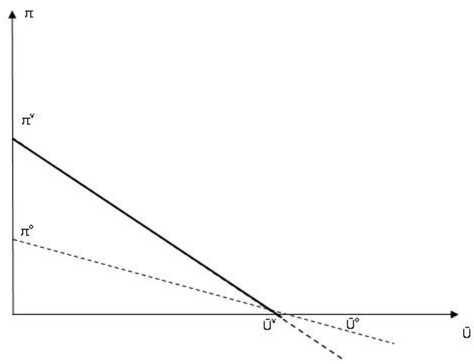
The last equation states that if the ratio of fixed costs is higher than the ratio of variable costs, the upper bound of the inefficiency level which can be borne by a vertical integrated firm is lower than the one borne by a disintegrated firm. Moreover, from Equation 3.17 it is easy to see that vertical integrated firms will have a profit function with a lower (negative) slope, due to the fact that $w_{Mv} < w_{Mo}$ (Assumption 1). We can represent this situation in Figure 3.2. In this case, more efficient firms will choose to produce with a vertical integrated structure because of the higher attainable profits, while less efficient firms will produce with a disintegrated structure, engaging in the outsourcing of the intermediate input. Moreover, a lower $\overline{U_v}$ implies a lower average inefficiency level for vertically integrated firms and a smaller variation of inefficiency (variance) among vertical integrated producers with respect to disintegrated producers.

Figure 3.1: Higher bound of inefficiency for disintegrated firms



Case 2 - $\overline{U_o} < \overline{U_v}$. If the difference between the organizational costs are negligible, while the difference in variable costs are still significant, all the firms would choose to produce as vertically integrated, given that it ensures higher profits than those endured to disintegrated structure, for each maximum inefficiency level. Figure 3.2 clarifies this situation. The first case seems more appropriate for the industry under

Figure 3.2: Similar bounds of inefficiency among firms



analysis: as we have clarified above, fixed costs of a vertical integrated firm are not negligible, and the observation of a dispersion of vertical integration choices among the Italian MT producers is also supported by the descriptive analysis of data, as showed in Section 3.5.2. Thus, we can formulate the following

Testable hypothesis: Vertically integrated firms are expected to show lower levels of inefficiency and to be located nearer to a common production frontier, with respect to disintegrated firms. The distribution of inefficiency for the vertically integrated firms will have a smaller variance with respect to the inefficiency distribution of the disintegrated firms.

3.4 The empirical strategy

We implement a stochastic production frontier model in order to investigate the relationship between firm efficiency and the choices regarding the vertical organization. This is an econometric model which estimates the best-practice production frontier, accounting for random factors not related to technical inefficiency, but which nonetheless affect the productive performance of the firm[11]. The stochastic frontier framework seems appropriate in our case, not only because is allows a direct estimation of the inefficiency level of each production unit, but also because it permits to conduct a one-step estimation of the parameters of the production function and of the coefficient of third variables related to inefficiency. This can be considered as an econometric advantage, which avoids more traditional two-step procedures in which a measure of performance obtained in the first step of the analysis (usually total factor productivity) is regressed on a set of covariates in the second step, likely generating problems of omitted variable bias and under-dispersion of the productive efficiency scores in the first step (see Wang and Schmidt, 2002, for detailed Monte Carlo evidence on this issue).

3.4.1 The stochastic frontier model

We start from the following stochastic production frontier model for panel data:

$$Y_{it} = f\left(\mathbf{X_{it}}, \beta\right) \cdot e^{\epsilon_{it}}, \tag{3.21}$$

where Y_{it} denotes production of the ith firm in the tth time period , $\mathbf{X_{it}}$ is the vector of N inputs used by the producer, β is the vector of technology parameters, and ϵ_{it} the composed error term. In the log-linear form, the stochastic frontier model can be rewritten as

$$y_{it} = f\left(\mathbf{x_{it}}, \beta\right) + \epsilon_{it}, \tag{3.22}$$

where

$$\epsilon_{it} = v_{it} - u_{it}. \tag{3.23}$$

[11] Aigner, Lovell, and Schmidt (1977) and Meeusen and van den Broeck (1977) proposed the stochastic frontier model, starting from the idea that deviations from the production frontier might not be fully under the firm's control.

Equations 3.22 and 3.23 combine to give

$$y_{it} = f\left(\mathbf{x_{it}}, \beta\right) + v_{it} - u_{it}. \tag{3.24}$$

The composed error consists of a component u_{it} which accounts for the difference of the actual level of production from the maximum attainable level, i.e. technical inefficiency, and a white noise component v_{it}, which accounts for random variations of the frontier across firms and measurement errors in y_{it}. The u_{it} component is assumed to follow an exponential distribution and the v_{it} component is assumed to be normally distributed; also, it is assumed that v_{it} and u_{it} are distributed independent of each other. Several distributions have been proposed in the relevant literature to model inefficiency: the half-normal, the exponential and the truncated normal are the three most widely used distributions both for tractability of the composed error term, and for the economic interpretation (see Kumbhakar and Lovell, 2000, p.74). The choice of an exponential distribution to model inefficiency is motivated, in our case, by three main reasons: first, it is a single-parameter distribution, thus easier to be estimated with comparison to more computationally burdensome distributions (like the gamma or the truncated normal)[12]; second, the single-parameter nature of the distribution implies that the variance and the mean of the inefficiency term vary in the same directions (i.e. a shrinkage in the variance corresponds to a reduction in the mean of the u_{it} distribution and vice versa): this perfectly adapts to the testable hypothesis we have advanced at the end of the theoretical Section 3.3; finally, the exponential distribution leads to a stochastic frontier model with the scaling property, and this property is particularly useful when the inefficiency term, u_{it}, is assumed to be a function of a set of firm-related variables as in our case:

$$u_{it}\left(\mathbf{z_{it}}, \gamma\right) \geq 0, \tag{3.25}$$

where $\mathbf{z_{it}}$ is a vector of the characteristics of the MT producers, including a measure of vertical integration and a set of control variables, and γ is a vector of parameters to be estimated indicating the relationship between these variables and u_{it}. The scaling property implies that changes in the values of the variables affecting inefficiency ($\mathbf{z_{it}}$), affect the *scale* but not the *shape* of the distribution of u_{it} (Wang and Schmidt, 2002; Alvarez, Amsler, Orea, and Schmidt, 2006). Formally,

$$u_{it}\left(\mathbf{z_{it}}, \gamma\right) = h\left(\mathbf{z_{it}}, \gamma\right) \cdot u_{it}*, \tag{3.26}$$

where $h(\mathbf{z_{it}}, \gamma) \geq 0$ is the scaling function and $u_{it}*$ is the basic distribution that

[12] Ritter and Simar (1997) propose a rather skeptical view on the use of the gamma and the truncated normal distribution in order to model the inefficiency term, because of problems in estimating the extra-parameter of the two distributions; Koop (2001) argues that the exponential distribution is able to capture a wide variety of inefficiency behaviour.

does not depend on the $\mathbf{z_{it}}$ vector[13]. The scaling property seems appealing in our context, because it allows to consider the effect of random firm characteristics, such as natural management skills (described by a basic random variable u) as distinct from the result of other firm characteristics (i.e. vertical integration) and the environmental 'constraints' under which it operates (for example some characteristics of the industry).

Different models have been proposed to take account of the effects of 'third variables' $\mathbf{z_{it}}$[14]. One method is to directly specify the distribution parameters of u_{it} as functions of the firm-related variables, and then to estimate all the parameters in the model (technology parameters of the frontier function plus all parameters of the inefficiency equation) via maximum likelihood (ML) estimation. In this paper, we hypothesize that the variance of u_{it} depends on the firm-specific degree of vertical integration and a set of firm controls and the variance of v_{it} (noise) is a function of firm size[15].

We can write these assumptions as

$$v_{it} \sim N(0, \sigma^2_{vit}), \tag{3.27}$$

and

$$u_{it} \sim Exp(\eta_{it}), \tag{3.28}$$

where η_{it} is the scale parameter of the exponential distribution, and

$$\eta^2_{it} = g(\mathbf{z_2}\gamma) \tag{3.29}$$

and

$$\sigma^2_{vit} = f(\mathbf{z_1}\delta), \tag{3.30}$$

where $\mathbf{z_2}$ includes the measures of firm vertical integration as well as several controls and $\mathbf{z_1}$ is a measure of firm size, while δ and γ are vectors of the parameters to be estimated. We have chosen to implement a double heteroskedastic frontier model not only because, as it has been said above, it is a way of looking at the relationship of inefficiency with a set of covariates of interest, but also because neglected heteroskedasticity in the two error components can bring to serious biases both in the technology parameters estimates and in the inefficiency estimates: in particular, as Kumbhakar and Lovell (2000) have noticed, (i) unmodeled heteroskedasticity in v_{it} leads to bias in the technical inefficiency estimates, while (ii) unmodeled heteroskedasticity in u_{it} causes bias in both the production frontier parameters and the

[13]It is easy to see that the exponential distribution enjoys this property, because an exponential distribution $u_{it} \sim Exp\left(\eta_{it}\left(\mathbf{z_{it}}, \gamma\right)\right)$, is equivalent to an exponential distribution $u_{it}* \sim Exp(1)$ times the parameter η_{it}.

[14]See Huang and Liu (1994); Battese and Coelli (1995); Caudill, Ford, and Gropper (1995); Wang (2003) among others.

[15]Heteroskedasticity depending on size of the firm usually arises because of the differences in scale.

technical inefficiency estimates[16].

Conditional on $\mathbf{z_{it}}$, u_{it} is assumed to be independent across i and t (u_{it}*s are independent across individuals and over time)[17]. With the above distributional assumptions on u_{it} and v_{it}, it is possible to write the density function of the composed error term $f(\epsilon_{it})$ as a generalization of the normal-exponential model presented by Meeusen and van den Broeck (1977) and Aigner, Lovell, and Schmidt (1977):

$$f(\epsilon_{it}) = \frac{1}{\eta_{it}} \cdot \Phi\left(-\frac{\epsilon_{it}}{\sigma_{vit}} - \frac{\sigma_{vit}}{\eta_{it}}\right) \cdot exp\left(\frac{\epsilon_{it}}{\eta_{it}} + \frac{\sigma_{vit}^2}{2\eta_{it}^2}\right), \tag{3.31}$$

where Φ is the standard normal cumulative distribution function, η_{it} is the standard deviation of the inefficiency component, σ_{vit} the standard deviation of the idiosyncratic part and $\epsilon_{it} = y_{it} - \boldsymbol{x_{it}}'\boldsymbol{\beta}$, is the vector of overall errors. Thus, the log-likelihood function $\ln L(y|\beta, \delta, \gamma)$ for an unbalanced panel of I firms, can be written as:

$$\sum_{i=1}^{I}\sum_{t=1}^{t\leq T}\left(-\log\left(\sqrt{g(\mathbf{z_2}\gamma)}\right)\right) + \sum_{i=1}^{I}\sum_{t=1}^{t\leq T}\log\left[\Phi\left(\frac{-\epsilon_{it}}{\sqrt{f(\mathbf{z_1}\delta)}} - \frac{\sqrt{f(\mathbf{z_1}\delta)}}{\sqrt{g(\mathbf{z_2}\gamma)}}\right)\right] + \\ + \sum_{i=1}^{I}\sum_{t=1}^{t\leq T}\frac{\epsilon_{it}}{\sqrt{g(\mathbf{z_2}\gamma)}} + \sum_{i=1}^{I}\sum_{t=1}^{t\leq T}\left(f\frac{(\mathbf{z_1}\delta)}{2g(\mathbf{z_2}\gamma)}\right), \tag{3.32}$$

where

$$\sigma_{it}^2 = \sigma_{vit}^2 + \eta_{it}^2 = f(\mathbf{z_1}\delta) + g(\mathbf{z_2}\gamma), \tag{3.33}$$

$$\lambda_i = \frac{\eta_{it}}{\sigma_{vit}} = \sqrt{\frac{g(\mathbf{z_2}\gamma)}{f(\mathbf{z_1}\delta)}}. \tag{3.34}$$

Equation 3.32 can be maximized to obtain estimates of $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$; the estimates of $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ in turn can be used to obtain estimates of η_{it} and σ_{vit}.

3.4.2 Model specification

In order to estimate the stochastic frontier model parameters via ML, we have to assume specific functional forms for Equations 3.24, 3.29 and 3.30. We adopt a translog specification for the production function with three inputs[18]:

$$y_{it} = \alpha_0 + \sum_n \beta_n \cdot (x_{nit}) + \frac{1}{2}\sum_n\sum_p \beta_{np} \cdot (x_{nit}x_{pit}) + \tau_t + \alpha_j + v_{it} - u_{it}, \tag{3.35}$$

[16]The issue of heteroskedasticity has captured the attention of several scholars in the field: see Reifschneider and Stevenson (1991), Caudill and Ford (1993), Caudill, Ford, and Gropper (1995), Hadri, Guermat, and Whittaker (2003).

[17]Note that ML estimates based on the assumption of independent observation are consistent even if observations are not independent; the requirement is the correct specification of the marginal distribution of each observation (Alvarez, Amsler, Orea, and Schmidt, 2006).

[18]The functional form adopted in the empirical analysis is a generalization of the simple Cobb-Douglas employed in the theoretical model. However, the basic prediction of the theoretical model does not depend on the specific functional form, while a more flexible function permits a better adaptation to the data.

where n, p=(*capital*, *labour*, *intermediates*). In order to control for unobserved heterogeneity among firms producing different typologies of machines, we include $(j-1)$ dummies α_j in the frontier, where $j = (1, \ldots, 9)$ refers to the type of machine produced by the firm; we control also for factors affecting all firms in the same way in a given year including $(t-1)$ year dummies τ_t [19]. It is also necessary to assume some specific functional forms for (3.29) and (3.30): following Hadri (1999), we employ an exponential function to model variances of the error components, in particular:

$$\eta_{it}^2 = exp\,(\mathbf{z_2}\gamma) = exp(\gamma_0 + \gamma_1 VDIS + \gamma_2 SIZE + \gamma_3 DOWNER + \gamma_4 DDIST + \gamma_5 DCYCLE), \quad (3.36)$$

where $\mathbf{z_2}$ denotes the degree of firm vertical (dis)integration, and includes controls for firm size, ownership type, agglomeration economies and the economic cycle (the explanation on how these variables have been measured is given in Section 3.5.1) and

$$\sigma_{vit}^2 = exp\,(\mathbf{z_1}\delta) = exp(\delta_0 + \delta_1 SIZE), \quad (3.37)$$

where $\mathbf{z_1}$ is a measure of firm size. ML estimation is implemented in order to obtain consistent and efficient estimates of the parameters in equations 3.35, 3.36 and 3.37, i.e. $\widehat{\alpha}$, $\widehat{\tau}$, $\widehat{\beta}$, $\widehat{\delta}$ and $\widehat{\gamma}$.

3.5 Data and descriptive analysis

We exploit an original database which has been compiled recovering data from several data sources. The list of MT producers is from Ucimu and includes information on firm's type production[20]. Information on output and inputs is from Bureau Van Dijk's AIDA database, which contains balance sheet information for firms with turnovers over €500,000. Information on the ownership status is from the Bureau Van Dijk's Ownership Database, and information on district location was obtained by comparing the locations of local firm units — contained in AIDA— with the list of Italian Labor Local Systems (LLS) regularly updated by the Italian National Institute of Statistics, ISTAT [21]. Deflators for output, intermediate inputs and capital stock respectively, were computed from the Value of Production and Investments series published by Istat annually at the sectoral level (2-digit level) [22].

[19] The inclusion of 'effects' in the stochastic frontier allows us to differentiate between unobserved heterogeneity and time-variant inefficiency.

[20] Note that the list does not include only Ucimu associates, it includes all firms covered by surveys and research questionnaires administered by the Association. There are almost 550 firms on this list.

[21] `http://www.istat.it.`

[22] `http://www.istat.it/conti/nazionali/.`

3.5.1 Description of the variables

Variables in the production frontier

The output (Y) is measured by the amount of revenues from sales and services at the end of the year, net of inventory changes and changes to contract work in progress. This measure is deflated in order to account for price variations during a year. The deflator was built at the 2-digit level (Ateco 2007 classification) and is equal to the ratio of the value of production at current prices, in a given year, over the corresponding value in the chained level series[23]. The measure is expressed in €'000.

The labour input (L) is measured as the total number of employees at the end of the year. Capital stock (K) in a given year is proxied by the nominal value of tangible fixed assets, which is deflated using the ratio of gross fixed investments at current prices over corresponding values in the chained level series. Given the unavailability of series at the 2-digit level, we use a common deflator for all firms (investments for aggregate C-D-E Ateco 2007 Industry sectors). The measure is expressed in €'000. Intermediate inputs (M) are measured as the sum of (i) costs of raw, materials consumed and goods for resale (net of changes in inventories) plus (ii) costs of services. The measure is deflated by the same deflator applied to output. It is expressed in €'000.

All inputs and the output have been normalized by mean-correction before including them in logs in the production frontier. In this way coefficients of the translog production function can be interpreted as output elasticities with respect to inputs for the average unit considered.

Vertical (dis)integration

We use a measure of vertical disintegration, ($VDIS$), and we build it as the ratio of intermediate inputs (M) over total costs of production for the year. For the ith firm in the tth time period, this can be written as:

$$VDIS_{it} = \frac{C_{RM,it} + C_{S,it}}{C_{RM,it} + C_{S,it} + C_{L,it} + C_{K,it} + C_{O,it}} \tag{3.38}$$

where $C_{RM,it}$ is the cost of raw, materials consumption and goods for resale (net of changes in inventories), $C_{S,it}$ is the cost of services, $C_{L,it}$ is total personnel costs, $C_{K,it}$ is total depreciation, amortization and write downs (thus it can be interpreted as the figurative cost of capital) and $C_{O,it}$ is a residual class, which is a negligible portion of the total costs of production and can be considered equal to zero for the purpose of the present analysis. This ratio is an indicator of the relative 'weight' of the factors of production external to the firm (i.e. acquired from other firms), over all factors of

[23]The base year for the chained series is 2000.

production including labour and capital[24]. This measure is related to that proposed by Adelman (1955), i.e. the ratio of value added to sales as a measure of vertical integration, however, we think about our measure as an improvement with respect to the Adelman index for several reasons.

Adelman's index has been criticized mostly for the problems involved in applying it in cross-industry studies[25] and its asymmetry[26]. However, our measure should not suffer the same problems in the case under analysis. First, the Italian MT industry is a quite narrowly defined industry so there should be no cross-industry problems. Second, even if the major drawback is that we do not have information on prices, and we cannot control explicitly for the likely different unitary costs which may be faced by different firms in the sample, it is relevant to note that as for labour, given the well known salary 'rigidities' in the Italian labour market, it is not restrictive to assume $\overline{w_{it}} = \overline{w_{jt}}$ for all firms $i \neq j$. For capital, it is reasonable to assume that the differences affecting variations in $C_{K,it}$ among firms, depend on the amount of machines and equipment acquired[27]. Finally, our measure is not sensitive to differences in the output price, which could simply result from different qualities in the output sold by the firm or different degrees of market power: these differences enter in the denominator of the Adelman index, but not in our measure of vertical disintegration. For these reasons, the measure we use appears to be the best available solution to capture the firm vertical organization given the available data, and in this context is preferred to Adelman's index. Nonetheless, we use the Adelman index as robustness check in the econometric analysis.

Control variables

In line with previous studies, we included a set of control variables in the vector **z2** in order to minimize the danger of capturing misleading spurious correlation between vertical disintegration and inefficiency.

We include a measure of firm size, ($SIZE$), which is defined as total number of employees at the end of the year. The relationship between size and efficiency has been debated in the empirical literature[28], but is still not clearcut: see Caves and Bar-

[24] A value of 1, means that the firm depends on external suppliers for almost all of its production inputs; values near 0 indicate that the firm bases its production on its own capital and labour, i.e. it is vertically integrated.

[25] The empirical literature on vertical integration has made some proposals to overcome these drawbacks, such as the use of other measures. See, e.g. the use of input/output tables proposed by Maddigan (1981) to build a 'vertical industry connection index' for all industries in which the firm operates, which was adapted by Acemoglu, Johnson, and Mitton (2009) to evaluate the determinants of vertical integration in a cross-country perspective.

[26] Holding the ratio(VA/Sales) constant, firms near the end of the production chain (and final consumers) appear less integrated (Davies and Morris, 1995).

[27] In fact, year quota of depreciations and amortizations are computed following fiscal deductibility purposes, using the coefficients established by the Ministry of Economy and Finance at sectoral level — and thus are common to all firms belonging to the same sector— in the Ministerial Decree 31.12.1988.

[28] The theme has also been deeply studied in the empirical literature regarding agricultural production.

ton (1990) for an investigation of US manufacturing; Gumbau and Maudos (2002), Taymaz (2005), Diaz and Sanchez (2008) for empirical investigations on Spanish and Turkish manufacturing; Badunenko, Fritsch, and Stephan (2008) for the relationship in German manufacturing. The contradictory results from these studies are an indication that single-industry studies are required in order to monitor the relationship between size and efficiency. Thus it is relevant to control for it, especially because it may be correlated with other non–observable firm characteristics such as degree of internationalization and quality of inputs, especially managerial staff.

Even if in the last years the geographical distribution of MT producers does not correspond to the typical industrial district, we include a control for firms localized in industrial districts, in order to take account of this kind of agglomeration economies: $DDIST$ is a time-invariant dummy variable that takes the value '1' if firms have at least one local unit (either headquarters or not) located in a mechanical engineering industrial district and '0' otherwise. It is well known that industrial districts are key socio-economic structures in the Italian industrial system (Becattini, 1990). Fabiani, Pellegrini, Romagnano, and Signorini (1998) found a positive relationship between efficiency and district location, in a sample of Italian manufacturing firms in the period 1982 to 1995, and Becchetti, Panizza, and Oropallo (2008) shows that industrial district firms demonstrate higher value added per employee and higher export intensity.

In the Italian MT industry, different decades are characterized by different ownership forms. The 1980s were characterized by a structural strengthening of the industry via external growth (Rolfo, 1993). This tendency slowed down in the first half of the 1990s, but was reinvigorated at the end of that decade, as MT builders tried to maintain control of the production process. During the second half of the 1990s, the mechanical engineering sector experienced a new wave of mergers (Rolfo, 1998), designed to cope better with risk and to exploit market and production complementarities. Thus the ownership structure is relevant for an analysis of firm efficiency: first, because it can be a substitute for vertical integration, and second, in line with Bottasso and Sembenelli (2004), because firm efficiency is heavily driven by managerial effort, and seriously affected by conflicts between ownership (shareholders) and control (management) (Shleifer and Vishny, 1986). To control for type of ownership we included a dummy variable $DOWNER$) that takes the value '1' if the firm belongs to an industrial group (either national or international)and '0' if the firm is independent: firms are considered as part of a group if they control or are controlled by other firms with a percentage of shares $\geq 50\%$[29].

[29]This may be a restrictive threshold. Control over other firms may be possible even at much lower shares; also, in the Italian MT industry there are informal groups which are linked not just by ownership of relevant shares quotas, but by familial links. However, this conservative measure of ownership control ensures a clear distinction between firms belonging to established groups and other firms (independent, or part of an informal group).

Finally we include a dummy, $DCYCLE$, for the years showing a downward trend in the value of production , i.e. 2002, 2003 and 2004. Given the cyclical nature of the MT industry, failing to control for the cycle could bias our results on the relationship between vertical disintegration and inefficiency. Moreover, the dummy variable allows us to look at the effect of the economic cycle on firm efficiency.

3.5.2 Descriptive statistics

Based on the reference list provided by Ucimu, we collected balance sheet data for 524 firms and 5,240 observations from Bureau Van Dijk's AIDA database. We discarded some observations after a preliminary analysis which revealed missing values and outliers. First, we excluded observations with missing values for output, inputs and the variables in the inefficiency model. The number of not usable observations is 1,467 (mostly due to the unavailability on the number of employees). Moreover, we excluded eleven observations because they presented negative values for output or inputs. In order to detect some possible outliers, we conducted an ordinary least squares (OLS) estimation of the translog production function, and found that the residuals-versus-fitted plot revealed five more observations which have not been included in the frontier analysis, due to their exceptional distance from the cloud of observations, i.e. observations with standardized residuals $> |5|$). These preliminaries reduced our final sample to an unbalanced panel amounting to 505 firms and 3,757 usable observations, for the period 1998 to 2007.

Table 3.3 presents descriptive statistics for the sample under analysis and Table 3.4 presents a breakdown of the obervations with respect to the production of the firm (i.e. the type of machine produced): the two largest product specializations are metal cutting machines (e.g. machining centers, lathes) and metal forming machines (presses, sheet metal deformation machines).

Overall, our sample depicts figures which are in line with general statistics on the industry that can be found in technical reports, as the one provided by UCIMU (Ucimu, 2007a). Almost 75% of machine producers in our sample invoice around €13.0 millions, while the top 10% of firms invoices (at least) more than two times of that amount: this claims for an high fragmentation among smaller and larger firms in terms of market shares, as already underlined in Section 3.2. If we compare the evidence contained in the technical report with our data (Table 3.5) our sample slightly over-represents medium firms and under-represents small firms (in terms of employees). This is basically confirmed when we look at the geographical distribution of the firms: it is well known that producers of machine tools in Emilia-Romagna are usually smaller than their counterparts located in Piemonte and Lombardia: that is why the sample under-represents the percentage of firms located in Emilia-Romagna and sightly over-represent the percentage of firms located in the other two regions. The descriptive evidence is also in line with previous studies on the industry. Firms in

Table 3.3: Descriptive statistics, 1998-2007

Variable	Notation	Unit	Mean	Std. Dev	Min	Max	p10	p25	p50	p75	p90	N firms	N obs
Gross output	Y	€'000	16854	57636	199	977748	1601	2837	5997	12961	30243	505	3757
Capital	K	€'000	2426	7687	.923	137786	76.3	225	790	2092	4884	505	3757
Labor	L	Number of workers	98.1	324	1	8158	11	20	41	86	185	505	3757
Intermediate inputs and services	M	€'000	11420	40855	119	679809	923	1716	3881	8697	19736	505	3757
Total costs of production	TC	€'000	17041	59643	267	1160910	1568	2818	6092	13014	30190	505	3757
Vertical disintegration	VDIS	Ratio	.67	.119	.172	1	.502	.594	.68	.757	.813	505	3757
Downward cycle	DCYCLE	Dummy	.328	.47	0	1	0	0	0	1	1	505	3757
Ownership	DOWNER	Dummy	.238	.426	0	1	0	0	0	0	1	505	3757
District location	DDIST	Dummy	.0615	.24	0	1	0	0	0	0	0	505	3757

Table 3.4: Breakdown of firms by the type of production

Product categories	**N firms**	**N obs**
Builders of metal cutting machines	175	1290
Builders of metal forming machines	124	898
Builders of unconventional machines	24	176
Builders of welding machines	2	13
Builders of measuring-control machines	15	111
Builders of heat treatment machines	19	141
Builders of mechanical devices	107	826
Builders of electric/electronic equipment	22	175
Builders of tools	17	127
Total	505	3757

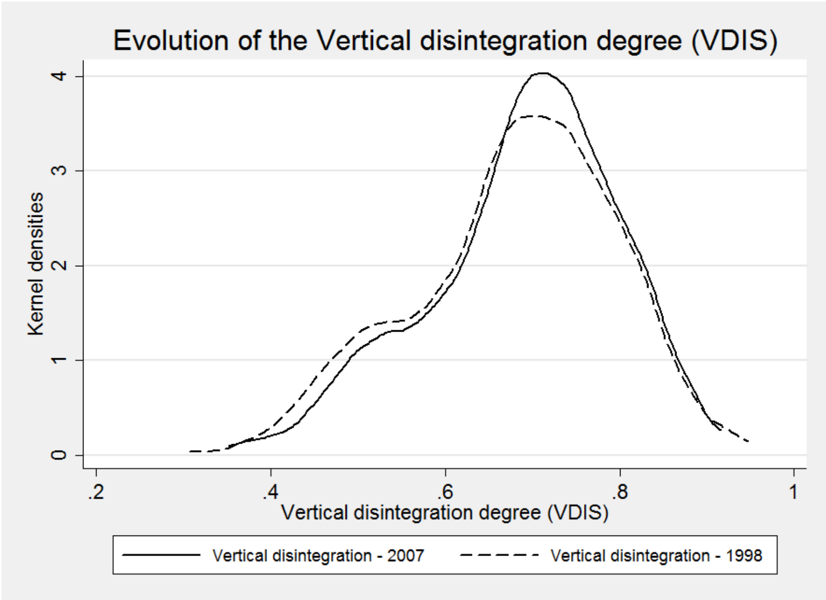
Table 3.5: Sample vs. Ucimu industry report

		Ucimu - industry report (2006)	**Sample (2006)**
		% on total number of firms	% on total number of firms
Size classes	≤ 50	63.10	57.11
	50:100	14.80	21.45
	>100	22.10	21.45
Regions	Lombardia	46.30	53.24
	Triveneto*	17.40	14.09
	Emilia-Romagna	16.10	10.42
	Piemonte	12.80	14.37
	Other regions	7.40	7.88

*Triveneto=Veneto+Friuli+Trentino Alto-Adige

our sample show high levels of vertical disintegrations (.67) on average, and this is in line with previous results, e.g. Arrighetti (1999) who provides an analysis of vertical integration among Italian manufacturing firms using the Adelman index, and shows an average degree of vertical integration of .35 for mechanical engineering firms . If we look at the distribution of levels of vertical (dis)integration in Figure 3.3 and we focus on its evolution from 1998 to 2007, for those firms which are observable in both years, two facts are evident: first the high heterogeneity in the vertical organization of MT producers which is stable as time has passed; second, the agreement with a general tendency toward a disintegration of production (outsourcing) in the past years, which occurred also in this industry. In fact, in the 2007 kernel density an higher number of observations are clustered around the .75 peak of the $VDIS$ distribution. The range of values is wide, showing the coexistence of vertically integrated firms with firms relying on external phases of productions (via acquired intermediate inputs). Rolfo (1998) underlines that from 1995 onwards, firms tried to strengthen their control over suppliers via external growth and the establishment of small industrial groups. In our sample almost 24% of firms belong to an industrial group (either a subsidiary or the holding company). Moreover, in our sample only a small proportion of firms (around 6%) are localized in a mechanics industrial district, that is in line with the studies referred to above. Given these preliminary evidences we are pretty confident

Figure 3.3: Vertical disintegration in 1998 and 2007



that our sample describes the industry under analysis in a fair way (maybe a little bit biased toward medium-sized firms), capturing a large set of relevant characteristics of it.

3.6 Econometric analysis

3.6.1 Baseline results

Our estimations are based on Stata 10.1 software[30]. In order to analyze the relationship between firm efficiency ant the vertical organization, we have run three specifications of the model. Below we describe the groupings; this makes the results easier to understand, and introduces the various statistical tests. All specifications (except M1, which has been estimated via OLS) are estimated via the ML method, which jointly estimates the frontier parameters in Equation 3.35, and the coefficients of variables in the models of variances in Equation 3.36 and 3.37. Table 3.6 presents the estimates for the frontier parameters and Table 3.7 presents the vector of coefficient estimates in Equations 3.36 and 3.37.

The specifications can be grouped as follows:

- **M1**: OLS average production function estimation, in which η_{it}^2 is assumed to be equal to zero; in other words, this model does not consider the possibility

[30]The estimation of the parameters of the stochastic frontier model has been performed using the `frontier` command.

Table 3.6: Frontier parameters estimation

Specification		**M1**	**M2**	**M3**	**M4**	**M5**
Variable	Coefficient					
lnK	β_k	0.0249***	0.0266***	0.0263***	0.0261***	0.0267***
		(0.0029)	(0.0029)	(0.0029)	(0.0029)	(0.0028)
lnL	β_l	0.2141***	0.2208***	0.2129***	0.2157***	0.2102***
		(0.0054)	(0.0054)	(0.0064)	(0.0067)	(0.0055)
lnM	β_l	0.7670***	0.7585***	0.7665***	0.7681***	0.7666***
		(0.0047)	(0.0047)	(0.0059)	(0.0060)	(0.0056)
(.5)(lnK)^2	β_{kk}	0.0071***	0.0089***	0.0087***	0.0087***	0.0087***
		(0.0020)	(0.0020)	(0.0021)	(0.0021)	(0.0021)
(.5)(lnL)^2	β_{ll}	0.1263***	0.1327***	0.1295***	0.1278***	0.1265***
		(0.0053)	(0.0053)	(0.0056)	(0.0056)	(0.0051)
(.5)(lnM)^2	β_{mm}	0.1218***	0.1268***	0.1246***	0.1245***	0.1238***
		(0.0056)	(0.0056)	(0.0057)	(0.0057)	(0.0056)
(lnK)·(lnL)	β_{kl}	-0.0037	-0.0038	-0.0037	-0.0040	-0.0027
		(0.0026)	(0.0026)	(0.0026)	(0.0026)	(0.0025)
(lnK)·(lnM)	β_{km}	-0.0033	-0.0056**	-0.0052**	-0.0052**	-0.0057**
		(0.0023)	(0.0023)	(0.0024)	(0.0024)	(0.0024)
(lnL)·(lnM)	β_{lm}	-0.1168***	-0.1208***	-0.1187***	-0.1180***	-0.1180***
		(0.0047)	(0.0047)	(0.0049)	(0.0049)	(0.0048)
Constant	α	0.0073	0.0517***	0.0529***	0.0534***	0.0560***
		(0.0070)	(0.0072)	(0.0073)	(0.0073)	(0.0077)
Year dummies	τ_t	Yes	Yes	Yes	Yes	Yes
Prod dummies	α_j	Yes	Yes	Yes	Yes	Yes
Log-likelihood		2787	2819	2823	2824	2843
Observations		3757	3757	3757	3757	3757

St. err. of coefficients in parentheses.
Significance levels: * 10%, ** 5%,*** 1%.
Year and Prod estimates omitted.
Complete table available from the authors upon request.

of existence of inefficiency in the sample. All firms are regarded as technical efficient, and all deviations from the frontier are due to noise.

- **M2**: Homoskedastic frontier; in this model variance of both error components — v_{it} and u_{it}— is assumed to be constant among the observations: the assumption can be formalized as $\sigma^2_{vit} = \sigma^2_v$ and $\eta^2_{it} = \eta^2$ for all i, t. In the case under analysis, the preference for this model would imply that MT producers' technical efficiency is not related to their degree of vertical disintegration and to other variables in **z2**, and noise is not heteroskedastic in firm size.

Table 3.7: Models of variance

Specification		**M2**	**M3**	**M4**	**M5**
$\ln(\eta^2)$ function					
VDIS	γ_1		2.0813**	2.0581**	2.6333***
			(0.8777)	(0.8790)	(0.9156)
SIZE	γ_2			0.0003*	0.0003**
				(0.0002)	(0.0002)
DOWNER	γ_3				-0.3313*
					(0.1992)
DDIST	γ_4				-1.1641**
					(0.5030)
DCYCLE	γ_5				-1.1523**
					(0.4989)
Constant	γ_0	-6.0947***	-7.5259***	-7.5413***	-7.6254***
		(0.1258)	(0.6471)	(0.6481)	(0.6719)
$\ln(\sigma^2_v)$ function					
SIZE	δ_1				-0.0006**
					(0.0003)
Constant	δ_0	-4.5340***	-4.5223***	-4.5236***	-4.4594***
		(0.0318)	(0.0336)	(0.0336)	(0.0379)
Year dummies		Yes	Yes	Yes	Yes
Prod dummies		Yes	Yes	Yes	Yes
Log-likelihood		2819	2823	2824	2843
Observations		3757	3757	3757	3757

St. err. of coefficients in parentheses
Significance levels: * 10%, ** 5%,*** 1%
Year and Prod estimates omitted.
Complete table available from the authors upon request.

- **M3-M5**: Heteroskedastic frontier specifications: the measure of vertical disintegration ($VDIS$) is introduced alone in specification M3, while a control for firm size enters in specification M4 and the full vector of controls is included in specification M5; this last specification should be the one in which spurious correlations between vertical disintegration and firm inefficiency are minimized.

Generalized likelihood ratio tests of the form $LR = -2\left[\ln L(H_0) - \ln L(H_1)\right] \sim \chi^2_J$[31] can be performed on the parameters of the frontier and on the coefficients of the inefficiency model in order to select the model that minimizes any misspecification

[31] J is the number of restrictions: see (Coelli, Rao, O'Donnell, and Battese, 2005, pp.258-259) for a useful introduction to statistical tests in stochastic frontier analysis.

bias. All test results are reported in Table 4.6. The translog specification seems an adequate representation of the technology: in fact, the likelihood ratio test, in the first row of the Table, strongly rejects the restrictions imposed by a nested Cobb-Douglas. Frontier models are preferred to the average production function model. If

Table 3.8: Generalized LR tests on the parameters of stochastic frontier model

Null Hypothesis	Conditions	χ^2 statistics	Critical values (5%)
Cobb-Douglas restrictions	$\beta_{n,p}$=0, for $n, p = K, L, M$	785.77	12.59
No inefficiency	η^2_{it}=0	65.57	2.71*
No time dummies	τ_t=0	161.08	16.92
No production dummies	α_j=0	236.13	15.50
Heteroskedastic vs. homoskedastic frontier	$\boldsymbol{\gamma}' = \delta_{SIZE}$=0	48.33	12.59
No vertical (dis)integration effect	γ_{VDIS}=0	9.48	3.84
No control variables effects	$\gamma_{controls} = \delta_{SIZE}$=0	41.81	15.09

*: the test is at the boundary of the parameter space η;
the critical value comes from the table provided by Kodde and Palm (1986)

we take Specification (M2), the homoskedastic frontier, we can test $\eta^2_{it} > 0$ versus the null hypothesis of $\eta^2_{it} = 0$: in the case in which the null hypothesis is accepted, the stochastic frontier model will reduce to an average production function model with symmetric errors, which could be consistently estimated by means of OLS. The second row in Table 4.6 definitely rejects the null hypothesis, thus confirming the presence of inefficiency in the sample and the adequacy of the stochastic frontier tool. Moving to specification (M5), both time dummies and production dummies result to be significant, showing that is relevant to control for the type of production of the firm and unobserved factors affecting all firms in a given year. Also, the heteroskedastic frontier specification (M5) is preferred to the homoskedastic frontier (M2): we tested the joint significance of all explanatory variables affecting the inefficiency variance and the null hypothesis is firmly rejected. This reassures us about the fact that measured inefficiency is a function of the chosen variables. We have tested also for the significance of the $VDIS$ variable, with respect to a specification that excludes it. The sixth row in Table 4.6 reports the results of this LR test, which show that the vertical organization of the firm, captured by the variable $VDIS$ is significant in explaining the inefficiency variance differences among MT producers. The last row in Table 4.6 shows the relevance of the controls.

A negative coefficient in Table 3.7 can be alternatively interpreted as a negative effect on the variance of inefficiency, or a positive relationship with firm efficiency. Results in specification (M5), which is our favorite given its better adaptation to data with respect to (M1-M4), show that after controlling for firm size, type of ownership, agglomeration economies and economic cycle, the higher degree of vertical disintegration is significantly related to an higher variance (and higher mean) of the inefficiency distribution, thus implying lower inefficiency for vertical integrated firms, *ceteris paribus*. The negative coefficient of $VDIS$ suggests that more integrated organizations are advantaged: firms that carry out more phases of the production process internally, enjoy advantages over less integrated producers. The result is confirmed by the significant negative value of the coefficient of the ownership dummy (DOWNER), in all of the specifications M6–M8. A group structure can substitute for vertical integration in some respects: both internal and external (through the group) vertical integration have positive effects on efficiency. The positive effect of group structure cancels out any potential negative outcomes of ownership–manager conflicts, such as the ones arising in the analysis conducted by Bottasso and Sembenelli (2004).

Overall, this result is pretty much in line with our theoretical model, predicting vertical integrated firms to be nearer to the technological frontier, with a lower upper bound level of inefficiency, due to higher fixed organizational costs. Given that the inefficiency distribution has been assumed as exponential, an lower threshold implies also a smaller variance of the inefficiency distribution, that is in line with what we find in the empirical analysis. However, even if the empirical results have captured a systematic pattern between firm efficiency and vertical integration, this result cannot be interpreted as a causal relationship: in fact, even controlling for a relevant set of firm characteristics and thus lowering the danger of misleading spurious correlations, we cannot control explicitly in this econometric framework for the reverse causality, i.e. the effect that goes from the vertical structure to firm efficiency. In the theoretical model, we have in mind a self-selection process of the most efficient firms to vertical integrated structures, but we cannot exclude that the regressions are capturing also a reinforcing phenomenon which runs in the opposite direction (a sort of learning channel): this could be explained by different factors, such as a greater coordination in production processes or a better adaptation (in terms of quality and quantity) of intermediate inputs to the final output which can be achieved by a firm which becomes vertical integrated.

The value of other parameters is worthy of comment. It should be noted that the measure of firm size is positively correlated with the inefficiency variance: this contrasts the commonly held view that a larger size can be used as a proxy for a better organization. However, it has been largely shown that the relationship between size and inefficiency is basically industry-specific: in our case it is relevant to control for it, as the significant coefficient demonstrates, in order to minimize dangers of spurious correlations[32]. A second robust result in the heteroskedastic frontier specification (M5), is the significant negative coefficient of the dummy for downward cycle: when the aggregate demand is low, the variance of inefficiency decreases. Taken together with the first result this means that down phases result in partial loss of the efficiency advantages from vertical integration and could suggest a sort of dynamic advantage among less integrated firms. Finally, the dummy for those firms localized in an

[32]Firm size is also significant in explaining differences in the variance of the noise term, thus it is necessary to include it in Equation 3.37.

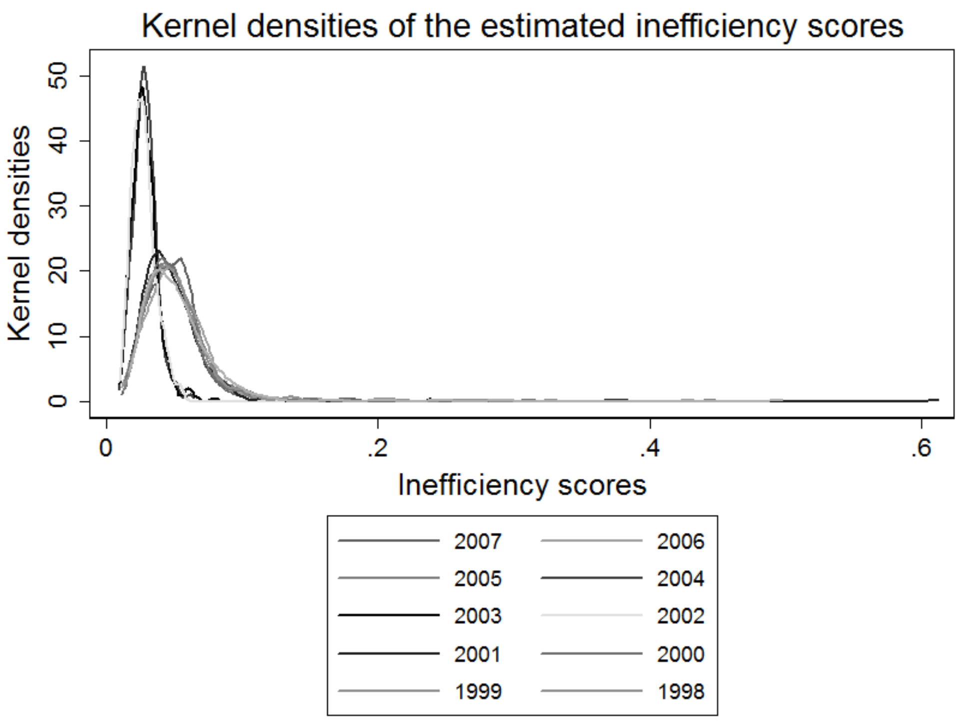
industrial district shows a negative coefficient: agglomeration economies seem to enhance the productive performance of firms in the Italian MT industry, showing a lower variance of the inefficiency distribution for firms localized in an mechanics industrial district.

It is possible to compute the firm and year-specific inefficiency scores via the following formula, which is an extension of the one proposed by Jondrow, Lovell, Materov, and Schmidt (1982) when u_{it} and v_{it} are heteroskedastic:

$$\hat{u}_{it} = E\left(u_{it}|e_{it}\right) = \hat{\sigma}_{vit}\left[\frac{\phi\left(\frac{e_{it}}{\hat{\sigma}_{vit}}+\frac{1}{\hat{\lambda}_{it}}\right)}{1-\Phi\left(\frac{e_{it}}{\hat{\sigma}_{vit}}+\frac{1}{\hat{\lambda}_{it}}\right)}-\left(\frac{e_{it}}{\hat{\sigma}_{vit}}+\frac{1}{\hat{\lambda}_{it}}\right)\right], \tag{3.39}$$

where e_{it} are the ML overall residuals. Figure 3.4 shows kernel densities of the efficiency scores from 1998 to 2007. It is possible to appreciate that in the year of

Figure 3.4: Inefficiency scores, 1998-2007



downward aggregate demand, the distribution of the inefficiency scores is more distributed around its central tendency, thus showing a lower variance, as the coefficient of the dummy $DCYCLE$ showed in Table 3.7.

3.6.2 Robustness checks

In the present Section we perform two types of robustness checks. First, we explore the sensitivity of the main result of our analysis —i.e. that vertically integrated firms delineate the technology frontier— to changes in the employed measure of vertical integration; second, we include the one-year lagged estimated variance in the skedastic Equation 3.36, in order to see if the variance of the inefficiency distribution is basically determined by its lag and just spuriously correlated with the vertical integration degree[33]. We do not report the frontier parameter estimates in order to save space, also because no significant changes are observed with respect to specification (M5), and we directly focus on the variance equations. In the first column of Table 3.9, we use the more traditional Adelman index (VI) as the measure for vertical integration. The index is equal to the ratio of value added over sales and higher values correspond to higher degrees of vertical integration: the coefficient is negative and significant showing lower variance of the inefficiency distribution to more vertically integrated firms, thus confirming the main result in the baseline specification (M5). In the second column of the Table we have substituted the $VDIS$ measure with its one-year lag and forward moving average, $VDIS_{mov,(i,t)} = (VDIS_{i,t-1} + VDIS_{i,t} + VDIS_{i,t+1}/3)$; this has been done in order to minimize undesirable variations in the vertical disintegration measure due to fluctuations in prices or cost shares which do not relate to the vertical structure of the firm, while to the economic situation in an year. The coefficient of the $VDIS_{mov}$ variable is pretty much in line with the estimated coefficient in specification (M5), thus reassuring us about the goodness of the employed measure. In the third column, we include the lagged estimated variances in the skedastic function of inefficiency performing a second round estimation of specification (M5). Overall, the magnitude of the coefficient of $VDIS$ raises with the inclusion of the lagged variance in Equation 3.36 and this is also partially due to sample selection (in fact the number of observations decreases from 3757 to 3031), but the sign of the relationship remains stable. More disintegrated firms show higher variance (and mean) of the inefficiency distribution, thus positioning further away from the stochastic production frontier with respect to more integrated ones. Moreover, given the negative coefficient of the lagged variance, it seems that firms with higher variance at time $t-1$ show a lower variance at time t, as a sort of 'converge to the frontier' phenomenon.

3.7 Concluding remarks and suggested further research

In this paper we have studied the relationship between vertical integration and firm efficiency in the Italian machine tool industry. We have first set up a theoretical model (in line with previous models on productivity heterogeneity and organizational

[33]We have also run specification (M5) on a sample made up of those firms which produce final good (machines) only, and not just components. The main result of the analysis is stable.

Table 3.9: Models of variance

Specification	**C1**	**C2**	**C3**
$\ln(\eta^2)$ function			
VI	-9.7555***		
	(0.7990)		
VDIS_mov		2.4751***	
		(0.8547)	
VDIS			6.8778***
			(1.7026)
$\ln(\widehat{\eta^2_{t-1}})$			-0.6573**
			(0.3220)
SIZE	0.0003	0.0003**	0.0005**
	(0.0002)	(0.0002)	(0.0002)
DOWNER	0.0655	-0.3334*	-1.4345***
	(0.1750)	(0.1951)	(0.4319)
DDIST	-1.0922***	-1.1660**	-2.1315**
	(0.4075)	(0.4976)	(0.9155)
DCYCLE	-1.2289***	-1.0680**	-1.8391***
	(0.2611)	(0.4298)	(0.6614)
Constant	-2.8902***	-7.4974***	-14.7113***
	(0.2228)	(0.6188)	(2.9983)
$\ln(\sigma^2_v)$ function			
SIZE	-0.0006***	-0.0007***	-0.0006***
	(0.0002)	(0.0003)	(0.0002)
Constant	-4.5223***	-4.4650***	-4.4399***
	(0.0306)	(0.0374)	(0.0376)
Year dummies	Yes	Yes	Yes
Prod dummies	Yes	Yes	Yes
Log-likelihood	2991	2843	2365
Observations	3757	3757	3031

St. err. of coefficients in parentheses
Significance levels: * 10%, ** 5%,*** 1%
Frontier parameter estimates omitted.
Complete table available from the authors upon request.

choices, as the one proposed by Antras and Helpman (2004)), in order to come up with a testable hypothesis: in our model more efficient firms decide to produce as vertically integrated, bearing higher (organizational) fixed costs while less efficient firms choose to outsource part of production process buying an intermediate input from other firms, thus reducing fixed costs but bearing higher marginal costs of production. In equilibrium, the two types of organizations coexist and the industry contemplates firms with different levels of efficiency. This theoretical result is pretty much in line with the previous quantitative and qualitative evidence on the industry, as the work by Zanfei and Gambardella (1994) who claim that in the Italian MT sector firms with different size, organization structures and sourcing strategies coexist, and complement each other in supplying the market all the varieties requested by a highly differentiated demand, or Wengel and Shapira (2004) who points to a dualistic structure of the industry. However, while previous work has stressed the general characteristic of 'size' as point of differentiation between the groups of firms in the industry, we think that the vertical structure better represents the different choices for the organization of production.

We empirically ground this result, conducting a stochastic frontier analysis on a sample of more than 500 machine tool producers. In this way it is possible to estimate the best practice technology frontier, measuring the distance to it as indicators of inefficiency (sub-optimal level of output, given the amount of inputs and the available technology). The empirical analysis shows that vertical integrated firms present a lower variance (and lower mean) of the inefficiency distribution, after having controlled for firm size, type of ownership, agglomeration economies and the economic cycle. Thus, vertical integrated firms are, *ceteris paribus* more efficient in the industry under analysis than disintegrated firms. An important clarification should be stressed: even if our theoretical model predict a self-selection mechanism of more efficient firms to vertical integrated structures, the empirical analysis cannot rule out the inverse direction of the relationship. In other words, there could be a positive effect which goes from vertical integration to firm efficiency, which have been supported by previous evidence in the management and industrial economics literature [34]. Thus, any kind of causal effect should be considered with caution. Nonetheless, the empirical results are a further evidence in line to our theoretical expectation and they result to be stable to several robustness check.

Overall, this paper contributes to a better understanding of the coexistence of heterogeneous firms characterized by different levels of efficiency and different organizational forms. Focusing on core competences and leaving some phases of the production to the 'outside' —that has been documented as one of the most relevant business practice in the last decades (see the evidence provided by Feenstra, 1998; Grossman and Helpman, 2005, among others)— may be a rational choice for less efficient firms in order to make positive operating profits and stay in the market. On the other hand, more efficient firms could exploit their efficiency advantage to control a greater part of the production chain in order to benefit from greater coordination among different phases and tailored intermediate inputs[35]. From a methodological point of view, the stochastic frontier framework allows us to estimate firm inefficiency as the distance from the technology frontier (the best practice) and to jointly estimate the relationship between the degree of vertical integration and inefficiency. This can be considered as an improvement with respect to previous works on the same topic, which rested on more traditional 2-step procedures which may lead up to omitted variable bias and under-dispersion of productive efficiency scores in the first step of the analysis.

Among the lines for future research, we highlight the following issues:

- A qualitative analysis of a small number of firms in the industry could be a natural complement to this study: the vertical organization heterogeneity that

[34] A greater coordination in the production process, a reduction in the transaction costs and the possibility of an optimal amount of specific investments have been advanced as key factors which may enhance the performance of a firm which becomes vertical integrated.

[35] This could further enhance the efficiency advantage of the most integrated firms, but we cannot asses this directly through our econometric analysis.

we detected through our econometric analysis could be grounded in a careful description of the stages of the production process which are actually kept in-house.

- Some econometric refinements may be possible. One of them is related to the 'simultaneity' problem, which, in our case, could stand for a reverse causality, from vertical integration to firm efficiency.

Chapter 4

Foreign Investments and Productivity: Evidence from European Regions[1]

4.1 Introduction

Regional competitiveness and social and economic cohesion have been crucial concerns for policy makers —especially in the European Union (EU)[2]— and have attracted a considerable amount of economic research. In particular, empirical works have focused on explaining differences in productivity among EU regions. Agglomeration economies, technology and human capital have been most often considered as the key dimensions to explain such differences[3]. With the notable exeception of Gambardella, Mariani, and Torrisi (2008), internationalization is rarely considered as a factor affecting regional productivity. This is probably due to the lack of accurate measures of a region's openness[4]. This lack of evidence is at odds with the increasing relevance of regions in the global economy, and in Europe in particular. As Krugman (1993) puts it, with the free movement of goods, capital and labour, it makes less and less sense to think about economic relations within Europe in terms of the standard paradigm of international trade. One should rather take a regional perspective and emphasize relations of sub-national units within the EU and with the rest of the world.

In this work, using a novel dataset on international investment projects, we are able to build unique measures of outward and inward foreign direct investment (FDI)[5]

[1]This Chapter draws on a joint work with Davide Castellani (University of Perugia).

[2]As documented by Fiaschi, Lavezzi, and Parenti (2009), 35% of the EU budget for the period 2007-2013 has been allocated to promote social and economic cohesion among the regions of its member states.

[3]See for example the empirical evidence on EU regions in Ciccone (2002), Paci and Usai (2000), Dettori, Marrocu, and Paci (2008).

[4]In fact, Gambardella, Mariani, and Torrisi (2008) introduce a generic measure of openness using the share of hotels in the population and the share of the population which speaks a second language.

[5]Following the definitions provided by the Organization for Economic Co-operation and Development

at the regional level (NUTS 2) for the countries of the Enlarged Europe (EU-27). This allows us to assess – for the first time – the extent to which regional productivity is associated with internationalization, and in particular with foreign investments by multinational enterprises (MNEs). It is worth mentioning that the European Union (EU) is a major home and host territory for FDIs. In particular, both inward and outward FDIs [6] are relevant in the EU: they account for almost the 4% of the EU GDP, but with very differentiated patterns across countries. For example outward FDIs, as a share of GDP, go from values close to zero in most New Member States, to around 1% in countries such as Italy and Greece and more than 5% in the UK, France and Spain; on the other hand, inward FDIs range from around 1% of GDP in Greece, Italy and Germany, to more than 10% of GDP in Bulgaria, Belgium and Estonia. Empirical works have also documented that inward FDIs are not uniformly distributed across regions within individual countries (Head and Mayer, 2004; Basile, Castellani, and Zanfei, 2008). Instead, evidence is lacking on the propensity of European regions to engage in outward FDIs and on how this relates to different regional productivity dynamics.

In order to investigate whether foreign investments actually affect regional productivity, we estimate regressions of (one-year) productivity growth as a function of one-year-lagged foreign investments. We find that inward FDIs have a positive and significant effect on regional productivity growth, but this effect is sizable only for relatively large number of investment projects. Conversely, outward FDIs are positively associated with productivity growth, even if this effect fades down with the number of projects, and may eventually become negative in regions with very large outward flows. These results are robust to a number of controls. In particular, we have added several regional characteristics (both in level and in growth rate), allowed for different technologies between regions belonging to the Old Member States and those which belong to the New Member States and accounted for spatial dependence.

This piece of evidence bears important implications for policy. In particular, it suggests, on the one hand, that fears of hollowing-out as a consequence of outward investments are not entirely founded, and local economies may in fact benefit from the fact that incumbent firms move some production abroad, and, on the other hand, that substantial investments may be needed to attract a amount of foreign investment sufficient to generate sizable effects on regional productivity growth.

The rest of the paper is organized as follows: Section 4.2 explains the theoretical

(OECD, 1996) and by the International Monetary Fund (IMF, 1993), a foreign direct investment is an investment in a foreign company which amounts to (at least) the 10% of the ordinary shares of the target company, and which aims at controlling it. Usually, FDIs entail a participation in the management of the controlled firm, which is frequently supported by the transmission of expertises and by the transfer of a part of the knowledge and the technology by the parent company. Firms involved in FDIs are known as multinational firms.

[6] Inward investments refer to incoming flows in a region/country, made by foreign companies, while outward investments are made by local companies investing abroad.

background and the link between foreign direct investments and regional productivity; Section 4.3 describes the empirical strategy we have set up in order to assess the effect of foreign investments on the productivity of EU regions; Section 4.4 details the characteristics of the original database, which has been recovered from different sources, then focusing on how the main variables of the analysis have been measured and Section 4.5 provide some descriptive evidence on them; Section 4.6 pass through the econometric results and the robustness checks which have been performed in order to validate the baseline results. Finally, Section 4.7 discuss the main results of the paper, underlying the novelties of the work and the policy implications. Two Data Appendixes, 4.8 and 4.9 follow.

4.2 Theory: foreign investments and productivity

From a theoretical point of view, the links between foreign investments and productivity of home and host countries are well known[7]. Extensive works have been done regarding the direct and indirect effects of inward foreign direct investments (FDI) on host economies. Direct effects refer to the fact that incoming multinationals tend be relatively more productive than domestic firms (Griffith, 1999; Harris and Robinson, 2002; Benfratello and Sembenelli, 2006), and to concentrate in higher productivity sectors (see Brainard, 1997, among others). Thus, entry of foreign multinationals changes the composition of the host economy –both within and between sectors– contributing to increasing aggregate productivity. Foreign multinationals may also have indirect effects, inducing pecuniary (Scitovsky, 1954; Görg and Strobl, 2005) and technological externalities (Blomström and Kokko, 1998; Lipsey, 2002; Castellani and Zanfei, 2006) but also determining a business stealing effect[8]. While the former usually have positive contribution to aggregate productivity, the latter may have opposite effects. Foreign multinationals usually possess some ownership advantages which make them more competitive than local firms (Dunning, 1993) which may in turn be forced to shrink their market share or exit from the market upon entry of foreign investors. To the extent that local firms are less productive than the foreign ones, this process may be beneficial for the aggregate productivity. On the other hand, if the sector is characterized by economies of scale, local firms which experience a shrinkage in market shares may increase their average costs, thus lowering their competitiveness (Aitken and Harrison, 1999): this phenomenon may have a negative effect at the aggregate level, at least in the short-run. Furthermore, if foreign multinationals keep only the low value added activities in the host region, while domestic firms carried out the whole production process in the region, the crowding-out effect may be detrimental for aggregate productivity dynamics.

[7]See Barba Navaretti and Venables (2006) for a recent review.

[8]It refers to crowding-out effect, that is the internationalized firm's expansion of its market shares at the expenses of its domestic competitors.

Outward investments have direct and indirect effects on the productivity of the home economy too. As for the direct effects, firms engaging in foreign activities (either through export or foreign investments) are more productive than purely domestic ones, since they need to overcome the cost of doing business abroad (Helpman, Melitz, and Yeaple, 2004). Thus, regions with a higher share of highly productive firms will also be more internationalized. Furthermore, by investing abroad firms may be able to access foreign knowledge and reaping the benefit of higher economies of scale (Cantwell, 1995; Fosfuri and Motta, 1999; Petit and Sanna-Randaccio, 2000). This will increase their knowledge capital and boost their growth, which will in turn contribute to raising aggregate productivity. Admittedly, outward investments may also be associated with a decrease in the size and productivity of home activities. This would occur when domestic firms relocate a substantial share of production, R&D or other activities. In this case, the competitiveness boost may not be able to compensate the offshored value-added activities.

Indirect effects associated with outward investments may also have consequences on the performance of local firms, which can contribute to a decrease or an increase in the aggregate productivity. On the one hand, an increase in size, productivity and/or knowledge of home multinationals may spill-over on other domestic firms through input-output relations and imitation. On the other hand, to the extent that investing firms move value-added creating activities, domestic suppliers along the value chain may be forced to shrink or to exit. At the same time, opportunities may arise in upstream or downstream sectors, for example in activities like logistics, R&D, design, and other business services. The overall effect of this process on aggregate productivity may be positive or negative, according to the balance between the productivity of firm entering (or increasing the market share) and exiting the market (or shrinking).

Overall, theoretical results do not predict clearcut effects of foreign direct investments on aggregate productivity. This boils down to empirical analyses in order to investigate the 'sign' of the relationship between foreign investments and productivity of local economies. Empirical works on inward FDIs and productivity have provided sound evidence that the entry of MNEs in a given territory is associated with a positive direct contribution to the productivity of host economies; moreover, multinational firms contribute to changes in the industrial mix towards relatively more knowledge and technology intensive sectors. Evidence on indirect effects is more mixed, and it seems to depend both on the characteristics of the multinational investments and those of firms in the host economy (Castellani and Zanfei, 2006). Econometric evidence on inward FDIs and productivity have been provided mainly with firm-level studies on one (or more) countries and with more aggregate cross-country studies (Barba Navaretti and Venables, 2006). A few empirical works have also taken a regional perspective within individual countries[9], but cross-country evidence of the effects of inward FDIs at a sub-national level is still lacking. This is unfortunate, considering that in the last decades stiffer competition have emerged among local territories (both within and across national boundaries) to attract foreign investors (Basile, Castellani, and Zanfei, 2008).

The literature on outward investments and productivity is more scattered, but has gained momentum in the last decade. Many studies in this field have provided evidence that firms investing abroad tend to be more productive than their home country counterparts (Greenaway and Kneller, 2007; Castellani and Zanfei, 2007): these results would predict that in regions with a larger share of highly productive firms (thus a higher average productivity) one would observe a higher number of firms investing abroad. Other studies have found that investing abroad may further reinforce productivity of investing firms (Debaere, Lee, and Lee, 2006; Hijzen, Inui, and Todo, 2007; Barba Navaretti, Castellani, and Disdier, 2010), while only a few works in this literature have addressed the indirect effects that firms investing abroad may have on their home country (Castellani and Zanfei, 2007; Vahter and Masso, 2007).

At the aggregate level, a small number of studies have been conducted on the relation between outward FDIs and productivity, and they also show mixed results. For example, van Pottelsberghe de la Potterie and Lichtenberg (2001), in a panel of 13 developed countries, find that outward investments are a more effective channels for international technology transfer among countries with respect to inward FDIs, while Braconier, Ekholm, and Knarvik (2001) find no effects of outward FDIs on domestic productivity in Sweden. More recently, Driffield, Love, and Taylor (2009) find that outward FDIs is positively related to productivity growth in UK, while Bitzer and Görg (2009), who examine the effect of outward and inward FDIs on domestic total factor productivity for 17 OECD countries, report that only the latter are positively related to a country productivity. Herzer (2010) find that outward FDIs have, on average, a positive long-run effect on total factor productivity in developing countries. To the best of our knowledge there are no studies at the sub-national level regarding the effects of outward FDIs on the productivity of local economies[10].

Overall, the regional level seems particularly appropriate to assess compositional as well as indirect (but geographically confined) effects of inward and outward investments.

[9]For example Altomonte and Colantone (2009) on Romanian regions, Driffield (2004) on UK regions.

[10]Mariotti, Mutinelli, and Piscitello (2003) analyze the effects of outward FDIs on the employment in the Italian regions, without a cross-country perspective and, furthermore, without analyzing the effects on the aggregate productivity.

4.3 The empirical model

In order to assess the effect of inward and outward foreign direct investments on regional productivity we start from the following econometric model:

$$y_{ij,t} = \gamma_{OUT} OFDI^{stock}_{ij,t-1} + \gamma_{INW} IFDI^{stock}_{ij,t-1} + \beta kl_{ij,t} + \mathbf{x_{ij,t}} \boldsymbol{\delta} + \mu_i + t \cdot \eta_j + \tau_t + \epsilon_{ij,t}, \quad (4.1)$$

where $y_{ij,t}$ is the (log of the) labour productivity of the ith region in the jth country at time t, and $OFDI^{stock}_{ij,t-1}$ and $IFDI^{stock}_{ij,t-1}$ are, respectively, (log of) the stocks of outward and inward foreign direct investments in the ith region at the $t-1$ time period. We make the hypothesis that foreign direct investments affect productivity with one-year lag[11]. We include a set of regional characteristics that economic theory has indicated as determinants of productivity and which are likely to be correlated with inflows and outflows FDI in European regions. Thus, the model is augmented with $kl_{ij,t}$, which indicates the (log of the) capital-labour ratio and $\mathbf{x_{ij,t}}$, which is a vector of (the log of) other regional characteristics, such as the level of human capital, the stock of technological capital, the regional industrial composition and the degree of concentration/diversification of the regional industry. We include a vector of regional effects, μ_i, to control for unobserved (and time invariant) regional characteristics which could be correlated both with the stocks of foreign direct investments (incoming or outgoing from the region) and with the regional productivity; a vector of time effects, τ_t, to control for factors affecting all regions in the same way in a given year; a set of country-specific interactions, $t \cdot \eta_j$, in order to capture the country-specific trends in labour productivity, which could be due, for example, to institutional characteristics affecting not only the level of productivity, but also the growth rate (Nicoletti and Scarpetta, 2003). The model has a four-parts error structure, and it allows for unobserved regional effect μ_i to be correlated with the foreign direct investment variables, $OFDI^{stock}_{ij,t-1}$ and $IFDI^{stock}_{ij,t-1}$, and the other regional characteristics $k_{ij,t}$ and $\mathbf{x_{ij,t}}$.

The choice of the control variables is based on previous theoretical and empirical works. We cross refer the reader to the Data Appendix 4.8 for a detailed discussion on the control variables and their measurement.

The model can be estimated either by means of the within-estimator or by the first-differenced estimator; we have chosen the second one, because of a constraint on available data. In fact, we have information on flows of foreign investments over the period 2003-2008. We could apply the PIM to this series and recover the stock of foreign investments but, in order to have a sensible measure of FDI stock, we would need to sum up at least 3 to 5 years of investments, and this would leave with no more than a cross-section. The obvious drawback of this solution is that we could

[11]This is explicitly tested against the hypothesis that FDI have a contemporaneous effect on productivity in Section 4.6.

not account for the unobserved heterogeneity which is likely to affect both regional productivity levels and FDI stocks. Consequently, the first differenced estimator seems the natural candidate with respect to our dataset[12].

The first differenced equation can be written as

$$\Delta y_{ij,t} = \gamma_{OUT} \Delta OFDI^{stock}_{ij,t-1} + \gamma_{INW} \Delta IFDI^{stock}_{ij,t-1} + \beta \Delta kl_{ij,t} + \Delta \mathbf{x_{ij,t}} \boldsymbol{\delta} + \eta_j + \tau_t + \Delta \epsilon_{ij,t}, \tag{4.2}$$

where Δ indicates the difference between the variable at time t and the variable at time $t-1$. With respects to the variables measuring foreign direct investments, differences are computed between the variable at time $t-1$ and the variable at time $t-2$.

The relationship between investments stocks and flows can be formalized, with some approximation, in the following way[13]:

$$\Delta_{(t-1,t-2)} OFDI^{stock}_{ij} \cong OFDI^{flows}_{t-1},$$

and

$$\Delta_{(t-1,t-2)} IFDI^{stock}_{ij} \cong IFDI^{flows}_{t-1}.$$

Knowing this fact, the differenced in Equation 4.2 can be re-written as

$$\Delta y_{ij,t} = \gamma_{OUT} OFDI^{flows}_{ij,t-1} + \gamma_{INW} IFDI^{flows}_{ij,t-1} + \beta \Delta kl_{ij,t} + \Delta \mathbf{x_{ij,t}} \boldsymbol{\delta} + \eta_j + \tau_t + \Delta \epsilon_{ij,t}. \tag{4.3}$$

Equation 4.3 has an appealing interpretation in our case, even besides the unobserved effects model illustrated in Equation 4.1: the parameters γ_{OUT} and γ_{INW}, which are the main focus of this work, explicitly consider the relationship between outward and inward flows of investments and the growth rate of the labour productivity.

4.4 Data

4.4.1 Data sources

We exploit an original database, which has been compiled recovering data from different sources. Data refer to the NUTS 2 level for the EU regions: this level of analysis has been chosen for four main reasons. First of all, it is suitable for taking into account the within-country heterogeneity (in terms of labour productivity, foreign direct investments and the other observed and unobserved characteristics); second, this sub-national level of analysis makes the appraisal of the indirect/compositional

[12]Of course, we are aware that the within-estimator is more efficient with respect to the first-differenced estimator if $\epsilon_{ij,t}$ are serially correlated, but the former is not a viable alternative due to data constraints.

[13]The approximation is due to the fact that change in the stock is given by the flow of investments plus the depreciation of the existing capital stock. Unfortunately the lack of the stock of investments forces us to rely on the approximation illustrated in the text.

effects of FDIs on regional productivity possible; third, it allows for comparable units across different countries; finally, a significant amount of information are available at this level of disaggregation.

Information on regional gross value added come from the *EU Regional Database* developed and maintained by Eurostat[14], while data on employment and capital investments at the regional level come from the *European Regional Database*, developed by Cambridge Econometrics (release 2006). We have used these information in order to build a measure of labour productivity and a measure of the capital-labour ratio at the regional level. Data on independent variables of main interest, i.e. outward and inward FDIs, come from *fDI Markets* an online database maintained by fDi Intelligence —a specialist division of the Financial Times Ltd—, which monitors crossborder investments covering all sectors and countries worldwide[15]. Only greenfield investments are recorded in the fDI Markets database: this is a particular class of investments which relates to the set-up of new company-related facilities and, consequently, mergers and acquisitions or mere expansions of existing companies are not considered. However, in Section 4.4.3 we provide evidence for a strong concordance between the patterns of greenfield investments in our database and the patterns of all classes of FDIs (both greenfield and other types, like M&A investments) in the European Countries. In this sense, we are confident about the representativeness of the employed data with respect to the total flows of FDIs in the European territories.

Gross valued added have been deflated using the series of price indexes which are available in the *Growth and Productivity Accounts* database developed by EU KLEMS[16] (releases 2008 and 2009). For further details on how each variable has been built we cross-refer the reader to Sections 4.4.2, 4.4.3 and to the Data Appendix 4.8.

4.4.2 Labour Productivity

The dependent variable is the labour productivity, which has been computed as the ratio of the regional gross valued added (VA_{ijt}) (at basic prices in millions of euro) obtained from the Regio database, to employment (thousands) in all sectors of the regional economy (L_{ijt}), which has been recovered from the European Regional Database. We have taken into account likely variations in prices during the considered period, multiplying the value added series by a deflator built using the series of price index (1995=100) of the value added ($I_{VA1995,jt}$) at the national level taken from the EU KLEMS database (release 2009). Given that index series for the gross value added are not available at the regional level, employing national level deflators has

[14]See the Eurostat web page
`http://epp.eurostat.ec.europa.eu/portal/page/portal/region_cities/`.
[15]See the web-page of the fDi Markets at
`http://fdiintelligence.com/index.cfm?page_name=markets`
[16]See the web page of the EU KLEMS project at `http://www.euklems.net/`

been considered as the best option to cope with variations in prices. Thus, for each region i, belonging to country j, the labour productivity at time t has been computed as

$$Y_{ijt} = \frac{VA_{ijt}}{L_{ijt}} \cdot \frac{100}{I_{VA1995,jt}}. \tag{4.4}$$

The last year for which information on value added are available in the Regio database is 2006. The variable has been included in logs in the performed econometric analysis, y_{ijt} [17].

4.4.3 Foreign investments

Data on inward and outward foreign direct investments flows ($IFDI^{flows}_{ijt}$, $OFDI^{flows}_{ijt}$) have been recovered from the fDi Markets database. This source tracked 60,301 world-wide greenfield investments projects announced by MNEs, in the period 2003-2008. The database collects information on the announced projects year by year. Each entry is a project, i.e. the investment has not been completed yet, but the database is carefully updated each year in order to check if projects have been actually 'completed' or not, and, in case, they are deleted from the database[18]. Projects, which are collected in the fDi Markets database, regard the major business activities and all the industrial sectors in which MNEs operate. For each project, information on the company which has undertaken the investment, the source and destination area (namely region, state and country), the business activity and the industry in which the investment has been made are available in the database. Thus, it is possible to count the number of inward and outward investment projects for each region in each year of the period under analysis (2003 to 2006), and that is the proxy for foreign direct investments flows:

$$wFDI^{flows}_{ijt} = \#\text{of projects in region } i \text{ belonging to country } j\text{, in year } t,$$

where $w = \{I, O\}$, are respectively inward and outward investments.

[17] Some remarks on the labour-productivity indicator should be made. First, data on the regional employment are drawn from the European Regional Database (ERD). We chose to use this source, since the employment series of the Regio database has a higher number of missing values which would have decreased the set of regions under analysis. The downside of this choice is that in the version of the ERD available to us, values for 2005 and 2006 were forecast. However, we checked that correlation with the actual (non missing) values, reported by the more updated Regio dataset is very high (0.95). Second, in order to build deflators for regions belonging to Cyprus, Estonia, Latvia, Lithuania and Malta (which are actually all single-region country) we have used the series of price index in the previous release of the EU KLEMS database (2008) given that they were not available in the last release yet. Third, for Bulgarian and Romanian regions we have used the 'Eurozone' series of price index, given that the national series were not available in the database.

[18] In this sense, data on the projects related to the first years of the series could be more reliable than the data regarding the last years of the series. As a matter of fact, we actually cannot use the last two years of data, so we are quite confident that our data on FDI project reflect realized projects. Furthemore, we show that the distribution of investment projects by European countries registered from the fDi Markets database are in line with the evidence –reported by the United Nations Conference on Trade and Development (UNCTAD)– on the actual FDIs flows in the same period.

In order to assess the reliability and the scope of the information which are available in the fDI Markets database, we provide some figures regarding some general patterns which can be found in our data. We can compare the number of investment projects in the database with the data provided by UNCTAD on FDIs flows at the country-level[19]. The high correlation coefficients (0.82 and 0.83) between the two pairs of series reassure us data on investment projects are actually a good proxy for FDI flows. A careful inspection reveals that the number of projects overestimates inward FDIs to some New Member States, such as Poland, Romania, Bulgaria, Hungary and Czech Republic, probably due to the fact such investments are relatively low capital-intensive. As can be seen from Table 4.1, almost 90% of EU outward investments are made from EU-15 countries[20], while inward investments are split more evenly among EU-15 and EU-12 countries[21].

United Kingdom, Germany and France result to be the leading countries both in terms of inward and outward FDIs in the period which goes from 2003 to 2006, which is the period under analysis. As for the inward investments, some New Member States (EU-12), like Poland, Romania, Hungary, Czech Republic and Bulgaria show a good performance in attracting foreign direct investments. Given this evidence, we use data on the projects as a proxy for foreign investments, and we refer to them simply as 'foreign direct investments'. More information on the patterns of investment projects can be found in the Data Appendix, Section 4.9.1.

4.5 Descriptive analysis

The time structure of the data imposes some constraints to the empirical analysis. In particular, regional productivity is observed only up to 2006, while information on foreign investments are available for the period 2003-2008. Thus, if we want to assess the econometric relationship between the latter and the former, we are left with four years of data: 2003, 2004, 2005 and 2006. Due to the lack of the information regarding some regional characteristics, regions belonging to Norway, Switzerland and Denmark cannot be taken considered. See the Data Appendix 4.9.3 for the detailed list of regions with all the relevant variables, that have been considered in the econometric analysis. In order to save space we have listed some descriptive statistics at the country level, while reproducing visual representations —maps— to provide information of the main characteristics of the regions under analysis. Moreover, given that we used a first differenced estimator, and given that the results of the econometric analysis can be interpreted as the effects of the investments flows on

[19]The comparison cannot be done at the NUTS 2 level, because data on FDIs flows are not available at that level of disaggregation.

[20]Italy, France, Netherlands, Luxumbourg, Belgium, Germay, United Kingdom, Denmark, Ireland, Greece, Spain, Portugal, Austria, Switzerland and Finland.

[21]Bulgaria, Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovakia and Slovenia.

Table 4.1: fDi Markets projects vs. UNCTAD Flows, 2003-2006

Outward			Inward		
Country	# proj.	flows	Country	# proj.	flows
Germany	22.2	11.7	United Kingdom	16.0	25.8
United Kingdom	20.3	16.3	France	9.2	15.2
France	13.8	17.6	Germany	8.3	8.1
Italy	6.3	5.7	Poland	6.5	3.0
Netherlands	5.9	13.7	Spain	6.2	7.2
Sweden	5.9	4.7	Romania	5.9	1.7
Austria	5.1	2.0	Hungary	5.4	1.4
Spain	4.6	11.7	Czech Republic	4.1	1.5
Finland	3.1	0.3	Bulgaria	4.1	1.1
Belgium	2.5	7.9	Ireland	4.1	-1.6
Denmark	1.9	1.4	Italy	3.9	5.9
Ireland	1.4	2.7	Sweden	3.2	3.4
Slovenia	1.1	0.1	Netherlands	3.1	5.1
Greece	0.9	0.4	Belgium	2.9	10.8
Latvia	0.9	0.0	Slovakia	2.6	0.8
Estonia	0.6	0.1	Lithuania	2.4	0.2
Portugal	0.5	1.2	Austria	2.2	1.9
Luxembourg	0.5	1.0	Denmark	1.9	1.2
Poland	0.5	0.7	Latvia	1.7	0.2
Czech Republic	0.5	0.1	Estonia	1.5	0.4
Hungary	0.4	0.4	Portugal	1.3	1.5
Lithuania	0.4	0.0	Greece	1.1	0.6
Cyprus	0.2	0.1	Finland	0.9	1.2
Romania	0.2	0.0	Slovenia	0.8	0.2
Slovakia	0.1	0.0	Luxembourg	0.4	2.7
Bulgaria	0.1	0.0	Cyprus	0.3	0.3
Malta	0	0.0	Malta	0.2	0.2
Total	100	100		100	100
Pearson corr. coefficient	0.82			0.83	

the labour productivity growth rates we report descriptive statistics for the variables both in levels and in growth rates. Table 4.2 provides some basic statistics for the variables used in the econometric analysis. With respect to foreign direct investments, the first column reports the cumulative number of outward and inward investments. This number is lower than the overall number of investments recorded by fDi Markets for the European regions, due to the fact that for some projects no information on the source or destination regions are available. On average, from each region about 14 outgoing investments and 10 incoming investments per year occur. However, the distribution of the number of investments is highly skewed: from more than 25% of regions no outward investment in one year would originate and more than 10% would not attract any inward investment.

Figure 4.1 provides a visual representation of the geographical distribution of the number of such investment projects at the regional (NUTS 2) level.

From Table 4.2, the skewness of the foreign investments variables is evident. This induce us to model their effect as a combination of a dummy taking value equal to '0' for region-year where no investments have taken place and a continuous variables taking the value equal to the log of the number of investments in the case of non-zero investments, and '0' otherwise[22]. In other words, investments variables enter the regressions as follow:

$$\mathrm{w}FDI(d)_{i,t} = \begin{cases} = 1 & \text{if \# of projects } {}^{w}_{i,t} > 0 \\ = 0 & \text{if \# of projects } {}^{w}_{i,t} = 0 \end{cases}$$

$$\mathrm{w}FDI(log)_{i,t} = \begin{cases} = log(\text{\# of projects } {}^{w}_{i,t}) & \text{if \# of projects } {}^{w}_{i,t} > 0 \\ = 0 & \text{if \# of projects } {}^{w}_{i,t} = 0 \end{cases}$$

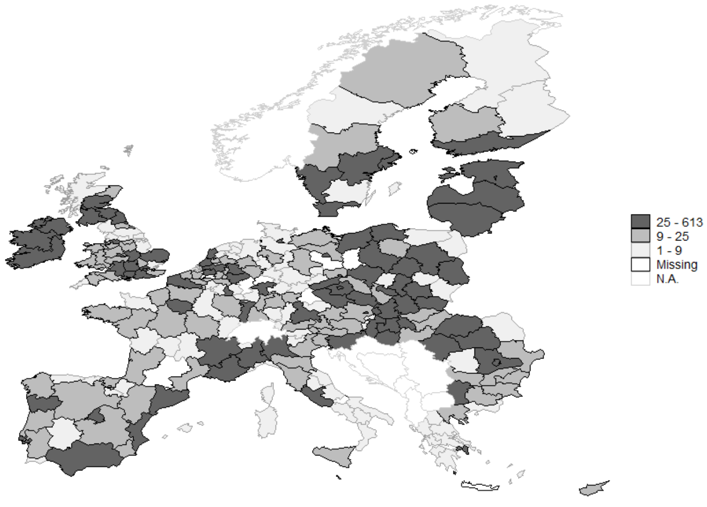
where $w = \{I, O\}$ are respectively inward and outward investments. In this way it is possible to distinguish the effect (for the region) of being generally involved in the internationalization process, which is captured by the dummy variable, from the effect of the intensity of the internationalization phenomenon, which is captured by the continuous variable in logs. Figure 4.2 provides a graphical representations of the variables measuring the labour productivity in levels and growth at the NUTS 2 level. Labour productivity (4.2) is clearly higher in the core regions of the EU-15, while it declines in Southern European regions and reach minimum values in the regions of EU-12 countries. As for the growth rates of labour productivity, Figure 4.3(b) shows that rather similar patterns are observed in the regions belonging to the same country, an this is true in particular in EU-12 countries, Italy, France and Spain, while Germany and UK show a much greater within-country variability. At

[22]We take the log of the number of investments so that we can interpret the coefficient of the continuous variable as an elasticity.

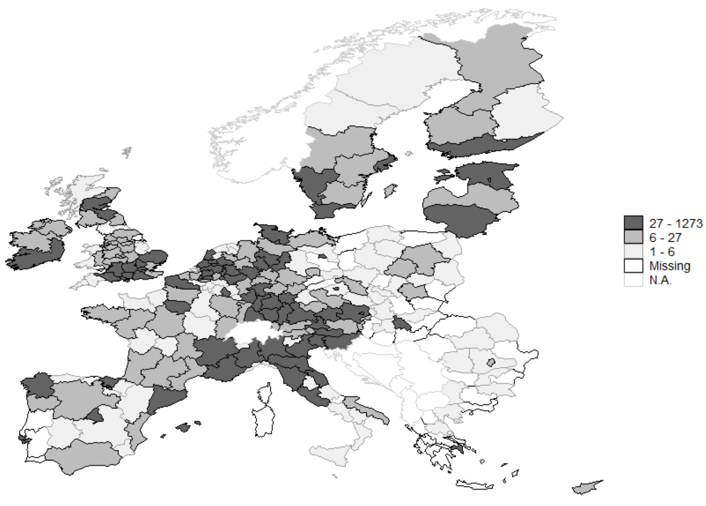
Table 4.2: Descriptive statistics, 2003-2006

Variable	Notation	Unit	Count	Obs.	Mean	Std. Dev	p10	p25	p50	p75	p90
Outward FDI	OFDI	count	14135	1032	13.697	39.710	0	0	3	11	30
Inward FDI	IFDI	count	10802	1032	10.467	19.825	0	1	4	12	24
Labour productivity	y	ratio (log)		1017	3.360	0.751	1.956	3.202	3.651	3.856	3.948
Capital-labour ratio	kl	ratio (log)		1036	4.148	0.863	2.714	3.949	4.387	4.753	4.923
Human capital	$hcap$	ratio (log)		1010	-1.468	0.378	-2.040	-1.728	-1.403	-1.189	-1.035
Herfindahl index	hhi	formula (log)		922	-1.377	0.177	-1.602	-1.514	-1.391	-1.246	-1.144
Innovation stock	$tech$	formula (log)		1036	-0.992	1.859	-3.721	-2.360	-0.416	0.397	0.982
Share of other industies	SH_EF	share		922	0.089	0.023	0.062	0.072	0.084	0.101	0.119
Share of High-tech man.	SH_HT	share		922	0.066	0.035	0.028	0.043	0.060	0.084	0.112
Share of Low-tech man.	SH_LT	share		922	0.125	0.046	0.068	0.088	0.122	0.153	0.191
Share of KI svcs	SH_KIS	share		922	0.316	0.088	0.212	0.254	0.309	0.379	0.431
Share of LKI svcs	SH_LKIS	share		922	0.336	0.047	0.280	0.312	0.338	0.364	0.392
Labour productivity-growth rate	$\Delta_{(t,t-1)}y$	ratio (log, differences)		1017	0.020	0.044	-0.017	0.003	0.017	0.034	0.059
Capital-labour ratio-growth rate	$\Delta_{(t,t-1)}k$	ratio (log, differences)		1036	0.022	0.026	-0.003	0.007	0.018	0.032	0.053
Human capital-growth rate	$\Delta_{(t,t-1)}hcap$	ratio (log, differences)		1002	0.039	0.072	-0.037	-0.002	0.036	0.074	0.119
Herfindahl index-growth rate	$\Delta_{(t,t-1)}hhi$	formula (log, differences)		891	0.009	0.035	-0.033	-0.009	0.008	0.028	0.048
Innovation stock-growth rate	$\Delta_{(t,t-1)}tech$	formula (log, differences)		1036	0.047	0.156	-0.079	-0.018	0.033	0.081	0.174
Share of other industries-growth rate	$\Delta_{(t,t-1)}$ SH_EF	share (differences)		891	0	0.009	-0.010	-0.004	0.001	0.006	0.012
Share of High-tech man.-growth rate	$\Delta_{(t,t-1)}$ SH_HT	share (differences)		891	0	0.009	-0.011	-0.006	-0.001	0.004	0.009
Share of Low-tech man.-growth rate	$\Delta_{(t,t-1)}$ SH_LT	share (differences)		891	0	0.011	-0.016	-0.009	-0.002	0.003	0.009
Share of KI svcs.-growth rate	$\Delta_{(t,t-1)}$ SH_KI	share (differences)		891	0	0.015	-0.012	-0.003	0.004	0.013	0.022
Share of LKI svcs.-growth rate	$\Delta_{(t,t-1)}$ SH_LKIS	share (differences)		891	0	0.016	-0.017	-0.009	0.002	0.010	0.020

Figure 4.1: Regional distribution of international investment projects, 2003-2006

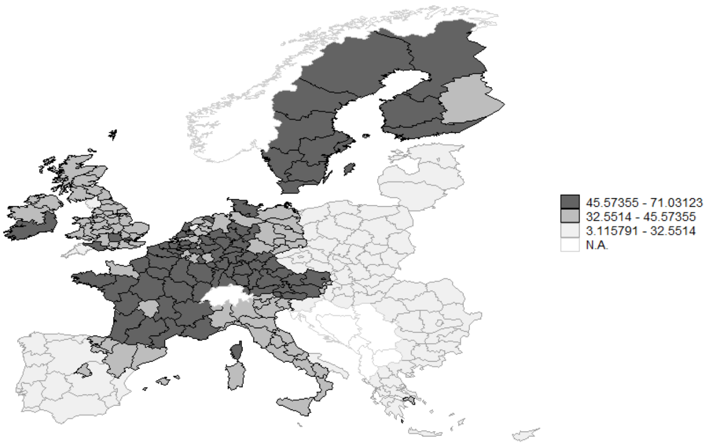


(a) Inward investments

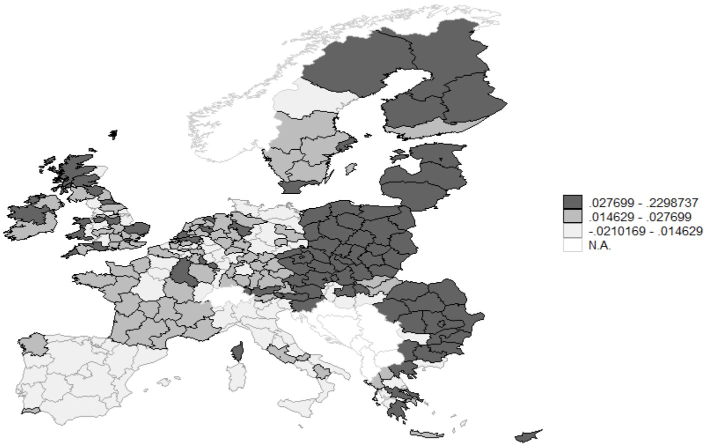


(b) Outward investments

Figure 4.2: Regional patterns of labour-productivity level and growth, 2003-2006 (average)



(a) Labour productivity (level)



(b) Labour productivity (growth)

the same time, higher growth rates are observed in EU-12 countries, supporting the hypothesis that some convergence is going on, but this does not appear as a common patterns, since in two relatively low productivity countries, such as Italy and Spain, growth rates are still below the median.

These insights are confirmed by results presented in Tables 4.3 and 4.4[23].

Table 4.3: Growth rates by country, EU15, 2003-2006

	Δy	Δkl	$\Delta hcap$	Δhhi	$\Delta tech$	ΔSH_EF	ΔSH_HT	ΔSH_LT	ΔSH_KIS	ΔSH_LKIS
Austria	0.022	0.015	0.018	0.001	0.039	0.000	0.001	-0.003	0.001	0.000
	(0.014)	(0.009)	(0.092)	(0.046)	(0.047)	(0.014)	(0.007)	(0.009)	(0.013)	(0.017)
	36	36	36	36	36	36	36	36	36	36
Belgium	0.014	0.014	0.021	0.002	0.009	0.001	-0.001	-0.002	0.002	-0.001
	(0.012)	(0.011)	(0.054)	(0.030)	(0.063)	(0.007)	(0.008)	(0.007)	(0.014)	(0.015)
	44	44	44	40	44	40	40	40	40	40
Germany	0.014	0.003	0.018	0.013	0.022	-0.002	-0.002	-0.002	0.005	0.000
	(0.013)	(0.015)	(0.059)	(0.036)	(0.052)	(0.009)	(0.011)	(0.010)	(0.012)	(0.016)
	156	156	148	140	156	140	140	140	140	140
Denmark	0.018	0.039	.	.	0.044	.	.	.	.	.
	(0.017)	(0.010)	.	.	(0.066)	.	.	.	.	.
	12	12	0	0	12	0	0	0	0	0
Spain	0.008	0.015	0.032	0.012	0.074	0.002	-0.002	-0.004	0.007	0.000
	(0.009)	(0.008)	(0.053)	(0.025)	(0.137)	(0.009)	(0.005)	(0.009)	(0.013)	(0.011)
	68	68	68	66	68	66	66	66	66	66
Finland	0.025	0.032	0.023	0.016	-0.037	0.000	-0.001	-0.003	0.006	0.000
	(0.016)	(0.010)	(0.070)	(0.016)	(0.117)	(0.004)	(0.005)	(0.006)	(0.005)	(0.010)
	20	20	20	16	20	16	16	16	16	16
France	0.017	0.024	0.031	0.016	0.027	0.000	-0.003	-0.004	0.004	0.002
	(0.014)	(0.010)	(0.107)	(0.054)	(0.066)	(0.012)	(0.012)	(0.017)	(0.022)	(0.023)
	88	88	88	84	88	84	84	84	84	84
Greece	0.021	0.066	0.062	0.008	0.092	0.002	0.000	-0.004	0.006	0.005
	(0.034)	(0.036)	(0.086)	(0.032)	(0.242)	(0.008)	(0.007)	(0.010)	(0.014)	(0.011)
	52	52	52	23	52	23	23	23	23	23
Ireland	0.032	0.081	0.047	0.017	0.023	0.007	-0.003	-0.005	0.004	0.002
	(0.023)	(0.013)	(0.037)	(0.015)	(0.075)	(0.005)	(0.004)	(0.006)	(0.006)	(0.009)
	8	8	8	8	8	8	8	8	8	8
Italy	0.005	0.016	0.056	0.008	0.019	0.001	0.000	-0.004	0.008	-0.002
	(0.019)	(0.011)	(0.060)	(0.029)	(0.084)	(0.009)	(0.006)	(0.010)	(0.020)	(0.015)
	84	84	84	80	84	80	80	80	80	80
Luxembourg	0.026	0.022	0.057	0.018	0.051	-0.003	0.001	-0.005	0.013	-0.005
	(0.022)	(0.003)	(0.329)	(0.017)	(0.039)	(0.003)	(0.002)	(0.003)	(0.011)	(0.009)
	4	4	4	4	4	4	4	4	4	4
Netherlands	0.027	0.034	0.049	0.008	0.006	0.000	-0.004	0.002	0.003	0.000
	(0.019)	(0.030)	(0.056)	(0.029)	(0.063)	(0.007)	(0.006)	(0.008)	(0.013)	(0.014)
	48	48	48	44	48	44	44	44	44	44
Portugal	0.010	0.029	0.070	0.007	0.130	-0.004	0.000	-0.004	0.008	0.002
	(0.015)	(0.024)	(0.125)	(0.016)	(0.192)	(0.005)	(0.005)	(0.009)	(0.014)	(0.010)
	20	20	20	15	20	15	15	15	15	15
Sweden	0.026	0.017	0.034	0.006	-0.040	0.002	-0.002	-0.003	0.001	0.002
	(0.024)	(0.023)	(0.020)	(0.024)	(0.076)	(0.004)	(0.004)	(0.007)	(0.011)	(0.009)
	32	32	32	32	32	32	32	32	32	32
United Kingdom	-0.004	0.019	0.032	0.014	-0.003	0.002	-0.003	-0.004	0.006	0.000
	(0.051)	(0.021)	(0.059)	(0.033)	(0.088)	(0.012)	(0.009)	(0.011)	(0.018)	(0.019)
	144	144	136	121	144	121	121	121	121	121
EU_15	**0.012**	**0.021**	**0.034**	**0.011**	**0.025**	**0.000**	**-0.002**	**-0.003**	**0.005**	**0.000**
	(0.028)	(0.024)	(0.074)	(0.035)	(0.107)	(0.010)	(0.009)	(0.011)	(0.016)	(0.016)
	816	816	788	709	816	709	709	709	709	709

Note: the average is reported in the first row; the standard deviation is reported in parentheses in the second row and the third row shows the number of observations (region/year)

[23]The growth rates in these tables are 'unweighted' means which have been computed in the following way: $\sum_t \sum_i \frac{\ln(v_{ij,t})-\ln(v_{ij,t-1})}{I\cdot T}$ for each j where v refers to the variable, j refers to the country, $t = (2003,\ldots,2006)$ refers to the considered years and i refers to regions belonging to country j. We have also computed 'weighted' average growth rates of the relevant variables by country, using the share of employment in a region over the employment in the relative country as the weight for the relative growth rates: however, weighted growth rates (which are available upon request) are in line with non-weighted figures.

Table 4.4: Growth rates by country, EU12, 2003-2006

	Δy	Δkl	$\Delta hcap$	Δhhi	$\Delta tech$	ΔSH_EF	ΔSH_HT	ΔSH_LT	ΔSH_KIS	ΔSH_LKIS
Bulgaria	0.023	-0.006	0.006	-0.001	0.112	0.006	0.001	0.002	-0.001	0.002
	(0.047)	(0.024)	(0.041)	(0.020)	(0.183)	(0.007)	(0.004)	(0.008)	(0.009)	(0.009)
	18	24	18	18	24	18	18	18	18	18
Cyprus	0.030	0.023	0.008	0.005	0.065	0.003	0.000	-0.004	0.005	-0.001
	(0.013)	(0.003)	(0.043)	(0.022)	(0.113)	(0.007)	(0.002)	(0.009)	(0.008)	(0.008)
	3	4	4	4	4	4	4	4	4	4
Czech Republic	0.066	0.003	0.030	0.004	0.088	0.000	0.004	-0.003	0.003	0.000
	(0.054)	(0.018)	(0.050)	(0.029)	(0.116)	(0.009)	(0.007)	(0.009)	(0.008)	(0.012)
	32	32	32	32	32	32	32	32	32	32
Estonia	0.030	0.069	0.029	-0.003	0.061	0.009	-0.001	-0.003	-0.005	0.003
	(0.008)	(0.005)	(0.044)	(0.038)	(0.059)	(0.005)	(0.013)	(0.008)	(0.025)	(0.008)
	3	4	4	4	4	4	4	4	4	4
Hungary	0.012	0.067	0.031	0.012	0.084	0.003	0.000	-0.007	0.003	0.004
	(0.048)	(0.029)	(0.050)	(0.028)	(0.120)	(0.007)	(0.010)	(0.010)	(0.012)	(0.014)
	28	28	28	28	28	28	28	28	28	28
Lithuania	0.073	0.047	0.061	0.009	0.280	0.006	0.000	0.000	0.002	0.008
	(0.004)	(0.013)	(0.018)	(0.007)	(0.262)	(0.008)	(0.003)	(0.003)	(0.006)	(0.002)
	3	4	4	4	4	4	4	4	4	4
Latvia	0.012	0.082	0.030	0.009	0.148	0.008	-0.001	-0.004	0.002	0.005
	(0.051)	(0.008)	(0.073)	(0.020)	(0.168)	(0.005)	(0.003)	(0.007)	(0.008)	(0.005)
	3	4	4	4	4	4	4	4	4	4
Malta	-0.012	0.015	0.060	0.021	0.095	0.002	-0.004	-0.006	0.006	0.004
	(0.042)	(0.010)	(0.072)	(0.027)	(0.110)	(0.009)	(0.015)	(0.008)	(0.005)	(0.014)
	3	4	4	4	4	4	4	4	4	4
Poland	0.034	0.032	0.087	-0.004	0.199	0.004	0.001	-0.002	0.002	0.002
	(0.085)	(0.018)	(0.051)	(0.022)	(0.275)	(0.008)	(0.006)	(0.011)	(0.012)	(0.014)
	64	64	64	32	64	32	32	32	32	32
Romania	0.139	0.016	0.056	-0.016	0.104	0.003	0.000	-0.001	0.003	0.009
	(0.078)	(0.024)	(0.087)	(0.054)	(0.451)	(0.009)	(0.010)	(0.013)	(0.009)	(0.013)
	24	32	32	32	32	32	32	32	32	32
Slovenia	0.026	0.073	0.092	0.011	0.143	0.000	-0.002	-0.007	0.008	0.000
	(0.025)	(0.004)	(0.066)	(0.029)	(0.102)	(0.002)	(0.009)	(0.005)	(0.005)	(0.010)
	4	4	4	4	4	4	4	4	4	4
Slovakia	0.074	0.029	0.067	0.007	0.082	0.003	0.003	-0.004	0.002	0.002
	(0.043)	(0.026)	(0.051)	(0.024)	(0.137)	(0.007)	(0.008)	(0.009)	(0.013)	(0.012)
	16	16	16	16	16	16	16	16	16	16
EU_12	**0.050**	**0.028**	**0.054**	**0.001**	**0.129**	**0.003**	**0.001**	**-0.003**	**0.002**	**0.003**
	(0.075)	(0.032)	(0.063)	(0.033)	(0.252)	(0.008)	(0.008)	(0.010)	(0.010)	(0.012)
	201	220	214	182	220	182	182	182	182	182

Note: the average is reported in the first row; the standard deviation is reported in parentheses in the second row and the third row shows the number of observations (region/year)

Regions belonging to EU-12 New Member States show (on average) an higher labour productivity growth rate (5%) with respect to regions belonging to 'Old' EU-15 countries (1.2%). This is in line with the literature that claims for the role of the economy restructuring and catching-up to the technological frontier as the main explanations for this phenomenon. Among the countries in the EU-15, it is possible to appreciate a certain amount of heterogeneity in growth rates. United Kingdom, Italy, Spain and Portugal show low performance in terms of labour productivity growth during the period 2003-2006. France and Germany show modest growth trends. Ireland shows the best performance on average, even showing a large standard error, which is likely due to the big difference between the region of Dublin (IE02) which saw a strong economic performance over the past number of years, and the other region (Border, Midland and Western, IE01); some North-European countries show fast growth rates, as the Netherlands (2.7%), Sweden (2.6%), and Finland (2.5%), which is in line with previous analysis at the country level (see O'Mahony, Rincon-Aznar, and Robinson, 2010, among others). Among the New Member States, Romania, Slovakia, Lithuania and Czech Republic show the best performance in terms of labour productivity growth. It is interesting to note the relative higher standard deviations in the growth rates of regions belonging to EU-12 with respect to regions belonging to the 'Old' member states. This is probably due to the fact that there is a considerable amount diversity in growth experience: for example in Romania, the capital region (RO32) shows the highest growth rate (0.169), while other regions perform differently (RO12, RO21, RO22); in the Czech Republic, Moravskoslezsko (CZ08) — which benefits from its location on the borders of Poland and Slovakia —, the Central Bohemian Region (CZ02) and the region of Prague (CZ01) show the best performance in terms of labour productivity growth, while the North East (CZ05) performs rather poorly.

4.6 Econometric analysis

4.6.1 Baseline results

As we have underlined in Section 4.3, we can interpret the econometric results of the empirical model presented in Equation 4.1 as the effect of flows of foreign direct investments —made in the previous year— on the (current) regional productivity growth rate, that is the interpretation we will give to the results in this and the following Sections. It would be highly desirable to specify differences longer than one-year for productivity growth but, given the short time span available in our data, this would reduce the number of observations, thus increasing measurement errors and reducing the precision of our estimates. After having introduced the variables regarding FDIs as explained in Section 4.5, the model specification becomes:

$$\Delta y_{ij,t} = \alpha + \sum_w \gamma_w^d wFDI(d)_{ij,t-1} + \sum_w \gamma_w^{log} wFDI(d)_{ij,t-1} \cdot wFDI(log)_{ij,t-1} + \beta \Delta kl_{ij,t} + \Delta \mathbf{x_{ij,t}} \delta + \eta_j + \tau_t + \Delta \epsilon_{ij,t}. \quad (4.5)$$

We estimate Equation 4.5 by OLS, and the results are showed in Table 4.5. In this case we are left with three pooled cross-sections of differenced equations: 2004-2003, 2005-2004 and 2006-2005. In this and the following regressions we report robust standard errors clustered by regions to control for the lack of independence of observations referring to the same region over time [24].

In Specification (1), we look at the effects of inward and outward foreign direct investments (made in year $t-1$) on productivity growth rates, taking into account the change in the capital-labour ratio but without controlling for the other regional characteristics (i.e. human capital, technological capital, the industrial mix and its degree of concentration/diversification). Results on the coefficient of the variables related to inward FDIs, γ_I^d and γ_I^{log}, suggest that for low levels of incoming investments the effect on regional productivity is negative, because the value of the coefficient of the dummy variable dominates with respect to the coefficient of the continuous variable. However, the effect of the continuous variable increases as the number of inward FDIs gets bigger: in other words, inward FDIs have a positive effect on regional productivity, only above a threshold number of investments. On the other hand, outward FDIs have a positive effect on regional productivity, but the effect decreases as the number of outward investments increases. In Specification (2) the change in the quality of the industrial mix is taken into account, together with changes in the level of human capital, in the technological capital stock and in the degree of concentration/diversification of the industrial mix. A non-negligible loss in the sample size occurs from (1) to (2), and this is mainly due to the lack of data for the sectoral employment shares in several regions: these missing values bring to corresponding loss of usable observations in the industrial mix variables (SH_{s*ijt}) and in the Herfindahl-Hirschman index (HHI_{ijt})[25]. To a lesser extent, few missing values are in the variables measuring the level of human capital and the technological capital. Despite the sizable reduction in sample size, results on coefficients of outward foreign investments do not change much, while the same is not true for inward in-

[24] All the regressions have been performed in Stata 10.1, except for those in Section 4.6.2, which have been run using the environment R.

[25] Data for employment shares are not available for the following regions in some (or all) of the three waves of growth rates: Belgium (BE34), Germany (DE30, DE41, DE42, DE50, DE60, DEB2, DED3, DEE0) Denmark (all regions; DK01, DK02, DK03), Spain (ES43), Finland (FI20), France (FR83), Greece (GR11, GR13, GR21, GR22, GR23, G25, GR42, GR43), Italy (ITC2), Netherlands (NL23), Poland (—just for the growth rate 2004-2003— all regions; PL11, PL12, PL21, PL22, PL31, PL32, PL33, PL34, PL41, PL42, PL43, PL51, PL52, PL61, PL62, PL63), Portugal (PT15), United Kingdom (UKE2, UKF3, UKK3, UKK4, UKM5, UKM6).

Table 4.5: Econometric results - Baseline (OLS)

		Specification		
		1	2	3
Variable	Coefficient			
OFDI$_{t-1}$ (dummy)	γ_O^d	0.0088***	0.0076***	0.0075**
		(0.0029)	(0.0029)	(0.0029)
OFDI$_{t-1}$(log. of n.inv)	γ_O^{log}	-0.0030***	-0.0027***	-0.0029***
		(0.0009)	(0.0009)	(0.0010)
IFDI$_{t-1}$(dummy)	γ_I^d	-0.0074***	-0.0024	-0.0072***
		(0.0026)	(0.0025)	(0.0027)
IFDI$_{t-1}$(log. of n.inv)	γ_I^{log}	0.0031***	0.0020*	0.0031***
		(0.0011)	(0.0011)	(0.0012)
$\Delta_{t,t-1}kl$	β	0.2401***	0.3592***	0.2392***
		(0.0839)	(0.1088)	(0.0842)
$\Delta_{t,t-1}$hcap	δ_{hcap}		-0.0120	0.0003
			(0.0164)	(0.0137)
$\Delta_{t,t-1}$hhi	δ_{hhi}		0.1975***	0.1577**
			(0.0616)	(0.0740)
$\Delta_{t,t-1}$tech	δ_{tech}		-0.0001	0.0008
			(0.0083)	(0.0100)
$\Delta_{t,t-1}$SH_EF	δ_{EF}		0.0420	0.1434
			(0.1434)	(0.1509)
$\Delta_{t,t-1}$SH_HD	δ_{HD}		0.0910	0.1638
			(0.1381)	(0.1416)
$\Delta_{t,t-1}$SH_LD	δ_{LD}		-0.1648	-0.1430
			(0.1438)	(0.1557)
$\Delta_{t,t-1}$SH_KIS	δ_{KIS}		-0.3420**	-0.1876
			(0.1325)	(0.1690)
$\Delta_{t,t-1}$SH_LKIS	δ_{LKIS}		-0.4560***	-0.3052*
			(0.1417)	(0.1751)
Constant	α	0.0272***	0.0212***	0.0270***
		(0.0039)	(0.0039)	(0.0039)
Country dummies	η_j	Yes	Yes	Yes
Year dummies	τ_t	Yes	Yes	Yes
Observations		755	662	746
Regions		258	238	255

Significance levels: * 10%, ** 5%,*** 1%
Cluster-robust standard errors in parentheses

vestments: the dummy variable, γ_I^d, becomes non significant and the coefficient of the continuous variable, γ_I^{log}, results to be poorly significant, even if the coefficient is rather stable in magnitude. The observed changes in the coefficients are the result of the sample-selection due to missing values in the sectoral employment shares. This fact is confirmed by Specification (3), in which we have filled in most of the missing values in the vector $\mathbf{x_{ij,t}}$. We have imputed the missing values in two steps. First, for the period 2002-2006, we assumed that missing values were equal to 'the last or the first available data' in the series[26]. Second, in the cases where no data was available or a given region throughout the 2003-2006 period, we imputed using national values.

Looking at Specification (3), in the third column of Table 4.5, it is possible to appreciate how these results are in line with those of Specification (1): in fact, by imputing missing values in the set of regional controls we have recovered almost all regions that were lost moving from Specification (1) to Specification (2); the reported coefficient of the capital-labour ratio is consistent with that of Specification (1). With respect to foreign direct investments variables, the dummy variable for inward investment, γ_I^d, shows a coefficient which is similar (both in magnitude and in statistical significance) to that in Specification (1). Thus, the result on inward investments is a bit sensitive to sample under analysis; nonetheless, outward investments variables show stable coefficients even with imputed data, thus reassuring us about the results. Overall, Specification (3) is our favorite one, both because it allows to control for an important set of regional characteristics without reducing the sample size. The cost for this choice is the use of variables with some imputed values for a limited number of observations: that is, the effect of the regional characteristics which we use as controls could be not consistent. However, there are no reasons to think that this unlucky event would affect the sign and the magnitude of the coefficients of main concern, i.e. those related to inward and outward FDIs variables. By the way, most of the coefficients of the controls result to be not significant for explaining the regional productivity growth. In particular, neither the contemporaneous change in the human capital, nor the change in the technology capital –even if they show the expected signs– seem to significantly explain the regional differences in productivity growth. However, the vector of controls is jointly significant, as reported in the first row of Table 4.6[27].

In Table 4.6 (second row) we report a test for the joint significance of foreign direct investments variables (based on Specification (3)): the null hypothesis of no effect by inward and outward foreign direct investments flows is tested and rejected.

[26]Take, for explanatory purpose, the employment share in the Hight-tech manufacturing: if the observation for the share of employees was missing in a given region in 2004 but it was observable in 2003, the value of the share in 2004 was set equal to that of 2003. On the other hand, if the observation for the share of employees was missing in a given region in 2002 but it was observable in 2003, the value of the share in 2002 was set equal to that of 2003. Thus, we assumed 'zero-changes' were information were not available.

[27]In the Data Appendix 4.9.4 we also report Specification (3), inserting regional controls one by one in the Equation.

Table 4.6: Tests on parameters of the baseline Specification (3)

Null Hypothesis (H0)	Conditions	F-Statistics	Critical value (5%)
No regional characteristics effects	$\beta = \delta' = 0$	2.92	1.92
No FDIs effects	$\gamma_w^{log} = \gamma_w^d = 0$	3.52	2.41
No country dummies effects	$\eta' = 0$	119.65	1.56

This confirms the significant role played by foreign direct investments in explaining differences in growth rates of labour productivity at the regional level, once a large set of regional characteristics together with unobserved country-specific trends in productivity have been taken into account. In the third row of Table 4.6, an F-test on the joint significance of country effects is carried out. The evidence of national trends in labour productivity captured by the national effects is clear: the country dummies result to be jointly significant and failing to account for them would bring us to neglect national patterns of growth[28].

Finally, let us comment on the threshold effects of inward and outward investments on productivity. From Equation 4.5, the marginal effect of an inward or outward investment on regional productivity growth can be computed as:

$$\frac{\partial \Delta y}{\partial wFDI} = \gamma_w^d + \gamma_w^{log} \cdot log(wFDI). \tag{4.6}$$

The marginal effect of one more investment will be positive as long as

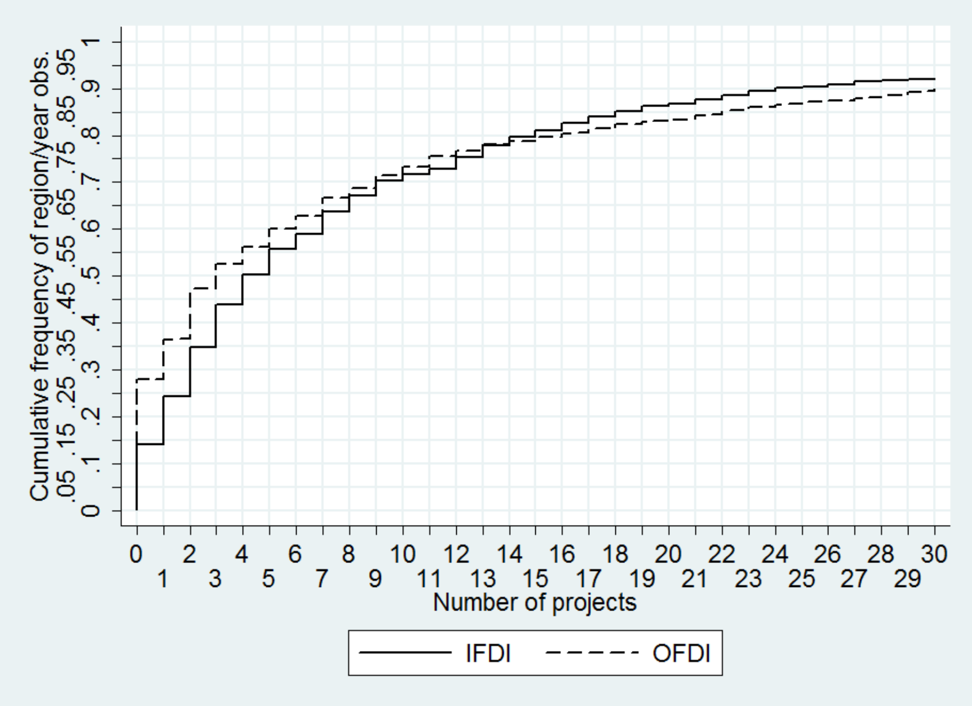
$$log(wFDI) > \frac{-\gamma_w^d}{\gamma_w^{log}}. \tag{4.7}$$

In particular, taking Specification (3) as a reference, with γ_I^d = -0.0072 and γ_I^{log} = 0.0031, the marginal effect of receiving inward investments would be positive for a number of investments greater or equal than $\exp^{\frac{0.0072}{0.0031}}$=10.2. For outward investments, with γ_O^d = 0.0075 and γ_O^{log} = -0.0029, the marginal effect will be positive up to $\exp^{\frac{-0.0075}{-0.0029}}$=13.3 investments.

Figure 4.3 allows to appreciate the extent to which inward and outward investments contribute to productivity growth of EU regions. The Figure plots the cumulative distribution of region/year observations by the number of inward and outward FDIs. The first thing to notice is that outward FDI is a twice more rare phenomenon

[28]We have also estimated Specification (3) without including the country dummies, for purposes of control, and results indicate that failing to account for them would bring to biased coefficients both of the foreign direct investments variables and of the other control variables. In particular, coefficients of the dummy variables of both inward and outward foreign direct investments result to be not significant any more, while that of the continuous variable measuring inward investments is positive and that of the outward investments results to be negative. The coefficient on the capital-labour ratio results to be not significant and it shows an implausible coefficient, (0.07). Interestingly enough, a larger number of control variables result to be significant with respect to Specification (3), i.e. the technological capital, the degree of concentration/diversification of the industrial mix and the industrial mix itself. This fact could suggest that patters referring to these variables are country-specific.

Figure 4.3: Cumulative frequency of region/year observation by number of inward and outward FDIs, 2003-2006



than inward FDI: 28% of region/year obervations have zero outgoing projects, as opposed to only 14% in the case of incoming investments. However, there is a sizable number of cases with a rather large number of outward investments, so that the cumulative distributions for OFDI and IFDI cross at 13 projects. To the extent that the threshold level of investments above which the effect is positive is 10.2, Figure 4.3 suggests that approximately 30% of region/year observations are above this threshold (and benefit from inward investments). This share could be higher if regions would attract more incoming multinationals. In the case of outward investments, 28% of regions would increase their productivity growth by 0.75% making one outgoing project, while about 22% are above the 13.3 threshold, and have thus lower productivity growth then non-internationalized. The remaining 50% are actually experiencing higher productivity growth, thanks to their international orientation.

4.6.2 Robustness checks

In the previous section we have argued that both inward and outward foreign investments can be a key determinant of differences in productivity growth among the European regions. However, this result may be the outcome of some specification error and omitted variable bias. In the present section we perform some robustness checks, in order to convince the reader that the previous results are not spurious correlations.

Regional controls in levels

Given the relevance of the country effects in Specification (3), we would like to exclude that our results are biased due to unobserved regional effects correlated to productivity trends at the regional level. In order to cope with this problem, we can augment Specification (3) by including the set of regional controls in levels at the beginning of the period for each cross-section.[29] Thus, the employed specification now becomes:

$$\Delta y_{ij,t} = \alpha + \sum_{w} \gamma_w^d wFDI(d)_{ij,t-1} + \sum_{w} \gamma_w^{log} wFDI(d)_{ij,t-1} \cdot wFDI(log)_{ij,t-1} + \beta \Delta kl_{ij,t} + \Delta \mathbf{x_{ij,t}} \delta + \mathbf{x_{ij,t-1}} \phi + \mathbf{z_{ij}} \varphi + \eta_j + \tau_t + \Delta \epsilon_{ij,t}, \quad (4.8)$$

where $\mathbf{x_{ij,t-1}}$ is the vector of regional controls at the beginning of the period.

To avoid that the variables measuring foreign investments capture a generic effect of the 'size' of the region, given that these are the sole variable non-standardized on the right-hand side of Equation 4.8, we include a measure of the total population of the region in the vector of regional controls at the beginning of the period. Moreover, we include the level of labour productivity at the beginning of the period, given that it could explain a significant part of the productivity growth rate (catching-up). Thus, the vector can be written as

$$\mathbf{x_{ij,t-1}} = \left(y_{ij,t-1}, kl_{ij,t-1}, hcap_{ij,t-1}, hhi_{ij,t-1}, tech_{ij,t-1}, pop_{ij,t-1}\right). \quad (4.9)$$

We further include in Equation 4.8 a vector of time invariant characteristics, $\mathbf{z_{ij}}\varphi$, which contains the following information:

- Two dummy variables for coastal ($COAST$) and capital ($CAPT$) regions, which take value '1', respectively, in the case in which the region lies on the coast or if it is the capital region of the country. The coastal dummy (information come from Salz, Buisman, Smit, and de Vos, 2006) should account for the general accessibility of a region, which should correlated with its productivity and the degree of internationalization, while the capital dummy is intended to capture agglomeration economies, which could certainly be a driver of productivity growth and which are generally associated with the economic activity and related services taking place in a country's capital.

- We also control for regions which are eligible for European structural funds. A dummy which takes value '1' has been included in Equation 4.8, when the region is indicated by the European Commission as eligible for 'Objective 1' funds[30].

[29] In principle, one could add regional fixed effects to the equation in first-differences but, one the one hand there is not clear theoretical motive to assume region-specific trends in productivity and, on the other hand, given the short time series, that would leave very little variation to identify our coefficients.

[30] The list of the eligible regions can be found at `http://ec.europa.eu/regional_policy/objective1/index_en.htm`.

Results are reported in Table 4.7.

Overall, the effects of inward and outward foreign direct investments on regional productivity are robust both after having taken into account the set of regional characteristics at the beginning of the period and the set of time-invariant regional characteristics. Specification (3contd) which is the more demanding, given the high number of covariates and the multicollinearity among them, shows that the coefficient of the dummy variable related to outward FDIs is still significant even if it decreases in magnitude, while the the continuous variable is not significant anymore. Thus, adding controls strengthen our finding of positive effects from $OFDI$, since the threshold effect disappears. Results on the inward FDIs variables are also robust: both the dummy and the continuous variable are significant and they do not change much in terms of magnitude with respect to Specification (3). The capital-labour ratio is stable across all different specifications, while the productivity level at the beginning of the period is never significant, even when it is included without regional controls, as in Specification (3conte). The measure of the total population of the region is not significant in Specification (3contd), thus reassuring us about the fact that results should not be sensitive to the inclusion of a generic measure of size of the region[31]. Interestingly enough, once the degree of concentration/diversification of the industrial mix at the beginning of the period is included in the regression, the coefficient δ_{hhi}, which relates to the change in the industrial mix quotas, comes to be not significant anymore. Regions with higher growth rates are those that at the beginning of the period presented more diversified industrial structures.

The effect of contemporaneous investments

As we explained in Section 4.3, we hypothesize that foreign direct investments (both inward and outward) show their effects in a given span of time, i.e. one year. In order to support this hypothesis, we also run Specification (3), substituting investments variables at time $t-1$ with investments variables at time t (contemporaneous investments), and including jointly in the same regression. Thus, we estimate two different specifications: the first one with the variables regarding contemporaneous investments only, which can be written as

$$\Delta y_{ij,t} = \alpha + \sum_w \lambda_w^d wFDI(d)_{ij,t} + \sum_w \lambda_w^{log} wFDI(d)_{ij,t} \cdot wFDI(log)_{ij,t} + \\ \beta \Delta kl_{ij,t} + \Delta \mathbf{x_{ij,t}} \delta + \eta_j + \tau_t + \Delta \epsilon_{ij,t}, \quad (4.10)$$

[31] Specification (3contd) has been also estimated substituting the FDIs continuous variables ($OFDI(log)$ and $IFDI(log)$) with the correspondent variables divided by the gross value added. This is another way of controlling for the size of the region: results, which are not reported in order to save space, are in line with Specification (3contd).

Table 4.7: Robustness check: regional characteristics (OLS)

		Specification					
		3	3conta	3contb	3contc	3contd	3conte
Variable	Coefficient						
OUT(dummy)$_{t-1}$	γ_O^d	0.0075**	0.0066**	0.0069**	0.0069**	0.0069**	0.0075**
		(0.0029)	(0.0030)	(0.0031)	(0.0031)	(0.0032)	(0.0029)
OUT(log. of n.inv)$_{t-1}$	γ_O^{log}	-0.0029***	-0.0021*	-0.0018	-0.0017	-0.0018	-0.0032***
		(0.0010)	(0.0011)	(0.0012)	(0.0011)	(0.0012)	(0.0010)
INW(dummy)$_{t-1}$	γ_I^d	-0.0072***	-0.0067**	-0.0064**	-0.0065**	-0.0065**	-0.0067**
		(0.0027)	(0.0028)	(0.0027)	(0.0027)	(0.0028)	(0.0027)
INW(log. of n.inv)$_{t-1}$	γ_I^{log}	0.0031***	0.0023*	0.0026**	0.0027**	0.0027**	0.0027**
		(0.0012)	(0.0012)	(0.0013)	(0.0013)	(0.0013)	(0.0012)
$\Delta_{t,t-1}kl$	β	0.2392***	0.2620***	0.2559**	0.2558**	0.2524**	0.2345***
		(0.0842)	(0.0970)	(0.0990)	(0.1003)	(0.1011)	(0.0825)
$\Delta_{t,t-1}$hcap	δ_{hcap}	0.0003	0.0020	0.0017	0.0012	0.0014	-0.0002
		(0.0137)	(0.0140)	(0.0140)	(0.0140)	(0.0140)	(0.0138)
$\Delta_{t,t-1}$hhi	δ_{hhi}	0.1577**	0.1181	0.1172	0.1182	0.1170	0.1464*
		(0.0740)	(0.0775)	(0.0775)	(0.0776)	(0.0778)	(0.0758)
$\Delta_{t,t-1}$tech	δ_{tech}	0.0008	0.0003	0.0005	0.0019	0.0018	0.0005
		(0.0100)	(0.0107)	(0.0107)	(0.0108)	(0.0108)	(0.0100)
y_{t-1}	ϕ_y					0.0053	0.0093
						(0.0134)	(0.0077)
kl_{t-1}	$\phi_{kl,t-1}$		0.0069	0.0058	0.0052	0.0039	
			(0.0062)	(0.0067)	(0.0067)	(0.0075)	
hcap$_{t-1}$	ϕ_{hcap}		0.0033	0.0031	0.0030	0.0035	
			(0.0056)	(0.0056)	(0.0058)	(0.0060)	
hhi$_{t-1}$	ϕ_{hhi}		-0.0379**	-0.0375**	-0.0366**	-0.0362**	
			(0.0184)	(0.0183)	(0.0182)	(0.0182)	
tech$_{t-1}$	ϕ_{tech}		0.0017	0.0018	0.0027	0.0025	
			(0.0016)	(0.0016)	(0.0016)	(0.0017)	
pop$_{t-1}$	ϕ_{pop}			-0.0013	-0.0022	-0.0021	
				(0.0015)	(0.0016)	(0.0016)	
COAST	φ_{COAST}				0.0028	0.0027	
					(0.0018)	(0.0017)	
CAPT	φ_{CAPT}				0.0044	0.0041	
					(0.0033)	(0.0034)	
OBJ1	φ_{OBJ1}				0.0041	0.0044	
					(0.0027)	(0.0027)	
Constant	α	0.0270***	-0.0597	-0.0385	-0.0158	-0.0228	-0.0084
		(0.0039)	(0.0376)	(0.0483)	(0.0507)	(0.0524)	(0.0296)
Country dummies	η_j	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	τ_t	Yes	Yes	Yes	Yes	Yes	Yes
Industrial mix*		Yes	Yes	Yes	Yes	Yes	Yes
Observations		746	746	746	746	746	746
Regions		255	255	255	255	255	255

* The industrial mix include both $\Delta_{t,t-1}SH_{s*}$ (differences) and $SH_{s*ji,t-1}$ (lagged)
Significance levels: * 10%, ** 5%,*** 1%
Cluster-robust standard errors in parentheses

where $wFDI(d)_{ij,t}$ and $wFDI(log)_{ij,t}$ are, respectively, the dummy variable and the continuous variables and $w = \{O, I\}$ are outward and inward investments; and the second one with both lagged and contemporaneous investments:

$$\begin{aligned}\Delta y_{ij,t} = \alpha + \sum_w \lambda^d_w wFDI(d)_{ij,t} + \sum_w \lambda^{log}_w wFDI(d)_{ij,t} \cdot wFDI(log)_{ij,t} + \\ \sum_w \gamma^d_w wFDI(d)_{ij,t-1} + \sum_w \gamma^{log}_w wFDI(d)_{ij,t-1} \cdot wFDI(log)_{ij,t-1} + \\ \beta \Delta kl_{ij,t} + \Delta \mathbf{x_{ij,t}} \delta + \eta_j + \tau_t + \Delta\epsilon_{ij,t}. \quad (4.11)\end{aligned}$$

Table 4.8: Robustness check: contemporaneous investments (OLS)

		Specification		
		3	3contf	3contg
Variable	Coefficient			
OUT(dummy)$_{t-1}$	γ^d_O	0.0075**		0.0097***
		(0.0029)		(0.0037)
OUT(log. of n.inv)$_{t-1}$	γ^{log}_O	-0.0029***		-0.0036*
		(0.0009)		(0.0018)
INW(dummy)$_{t-1}$	γ^d_I	-0.0072***		-0.0067**
		(0.0027)		(0.0029)
INW(log. of n.inv)$_{t-1}$	γ^{log}_I	0.0031***		0.0027
		(0.0012)		(0.0017)
OUT(dummy)$_t$	λ^d_O		-0.0032	-0.0067
			(0.0035)	(0.0042)
OUT(log. of n.inv)$_t$	λ^{log}_O		-0.0009	0.0010
			(0.0008)	(0.0018)
INW(dummy)$_t$	λ^d_I		-0.0009	0.0002
			(0.0030)	(0.0032)
INW(log. of n.inv)$_t$	λ^{log}_I		0.0019*	0.0008
			(0.0011)	(0.0015)
$\Delta_{t,t-1}kl$	β	0.2392***	0.2491***	0.2444***
		(0.0842)	(0.0825)	(0.0850)
$\Delta_{t,t-1}$HCAP	δ_{HCAP}	0.0003	0.0004	-0.0005
		(0.0137)	(0.0138)	(0.0137)
$\Delta_{t,t-1}$HHI	δ_{HHI}	0.1577***	0.1666**	0.1519**
		(0.0740)	(0.0737)	(0.0730)
$\Delta_{t,t-1}$INNOV	δ_{INNOV}	0.0008	-0.0002	0.0001
		(0.0100)	(0.0104)	(0.0098)
Constant	α	0.0270***	0.0288***	0.0293***
		(0.0039)	(0.0039)	(0.0039)
Country dummies	η_j	Yes	Yes	Yes
Year dummies	τ_t	Yes	Yes	Yes
Industrial mix	$\delta_{SH_{s*}}$	Yes	Yes	Yes
Observations		746	746	746
Regions		255	255	255

Significance levels: * 10%, ** 5%,*** 1%
Cluster-robust standard errors in parentheses

Results, which are reported in Table 4.8, definitely support our *a priori* on the span of time which is necessary to foreign investments to show their effects on productivity growth. Contemporaneous investments do no show significant effects on productivity, except for a small effect by inward investments, as Specification (3contf) shows. Moreover, in the third Column of Table 4.8, once we introduce both contemporaneous and lagged investments, only the last ones show a significant effect on productivity growth. The specification with lagged investments is also more robust to endogeneity problems with respect to that with contemporaneous investments: if shocks to current productivity growth would also determine a larger number of inward and outward investment projects, Specification (3contf) may be more sensitive to the simultaneity issue and the use of lagged investments should lessen this problem.

Different technological regimes

One possible source of bias could be that we impose the same technology to very different economies, such as regions belonging to EU-15 and EU-12 countries. If inward and outward foreign direct investment variables were jointly determined with the choice of the production technology we might estimate biased coefficients. To avoid this problem, we have further augmented Specification (3) including the interaction between the capital-labour ratio and a dummy variable ($eu12$) which is equal to '1' for all regions belonging to countries of the EU-12. The new Specification becomes:

$$\Delta y_{ij,t} = \alpha + \sum_w \gamma_w^d wFDI(d)_{ij,t-1} + \sum_w \gamma_w^{log} wFDI(d)_{ij,t-1} \cdot wFDI(log)_{ij,t-1} + \beta \Delta kl_{ij,t} + \beta_{kl*EU12} \Delta(kl_{ij,t} \cdot eu12) + \mathbf{x_{ij,t}}\delta + \eta_j + \tau_t + \Delta\epsilon_{ij,t}. \quad (4.12)$$

Results are reported in Table 4.9. We find a significant difference in the technological regimes of the two groups of regions with the regions belonging to the EU-15 showing a larger coefficient for the change in the capital-labour ratio. However, results of both inward and outward investments are stable with respect to Specification (3).

Accounting for spatial dependence

In the previous paragraphs we have made the implicit assumption that spatial interactions among regions in terms productivity growth are fully captured by the inclusion of country effects. This could be a too restrictive assumption for a number reason: first, spatial interactions could be at work also among regions which belong to different countries; second, they can be time-variant; third, benefits from being localized nearer to more productive regions can be differentiated even within a country (different intensities of spatial interactions). These uncontrolled spatial effects could invalidate the OLS estimation of the Specification (3), that neglects the role of proximity in explaining the regional productivity growth as a function of foreign direct investments.

Table 4.9: Robustness check: different technological regimes (OLS)

		Specification	
		3	3conth
Variable	Coefficient		
OUT(dummy)$_{t-1}$	γ_O^d	0.0075**	0.0073**
		(0.0029)	(0.0029)
OUT(log. of n.inv)$_{t-1}$	γ_O^{log}	-0.0029***	-0.0031***
		(0.0010)	(0.0010)
INW(dummy)$_{t-1}$	γ_I^d	-0.0072***	-0.0064***
		(0.0027)	(0.0026)
INW(log. of n.inv)$_{t-1}$	γ_I^{log}	0.0031***	0.0031***
		(0.0012)	(0.0012)
$\Delta_{t,t-1}kl$	β	0.2392***	0.3091***
		(0.0842)	(0.1009)
$\Delta_{t,t-1}kl \cdot eu12$	$\beta_{kl \cdot EU12}$		-0.3135*
			(0.1865)
$\Delta_{t,t-1}$hcap	ϕ_{hcap}	0.0003	-0.0006
		(0.0137)	(0.0137)
$\Delta_{t,t-1}$hhi	ϕ_{hhi}	0.1577**	0.1696**
		(0.0740)	(0.0735)
$\Delta_{t,t-1}$tech	δ_{tech}	0.0008	0.0013
		(0.0100)	(0.0099)
Constant	α	0.0270***	0.0259***
		(0.0039)	(0.0039)
Country dummies	η_j	Yes	Yes
Year dummies	τ_t	Yes	Yes
Industrial mix	$\delta_{SH_{s*}}$	Yes	Yes
Observations		746	746
Regions		255	255

Significance levels: * 10%, ** 5%,*** 1%
Cluster-robust standard errors in parentheses

Anselin (1988) originally proposed two alternative ways of representing units interactions in the space for a cross-section, and Elhorst (2010) has recently provided a review of these methods in the case of panel data. The most used frameworks by which regional interactions can be modeled are the spatial autoregressive (or spatial lag) model and the spatial error model. The first model assumes that the productivity growth of each region is influenced by that of the neighboring regions. The differenced Equation 4.3 can be rewritten in the following way, in order to account for spatial interactions in the dependent variable:

$$\Delta y_{ij,t} = \gamma_{OUT} OFDI^{flows}_{ij,t-1} + \gamma_{INW} IFDI^{flows}_{ij,t-1} + \lambda W \Delta y_{ij,t} + \beta \Delta kl_{ij,t} + \Delta \mathbf{x_{ij,t}} \boldsymbol{\delta} + \eta_j + \tau_t + \Delta\epsilon_{ij,t}, \tag{4.13}$$

where W represents the spatial weight matrix, $W\Delta y_{ij,t}$ is the spatially lagged dependent variable, λ is called the spatial autoregressive coefficient, and the other explanatory variables remain unchanged with respect to the baseline Specification (3). The spatial weight matrix can be specified as:

$$W = \begin{bmatrix} 0 & w_{12} & \dots & w_{1j} & \dots & w_{1N} \\ w_{21} & 0 & \dots & \dots & \dots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ w_{j1} & \dots & \dots & \ddots & \dots & w_{jN} \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ w_{N1} & w_{N2} & \dots & w_{Nj} & \dots & 0 \end{bmatrix}$$

Each off-diagonal element of the matrix, w_{ij}, can either be a inverse mesaure of distance between region i and j, or can take value '1' or '0', if regions i and j are neighbors or not, respectively. In the latter case, we have binary contiguity matrix. In this work we adopt this type of weighting matrix and we define as neighbours all the regions within a 392 km radius of the region centroid[32]. Equation 4.13 can be estimated by Maximum Likelihood (ML) and, contrary to an OLS estimation that neglects significant spatial interactions, it allows one to obtain unbiased and consistent parameters.

A different specification of the spatial dependence is the spatial error model, which posits that, conditional on regressors, the error terms are correlated in space. In our case, the spatial error model can be written as

$$\Delta y_{ij,t} = \gamma_{OUT} OFDI^{flows}_{ij,t-1} + \gamma_{INW} IFDI^{flows}_{ij,t-1} + \beta \Delta kl_{ij,t} + \Delta \mathbf{x_{ij,t}} \boldsymbol{\delta} + \eta_j + \tau_t + \rho W \Delta u_{ij,t} + \Delta\epsilon_{ij,t}, \tag{4.14}$$

where W represents the spatial weight matrix defined as above, $\Delta u_{ij,t}$ reflects the

[32]This threshold have bee computed as the minimum distance that allow each region to have at least one neighbor, i.e. at least one out-of-diagonal element is equal to one. However, taking a larger radius does not affect the results.

spatially autocorrelated error term, and ρ is the spatial autocorrelation coefficient. Using ML estimation one avoids to incur in inefficient estimates yielded by OLS which do not account for spatial dependence in the error term.

The main difference between the two models is that, in the spatial-lag case, productivity growth of neighbor regions is the channel trough which externalities are trasmitted in space, while in the spatial-error model one assumes that the regional dependence arises from the spatial propagation of idiosyncratic shocks (Sterlacchini and Venturini, 2009). Since we do not have an *a priori* on the shape of regional interactions, we estimate both Equation 4.13 and Equation 4.14 by ML, using the the routine developed by Millo and Piras (2009) for the environment R and applying the spatial contiguity matrix previously defined. Results of the estimation are reported in Table 4.10. Since the routine has been programmed for balanced panel data, we loose some observations in order to balance our panel dataset: the final sample consists of 702 observations and 234 regions. In the first column of Table 4.10, we report the baseline model –which does not account for spatial interactions– estimated for the balanced panel by OLS (Specification 3_res). It is possible to compare it with Specification (3) in Table 4.5, noting that there all the coefficients of the FDI variables shrink, both in absolute values and in their statistical significance, due to the sample selection. Along the same lines, the capital-labour ratio shifts from (0.23) in Specification (3) to (0.20) in Specification (3_res). However, the positive effects of inward and outward FDIs (as well as the threshold effects) do not disappear.

Estimating the spatial lag model (Specification (3_splag)) we obtain a spatial autoregressive coefficient (λ) equal to 0.68, supporting the existence of significant spatial dependence. Nonetheless, all the FDI variables of the model are significant, and comparing the coefficients with those in Specification (3_res) there are no dramatic changes in the magnitude of the coefficients. We observe a slight drop both in the magnitude and in its statistical significance of the coefficients related to outward investments, γ_O^d and γ_O^{log}. This result could be explained by the fact that the coefficient of the dummy variable related to outward investments could capture the tendency of experiencing higher productivity growth rates which may be related to a larger proportion of multinational enterprises localized in the territory, and given the phenomenon of clustering of the higher productive regions in the EU (see Fiaschi, Lavezzi, and Parenti, 2009, among others), the inclusion of the spatial autoregressive term may clean this bias off, reducing the magnitude of the dummy variable for outward FDIs.

Specification (3_splag_nocd) reports the estimation of the spatial lag model without the inclusion of the country dummies: interestingly enough, those seem to capture country-specific spatial characteristics which are time-invariant and which cannot be captured by the spatial autoregressive term: natural candidates are institutional characteristics. However, the specification with the country dummies should be preferred:

Table 4.10: Spatial lag and spatial error models (ML)

		Specification				
		3_res	3_splag	3_splag_nocd	3_sperr	3_sperr_nocd
Variable	Coefficient					
OUT_{t-1}(dummy)	γ_O^d	0.0058**	0.0043*	0.0066**	0.0045*	0.0051*
		(0.0027)	(0.0024)	(0.0026)	(0.0023)	(0.0026)
OUT_{t-1}(log. of n.inv)	γ_O^{log}	-0.0021*	-0.0018*	-0.0043***	-0.0020**	-0.0040***
		(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
$IFDI_{t-1}$(dummy)	γ_I^d	-0.0057*	-0.0053**	-0.0045	-0.0049*	-0.0041
		(0.0026)	(0.0026)	(0.0029)	(0.0026)	(0.0029)
$IFDI_{t-1}$(log. of n.inv)	γ_I^{log}	0.0021*	0.0020*	0.0047***	0.0023**	0.0046***
		(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)
$\Delta_{t,t-1}kl$	β	0.2070***	0.2372***	0.0839**	0.2208***	0.1624***
		(0.0782)	(0.0477)	(0.0389)	(0.0460)	(0.0444)
$\Delta_{t,t-1}$hcap	δ_{hcap}	0.0056	0.0039	0.0182	-0.0010	0.0163
		(0.0135)	(0.0129)	(0.0142)	(0.0124)	(0.0141)
$\Delta_{t,t-1}$hhi	δ_{hhi}	-0.0178	0.0750	0.0592	0.0814	0.0701
		(0.1105)	(0.0640)	(0.0706)	(0.0630)	(0.0725)
$\Delta_{t,t-1}$tech	δ_{tech}	0.0057	0.0115	0.0191**	0.0106	0.0224***
		(0.0150)	0.0077)	(0.0081)	(0.0073)	(0.0083)
Constant	α	0.0231***	-0.0031	-0.0005	0.0219***	0.0175*
		(0.0060)	(0.0055)	(0.0032)	(0.0082)	(0.0091)
(Spatial autoregressive coefficient)	λ		0.6786***	0.7807***		
			(0.0467)	(0.0357)		
(Spatial autocorrelation coefficient)	ρ				0.7556***	0.8137***
					(0.0406)	(0.0332)
Country dummies		Yes	Yes	No	Yes	No
Year dummies		Yes	Yes	Yes	Yes	Yes
Industrial mix		Yes	Yes	Yes	Yes	Yes
Observations		702	702	702	702	702
Regions		234	234	234	234	234
Significance levels: * 10%, ** 5%,*** 1%						

first, a non-negligible number of country dummies (5 over 19) are significant in Specification (3_splag) and the null hypothesis that they are jointly significant cannot be rejected; second, the model without country dummies (Specification 3_splag_nocd) shows an uncredible coefficient of the capital-labour ratio (0.08); third the spatial autoregressive coefficient is larger in the model without the country dummies (0.78), thus indicating their ability in capturing state-specific spatial dependence.

The results of the spatial error model, confirm the presence of spatial dependence, which is indicated by the high and significant spatial autocorrelation coefficient, ρ=(0.75). In line with the spatial lag model, the coefficient of the dummy variable related to outward investments shrinks with respect to Specification (3_res)— from (0.0058) to (0.0045)— and the same is true for the coefficient of the dummy variable of the inward investments —from (0.0057) to (0.0049). This result can be explained by the fact that in the spatial error model, the spatial parameter could pick up the well-known geographical agglomeration phenomenon of the inward foreign investments. The spatial error model has been estimated without the inclusion of the country dummies, and the results are reported in the last column of the Table. However, the reduction in the autocorrelation coefficient and the more credible coefficient of the capital-labour ratio support the model with country dummies.

Overall, the estimation of models which account for spatial dependence do not change the basic results on the positive effects of foreign investments (both inward and outward) on regional productivity growth.

4.7 Concluding remarks

Despite the increasing evidence of integration of sub-national economies in the global arena, and the positive role of multinational firms for economic prosperity in local economies documented in a number of recent studies, evidence on the relationship between foreign investments and regional performance is lacking. Exploiting an original and extensive dataset on FDIs, we investigate the relationship between FDIs and productivity in a sample of European regions. The results of the econometric analysis support that both inward and outward foreign direct investments have positive effects on productivity growth at the regional level, after controlling for a relevant set of regional characteristics, such as human capital, technology capital and industry mix. This is an interesting result, given the increasing role of regions in the European context and the relevance –in terms of GDP– of inward and outward FDIs in the European Union. The econometric analysis has provided –to our knowledge for the first time– a robust evidence of positive effects in a large set of NUTS2 regions in almost all countries of the Enlarged Europe (EU-27). This is an original contribution to the international economics literature in several dimensions: previous studies with a regional perspective, as Driffield (2004) and Altomonte and Colantone (2009), have focused on comparisons within single countries and have addressed only the role of 'inward' investments as a driver of increasing local performance. Moreover, those few studies which have attempted to assess the specific role of outward investments on productivity (Bitzer and Görg, 2009; Driffield, Love, and Taylor, 2009; Herzer, 2010) have taken a country perspective, almost neglecting the sub-national level of analysis. This is unfortunate, given that the regional level of analysis is much more appropriate in order to capture those indirect effects introduced in Section 4.2. Our results are consistent with the idea that direct effects of MNEs on productivity and positive indirect effects (i.e. pecuniary and technology externalities) prevail over negative indirect effects (crowding-out and business stealing effects), thus resulting in a positive effect on aggregate productivity. This is in line with previous empirical literature on the entry of MNEs, finding a positive direct contribution to the productivity of the host economy (Barba Navaretti and Venables, 2006); moreover, it reinforces the (scatter) previous evidence on the positive effects of having a larger number of 'domestic' MNEs localized in a territory (van Pottelsberghe de la Potterie and Lichtenberg, 2001).

Our specification allows to add an important qualification to previous results. In particular, inward foreign investments have a positive effect on regional productivity only above a certain threshold level. This result can be explained by the fact that, even large firms, such as multinationals, produce a relatively small value added in the host country with respect to the economy of a NUTS2 region. Therefore, entry of one or few multinationals make a relatively small contribution to the aggregate productivity, and it requires several foreign entries, to make a appreciable direct effect. On the other hand outward investments seem to have a positive effects up to a certain threshold, which is however very high in our sample. Results from our preferred specification suggest that about 30% of regions have higher productivity growth, thanks to the relatively large flows of inward investments, while in 50% of case productivity growth is higher due to outward investments.

These results have been showed to be robust to different specifications of the econometric model, like the inclusion of regional characteristics (in levels and growth) and the diversity in technological regimes between regions belonging to the EU-15 and regions belonging to the EU-12. We also controlled for the well-known spatial dependence in labour productivity across European regions (see Basile, 2007, among others), by estimating both the spatial lag model and the spatial error model: the positive effects of inward and outward FDIs are robust and quite stable also in terms of absolute values of the coefficients.

In conclusion we can say that both inward and outward FDIs can bring significants benefits to regional economies by increasing productivity growth. This has important implications for local and national policy. One the one hand, policies to attract inward FDIs conducive to higher productivity growth, but the effort must be substantial, so that foreign entries reach the threshold level required to determine positive effects. On the other hand, the fear of hollowing-out European knowledge which has accompanied measures aimed at reducing outward investments is not completely founded. Our results suggest that up to a certain point it is good for a region that local firms invest abroad. Thus, this calls for policies aimed at removing the obstacles to foreign investments.[33]

Further work can be done along the lines of the present analysis. First, some important regional characteristics need to be added. In particular, foreign investments benefit from better local infrastructures, which may also be associated with higher productivity (Mastromarco and Woitek, 2006), or may signal a more general association between openness and productivity (Gambardella, Mariani, and Torrisi, 2008) thus efforts need to be done to add further controls in these directions. Second, following the recent trade theory with heterogeneous firms, one may want to control for the effect of the number of investing firms (extensive margin) and the effect of the average investment of the volume of investment (intensive margin).

[33]Admittedly, many policies limiting outward investments were also motivated by the fear of job losses. While we cannot say anything on the effect on regional employment here, we argue that higher productivity growth is likely to increase jobs in the medium-run, whatever the displacement effect in the short run (Barba Navaretti and Venables, 2006).

4.8 Data Appendix: Control Variables

This section discusses the main variables that economic theory suggests to introduce as determinants of aggregate labour productivity and the actual measures used in this paper to proxy for those determinants.

4.8.1 Theory

The capital intensity needs to be taken into account, in order to control for the combination of factors (physical capital and labour) in each region. In fact, labour productivity is positively related to capital intensity in the standard theory; nonetheless, the relative endowment of production factors may be related to the amount of incoming investments, as in the case of multinational enterprises which seeks for cheap labour, thus making investments in regions with relative abundance of it. To a lesser extent, higher capital-intensive regions may be home to an higher number of enterprises which invest abroad, given that multinational enterprises are usually more capital-intensive than firms which sell their products in the domestic market only.

With respects to other driving forces of productivity at the regional level, which enter in Equation 4.1 via the vector $\mathbf{x_{ij,t}}$, three main factors have been taken into account: the level of human capital, the stock of technological capital, and the regional industrial mix. First, the positive role of human capital on productivity have been underlined by several scholars (see Mankiw, Romer, and Weil, 1992; Benhabib and Spiegel, 1994, among others). Second, both from a theoretical (Lucas, 1990) and an empirical point of view (Noorbakhsh, Paloni, and Youssef, 2001), a higher availability of well educated workers has been documented to be one of the key determinants of investment choices by multinational enterprises in a given territory. Finally several analyses, conducted at the sectoral or firm level, have underlined the positive role of outward foreign investments on the demand of high-skilled workers at home (see Head and Ries, 2002; Hansson, 2005, among others). Thus, it is relevant to include in the vector of regional characteristics a measure of human capital, which can be correlated both to regional productivity and to foreign direct investments.

The effect of technology on aggregate productivity is well known since Griliches (1979) who suggested to include a direct measure of the technology in the production function model. The idea of the technology-capital model, at the macro level, is based on the idea that technology is partly a public good and firms localized in a certain area can benefit (in terms of higher productivity) from the degree of knowledge that is available there. In a recent work on the determinants of productivity of European regions, Dettori, Marrocu, and Paci (2008) include a measure of 'technological capital' in order to investigate if observed differences in productivity are explained, *ceteris paribus*, by an higher stock of technology in the region. Nonetheless, multinational enterprises may, on the one hand, take into account positive externalities due to the average regional propensity to do research and to innovate in their decision regarding the location of the investment (Cantwell and Piscitello, 2005). On the other hand, firms investing abroad have a higher propensity to accumulate technology and human capital, thus a region with a more advanced technological base are more likely to be home to outward investing firms Thus, regional knowledge capital can be both correlated to regional productivity and foreign direct investment.

Finally, it is necessary to control for the regional industrial mix. The industry mix can be viewed as one of the structural characteristics of the region which changes in the long run. Several studies have tried to evaluate its relevance in explaining regional performance, but results are mixed. For example, Bracalente and Perugini (2008), analyzing the components of development disparities among the EU regions, find that the industry mix is relevant in explaining per capita GDP differences for regions of Eastern and Central Europe; on the other hand, Esteban (2000) finds that differences in productivity can be fully explained by the existence of region-specific productivity differentials which are uniform across sectors (e.g. human capital and infrastructures), while the regional industrial mix comes out to have a very minor role. The industrial composition can also be related to the stocks of inward and outward foreign direct investment. As for inward investments, multinational enterprises may decide to invest where particular kind of knowledge intensive services (e.g. business services) count for a large share of the economy, or where certain types of intermediate input can be easily provided. On the other hand, multinational enterprises usually belong to the most productive sectors of the economy (i.e. high-tech manufacturing and knowledge intensive services), thus an higher share of these sectors in certain regions could imply an higher share of enterprises which invest abroad. Consequently, if we do not take into account these relationships, we may incur in omitted variable biases. In the present work, the industrial mix has been taken into account both in terms of its 'quality', including weights of six different broad sectors in the regional economy, and in terms of the degree of specialization/diversification in these sectors. In the next section we offer a detailed explanation on how each variable in Equation 4.1 has been measured.

4.8.2 Measurement

Capital-labour ratio

We have included the capital-labour ratio (KL_{ijt}) in Equation 4.1, in order to control for the regional factor share. The variable has been computed as the ratio of the regional capital stock (K_{ijt}) to employment (thousands) in all sectors of the regional economy (L_{ijt}):

$$KL_{ijt} = \frac{K_{ijt}}{L_{ijt}}. \tag{4.15}$$

We have computed a measure of the capital stock at the regional level, applying the perpetual inventory method (PIM) to the series of capital investments in all the sectors of the regional economy (at 1995 prices in millions of euro)[34] taken from the European Regional Database. As for the employment series, capital investments' information for 2005 and 2006 are forecast.

We followed Hall and Mairesse (1995), and the capital stock at the beginning of the first year has been defined as below:

$$K_{ij,t=1} = \frac{I_{ij,t=1}}{g_{ij} + \delta}, \tag{4.16}$$

where $I_{ij,t=1}$ is the amount of capital investments taken by the region i in the first year of the series[35], g_{ij} is the rate of growth of capital investments observed in the region in a given span of time (in this case is from 1995-2002[36]), and δ is depreciation rate which has been set equal to 7.5%[37]. Capital stock from the second year onward has been computed using the following formula:

$$K_{ij,t} = (1 - \delta) \cdot K_{ij,t-1} + I_{ij,t}. \tag{4.17}$$

The variable has been included in logs in the econometric analysis, kl_{ijt}.

Other regional characteristics

In this Section, we detail how regional characteristics — i.e. the level of human capital, the technological capital and the regional industrial mix — have been measured.

- Human capital ($HCAP_{ijt}$) has been proxied by the (log of the) share of population aged 25 or more (thousands) with tertiary-type education degree (ISCED 5-6) in each region. Information come from the EU Regional Database, maintained by Eurostat.

- The regional technological capital ($TECH_{ijt}$) has been proxied by the ratio of the stock of patents applications to the total population (thousands) in the region (POP_{ijt}). More precisely:

$$TECH_{ijt} = \frac{INNOV_{ijt}}{POP_{ijt}} \tag{4.18}$$

[34] The series comprehend aggregate investments by the following sectors: agriculture, total energy and manufacturing, construction, market and non-market services.

[35] We start computing the capital stock series at 1995 up to 2006, even if in the econometric analysis we use the values from 2002 to 2006. The main motivation relates to the possibility to rest on a more reliable capital stock at the left hand side of Equation 4.17 for the years under analysis.

[36] For Romanian regions the investments' growth rate has been computed for the period 1998-2002, given the lack of data for the years 1995, 1996 and 1997.

[37] As robustness checks we also computed the capital stock assuming depreciation rate of 5% and 10%, and we did not register significantly different results.

The stock has been recovered using information on the number of patent applications to the European Patent Office (EPO) coming from each European region, which are available in the database maintained by Eurostat[38]. Data on total population comes from the database developed by Cambridge Econometrics. The stock for the years $t = (2003{,}2004{,}2005{,}2006)$ has been computed as the sum of the patent applications in all sectors in the previous five years ($PATAPP_{ijt}$), plus the current year applications:

$$INNOV_{ij,t} = \sum_{t=t-5}^{t} PATAPP_{ijt}. \tag{4.19}$$

The ratio has been included in logs in the econometric analysis, $tech_{ijt}$.

- We have taken into account the regional industrial mix (SH_{s*ijt}), by introducing the share of employment in six broad sectors s^* of the regional economy: Agriculture, hunting, forestry and fishing (AC), Electricity, gas, water supply and Constructions (EF), High-tech manufacturing & Medium high-tech manufacturing (HD), Medium low-tech manufacturing & Low-tech Manufacturing (LD), Knowledge-intensive services (KI) and Less knowledge-intensive (LKI) services. Each share has been computed in the following way:

  $$SH_{s*ijt} = \frac{L_{s*ijt}}{L_{ijt}}$$

  where L_{ijt} and L_{s*ijt} denote, respectively, total employment in the region i which belongs to country j (thousands), and employees belonging to the sector $s*$. To avoid multicollinearity we introduced five coefficients in the regressions. The excluded sectoral share is the AC sector (Agriculture, hunting, forestry, fishing, mining and quarrying). Data regarding employees in each sector come from the database maintaned by Eurostat.

  The data on employment by sectors, showed a given amount of missing observations (region/year); in order not to loose them, we have used linear interpolation to fill the gaps for all the observations that were 'missing', but which had 'non-missing' observations the year before and the year after the missing ones. We further filled in a small amount of missing observations in the High-tech manufacturing sector (which showed the highest number of missing observations among the considered sectors) as the difference between total regional employment and the sum of employees in all the others sectors (AC, EF, Medium-high tech manufacturing, Medium-low tech manufacturing, Low-tech manufacturing, KI, LKI).

[38] Data on patent applications are regionalised on the basis of the investors' residence: in the case of multiple investors proportional quotas have been attributed to each region.

- We have controlled for the degree of concentration/diversification of the regional industrial mix. Following the literature (see Cingano and Schivardi, 2004; Bracalente and Perugini, 2008, among others), we have used the Herfindahl-Hirschman index as a proxy for concentration/diversification computed as follows:

$$HHI_{ijt} = \sum_s SH_{sijt}^2 = \sum_s \left(\frac{L_{sijt}}{L_{ijt}}\right)^2, \tag{4.20}$$

where SH_{sijt} are a more detailed disaggregation of the employment shares defined above. In fact, as elements of the HHI we take into account 8 broad sectors, s: Agriculture, hunting, forestry and fishing (AC), Electricity, gas, water supply and Constructions (EF), High-tech manufacturing (HTD), Medium high-tech manufacturing(MHTD), Medium low-tech manufacturing (MLTD), Low-tech Manufacturing (LTD), Knowledge-intensive services (KI) and Less knowledge-intensive (LKI) services. In particular, we consider the HTD and the MHTD as two separate sectors here, and the same holds for the LTD and the MLTD which are considered separate elements of the HHI[39]. The HHI index, which is equal to '1' for regions with all employees in one sector and which goes toward '0' for more diversified regional structures, allows us to control for the sectoral concentration/variety of the region, while by introducing the SH_{s^*it} ratios, we account for the different 'quality' of the industrial mix. For any given level of HHI we expect regional productivity to be higher in regions where the share of high-value added activities (such as High-tech Manufacturing and Knowledge-intensive services) is higher[40].

The HHI enters in logs in the econometric analysis, hhi.

4.9 Data Appendix: Other information on the database

4.9.1 Foreign investments; source: fDi Markets database

We can gain more insights on the scope and the reliability of the data on investment projects, looking at the distribution of projects by business activity: Table 4.11 shows the breakdown of investments by business activity using information contained in the fDi Markets database, and Table 4.12 shows some general statistics of investment projects directed towards the EU and originated by firms located in the EU.

Seven major categories of business activities can be derived from a more disaggregated taxonomy of nineteen categories, which is contained in the fDi Markets

[39] The detailed taxonomy of sectors s is presented in Table 4.13 of the Appendix 2 4.9.2.

[40] The use of different levels of aggregation in the HHI with respect to these employments shares is motivated both by the achieved greater precision of the Herfindahl-Hirschman index, which aims at capturing the variability in the regional industrial mix, and –on the contrary– by the attempt to minimize over-specification in the estimates of the coefficients of the sectoral employment shares.

Table 4.11: Taxonomy of investments by business activity

Categories	Business Activities
Manufacturing	Manufacturing
Business Services	Business Services
Headquarters	Headquarters
R&D	Design & Related Activities
	Research & Development
Sales	Sales & Marketing
	Retail
Other industries	Construction
	Electricity
	Extraction
	Recycling
Other services	Customer Contact Centres
	Education & Training
	ICT and Internet Infrastructure
	Logistics
	Maintenance
	Shared Service Centers
	Technical Support Centers

database. As one would expect, more than 95% of EU outward investments are made

Table 4.12: Outward and inward international investment projects in Europe, by business activity, 2003-2008

	Area of destination			Area of origin		
	EU-12(%)	EU-15(%)	Total(%)	EU-12(%)	EU-15(%)	Total(%)
Manufacturing	61.2	38.8	100	4.0	96.0	100
R&D	21.2	78.8	100	1.2	98.8	100
Sales	27.1	72.9	100	5.4	94.6	100
Business Services	23.8	76.2	100	5.5	94.5	100
Headquarters	7.7	92.3	100	0.8	99.2	100
Other industries*	49.0	51.1	100	6.5	93.5	100
Other services**	32.6	67.4	100	3.4	96.6	100
Total	33.8	66.2	100	4.7	95.3	100

*Construction, extraction, electricity and recycling
**Logistics, ICT, Customer Contact Cent. and maintenance

from EU-15 countries[41], and this share is even higher when R&D or Headquarter activities are concerned. Inward investments are split more evenly among EU-15 and EU-12 countries[42], but significant differences emerge when different business activities are considered. EU-12 countries attract the majority of new Manufacturing plants (61%) and a half of investments in Construction, extraction, elextricity and recyclying, while countries in the EU-15 attract almost all the investments aimed at the creation of Headquarters [43], and a large share of R&D, Business Services, Sales

[41] Italy, France, Netherlands, Luxumbourg, Belgium, Germay, United Kingdom, Denmark, Ireland, Greece, Spain, Portugal, Austria, Switzerland and Finland.

[42] Bulgaria, Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovakia and Slovenia.

[43] Investments in headquarters do not mean that the firm is moving its headquarters, but rather they

and Logistics, ICT, Customer Contact Cent. and Maintenance plants.

4.9.2 Shares of employment by sectors; source: EU Regional Database by Eurostat

The taxonomy of broad sectors —which have been used in order to build the Herfindahl index of diversification and the shares of employment which proxy the regional industrial mix— has been taken from the list which has been proposed by Eurostat in the EU regional database. We cross-refer the reader to the technical repost by Felix (2006) for further details on the employed taxonomy. Sectors are presented in Table 4.13.

4.9.3 List of regions

The list of the NUTS 2 regions which have been considered in the baseline Specification (3) is reported in Table 4.14. Overall, we can account for 255 regions (and 746 observations) belonging to the EU in our analysis, for the period 2003-2006.

4.9.4 Inserting regional controls one by one in Specification (3)

In this Section, the variables related to the regional carachteristics are introduced one by one in the econometric Specification (3). Overall, variables of main concern, i.e. FDI variables and the capital-labour ratio are stable both in terms of magnitude and in terms of statistical significance. The sole coefficient which faces a small shrinkage in its absolute value once the industrial mix has been taken into account is the one related to the dummy variable of outward investments, γ_O^d. That could be explained by the fact that multinational enterprises are concentrated in more productive sectors of the economy, likely the high-tech manufacturing sectors (HD) and the knoledge intensive services sectors. Interestingly enough in the estimation of Specification (3), which is reported in the last column of Table 4.15, two controls only show significant coefficients, namely the positive effect of a change in the degree of concentration in the industrial structure, which is captured by the δ_{hhi} coefficient, and the negative effect of an increase of the share of less knowledge intensive services sectors in the regional economy, measured by the δ_{LKIS} coefficient. Neither the coefficient of the growth rate of the human capital, δ_{hcap}, nor the one of the technological capital, δ_{tech}, are significant in any column even in Specifications (1a) and (1b) in which they are introduce separately from the other controls.

may be creating a regional or functional headquarter abroad.

Table 4.13: Breakdown of sectors (Nace Rev. 1.1 codes)

Agriculture, hunting, forestry and fishing	01 to 05 Agriculture, hunting, forestry and fishing
Electricity, gas, water supply and constructions	40 to 41; 45 Electricity, gas, water supply and constructions
High-tech Manufacturing	30 Manufacture of office machinery and computers 32 Manufacture of radio, television and communication equipment and apparatus 33 Manufacture of medical, precision and optical instruments, watches and clocks
Medium High-tech Manufacturing	24 Manufacture of chemicals and chemicals products 29 Manufacture of machinery and equipment n.e.c. 31 Manufacture of electrical machinery and apparatus n.e.c. 34 and 35 Manufacture of transport equipment
Low and medium-low-tech Manufacturing	15 to 22 Manufacture of food products, beverages and tobacco; textiles and textile products; leather and leather products; wood and wood products; pulp, paper and paper products,publishing and printings 23 Manufacture of coke, refined petroleum products and nuclear fuel 25 to 28 Manufacture of rubber and plastic products; basic metals and fabricated metals product; other non-metallic mineral products 36 to 37 Manufacturing n.e.c.
Knowledge-intensive services	61 Water Transport 62 Air Transport 64 Post and telecommunications 65 to 67 Financial intermediation 70 to 74 Real estate, renting and business activities 80 Education 85 Health and social work 92 Recreational, cultural and sporting activities
Less knowledge-intensive services	50 to 52 Motor trade 55 Hotels and restaurants 60 Land transport ; transport via pipelines 63 Supporting and auxiliary transport activities; activities of travel agencies 75 Public administration and defence; compulsory social security 90 Sewage and refuse disposal, sanitation and similar activities 91 Activities of membership organization n.e.c. 93 Other service activities 95 Activities of households as employers of domestic staff 99 Extra-territorial organizations and bodies

Table 4.14: List of the 255 regions considered in the present study, by country

NUTS	Name	# regions	NUTS	Name	# regions	NUTS	Name	# regions	NUTS	Name	# regions
Austria		9	**Germany**		39	ITD3	Veneto		ES41	Castilla y León	
AT11	Burgenland		DE11	Stuttgart		ITD4	Friuli-Venezia Giulia		ES42	Castilla-la Mancha	
AT12	Niederösterreich		DE12	Karlsruhe		ITD5	Emilia-Romagna		ES43	Extremadura	
AT13	Wien		DE13	Freiburg		ITE1	Toscana		ES51	Cataluña	
AT21	Kärnten		DE14	Tübingen		ITE2	Umbria		ES52	Comunidad Valenciana	
AT22	Steiermark		DE21	Oberbayern		ITE3	Marche		ES53	Illes Balears	
AT31	Oberösterreich		DE22	Niederbayern		ITE4	Lazio		ES61	Andalucia	
AT32	Salzburg		DE23	Oberpfalz		ITF1	Abruzzo		ES62	Región de Murcia	
AT33	Tirol		DE24	Oberfranken		ITF2	Molise		ES70	Canarias (ES)	
AT34	Vorarlberg		DE25	Mittelfranken		ITF3	Campania		**Sweden**		8
Belgium		11	DE26	Unterfranken		ITF4	Puglia		SE11	Stockholm	
BE10	Brussels		DE27	Schwaben		ITF5	Basilicata		SE12	Östra Mellansverige	
BE21	Prov. Antwerpen		DE30	Berlin		ITF6	Calabria		SE21	Småland med öarna	
BE22	Prov. Limburg (B)		DE41	Brandenburg-Nordost		ITG1	Sicilia		SE22	Sydsverige	
BE23	Prov. Oost-Vlaanderen		DE42	Brandenburg-Südwest		ITG2	Sardegna		SE23	Västsverige	
BE24	Prov. Vlaams Brabant		DE50	Bremen		**Latvia**		1	SE31	Norra Mellansverige	
BE25	Prov. West-Vlaanderen		DE60	Hamburg		LV00	Latvia		SE32	Mellersta Norrland	
BE31	Prov. Brabant Wallon		DE71	Darmstadt		**Lithuania**		1	SE33	Övre Norrland	
BE32	Prov. Hainaut		DE72	Gießen		LT00	Lithuania		**United Kingdom**		36
BE33	Prov. Liège		DE73	Kassel		**Luxembourg**		1	UKC1	Tees Valley and Durham	
BE34	Prov. Luxembourg (B)		DE80	Mecklenburg-Vorpommern		LU00	Luxembourg		UKC2	Northumberland, Tyne and Wear	
BE35	Prov. Namur		DE91	Braunschweig		**Malta**		1	UKD1	Cumbria	
Bulgaria		6	DE92	Hannover		MT00	Malta		UKD2	Cheshire	
BG31	Severozapaden		DE93	Lüneburg		**Poland**		16	UKD3	Greater	
BG32	Severen tsentralen		DE94	Weser-Ems		PL11	Lódzkie		UKD4	Lancashire	
BG33	Severoiztochen		DEA1	Düsseldorf		PL12	Mazowieckie		UKD5	Merseyside	
BG34	Yugoiztochen		DEA2	Köln		PL21	Malopolskie		UKE1	East Riding and North Lincolnshire	
BG41	Yugozapaden		DEA3	Münster		PL22	Slaskie		UKE2	North Yorkshire	
BG42	Yuzhen tsentralen		DEA4	Detmold		PL31	Lubelskie		UKE3	South Yorkshire	
Cyprus		1	DEA5	Arnsberg		PL32	Podkarpackie		UKE4	West Yorkshire	
CY00	Cyprus		DEB1	Koblenz		PL33	Swietokrzyskie		UKF1	Derbyshire and Nottinghamshire	
Czech Republic		8	DEB2	Trier		PL34	Podlaskie		UKF2	Leicestershire, Rutland and Northants	
CZ01	Praha		DEB3	Rheinhessen-Pfalz		PL41	Wielkopolskie		UKF3	Lincolnshire	
CZ02	Stredni Cechy		DEC0	Saarland		PL42	Zachodniopomorskie		UKG1	Herefordshire, Worcestershire and Warks	
CZ03	Jihozápad		DED1	Chemnitz		PL43	Lubuskie		UKG2	Shropshire and Staffordshire	
CZ04	Severozápad		DED2	Dresden		PL51	Dolnoslaskie		UKG3	West Midlands	
CZ05	Severovýchod		DED3	Leipzig		PL52	Opolskie		UKH1	East Anglia	
CZ06	Jihovýchod		DEE0	Sachsen-Anhalt		PL61	Kujawsko-Pomorskie		UKH2	Bedfordshire, Hertfordshire	
CZ07	Stredni Morava		DEF0	Schleswig-Holstein		PL62	Warminsko-Mazurskie		UKH3	Essex	
CZ08	Moravskoslezko		DEG0	Thüringen		PL63	Pomorskie		UKI	London	
Estonia		1	**Greece**		12	**Portugal**		5	UKJ1	Berkshire, Bucks and Oxfordshire	
EE00	Estonia		GR11	Anatoliki Makedonia, Thraki		PT11	Norte		UKJ2	Surrey, East and West Sussex	
Finland		5	GR12	Kentriki Makedonia		PT15	Algarve		UKJ3	Hampshire and Isle of Wight	
FI13	Itä-Suomi		GR13	Dytiki Makedonia		PT16	Centro (PT)		UKJ4	Kent	
FI18	Etelä-Suomi		GR14	Thessalia		PT17	Lisboa		UKK1	Gloucestershire, Wiltshire and North Somerset	
FI19	Länsi-Suomi		GR21	Ipeiros		PT18	Alentejo		UKK2	Dorset and Somerset	
FI1A	Pohjois-Suomi		GR22	Ionia Nisia		**Romania**		8	UKK3	Cornwall and Isles of Scilly	
FI20	Åland		GR23	Dytiki Makedonia		RO11	Nord-Vest		UKK4	Devon	
France		22	GR24	Sterea Ellada		RO12	Centru		UKL1	West Wales and The Valleys	
FR10	Île de France		GR25	Peloponnisos		RO21	Nord-Est		UKL2	East Wales	
FR21	Champagne-Ardenne		GR30	Attiki		RO22	Sud-Est		UKM2	Eastern Scotland	
FR22	Picardie		GR42	Notio Aigaio		RO31	Sud		UKM3	South Western Scotland	
FR23	Haute-Normandie		GR43	Kriti		RO32	Bucuresti		UKM5	North Eastern Scotland	
FR24	Centre		**Hungary**		7	RO41	Sud-Vest		UKM6	Highlands and Islands	
FR25	Basse-Normandie		HU10	Közép-Magyarország		RO42	Vest		UKN0	Northern Ireland	
FR26	Bourgogne		HU21	Közép-Dunántúl		**Slovakia**		4	**Netherlands**		12
FR30	Nord-Pas-de-Calais		HU22	Nyugat-Dunántúl		SK01	Bratislavský		NL11	Groningen	
FR41	Lorraine		HU23	Dél-Dunántúl		SK02	Západné Slovensko		NL12	Friesland(NL)	
FR42	Alsace		HU31	Észak-Magyarország		SK03	Stredné Slovensko		NL13	Drenthe	
FR43	Franche-Comté		HU32	Észak-Alföld		SK04	Východné Slovensko		NL21	Overijssel	
FR51	Pays de la Loire		HU33	Dél-Alföld		**Slovenia**		1	NL22	Gelderland	
FR52	Bretagne		**Ireland**		2	SI0	Slovenia		NL23	Flevoland	
FR53	Poitou-Charentes		IE01	Border, Midlands and Western		**Spain**		17	NL31	Utrecht	
FR61	Aquitaine		IE02	Southern and Eastern		ES11	Galicia		NL32	Noord-Holland	
FR62	Midi-Pyrénées		**Italy**		21	ES12	Principado de Asturias		NL33	Zuid-Holland	
FR63	Limousin		ITC1	Piemonte		ES13	Cantabria		NL34	Zeeland	
FR71	Rhône-Alpes		ITC2	Valle d'Aosta/Vallée d'Aoste		ES21	Pais Vasco		NL41	Noord-Brabant	
FR72	Auvergne		ITC3	Liguria		ES22	Comunidad Foral de Navarra		NL42	Limburg(NL)	
FR81	Languedoc-Roussillon		ITC4	Lombardia		ES23	La Rioja				
FR82	Provence-Alpes-Côte d'Azur		ITD1	Provincia Autonoma Bolzano-Bozen		ES24	Aragón				
FR83	Corse		ITD2	Provincia Autonoma Trento		ES30	Comunidad de Madrid		**Total**		255

Table 4.15: Entering one by one the regional controls in the preferred Specification (3)

		Specification						
		1	1a	1b	1c	1d	1e	3
Variable	Coefficient							
OUT_{t-1}(dummy)	γ_O^d	0.0088***	0.0088***	0.0088***	0.0087***	0.0083***	0.0077***	0.0075**
		(0.0029)	(0.0029)	(0.0029)	(0.0029)	(0.0029)	(0.0029)	(0.0029)
OUT_{t-1}(log. of n.inv)	γ_O^{log}	-0.0030***	-0.0030***	-0.0030***	-0.0030***	-0.0030***	-0.0030***	-0.0029***
		(0.0009)	(0.0010)	(0.0010)	(0.0009)	(0.0009)	(0.0009)	(0.0010)
IFDI_{t-1}(dummy)	γ_I^d	-0.0074***	-0.0074***	-0.0073***	-0.0073***	-0.0073***	-0.0071***	-0.0072***
		(0.0026)	(0.0026)	(0.0027)	(0.0027)	(0.0026)	(0.0027)	(0.0027)
IFDI_{t-1}(log. of n.inv)	γ_I^{log}	0.0031***	0.0032***	0.0031***	0.0031***	0.0032***	0.0032***	0.0031***
		(0.0011)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
$\Delta_{t,t-1}kl$	β	0.2401***	0.2390***	0.2383***	0.2369***	0.2409***	0.2361***	0.2392***
		(0.0839)	(0.0842)	(0.0842)	(0.0846)	(0.0853)	(0.0831)	(0.0842)
$\Delta_{t,t-1}$hcap	δ_{hcap}		0.0004	0.0004	-0.0012	-0.0004	-0.0030	0.0003
			(0.0138)	(0.0138)	(0.0140)	(0.0141)	(0.0145)	(0.0137)
$\Delta_{t,t-1}$tech	δ_{tech}			0.0028	0.0025	0.0024	0.0024	0.0008
				(0.0105)	(0.0103)	(0.0103)	(0.0105)	(0.0100)
$\Delta_{t,t-1}$hhi	δ_{hhi}				0.0429*	0.0609**	0.0786	0.1577**
					(0.0258)	(0.0288)	(0.0495)	(0.0740)
$\Delta_{t,t-1}$SH_EF	δ_{EF}					0.1769	0.2131	0.1434
						(0.1278)	(0.1478)	(0.1509)
$\Delta_{t,t-1}$SH_HD	δ_{HD}						0.2298*	0.1638
							(0.1336)	(0.1416)
$\Delta_{t,t-1}$SH_LD	δ_{LD}						-0.0815	-0.1430
							(0.1564)	(0.1557)
$\Delta_{t,t-1}$SH_KIS	δ_{KIS}							-0.1876
								(0.1690)
$\Delta_{t,t-1}$SH_LKIS	δ_{LKIS}							-0.3052*
								(0.1751)
Constant	α	0.0272***	0.0271***	0.0271***	0.0274***	0.0276***	0.0269***	0.0270***
		(0.0039)	(0.0039)	(0.0039)	(0.0039)	(0.0038)	(0.0039)	(0.0039)
Observations		755	746	746	746	746	746	746
Regions		258	255	255	255	255	255	255

Significance levels: * 10%, ** 5%,*** 1%
Cluster-robust standard errors in parentheses

Chapter 5

Conclusions

5.1 Concluding remarks and step for further research

The analysis of performance is a wide field of economic research which embraces rather different issues and levels of investigation. Two main concepts are at the core of measuring the performance of production units (firms and organizations, industries, regions and countries): *productivity* and *efficiency*. This thesis deals with these two concepts looking at different levels of analysis, and making use of different methods.

In Chapter 1, productivity and efficiency are defined and grounded on production theory in economics; subsequently, I introduce a unified framework of analysis in which the two concepts are presented in a coherent and up-to-date way. The presented framework has the advantage to consider both of them to be possible in economic relevant settings. A long discussion on the methods which are available to the researcher follows, and I have tried to detail strengths and weakness of each method reviewing empirical studies which have employed real and simulated data. A survey of relevant literatures —micro and macro— on the determinants of productivity and efficiency concludes the introductory Chapter.

Chapter 2 focuses on stochastic frontier models. In this framework each observed unit is considered to lie on or below its production frontier, which is defined by the 'best-practice' units. One of the typical characteristic of these models is the need to assume a distributional form for the unobservable inefficiency term. The relevant literature has suggested different one-sided distributional form, basically the half-normal, the exponential and the truncated normal. Scholars of the field have questioned whether the assumption on the specific distribution of the inefficiency term is relevant and may actually drive the results of the analysis: a common practice is to compare the results obtained by estimating differing —in the specification of the inefficiency distribution— stochastic frontier models from the same sample of production units; previous evidence indicates general concordance among set of estimated inefficiency scores. However, an extensive exercise on this issue is still lacking in the literature. Using Monte Carlo simulations, I have showed that for each of the six inefficiency distributions considered, the three estimated models —i.e., normal-half-normal, normal-exponential and normal-truncated normal—, reproduce the inefficiency ranking with the same precision. This is true for both correctly specified models (those which assume the correct inefficiency distribution) and for misspecified ones. It is important to note that the results are robust to various types of misspecifications. This is a useful piece of evidence for practitioners, because if the ranking of inefficiency is the main object of the analysis, the three most frequently employed models give the same results. Conversely, if one is interested in the inefficiency value *per se*, it is important to specify the correct distribution of the inefficiency term: for each of the estimated models the average difference between true inefficiency scores and estimated ones is lower when the model is correctly specified than when it is misspecified. From a methodological point of view, this paper examined the role of variance of error components. In previous experiments on the misspecification of the inefficiency distribution, like those of Ruggiero (1999) and Jensen (2005), the results 'suffered' from the fact that the authors compared inefficiency distributions with the same value of σ_u, but different values of $Var(u)$. In this paper we performed all experiments in two different settings: the first was similar to that of previous studies for purposes of comparison, and the second, in order to check the robustness of the results, kept the variance of the overall error term fixed and move the (square root of) the ratio of variances ($Var(u)/Var(v)$): the main result of the paper remain stable in the two settings. A further development of this study could be to examine misspecification of noise term, v_i, on the correct estimation of the inefficiency scores: this is a type of misspecification which has been almost completely neglected in previous works and which was briefly considered only by Jensen (2005). It would be also interesting to explore the consequences of the correct estimation of the inefficiency scores of a neglected correlation between the two random terms, u_i and v_i, which are always assumed to be uncorrelated.

In Chapter 3 I study the relationship between vertical integration and firm efficiency in the Italian machine tool industry. I have first set up a theoretical model, in order to come up with a testable hypothesis: more efficient firms decide to produce as vertically integrated, bearing higher (organizational) fixed costs while less efficient firms choose to outsource part of production process buying an intermediate input from other firms, thus reducing fixed costs but bearing higher marginal costs of production. In equilibrium, the two types of organizations coexist and the industry contemplates firms with different levels of efficiency. This theoretical result is pretty much in line with the previous quantitative and qualitative evidence on the industry, as the work by Zanfei and Gambardella (1994) who claim that in the Italian MT sector firms with different size, organization structures and sourcing strategies coexist, and complement each other in supplying the market all the varieties requested by a highly differentiated demand, or Wengel and Shapira (2004) who points to a dualistic structure of the industry. However, while previous work has stressed the general characteristic of 'size' as point of differentiation between the two groups we think that the vertical structure better represents the different choices for the organization of production. I have empirically strengthened this result, conducting a stochastic frontier analysis on a sample of more than 500 machine tool producers. In this way it is possible to estimate the best practice technology frontier, measuring the distance to it as indicators of inefficiency (sub-optimal level of output, given the amount of inputs and the available technology). The empirical analysis shows that vertical integrated firms present a lower variance (and lower mean) of the inefficiency distribution, after having controlled for firm size, type of ownership, agglomeration economies and the economic cycle. Thus, vertical integrated firms are, *ceteris paribus* more efficient in the industry under analysis than disintegrated firms. Overall, this paper contributes to a better understanding of the coexistence of heterogeneous firms characterized by different levels of efficiency and different organizational forms. Moreover, the stochastic frontier framework allows me to estimate firm inefficiency as the distance from the technology frontier (the best practice) and to jointly estimate the relationship between the degree of vertical integration and inefficiency. This can be considered as an improvement with respect to previous works on the same topic, which rested on more traditional 2-step procedures which may lead up to omitted variable bias and under-dispersion of productive efficiency scores in the first step of the analysis. Among the lines for future research, I highlight that: (i) a qualitative analysis of a small number of firms in the industry could be a natural complement to this study: the econometric analysis could be grounded in a careful description of the stages of the production process which are actually kept in-house; (ii) some econometric refinements may be possible, as a direct attempt to account for a reverse causality, from vertical integration to firm efficiency.

In Chapter 4 I move to the regional level of analysis and exploiting an original and extensive dataset on FDIs, I investigate the relationship between FDIs and productivity in in a large set of NUTS2 regions in almost all countries of the Enlarged Europe (EU-27). Despite the increasing evidence of integration of sub-national economies in the global arena, and the positive role of multinational firms for economic prosperity in local economies documented in a number of recent studies, evidence on the relationship between foreign investments and regional performance is lacking. The results of the econometric analysis support that both inward and outward foreign direct investments have positive effects on productivity growth at the regional level, after controlling for a relevant set of regional characteristics, such as human capital, technology capital and industry mix. This is an interesting result, and it is an original contribution to the international economics literature in several dimensions: previous studies with a regional perspective have focused on comparisons within single countries and have addressed only the role of 'inward' investments as a driver of increasing local performance. Moreover, those few studies which have attempted to assess the specific role of outward investments on productivity have taken a country perspective, almost neglecting the sub-national level of analysis. These results have been showed to be robust to different specifications of the econometric model, like the inclusion of regional characteristics (in levels and growth) and the diversity in technological regimes between regions belonging to the EU-15 and regions belonging to the EU-12. I also controlled for the well-known spatial dependence in labour productivity across European regions, by estimating both the spatial lag model and the spatial error model: the positive effects of inward and outward FDIs are robust and quite stable also in terms of absolute values of the coefficients. These results have important implications for local and national policy. One the one hand, policies to attract inward FDIs conducive to higher productivity growth, but the effort must be substantial, so that foreign entries reach the threshold level required to determine positive effects. On the other hand, the fear of hollowing-out European knowledge which has accompanied measures aimed at reducing outward investments is not completely founded. Results suggest that up to a certain point it is good for a region that local firms invest abroad. Thus, this calls for policies aimed at removing the obstacles to foreign investments. Further work can be done along the lines of the present analysis. First, some important regional characteristics need to be added. In particular, foreign investments benefit from better local infrastructures, which may also be associated with higher productivity, or may signal a more general association between openness and productivity thus efforts need to be done to add further controls in these directions. Second, following the recent trade theory with heterogeneous firms, one may want to control for the effect of the number of investing firms (extensive margin) and the effect of the average investment of the volume of investment (intensive margin).

Bibliography

Abramovitz, M. (1956): "Resorce and Output Trends in the United States Since 1870," *American Economic Review*, 46, 5–23.

——— (1979): "Rapid Growth Potential and its Realization: The Experience of Capitalist Economies in the Postwar Period," in *Economic Growth and Resources: The Major Issues*, ed. by E. Malinvaud, vol. 1. London: Macmillan.

Acemoglu, D., S. Johnson, and T. Mitton (2009): "Determinants of Vertical Integration: Financial Development and Contracting Costs," *Journal of Finance*, 64, 1251–1290.

Adelman, M. A. (1955): "Concept and Statistical Measurement of Vertical Integration," in *Business Concentration and Price Policy*, ed. by U.-N. Bureau, pp. 279–328. Princeton University Press.

Afriat, S. N. (1972): "Efficiency Estimation of Production Functions," *International Economic Review*, 13, 568–598.

Aghion, P., and P. Howitt (1992): "A Model of Growth Through Creative Destruction," *Econometrica*, 60, 323–351.

Aigner, D., K. Lovell, and P. Schmidt (1977): "Formulation and Estimation of Stochastic Frontier Production Function Models," *Journal of Econometrics*, 6, 21–37.

Aigner, D. J., and S. F. Chu (1968): "On Estimating the Industry Production Function," *American Economic Review*, 58, 826–39.

Aitken, B., and A. Harrison (1999): "Do domestic firms benefit from foreign investment? Evidence from Venezuela," *American Economic Review*, 89, 605–618.

Alchian, A. A. (1965): "Some Economics of Property Rights," *Il Politico*, 30, 816–829.

Alchian, A. A., and R. A. Kessel (1962): "Competition, Monopoly, and the Pursuit of Money," in *Aspects of Labor Economics*. Princeton: Princeton University Press for the National Bureau of Economic Research.

Altomonte, C., and I. Colantone (2009): “Firm Heterogeneity and Endogenous Regional Disparities”,” *Journal of Economic Geography.*

Alvarez, A. (2007): “Decomposing regional productivity growth using an aggregate production frontier,” *The Annals of Regional Science*, 41, 431–441.

Alvarez, A., C. Amsler, L. Orea, and P. Schmidt (2006): “Interpreting and testing the scaling property in models where inefficiency depends on firm characteristics,” *Journal of Productivity Analysis*, 25, 201–212.

Anselin, L. (1988): *Spatial econometrics: methods and models.* Dordrecht: Kluwer.

Antras, P., and E. Helpman (2004): “Global Sourcing,” *Journal of Political Economy*, 112, 552–580.

Arrighetti, A. (1999): “Integrazione verticale in Italia e in Europa: tendenze ed ipotesi interpretative,” in *La 'questione dimensionale' nell'industria italiana*, ed. by F. Traù. Il Mulino.

Aw, B. Y., S. Chung, and M. J. Roberts (2000): “Productivity and Turnover in the Export Market: Micro-level Evidence from the Republic of Korea and Taiwan (China),” *The World Bank Economic Review*, 14, 65–90.

Badunenko, O., M. Fritsch, and A. Stephan (2008): “What drives the productive efficiency of a firm?,” DIW Berlin, Discussion Papers 2008.

Baldwin, R. E. (2006): “Globalisation: the great unbundling(s),” in *Globalisation challenges for Europe*, ed. by S. of the Economic Council. Finnish Prime Ministerś Office.

Barba Navaretti, G., D. Castellani, and A.-C. Disdier (2010): “How Does Investing in Cheap Labour Countries Affect Performance at Home? France and Italy,” *Oxford Economic Papers*, doi:10.1093/oep/gpp010.

Barba Navaretti, G., and A. J. Venables (2006): *Le multinazionali nell'economia mondiale.* il Mulino.

Bartelsman, E., S. Scarpetta, and F. Schivardi (2005): “Comparative analysis of firm demographics and survival: evidence from micro-level sources in OECD countries,” *Industrial and Corporate Change*, 14(3), 365–391.

Bartelsman, E. J., and M. Doms (2000): “Understanding Productivity: Lessons from Longitudinal Microdata,” *Journal of Economic Literature*, 38(3), 569–594.

Basile, R. (2007): “Productivity Polarization Across Regions in Europe,” *Quaderni del Dipartimento di Economia, Finanza e Statistica, Università degli Studi di Perugia*, 31.

Basile, R., D. Castellani, and A. Zanfei (2008): "Location Choices of Multinational Firms in Europe: The Role of EU Cohesion Policy," *Journal of International Economics*, 74, 328–340.

Battese, G. E., and T. Coelli (1992): "Frontier production functions, technical efficiency and panel data: with applications to paddy farmers in India," *Journal of Productivity Analysis*, 3, 153–169.

——— (1995): "A model for technical inefficiency effects in a stochastic frontier production function for panel data," *Empirical Economics*, 20, 325–332.

Battese, G. E., and T. J. Coelli (1988): "Prediction of Firm-Level Technical Efficiences with a Generalized Frontier Production Function and Panel Data," *Journal of Econometrics*, 38, 387–399.

Becattini, G. (1990): "The marshallian industrial district as socio-economic notion," in *Industrial Districts and Inter-firm Cooperation in Italy*, ed. by F. Pyke, G. Becattini, and W. Sengenberger. International Institute of Labour Studies.

Becchetti, L., A. Panizza, and F. Oropallo (2008): "Il ruolo dei distretti industriali nella performance in termini di export e di valore aggiunto: evidenze empiriche dall'universo delle imprese italiane," in *Competitività ed efficienza dell'economia italiana: fattori sistemici e valutazioni quantitative*, ed. by G. Cella, and A. Zago, pp. 291–335. Il Mulino.

Benfratello, L., and A. Sembenelli (2006): "Foreign ownership and productivity: Is the direction of causality so obvious?," *International Journal of Industrial Organization*, 4, 733–751.

Benhabib, J., and M. Spiegel (1994): "The role of Human Capital in Economic Development: evidence from aggregate cross-country data," *Journal of Monetary Economics*, 34, 143–174.

Berger, S. (2006): *How we compete. What companies around the world are doing to make it in today's global economy.* Doubleday Random House.

Bernard, A. B., and J. B. Jensen (1995): "Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987," *Brookings Papers on Economic Activity: Microeconomics*, pp. 67–119.

——— (1999): "Exceptional exporter performance: cause, effect, or both?," *Journal of International Economics*, 47, 1–25.

Bitzer, J., and H. Görg (2009): "Foreign direct investment, competition and industry performance," *The World Economy*, 31(2), 221–233.

Blomström, M., and A. Kokko (1998): "Regional Integration and Foreign Direct Investment," *NBER working papers*, 6091.

Blundell, R., and S. Bond (2000): "GMM estimation with persistent panel data: an application to production functions," *Econometric Reviews*, 19, 321–340.

Bottasso, A., and A. Sembenelli (2004): "Does Ownership affect firms' efficiency?," *Empirical Economics*, 29, 769–786.

Bracalente, B., and C. Perugini (2008): "The components of regional disparities in Europe," *The Annals of Regional Science*.

Braconier, H., K. Ekholm, and K. H. M. Knarvik (2001): "Does FDI work as a channel for R&D spillovers? Evidence based on Swedish data," *Review of World Economics*, 137, 644–665.

Brainard, S. L. (1997): "An empirical assessment of the proximity-concentration trade-off between multinational sales and trade," *American Economic Review*, 87, 520–544.

Broedner, P., S. Kinkel, and G. Lay (2009): "Productivity effects of outsourcing. New evidence on the strategic importance of vertical integration decisions," *International Journal of Operations & Production Management*, 29, 127–150.

Cantwell, J. A. (1995): "The Globalisation of Technology: What Remains of the Product Cycle Model?," *Cambridge Journal of Economics*, 19(1), 155–74.

Cantwell, J. A., and L. Piscitello (2005): "The recent location of foreign R&D activities by large MNCs in the European regions. The role of spillovers and externalities," *Regional Studies*, 39, 1–16.

Carree, M. A. (2002): "Technological inefficiency and the skewness of the error component in stochastic frontier analysis," *Economics Letters*, 77, 101–107.

Castellani, D. (2002): "Export Behavior and Productivity Growth: Evidence from Italian Manufacturing Firms.," *Review of World Economics*, 138, 605–628.

Castellani, D., and A. Zanfei (2006): *Multinational Firms, Innovation and Productivity*. Edward Elgar.

——— (2007): "Internationnalisation, Innovation and Productivity: How do Firms Differ in Italy?," *The World Economy*, 30, 156–176.

Caudill, S., and J. Ford (1993): "Biases in Frontier Estimation Due to Heteroskedasticity," *Economics Letters*, 41, 17–20.

Caudill, S., J. Ford, and D. Gropper (1995): "Frontier Estimation and Firm Specific Inefficiency Measures in the Presence of Heteroscedasticity," *Journal of Business and Economic Statistics*, 13, 105–111.

Caves, D. W., L. R. Christensen, and E. W. Diewert (1982a): "The Economic Theory of Index Numbers and Measurement of Input, Output, and Productivity," *Econometrica*, 50, 1393–1414.

——— (1982b): "Multilateral Comparisons of Output, Input, and Productivity using Superlatives Index Numbers," *Economic Journal*, 92, 73–86.

Caves, R. E., and D. Barton (1990): *Efficiency in U.S. Manufacturing Industries*. The MIT Press.

Ciccone, A. (2002): "Agglomeration effects in Europe," *European Economic Review*, 46(2), 213–227.

Cingano, F., and F. Schivardi (2004): "Identifying the Sources of Local Productivity Growth," *Journal of the European Economic Association*, 2(4), 720–742.

Clerides, K. S., S. Lach, and J. R. Tybout (1998): "Is Learning by Exporting Important? Micro-dynamic Evidence from Colombia, Mexico, and Morocco.," *Quarterly Journal of Economics*, 113, 903–947.

Coelli, T. (1995): "Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis," *Journal of Productivity Analysis*, 6, 247–268.

Coelli, T. J., D. S. P. Rao, C. J. O'Donnell, and G. E. Battese (2005): *An introduction to efficiency and productivity analysis*. Springer, New York, NY, 2. ed. edn.

Cornwell, C., P. Schmidt, and R. C. Sickels (1990): "Production Frontier with cross-sectional and time-series variation in efficiency levels," *Journal of Econometrics*, pp. 185–200.

Davies, S. W., and C. Morris (1995): "A new index of vertical integration," *International Journal of Industrial Organization*, 13, 151–177.

Debaere, P., H. Lee, and J. Lee (2006): "Does Where You Go Matter? The Impact of Outward Foreign Direct Investment on Multinationals' Employment at Home," CEPR Discussion Papers 5737, C.E.P.R. Discussion Papers.

Debreu, G. (1951): "The Coefficient of Resource Utilization," *Econometrica*, 19, 273–292.

Degasperi, M., and T. Fredholm (2010): "Productivity Accounting based on Production Prices," *Metroeconomica*, 61, 267–281.

Del Gatto, M., A. Di Liberto, and C. Petraglia (2010): "Measuring Productivity," *Journal of Economic Surveys.*

Denison, E. F. (1962): "The sources of economic growth in the U.S. and the alternatives before us," Supplementary Paper No 13, New York: Committee for Economic Development.

Deprins, D., L. Simar, and H. Tulkens (1984): "Measuring labor inefficiency in post offices," in *The Performance of Public Enterprises: Concept and measurements*, ed. by M. Marchand, P. Pestieau, and H. Tulknes, pp. 243–267. Amsterdam, North-Holland.

Dettori, B., E. Marrocu, and R. Paci (2008): "Total Factor Productivity, Intangible Assets and Spatial Dependence in The European Regions," Discussion Paper 23, CRENOS Working Paper Series.

Diaz, M. A., and R. Sanchez (2008): "Firm size and productivity in Spain: a stochastic frontier analysis," *Small Business Economics*, 30, 315–323.

Doms, M. E., T. Dunne, and K. Troske (1997): "Workers, Wages and Technology," *Quarterly Journal of Economics*, 112, 253–290.

Dosi, G., M. Grazzi, C. Tomasi, and A. Zeli (2010): "Turbolence underneath the big calm: Exploring the micro-evidence behind flat trend of manufacturing productivity in Italy," *LEM Working Paper Series*, 03, mimeo.

Driffield, N. (2004): "Regional policy and spillovers from FDI in the UK," *The Annals of Regional Science*, 38, 579–594.

Driffield, N., J. H. Love, and K. Taylor (2009): "Productivity and labor demand effects of inward and outward FDI on UK industry," *The Manchester School*, 77, 171–203.

Dunne, T. (1994): "Patterns of Technology Usage in U.S. Manufacturing Plants," *Rand Journal of Economics*, 25, 488–499.

Dunning, J. H. (1993): *Multinational Enterprises and the Global Economy.* Addison-Wesley.

Elhorst, J. P. (2010): "Spatial Panel Data Models," in *Handbook of Applied Spatial Analysis*, ed. by M. Fischer, and A. Getis. Springer, Berlin.

Esteban, J. (2000): "Regional convergence in Europe and the industry mix: a shift-share analysis," *Regional Science and Urban Economics*, 30, 353–364.

FABIANI, S., G. PELLEGRINI, G. ROMAGNANO, AND L. F. SIGNORINI (1998): "Efficiency and localisation: the case of industrial districts," in *The Competitive Advantage of Industrial Districts: Theoretical and Empirical Analysis*, ed. by M. Bagella, and L. Becchetti. Physica-Verlag.

FAGERBERG, J. (1991): "Innovation, Catching Up and Growth," in *Technology and Productivity: The Challenge for Economic Policy*, pp. 37–46. OECD.

——— (1995): "Convergence or Divergence - The Impact of Technology," Norwegian Institute of International Affairs (NUPI).

FARRELL, M. J. (1957): "The Measurement of Productive Efficiency," *Journal of the Royal Statistical Society*, Series A, 253–81.

FEDERICO, S. (2010): "Outsourcing versus integration at home and abroad and firm heterogeneity," *Empirica*, 37, 47–63.

FEENSTRA, R. C. (1998): "Integration of Trade and Disintegration of Production in the Global Economy," *Journal of Economic Perspective*, 12, 31–50.

FELIX, B. (2006): "Statistics in Focus - Science and Technology," Discussion paper, Eurostat.

FIASCHI, D., A. M. LAVEZZI, AND A. PARENTI (2009): "Productivity Dynamics across European Regions: the Impact of Structural and Cohesion Funds," Discussion Papers 2009/84, Dipartimento di Scienze Economiche (DSE), University of Pisa, Pisa, Italy.

FOSFURI, A., AND M. MOTTA (1999): "Multinationals without advantages," *Scandinavian Journal of Economics*, 101, 617–630.

FRIED, H. O., C. A. K. LOVELL, AND S. S. SCHMIDT (2008): "Productivity and Efficiency," in *The Measurement of Productive Efficiency and Productivity Change*, ed. by H. O. Fried, C. A. K. Lovell, and S. S. Schmidt, chap. 1, pp. 1–105. Oxford University Press.

FRISCH, R. (1965): *Theory of Production.* Dordrecht: D. Reidel publishing company.

GABRIELSEN, A. (1975): "On Estimating Efficient Production Functions," Working Paper No. A-85, Chr. Michelsen Institute, Department of Humanities and Social Sciences, Bergen, Norway.

GAMBARDELLA, A., M. MARIANI, AND S. TORRISI (2008): "How 'Provincial' is your Region? Openness and Regional Performance in Europe," *Regional Studies*, pp. 1–13.

Georgescu Roegen, N. (1966): *Analytical Economics: issues and problems.* Cambridge, Massachussets: Harvard University Press.

Gerschenkron, A. (1962): *Economic Backwardness in Historical Perspective.* Cambridge(USA): The Belknap Press.

Giannakas, K., K. C. Tran, and V. Tzouvelekas (2003): "Predicting technical efficiency in stochastic production frontier models in the presence of misspecification: a Monte-Carlo analysis," *Applied Economics*, 35, 153–161.

Gong, B.-H., and R. C. Sickles (1992): "Finite sample evidence on the performance of stochastic frontiers and data envelopment analysis using panel data," *Journal of Econometrics*, 51(1-2), 259–284.

Görg, H., and E. Strobl (2005): "Foreign direct investment and local economic development: Beyond productivity spillovers," in *Does Foreign Direct Investment Promote Development?*, ed. by T. Moran, E. Graham, and M. Blomström. Institute for International Economics, Washington DC.

Gould, W., J. Pitblado, and W. Sribney (2006): *Maximum likelihood estimation with stata.* Stata Press.

Greenaway, D., and R. Kneller (2007): "Firm heterogeneity, exporting and foreign direct investment," *Economic Journal*, 117(517), F134–F161.

Greene, W. H. (1990): "A Gamma-Distributed Stochastic Frontier Model," *Journal of Econometrics*, 46, 141–164.

——— (2005): "Reconsidering heterogeneity in panel data estimators of the stochastic frontier model," *Journal of Econometrics*, 126, 269–303.

——— (2008): "The Econometric Approach to Efficiency Analysis," in *The Measurement of Productive Efficiency and Productivity Change*, ed. by H. O. Fried, C. A. K. Lovell, and S. S. Schmidt, chap. 2, pp. 92–250. Oxford University Press.

Griffith, R. (1999): "Using the ARD establishment level data to look at foreign ownership and productivity in the United Kingdom," *Economic Journal*, 109, 416–442.

Griliches, Z. (1960): "Measuring Inputs in Agriculture; A Critical Survey," *Journal of Farm Economics*, 42, 1411–1433.

——— (1963a): "The Sources of Measured Productivity Growth; U.S. Agriculture 1940-1960," *Journal of Political Economy*, 81, 331–346.

——— (1979): "Issues in Assessing the Contribution of Research and Development to Productivity Growth," *The Bell Journal of Economics*, 10, 92–116.

——— (1994): "Productivity, R&D, and the Data Constraint," *American Economic Review*, 84, 1–22.

Griliches, Z., and J. Mairesse (1983): "Comparing Productivity Growth: An Exploration of French and U.S. Industrial and Firm Data," *European Economic Review*, 21, 89–119.

Grossman, G. M., and E. Helpman (1991): *Innovation and Growth in the Global Economy*. Cambridge(USA): The MIT Press.

Grossman, M. G., and E. Helpman (2005): "Outsourcing in a Global Economy," *Review of Economic Studies*, 72, 135 – 159.

Grossman, S. J., and O. D. Hart (1986): "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94, 691–714.

Guan, Z., S. C. Kumbhakar, R. J. Myers, and A. O. Lansink (2009): "Measuring Excess capital capacity in Agricultural Production," *American Journal of Agricultural Economics*, 91, 765–776.

Gumbau, M., and J. Maudos (2002): "The Determinants of Efficiency: The Case of the Spanish Industry," *Applied Economics*, 34, 1941–1948.

Hadri, K. (1999): "Estimation of a Doubly Heteroscedastic Stochastic Frontier Cost Function," *Journal of Business and Economic Statistics*, 17, 359–363.

Hadri, K., C. Guermat, and J. Whittaker (2003): "Estimation of technical inefficiency effects using panel data and doubly heteroscedastic stochastic production frontiers," *Empirical Economics*, 28, 203–222.

Hall, B. H., and J. Mairesse (1995): "Exploring the relationship between R&D and productivity in French manufacturing firms," *Journal of Econometrics*, 65, 263–293.

Hansson, P. (2005): "Skill Upgrading and Production Transfer Within Swedish Multinationals," *Scandinavian Journal of Economics*, 107 (4), 673–692.

Harris, R., and C. Robinson (2002): "The effect of foreign acquisitions on total factor productivity: Plant-level evidence from UK manufacturing, 1987-1992," *Review of Economics and Statistics*, 84, 562–568.

Head, K., and T. Mayer (2004): "Market Potential and the Location of Japanese Investment in the European Union," *The Review of Economics and Statistics*, 86(4), 959–972.

Head, K., and J. Ries (2002): "Offshore production and skill upgrading by Japanese manufacturing firms," *Journal of International Economics*, 58, 81–105.

Helpman, E., M. Melitz, and S. Yeaple (2004): "Export versus FDI with heterogeneous firms," *American Economic Review*, 94(1), 300–316.

Henderson, D. J., and R. R. Russell (2005): "Human Capital and Convergence: A Production-Frontier Approach," *International Economic Review*, 46, 1167–1205.

Herzer, D. (2010): "The Long-Run Relationship between Outward FDI and Total Factor Productivity: Evidence for Developing Countries," *Discussion Papers, Ibero-America Institute for Economic Research, University of Göttinen*, 199.

Heshmati, A. (2003): "Productivity Growth, Efficiency and Outsourcing in Manufacturing and Service Industries," *Journal of Economic Surveys*, 17, 79–112.

Hijzen, A., T. Inui, and Y. Todo (2007): "The Effects of Multinational Production on Domestic Performance: Evidence from Japanese Firms," Discussion papers 07006, Research Institute of Economy, Trade and Industry (RIETI).

Hitt, L. M. (1999): "Information Technology and Firm Boundaries: Evidence from Panel Data," *Information Systems Research*, 10, 134–149.

Hortaçsu, A., and C. Syverson (2009): "Why do firms own production chain?," *US Census Bureau Center for Economic Studies Paper*, 09-31.

Huang, C. J., and J. T. Liu (1994): "Estimation of a non-neutral stochastic frontier production function," *Journal of Productivity Analysis*, 5, 171–180.

Huang, R. (2004): "Estimation of Technical Inefficiencies with Heterogenous Technologies," *Journal of Productivity Analysis*, 21, 277–296.

IMF (1993): *Balance of Payments Manual*. Washington, D. C.

Islam, N. (1995): "Growth Empirics: a Panel data Approach," *Quarterly Journal of Economics*, 110, 1127–1170.

Jensen, U. (2005): "Misspecification Preferred: The Sensitivity of Inefficiency Rankings," *Journal of Productivity Analysis*, 23, 223–244.

Jondrow, J., C. A. K. Lovell, S. Materov, and P. Schmidt (1982): "On the Estimation of Technical Efficiency in the Stochastic Frontier Production Function Model," *Journal of Econometrics*, 19, 233–238.

Jorgenson, D. W., and Z. Griliches (1967): "The Explanation of Productivity Change," *Review of Economic Studies*, 34, 249–283.

KLETTE, T. (1999): "Market Power, Scale Economies and Productivity: Estimates from a Panel of Establishment Data," *Journal of Industrial Economics*, 47, 451–476.

KNIGHT, F. H. (1933): *The Economic Organization*. New York: Harper & Row, 2nd [1965] edn.

KODDE, D. A., AND F. C. PALM (1986): "Wald Criteria for Jointly Testing Equality and Inequality Restrictions," *Econometrica*, 54, 1243–1248.

KOOP, G. (2001): "Cross-Sectoral Patterns of Efficiency and Technical Change in Manufacturing," *International Economic Review*, 42, 73–103.

KOOPMANS, T. C. (1951): "An Analysis of Production as an Efficient Combination of Activities," in *Activity Analysis of Production and Allocation*, ed. by T. C. Koopmans, no. 13. Wiley, New York.

KRUGMAN, P. (1993): *Geography and Trade*. MIT Press.

KUMAR, S., AND R. R. RUSSELL (2002): "Technological Change, Technological Catch-Up, and Capital Deepening: Relative Contributions to Growth and Convergence," *American Economic Review*, 92, 527–548.

KUMBHAKAR, S., AND C. A. K. LOVELL (2000): *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.

KUMBHAKAR, S., AND H. J. WANG (2005): "Estimation of growth convergence using a stochastic production frontier approach," *Economics Letters*, 88, 300–305.

LAFONTAINE, F., AND M. SLADE (2007): "Vertical Integration and Firm Boundaries: The Evidence," *Journal of Economic Literature*, 45, 629–685.

LEIBENSTEIN, H. (1966): "Allocative Efficiency vs. X-Efficiency," *American Economic Review*, 56, 392–415.

——— (1976): *Beyond Economic Man*. Cambridge, MA: Harvard University Press.

LEVINSOHN, J., AND A. PETRIN (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70, 317–342.

LICHTENBERG, F. R. (1992): *Corporate Takeovers and Productivity*. Cambridge: MIT Press.

LICHTENBERG, F. R. (1996): "The Output Contributions of Computer Equipment and Personnel: A Firm- Level Analysis," *NBER working papers*, 4540.

LIPSEY, R. E. (2002): "Home and Host Country Effects of FDI," *NBER working papers*, 4927.

Lovell, C. A. K. (1993): "Production Frontiers and Productive Efficiency," in *The Measurement of Productive Efficiency: Tecniques and Applications*, ed. by F. H. O., C. A. K. Lovell, and S. S. Schmidt. Oxford University Press.

Lucas, R. E. (1988): "On the Mechanisms of Economic Development," *Journal of Monetary Economics*, 22, 3–42.

——— (1990): "Why doesn't capital flow from rich to poor countries," *American Economic Review*, 80, 92–96.

Maddigan, R. J. (1981): "The Measurement of Vertical Integration," *The Review of Economics and Statistics*, 63, 328–335.

Mairesse, J., and Z. Griliches (1990): "Heterogeneity in Panel Data: Are There Stable Production Functions?," in *Essays in Honor of Edmond Malinvaud*, ed. by P. Champarus, pp. 125–147. MIT Press, Cambridge, Massachussets.

Mankiw, N. G., D. Romer, and D. N. Weil (1992): "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107, 407–437.

Mariotti, S., M. Mutinelli, and L. Piscitello (2003): "Home country employment and foreign direct investment: evidence from the Italian case," *Cambridge Journal of Economics*, 27(3), 419–431.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995): *Microeconomic Theory*. Oxford University Press.

Mastromarco, C., and U. Woitek (2006): "Public infrastructure investment and efficiency in Italian regions," *Journal of Productivity Analysis*, 25, 57–65.

McFadden, D. (1966): "Cost, Revenue and Profit Functions: A cursory review," Working paper n.86, University of California Berkeley.

McGuckin, R. H., and S. V. Nguyen (1995): "On Productivity and Plant Ownership Change: New Evidence form the LRD," *Rand Journal of Economics*, 26, 257–276.

Meeusen, W., and J. van den Broeck (1977): "Efficiency Estimation for Cobb-Douglas Production Functions with Composed Error," *International Economic Review*, 18, 435–444.

Melitz, M., and G. Ottaviano (2008): "Market Size, Trade and Productivity," *Review of Economic Studies*, 75, 295–316.

Millo, G., and G. Piras (2009): "`splm`: spatial panel data models in `R`," .

NELSON, R. R. (1981): "Research on Productivity Growth and Productivity Differences: Dead Ends and New Departures," *Journal of Economic Literature*, 19, 1029–1064.

NICOLETTI, G., AND S. SCARPETTA (2003): "Regulation, productivity and growth: OECD evidence," *Economic Policy*, 18(36), 9–72.

NOORBAKHSH, F., A. PALONI, AND A. YOUSSEF (2001): "Human Capital and FDI inflows to developing countries: new empirical evidence," *World Development*, 29, 1593–1610.

O'DONNELL, C., D. RAO, AND G. BATTESE (2008): "Metafrontier frameworks for the study of firm-level efficiencies and technology ratios," *Empirical Economics*, 34(2), 231–255.

OECD (1996): *Benchmark definition of Foreign Direct Investments*. Paris.

OLLEY, S., AND A. PAKES (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263–1295.

OLSEN, K. B. (2006): "Productivity Impact of Offshoring and Outsourcing: a Review," Discussion Paper 2006/1, OECD Directorate for Science, Technology and Industry.

OLSON, J. A., P. SCHMIDT, AND D. M. WALDMAN (1980): "A Monte Carlo study of Estimators of Stochastic Frontier Production Functions," *Journal of Econometrics*, 13, 67–82.

O'MAHONY, M., A. RINCON-AZNAR, AND C. ROBINSON (2010): "Productivity Growth in Europe and the US: a Sectoral Study," *Review of Economics and Institutions*, 1(1).

ONDRICH, J., AND J. RUGGIERO (2001): "Efficiency measurement in the stochastic frontier model," *European Journal of Operational Research*, 129, 434–442.

PACI, R., AND S. USAI (2000): "Technological Enclaves and Industrial Districts: An Analysis of the Regional Distribution of Innovative Activity in Europe," *Regional Studies*, 34(2), 97–114.

PARETO, V. (1927): *Manual of Political Economy*. August M. Kelley (1971 translation).

PAVITT, K., AND L. G. SOETE (1982): "International Differences in Economic Growth and the International Location of Innovation," in *Emerging Technologies: Consequences for Economic Growth, Structural Change, and Employment*, ed. by H. Giersch. Tubingen: J.C.B.Mohr (Paul Siebeck).

PERRY, M. K. (1989): "Vertical integration: Determinants and effects," in *Handbook of Industrial Organization*, ed. by R. Schmalensee, and R. Willig, vol. 1, chap. 4, pp. 183–255. Elsevier.

PESTIEAU, P., AND H. TULKENS (1993): "Assessing and Explaining the Performance of Public Enterprises," *Finanz Archiv*, 50, 293–323.

PETIT, M.-L., AND F. SANNA-RANDACCIO (2000): "Endogenous R&D and foreign direct investment in international oligopolies," *International Journal of Industrial Organization*, 18(2), 339 – 367.

POLEDRINI, S. (2008): "From modular innovation to Systems Integration: the main experiences from some Italian machine tool builders," in *8th global Conference on Business & Economics Proceedings*.

POZZANA, R., AND E. ZANINOTTO (1989): *Efficienza produttiva e concorrenza : il caso della distribuzione commerciale*. F. Angeli.

REIFSCHNEIDER, D., AND R. STEVENSON (1991): "Systematic departures from the frontier: A framework for the analysis of firm inefficiency," *International Economic Review*, 32, 715–723.

RICHMOND, J. (1974): "Estimating the Efficiency of Production," *International Economic Review*, 15, 515–521.

RITTER, C., AND L. SIMAR (1997): "Pitfalls of Normal-Gamma Stochastic Frontier Models," *Journal of Productivity Analysis*, 8, 167–182.

ROBERTS, M. J., AND J. R. TYBOUT (1997): "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs," *American Economic Review*, 87(4), 545–64.

ROLFO, S. (1993): "The Italian Machine Tool Industry," *Ceris-CNR working paper*, 4.

——— (1998): "L'Industria italiana della meccanica strumentale di fronte alla globalizzazione: opportunità e limiti," *L'Industria*, 4, 881–894.

——— (2000): "The Italian Machine Tool Industry Towards Production Development Networks," in *New product development and production networks*, ed. by H. Juergens, chap. 8. Springer.

ROLFO, S., AND G. CALABRESE (2006): "Struttura industriale e profili di competitività nella meccanica strumentale," *L'Industria*, 4, 603–622.

ROMER, P. M. (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98, 1002–1037.

RUGGIERO, J. (1999): "Efficiency estimation and error decomposition in the stochastic frontier model: A Monte Carlo analysis," *European Journal of Operational Research*, 115, 555–563.

SALTER, W. E. G. (1960): *Productivity and Technical Change*. Cambridge: Cambridge University Press, second edition: 1966.

SALZ, P., E. BUISMAN, J. SMIT, AND B. DE VOS (2006): "Employment in the fisheries sector: current situation," Final report, Eurostat.

SAMUELSON, P. (1956): *Foundations of Economic Analysis*. Harvard University Press.

SCITOVSKY, T. (1954): "Two Concepts of External Economies," *Journal of Political Economy*, 62, 143.

SERTI, F., AND C. TOMASI (2008): "Self-Selection and Post-Entry Effects of Exports: Evidence from Italian Manufacturing Firms," *Review of World Economics*, 144(4), 660–694.

SHEPARD, R. W. (1953): *Cost and Production Functions*. Princeton University Press, Princeton.

SHLEIFER, A., AND R. VISHNY (1986): "Large Shareholders and Corporate Control," *Journal of Political Economy*, 94, 461–488.

SIMAR, L., AND P. WILSON (1999): "Estimating and Bootstrapping Malmquist Indices," *European Journal of Operational Research*, 115, 459–471.

SIMAR, L., AND P. W. WILSON (1998): "Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models," *Management Science*, 44, 49–61.

SIMON, H. (1955): "A Behavioural Model of Rational Choice," *Quarterly Journal of Economics*, 69, 99–118.

SOLOW, R. (1957): "Technical Change and the Aggregate Production Function," *The Review of Economics and Statistics*, 39, 312–320.

——— (1960): "Investment and Technical Progress," in *Mathematical Models in Social Sciences*, ed. by K. Arrow, S. Karlin, and P. Suppes. Stanford University Press.

SRAFFA, P. (1960): *Production of Commodities by Means of Commodities*. Cambridge: Cambridge University Press.

STERLACCHINI, A., AND F. VENTURINI (2009): "Knowledge capabilities and regional growth: an econometric analysis for European developed regions," *Scienze Regionali/Italian Journal of Regional Science*, 8(2), 45–70.

STEVENSON, R. (1980): "Likelihood Functions for Generalized Stochastic Frontier Estimation," *Journal of Econometrics*, 13, 58–66.

STIGLER, G. J. (1976): "The Xistence of X-Efficiency," *American Economic Review*, 66, 213–216.

STONE, R. (1980): "Whittling Away the Residual: Some Thoughts on Denison's Growth Accounting," *Journal of Economic Literature*, 18, 1539–1543.

SYVERSON, C. (2004): "Product Substitutability and Productivity Dispersion," *The Review of Economics and Statistics*, 86, 534–550.

——— (2010): "What determines productivity?," *Journal of Economic Literature*, forthcoming.

TAYMAZ, E. (2005): "Are Small Firms Really Less Productive?," *Small Business Economics*, 25, 429–445.

UCIMU (2007a): "Industry Report," Ucimu-Centro Studi.

——— (2007b): "L'industria italiana di settore: evoluzione dal dopoguerra a oggi," Ucimu.

VAHTER, P., AND J. MASSO (2007): "Home versus Host Country Effects of FDI: Searching for New Evidence of Productivity Spillovers," *Applied Economics Quarterly (formerly: Konjunkturpolitik)*, 53(2), 165–196.

VAN BIESEBROECK, J. (2003): "Exporting Raises Productivity in Sub-Saharan African Manufacturing Plants," NBER Working Paper 10020.

——— (2007): "Robustness of Productivity Estimates," *Journal of Industrial Economics*, 55(3), 529–569.

——— (2008): "The Sensitivity of Productivity Estimates," *Journal of Business & Economic Statistics*, 26, 311–328.

VAN POTTELSBERGHE DE LA POTTERIE, B., AND F. LICHTENBERG (2001): "Does Foreign Direct Investment Transfer Technology across Borders?," *Review of Economics and Statistics*, 82, 490–97.

WALDMAN, D. M. (1984): "Properties of Technical Efficiency Estimators in the Stochastic Frontier Model," *Journal of Econometrics*, 25, 353–364.

WALRAS, L. (1874): *Element of Pure Economics: Or the theory of social wealth.*

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.

WANG, H.-J., AND P. SCHMIDT (2002): "One-Step and Two-Step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels," *Journal of Productivity Analysis*, 18, 129–144.

WANG, W. S., AND P. SCHMIDT (2009): "On the distribution of estimated technical efficiency in stochastic frontier models," *Journal of Econometrics*, 148, 36–45.

WENGEL, J., AND P. SHAPIRA (2004): "Machine tools: the remaking of a traditional sectoral innovation system," in *Sectoral systems of Innovation*, ed. by F. Malerba. Cambridge: Cambridge University Press.

WILLIAMSON, O. E. (1971): "The Vertical Integration of Production: Market Failure Considerations," *American Economic Review*, 61(2), 112–23.

WILLIAMSON, O. E. (1975): *Market and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.

WINSTEN, C. (1957): "Discussion on Mr. Farrell's Paper," *Journal of the Royal Statistical Society*, 120, 282–284.

ZANFEI, A., AND A. GAMBARDELLA (1994): "The Italian Machine-Tool Industry," in *The Decline of the U.S. Machine-Tool Industry and Prospects for Its Sustainable Recovery, Vol. 2 - Appendices*, ed. by D. M. Adamson. RAND, Critical Technologie Institute.