Title

A comprehensive approach on the application of cost efficiency methods to network industries: special focus on the postal sector

Author

Gori, Stefano

Publication Date

2013-07-28

A comprehensive approach on the application of cost efficiency methods to network industries. Special focus on the postal sector [i]

Stefano Gori[‡]

First Draft 28th July, 2013

1. Introduction

The 2009 Working Paper of the Institute of Governmental Studies by Gori and Visco "Antitrust law and public services performances with reference to the postal industry" *(Gori & Visco, 2009)* addresses the issue of State Aid legislation infringement arising from the compensation of universal service cost burdens in network industries through public subsidies. More specifically, it analyzes the impact on regulation of network industries played by the European Court of Justice (ECJ) with its *Altmark* decision, which defined the conditions so that a compensation for public services is not considered state aid. The analysis address specifically on the fourth condition which applies whenever the undertaking is not chosen in a public procurement: compensation needs to be determined by benchmarking the operations of the public service provider against market determined standards *(Gori & Visco, 2009:1)*.

The principle set by the ECJ is interesting also for other industrialized countries outside Europe because it sets a standard. It is based on the idea that when providing services at the lowest cost to the community, the level of compensation must be determined on the basis of an analysis of the costs which a typical well run undertaking would have incurred in discharging those same obligations. This approach would have to take into account the revenues and a reasonable profit for discharging these obligations. Thus, compensation needs to be determined by benchmarking the operations of the public service provider against specific market standards. A number of questions arise: What is an average, well-managed company providing public services? What if in a market or in a country there is no company operating in a comparable market, would it be necessary to carry out a cross-country benchmark? *(Hansen et. al, 2003)*.

The working paper from Gori and Visco concluded that at first sight a cross-country efficiency benchmarking exercise between national companies involved in network industries seems to be inappropriate or at best misleading given the peculiarities of certain network industries, for example

[‡] Bristol Business School

highly labor intensive with natural monopoly characteristics and providing merit goods such as the postal industry *(Gori et al., 2002, Gori & Marcone, 2007).* However, they believe that this exercise not only is feasible but it is also useful. They believe that comparisons based on reasonable causal chains with exogenous variables (e.g. factors independent from provider's responsibility) and modeled by specific econometric techniques allow building a benchmarking framework able to account for efficiency differences *(Gori & Visco, 2009:10).*

The objective of this paper is in fact to identify innovative ways of addressing the issue raised by the 2009 paper on efficiency in network industries and more specifically in the postal sector and present original approaches which have not yet been utilized in this industry. Furthermore, the aim is the development of a comprehensive method to be used not only in the postal sector but in other network industries with similar characteristics, able to be empirically applied not only by economists, but also by law scholars such as national regulatory authorities and national courts not only in Europe.

This comprehensive framework is the result of the research carried out at Bristol Business School with Professor Don Webber and it is based on a three step approach. In the first step the network industry needs to be subdivided in sub-phases and after that it is important to identify which phase or sub-phase needs to be scrutinized from an efficiency point of view. If the one or more phase or sub-phase involved concern a highly competitive sector with feasible entrance and exit from the sector then the analysis goes to the second step, where a survivor technique is used to analyze the market outcome and identify the long and short term cost curves. While if the segment of the sector under scrutiny is less competitive and there is a monopoly or few market players and exit and entrance is not easy then the second step is skipped and the analysis goes to the third step. In this step a mix of econometric methods are utilized to carry out a benchmarking exercise.

The methods used are both parametric and nonparametric and the comparison of the results obtained and the ranking between operators benefits from the technique Spearman correlation coefficient.

2. Step 1- Identify the segment of the market under scrutiny

Services of General Economic Interest (SGEI) play a fundamental role in the shared values of a nation. These services are essential for the daily life of citizens and enterprises, and play a major role in ensuring social and economic cohesion and reflect a model of society. They tend to be offered by industries of national interest (such as public transportation, telecommunications, postal services, local government services, water supply, waste management, health and social services) often holding special obligations in fulfilling public tasks not performed by usual market mechanisms. If a SGEI has a common network structure in economic literature it is called a network industry, and generally the sectors that fit into this category tend to be regulated. These industries share an extensive distribution system of lines, pipes, or routes requiring the use of public rights of way, often with strong operational links between component parts leading to intertwined structures *(Gori & Visco, 2009:4)*. These networks often exhibit economies of scale and involve substantial sunk costs, hence raising the issue of natural monopoly and for this reason in some cases the state at both local and national level directly owns a part or the whole infrastructure *(Geddes, 1999, European Commission, 2000).*

The industries providing these services have been at the center of political debate in the past few decades in industrialized countries. The gradual opening up of these sectors to competition has evolved with the definition of a number of public service obligations for each sector, covering aspects such as universal service, quality of service, consumer rights, special conditions for access to the network and health and safety concerns.

An increasing number of sectorial regulatory frameworks have been put in place to better specify the scope of public policy intervention in regulating these networks, with particular reference to the role of national sectorial regulatory authorities. The task of reconciling the social and economic cohesiveness of such industries and the workings of the single market to augment competition (generating economic growth by increasing efficiency, reducing prices and increasing quality of service) has enhanced the use of economic and econometric tools for the analysis of the performance of these sectors in light of the liberalization processes.

These regulated network industries tend to create a good or service at one location, and then distribute it over a network where it is delivered to numerous customers for end use. For the purpose of simplification, the activities of utilities can be broken down into three components: production, transmission and distribution. In some industries the firms are fully vertically integrated into all three activities, while in others different firms may perform the production and transmission/distribution functions, hence the critical issue is how to reconcile the issue of the unbundling of these functions, the degree of competition in the different phases and the operational interconnection between the different phases table 1 adapted from Crandall and Ellig*, (1997:70),* and Geddes *(1999:1064).*

Table 1 Different phases of network industries

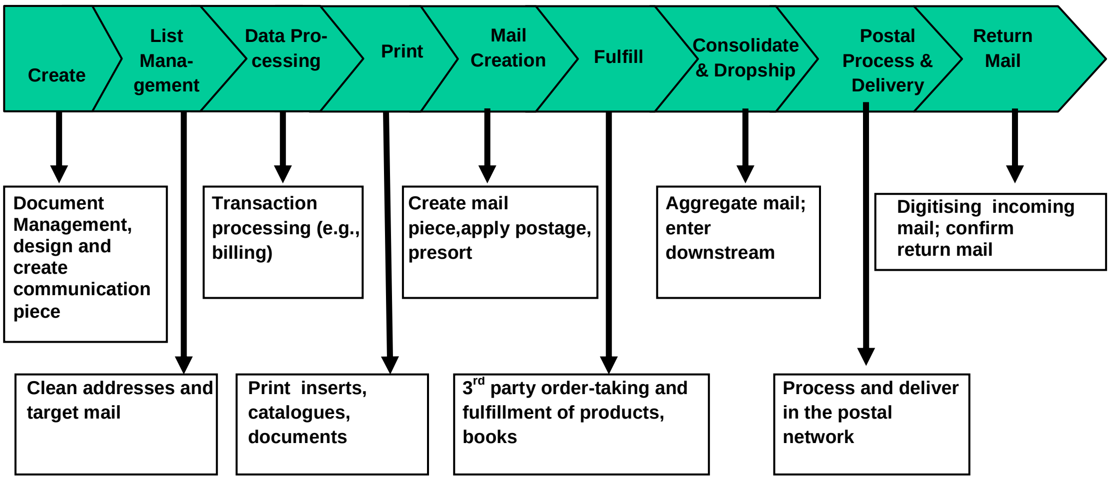
Industry	Production	Transmission	Distribution
Airlines	Airplanes	Air Traffic control	Airports
Trucking	Trucks	Highways	Local streets, distribution centers
Telecoms	Telecom terminal equipment	Long-distance cos. And local telecoms	Local telecoms
Electricity	Generating plants	High voltage lines	Local power lines
Natural gas	Gas wells	Interstate pipelines	Local distribution companies
Postal	Printing/writing of letters	Transportation (trucks, planes)	Postmen (last mile)
Railroads	Trains	Trunk lines	Local sidings

Source: based on Crandall & Ellig *(1997),* Geddes *(1999)* with addition of Postal services

The coexistence of multiple constraints raises the issue of efficiency optimization of SGEI taking into account the environmental factors impacting the design of the network. For example, there is a great difference in measuring the efficiency of the local railway system in the mountainous area of Switzerland compared with national railway grid of the Dutch Railway company which operates in practically flat terrain. The importance of identifying the postal phase or sub-phase being put into scrutiny is very important because it helps defining the scope of the inquiry also defines the market of reference for a multiproduct and multiphase business.

More specifically, the three phases described in Table 1 (adapted from *Crews, Gori et al., 2006*) can be expanded for the postal sector in 9 sub-phases in Figure 1 where production includes all phases from creation to fulfillment, where transmission incorporates Consolidation, Dropship and the first part of the postal process while distribution is Delivery and managing the return mail.

Figure 1 - Mailstream

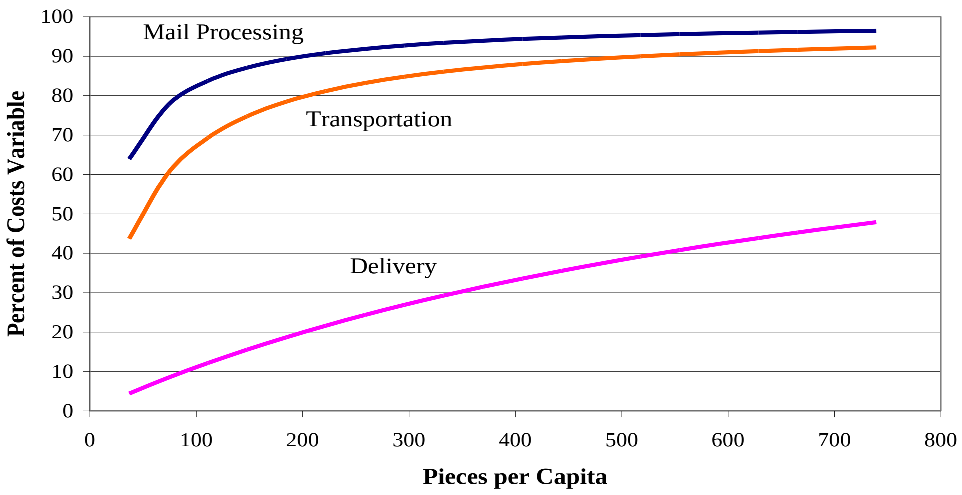


Source: Crews, Gori et al. (2006)

The postal sector presents some specific characteristics that it makes it even harder to assess efficiency and effectiveness also because in the same platform several products, with very distinctive characteristics and revenue and cost models, are distributed simultaneously.

For this sector the key phase is the distribution (in postal terms the delivery phase) which is a labor intensive phase and it is a phase in which the postal operator faces less variable costs than other phases (Figure 2: Cost elasticity or Variability Curves by Activity from *Cohen et al., 2004*). All phases upstream from this phase tend to be more open to competition while the delivery phase and the other phase more downstream tend to face stiffer regulatory and economic conditions which lead to de facto monopolies. This step is a very important one because it helps to identify the scope of the inquiry and to define the market of reference. The postal service is a multiproduct and multiphase business and some products and phases are under regulatory scrutiny in various countries, for example quality of service standards, price caps for retail mail and monopoly of delivery of certain products. Thus it is crucial to address the right product or set of products, phase or subset of phases to properly address the issue of efficiency.

Figure 2 Cost Elasticity or Variability Curves by activity



Source: Cohen et al. 2004

3. Step 2- Using the Survivor technique for competitive market segments

If the efficiency analysis concerns more upstream market phases (production and transmission in regulated industries terminology, from creation to consolidation in the mailstream) where there is a more competitive environment then it is more appropriate to use the survivor technique as a robust method to assess the efficiency of an operator. The assumption of the survivor technique is that firms compete vigorously in their markets and they adopt alternative scales of production. The greater the competition, the more likely it would be for efficient scales to be chosen. ‘An efficient size of firm.....is one that meets all and any problems the entrepreneur actually faces: strained labor relations, rapid innovation, government regulation...and what not. This is, of course, is the decisive meaning of efficiency from the viewpoint of the enterprise’ (*Stigler, 1958:56*, discussed in *Giordano 2008:361*). There are multiple objectives of this method: a) identify the long run average cost curve (LRAC) and the short run average cost (SRAC), b) attempt to identify the optimum size of the firm and c) identify the relative positions of firms and their likely evolution in the future and how they got to this position in the past. The LRAC curve depicts the cost per unit of output in the long run and is created as an envelope of an infinite number of short-run average total cost curves, each based on a particular mix of inputs. All points on the line represent an optimal combination and is typically U-shaped. In the negatively sloped section of curve it signals increasing returns of scale, constant returns in the horizontal part and decreasing returns in the positively sloped segment. Even if all points on the LRAC are optimal combinations, in the long-run and in a perfectly competitive environment, the equilibrium level of output corresponds to the minimum efficient scale *(Himmelweit, Simmonetti & Trigg, 2001)*.

3.1 Optimum size

The other objective is to identify the optimum size. It can not only be an exercise to identify the efficient transformation of inputs into outputs but it also needs to take into account the introduction of new products and future demand. Thus it should ‘include the demand conditions facing the firm, the supply conditions of the factors of production facing the firm, any taxes, subsidies or other form of government interference (both real or potential), and any other factor which may affect the economic operation of the firm. These factors must be included in a definition of optimum size since they influence the average costs of production of both the plant and the firm. The optimum size…..may not be the social optimum since the consideration of everything at market price may either under or overestimate the social costs of production’ *(Saving, 1961:569-570).* Most of the empirical literature on optimum size can be subdivided in two branches: the one focused on identifying the optimum size of a plant and the one analyzing the optimal size of a firm.

3.1.1 Optimum size of plant and of the firm

In static theory, size of plant is synonymous with the output of the plant. The size of plant becomes a multidimensional concept if we drop the assumption of product homogeneity and introduce varying amounts of vertical integration. This new concept of plant size encompasses: 1) the variety of products which are produced; 2) the rate of production of each final product; 3) the degree of vertical integration *(Saving, 1961:570).*

Conventionally it is believed that the optimum size of plant will be larger, the greater: 1) the size of the market; 2) the complexity of the production process; 3) the capital intensiveness of production; 4) the stability of demand. For Saving, a fifth determinant of optimum plant size is the rate of growth of the industry *(Saving, 1961:587)*. Thus firms have an incentive to concentrate production in plants of optimal size and to build new plants of the emerging optimal size taking into account the dynamism of the sector and of competition *(Rees, 1973:394).*

The study of Optimum size of plant is purported to be a causal factor in non-competitive industry behavior for the following reasons: 1) the optimum size of plant may be so large as to necessitate high levels of firm concentration in an industry *(Bain, 1959:148)*; 2) the minimum size may be large from a percentage of the industry standpoint and, hence, may result in a substantial barrier to the entry of new plants into the industry *(Bain, 1956)*. Both depend on large optimum size of plant and both result in oligopolistic industry behavior *(Saving, 1961:596-597).*

The empirical works identifying the optimal size survivor technique involve the comparison of the distribution of plants in a specific industry at two or more points in time to test the conventional wisdom on optimal size. The most important empirical research on optimal plant was carried out by Saving *(1961)* where the data used was the four-digit manufacturing industry data classified by size of plant from the U.S. Bureau of the Census for the years 1947 and 1954. From these data a random sample of 200 industries was chosen *(Saving, 1961:575)*. The strong assumption of Saving was that changes in the distribution would indicate the optimum size based on the assumption that existing plants and new plants tend toward that size which has minimum average cost *(Saving, 1961: 573).*

The analysis of optimum plant size and optimum firm size are connected because firm size is determined by two factors: 1) the optimum size of plant; 2) the extent of economies of multi-plant operations *(Saving, 1961:587).*

3.2 Literature on the use of the Survivor Technique

There is an extensive literature on how the survivor technique has been applied to various industries except the postal sector which has only recently, and in selected regions and countries, been opened to competition.

Willard L. Thorp used the survivor technique in 1924 in a census monograph *(Thorp, 1924)* to point to the trend in the size of plant under different patterns of industry growth but he did not use it in the estimation of optimum size. John Stuart Mill *(1929:134)* was the first to introduce the concept that the survival of plants and firms is the ultimate test of their efficiency, while Stigler was the first to make actual use of this technique to estimate optimum size of enterprise and applied it to the steel, auto, and oil refining industries (*Stigler, 1958* discussed in *Saving, 1961: 573*). Since then the technique has been put to numerous applications the main ones being Saving *(1961)*, Weiss *(1964)*, Rees *(1973)*, Frech and Ginsburg *(1974)*, Blair and Vogel *(1978)*, Bays *(1986)*, Norton and Norton *(1986)*, Elzinga *(1990)* and Giordano *(1995, 1997, 2003, 2008)*.

Nonetheless, the interpretation of survivor estimates has faced various criticisms especially by Shepherd *(1967)* and Bain *(1969)*. They criticize the fundamental logic and express doubts about its empirical reliability *(Giordano, 2008: 361)*. Moreover, the competitive outcome often depends on factors other than those which would normally represent genuine welfare gains. These factors include, for example, restrictive agreements between players in the market and the control of scarce resources *(Rees, 1973: 394)*. Giordano *(2003)* evaluates most of these criticisms and he reaches the conclusion that none of which seriously disqualifies the technique.

The empirical works identifying the optimal size survivor technique involve the comparison of the distribution of plants in a specific industry at two or more points in time to test the conventional wisdom on optimal size. The most important empirical research on optimal plant was carried out by Saving on 1961 where the data used was the four-digit manufacturing industry data classified by size of plant from U.S. Bureau of the Census for the years 1947 and 1954. From these data a random sample of 200 industries was chosen *(Saving, 1961:575)*. The strong assumption of Saving was that changes in the distribution would indicate the optimum size based on the assumption that existing plants and new plants tend toward that size which has minimum average cost. (Saving, 1961:573). From the analysis it emerged that: 1) both the mean optimum size and the minimum optimum size are usually small when compared with their respective industry sizes; 2) the range of optimum size is usually large relative to its respective mean optimum size; 3) in those industries in which the plants compete in national markets, optimum size is rarely so large as to necessitate non-competitive industry behavior; 4) the primary determinants of optimum size are the industry size and capital intensiveness; 5) the elasticity of optimum size (both mean and minimum, with respect to industry size) is approximately 0.5, hence large industries are likely to have relatively small optimum sizes (where optimum size is measured as a percentage of the industry total value added); 6) the range of optimum size is primarily statistically determined by the mean optimum size *(Saving, 1961: 596-597)*.

Other important study in this domain was the one carried out by Bays in 1986 where he applied the optimal plan techniques to hospitals *(Bays, 1986:359)*. The paper develops estimates of hospital size by studying changes in the size distribution of US short term general hospitals over the period 1971-77 *(Bays, 1986: 359)*. Bays believes that survivor technique is a useful technique for the study of hospitals in that it does not require data on costs *per se* and it thus ameliorates the problem of accounting for the input of admitting physicians in the production of hospital care *(Bays, 1986: 362)*. However, there are other factors that influence the optimal size range for hospitals such as administrative factors that impact entry, exit prices, and products in the industry and widespread hospitalization insurance further inhibits the market survival mechanism *(Bays, 1986: 362)*. From the analysis it emerges that there are substantial differences in survival sizes among geographic regions *(Bays, 1986: 359)*.

An important empirical paper which uses the survivor technique to identify the optimum firm size is the 2008 paper from Giordano of the Less than truckload industry (here after LTL). It analyses the post-deregulation market structure question with evidence on the extent of economies of scale and market concentration for the industry's LTL segment from 1981 to 2001 using along the survivor technique a trans-log cost function *(Giordano, 2008:358)*. It finds that economies of scale do extend across the entire spectrum of firm sizes in LTL. From the analysis using the Long-run average cost it appears to decline mildly and at a diminishing rate with increases in firm size, however, such that any cost advantage for larger firms has been insufficient to eliminate new entry and competition from smaller rivals *(Giordano, 2008:357)*. This paper also provides a detailed description of how to determine whether or not the observed shifts from one year to the other are merely the result of random fluctuation, χ^2 test of statistical significance is performed *(Giordano, 2008:363-365).*

This method has been applied for the first time in the postal sector only recently *(Gori, 2013b)* where both an analysis was carried out on the UK postal market both on the End to End (E2E) mailstream sector and on the upstream phases. While the results from the E2E market were not satisfactory due to very static market dynamic, the ones from the upstream market with access on the downstream phase market has showed a more interesting evolution in the same period. The survivor technique type diagrams *(Elzinga and Page, 2009)* have emerged as a valid instrument to capture the trends, where the negative skewed curve signals the shift toward higher volume classes thus leading to larger optimal size of firms *(Gori, 2013b:13)*.

Because the results from the survivor technique are more qualitative than quantitative and it is not likely that the survivor technique will fully substitute the cost function methodologies, however this technique offers a credibility check on evidence produced by the more traditional econometric procedures. Furthermore, it does identify the scales of market survivors and the margins by which they prevail *(Giordano, 2008:368-369)*.

4. Step 3- Econometric approach

The three main "families" of methods to carry out comparative efficiency using econometric tools analysis are parametric, non-parametric and semi-parametric (e.g. mixed two stage, one non parametric and one parametric) methods. Before applying the different methods it is important to identify the main determinants of the demand function.

4.1. Demand side of the postal sector

It is important to identify the main determinants of the demand function faced by a network industry prior to carrying out efficiency analysis. These determinants are closely linked to the output generated by a firm or a sector which is itself an important variable of the cost efficiency analysis. The identification of the different determinants and drivers of demand is often performed using panel data econometric estimates. Panel data estimation is often considered to be an efficient analytical method in handling econometric data.

In this section, a brief description of the panel data technique to analyze the demand function of a sector will be presented and subsequently one of the techniques will be scrutinized, the Principal Component Analysis (PCA).

The PCA technique has been applied in many area of research not only in social sciences but as well as biology, medicine, chemistry, meteorology and geology (for a more detailed analysis of the empirical analysis, *Dunteman, 1989:8*).

Concerning the use of this technique in regulated industries in the postal sector only marginally used *(Harding, 2006)* while there is an extensive literature in the water sector and in the electricity sector. In the water sector, Principal component has been used to estimate regional water demand *(Kim et al., 2005)* and to develop indexes to assess the availability of water *(Ali, 2008)*. The later paper uses a multivariate model, based on the principal component analysis, to develop the Arab Water Sustainability Index (AWSI) with the objective to incorporate a variety of physical, socio-economic and environmental factors determining the availability of water in the Arab region.

Electricity-supply planning utilizes this technique mainly for two purposes efficient management of existing power systems and optimization of the decisions concerning additional capacity and to put in place early warning systems to address misalignments between supply and demand *(Li Jinchao, 2011)*. Demand prediction is an important aspect in the development of any model for electricity planning more specifically to capture the intra-day variation in electricity demand *(Taylor et al., 2006: 1 and 11, Ramanathan et al., 1997 and Carnero et al., 2003).*

Hence, there is enough theoretical and literature supporting material to extend this technique to network industries and more specifically to the postal sector because the two utilities sectors (water and electricity) mentioned previously share with the postal sector several characteristics. Mainly the nature of the service which is home delivery and furthermore all three sectors face the same issue of retail consumption by households heavily influenced by cultural and environmental factors.

4.1.1 Theoretical framework on the use of panel data

The panel data matrix set consists of a time series for each cross sectional member in the data set, and offers the opportunity to use several estimation methods. The basic idea behind panel data analysis is known as the pooling assumption and is based on the notion that the individual relationships will all have the same parameters. It groups all the individuals into one dataset and it imposes a common set of parameters across them. If the pooling assumption is appropriate then it offers several advantages versus a single individual regression, such as the possible problem of omitted variables (causing biased estimates) in individual regressions are easier to control in panel data. Even in case of a heterogeneous panel (where parameters are different across individuals) it is expected that the panel data estimator gives some representative average estimate of the individual parameters *(Asteriou & Hall, 2007:344).*

Starting point from a linear panel with a sample containing N cross sectional units (i= 1,2,…, N sections) that are observed at T time periods (t=1,2,…, T) with one explanatory variable given by:

$$Y_{it} = a + \beta X_{it} + u_{it} \qquad (1)$$

Simple linear panel data models can be generally estimated using three different methods: a) incorporating a common constant in the equation, b) taking into account for fixed effects and c) allowing for random effects. The common constant method of estimation presents results under the restrictive assumption that there are no differences between the estimated cross sections. In the fixed effects method the constant is treated as group specific allowing for different constants for each group. The standard F-test can be used to check whether fixed effects should be included in a model in place of a simple constant OLS method. The null hypothesis of this test is the homogeneity condition which entails that all constants are the same. To capture any effects which vary over time but are common across the whole panel it is possible to include a set of time dummies this method is known as the ‘Two way Fixed effect Model’. The third method is the random effects model which handles the constants for each section not as fixed, but as random parameters. The random effects model has several advantages such as that it has fewer parameters to estimate compared to the fixed effects method and it allows for additional explanatory variables that have equal value for all observations within a group. However, it has also several disadvantages. There is a need to make specific assumptions about the distribution of the random component. Furthermore, if the unobserved group specific effects are correlated with the explanatory variables, then the estimates will be biased and inconsistent.

The main difference between the fixed effects versus the random effect models is that fixed effects model assumes that each observed institution differs in its intercept term, whereas the random effects model assumes that the difference is in the error term. The Hausman test *(Hausman, 1978)* is used in choosing between the two approaches. It investigates whether random effects estimation of a panel data model is as good as a model where fixed effects are appropriate *(Asteriou & Hall, 2007:345-49).*

Estimation of the determinants of demand needs to take into account both variation over time and across countries, modeling unobserved heterogeneity as country specific fixed effects, and allow estimating within country determinants. It is important to isolate those country specific fixed effects which are constant or evolve very slowly over time and capture characteristics that include institutions, culture, history and business practices *(Harding et al., 2008:79).*

Often, in order to capture these characteristics there is a need for a large number of variables and often these variables are somehow intercorrelated leading to the so called "multicollinearity problem". The best indicators of this problem are the standard errors t-ratios of the individual coefficients. However multicollinearity needs not to be a problem without solution. One of the ways Maddala *(2005:268)* suggests to solve this problem is through the Principal Component Analysis (PCA) which is a method that is used for reducing to a smaller set of variables the dimension of multivariate data sets. This method is based on the intuition that a smaller set of uncorrelated variables is better than a larger set of correlated variable an idea which was independently conceived by Pearson *(1901)* and Hotelling *(1933)* (discussed by *Dunteman, 1989:7*). The PCA will be thoroughly analyzed in the next section.

4.1.2 Principal Component Analysis

The PCA is a useful method especially when variables are highly correlated. The new variables (or factors) are called principal components and they are linear combinations of the original variables, which are uncorrelated and explain most of the variation in the data. In a regression context, one can therefore focus on this smaller number of independent variables, rather than dealing with the large number of original variables with complex interrelationships *(Chaterjee et al., 1999* described by *Taylor et al., 2006:1).*

Based on the formal derivation of the principal component from Maddala *(Maddala, 2005:281-282)*, there are k explanatory variables with k linear functions:

$$\begin{aligned} z_1 &= a_1 x_1 + a_2 x_2 + \dots + a_k x_k \\ z_2 &= b_1 x_1 + b_2 x_2 + \dots + b_k x_k \end{aligned} \qquad (2)$$

and the choice of the a's is based on the maximization of z_1 subject to the normalization condition:

$$a_1^2 + a_2^2 + \dots + a_k^2 = 1 \qquad (3)$$

The process of maximizing the variance of the linear functions subject to the condition that the sum of squares of the coefficients of x's is equal to 1 produces k solutions which are linked to the k linear functions (2.2). If z_1 is the linear function of the x's that has the highest variance and it is called the "first principal component" and z_2 the next highest variance we can order so that:

$$\operatorname{var}(z_1) > \operatorname{var}(z_2) > \dots > \operatorname{var}(z_k) \qquad (4)$$

These components have several properties:

First of all, $\operatorname{var}(z_1) + \operatorname{var}(z_2) + \dots + \operatorname{var}(z_k) = \operatorname{var}(x_1) + \operatorname{var}(x_2) + \dots + \operatorname{var}(x_k)$ (5)

Secondly, there is zero multicollinearity among the z's while the x's are correlated *(Maddala, 2005: 282).*

As suggested by the research carried out with Dr. Emiliano Piccinin in the spring of this year, this method is carried out empirically in several steps. The first step is to carry out a covariance analysis, where the correlation coefficients of the variables in the dataset are obtained. The correlations are useful to cluster the factors. The strong linear links within the variables mean that it is possible to represent the information in the database with a lower amount of dimensions. The second step is to select the number of factors through eigenvalues which identify the variance explained by the principal components. The largest eigenvalues correspond to the factors that are associated with most of the covariability; more specifically the first factor describes most of variability. The third step is to assess the communality, which is the proportion of the variance explained by the common factors *(Garrett-Mayer, 2006:29).* In the fourth step, the analysis of the factor pattern takes place; it is often referred as "the factor loading matrix in factor analysis" where the elements in the loading matrix are called factor loadings. Each loading can be interpreted as a correlation between an observed variable and a factor, provided that the factor solution is an orthogonal one (factors are uncorrelated), such as the current initial factor solution. Hence, the factor loadings indicate how strongly the variables and the factors or components are related. If a

Principal Component presents a strong correlation with a certain variable, it means that this component represents that variable. The objective in this phase is to improve the interpretability of factors by spreading the variability more evenly among factors through the process called "rotation" *(Garrett-Mayer, 2006:35 and 40).* To do so one needs to rotate the factor loads to obtain a clearer picture of the relevance of each variable in the factor. There are five types of rotations: three are orthogonal the varimax, the quartimax and equamax whereas direct oblimin and promax are oblique rotations. The choice amongst these five methods depends on the relation between the underlying factors, if the factors are expected to be independent then the three orthogonal rotations are more appropriate *(Field, 2005:3)*. More specifically, the Varimax (change of coordinates that maximizes the sum of the variances of the squared loadings, *Kaiser, 1958*) is recommended when the objective is to create new variables without inter-correlated components *(Torres-Reyna, 2010:4, Field, 2005:3).*

In the second to last step it is important to verify through a regression if these new factors have an impact on the original variable. The last step is to analyze the quality of the results obtained using the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) and Bartlett's test of sphericity. The KMO is a measure of sample adequacy both for each variable and overall *(Kaiser, 1970)* and the statistic varies between 0 and 1, where 0 indicates diffusion in the pattern of correlations while a value close to 1 indicates that patterns of correlations are relatively compact *(Field, 2005:6)*. The value of KMO should be greater than 0.5 for the sample to be considered adequate *(Kaiser, 1974, Field, 2005:3,* for more detail on the different thresholds *Hutchenson and Sofroniou, 1999:224-225)* The Bartlett's test of sphericity tests the null hypothesis that the original correlation matrix is an identity matrix (all correlation coefficients would be zero). For the success of the analysis it is important that there is some relationships between the variables. The null hypothesis is rejected when the test is significant (ideally a significance of less than 0.05) *(Field, 2005:6).*

4.2 The parametric methods

4.2.1 Corrected Ordinary Least Squares

In standard econometric analysis, residuals represent the deviation between observed data and the data predicted and can be interpreted as statistical error, caused by measurement inaccuracies or unobservable heterogeneity *(Jacobs et al., 2006: 50–57)*. Farrell *(1957)* suggests that residuals can be used to describe the difference between observed and predicted costs and signals the presence of inefficiency or efficiency. Corrected ordinary least squares (COLS) begins with the estimation of a standard cost function *(Aigner et al., 1977; Meeusen and van Den Broeck, 1977)*:

$$\ln C(y_i, w_i) = \alpha_0 + \eta_1 \ln y_i + \beta_1 \ln w_i + \varepsilon_i \qquad (10)$$

where $\varepsilon_i = v_i + u_i$ (11)

and *yi* represents the quantity produced for the *i*th unit under investigation, *wi* represents input prices, v_i is a random error term, and u_i captures non-random departures from efficiency (*Gori, Pierleoni 2013:262).*

4.2.2 Modified Ordinary Least Squares

An alternative method known as modified ordinary least squares (MOLS) is based on the argument from Schmidt *(1976)* that COLS is inadequate if the error term follows a one-sided distribution such as exponential or half- normal *(Uğur, 2004)*. MOLS consists of correcting the intercept with the expected value of the error term (e*i*) and adopting OLS to get a consistent estimate. A problem with the MOLS technique is that the estimates can take on values which have no statistical meaning (Mastromarco, 2008). It does not ensure that all units are bounded from above/below by the estimated production/cost frontier. If a unit has a large positive OLS residual then it is possible that

$$u_i - \mu_{\varepsilon i} \gtrless 0 \qquad (12)$$

thus the technical efficiency score is greater/less than unity. The most important weakness for both COLS and MOLS is the issue of consistency. The estimators are efficient but not consistent and they are applied mainly in cross-sections *(Gori, Pierleoni 2013:262).*

4.2.3 Ordinary Least Squares with Fixed Effect

If unobservable factors are time invariant, then fixed effect regression will eliminate omitted variable bias. Including fixed effects in efficiency analysis makes it possible to control for the average differences across operators in any observable or unobservable predictors. The fixed effect coefficients absorb all the across- group action; what is left over is the within-group action. The fixed effect offers several advantages: for example, it avoids obtaining distorted estimates of the parameters due to possible correlation between individual effects and the observed regressors *(Mundlak, 1978).* The disadvantage of using the fixed- effect model is that the regressors, which are invariant to time, are eliminated from the specified model and their effect is captured by the individual dummy variable. This does not allow the introduction in the cost model time-invariant variables. Another disadvantage is that the coefficients of the dummy variables could represent not only the efficiency levels but also the effect of the omitted variables *(Dranove, 2010; Vartanian and Buck, 2005, Gori, Pierleoni 2013:262).*

4.2.4 Generalized Least Squares Random Effect

In the random effects approach, it is assumed that the 'cross- specific effects' and the 'period effect' are modeled like random variables. When these components are independent from the other regressors it is possible to estimate a cost or production model using generalized least squares (GLS). The advantage is that it allows the use of time- invariant regressors in the model. In the random effects (RE) model, the *ui* are assumed to be randomly distributed with a constant mean and variance and they are assumed to be independent from both the random errors and the regressors. Given that inefficiency can only take non- negative values, the distribution of *ui* is often assumed to be half- normal, truncated normal, gamma or exponential. RE models are used in the analysis of panel data when one assumes no fixed effects. It uses the weighted average of the residuals of the GLS regression with random effects for each country across the years *(Gori, Pierleoni 2013:263).*

3.2.1.5 Stochastic frontier

Stochastic frontier models date back to Aigner, Lovell & Schmidt *(1977)* and Meesen and Van Den Broek *(1997),* who developed a stochastic frontier production function with a two-part 'composed' error term. This error is composed of a standard random error term, representing measurement error and other random *factors (Lovell & Schmidt, 1993 and Coelli, Rao and Battese, 1999),* and a one-sided random variable representing what Farrell *(1957)* called 'technical inefficiency', which is the gap between the observation and the production frontier. This notion of technical efficiency reflects the ability to obtain maximal output from a given set of inputs. It is measured by the output of the firm relative to that which it could attain if it were perfectly efficient, and thus lay on the frontier itself. When one combines this with allocative efficiency, the ability of the operator to use the inputs in optimal proportions, given their respective prices, one has a measure of total economic efficiency *(Gori & Pierleoni, 2013).* The functional forms more used in the literature and the Cobb Douglas *(Meeusen and Van Den Broek, 1997)* and the translogarithmic *(Christensen, Jorgenson and Lau 1971 and 1973,* discussed in *Gori et al., 2006).* The former one is more easily interpreted while the latter one is more flexible.

Using duality in production, one can consider cost efficiency: a typical stochastic cost frontier would be : $c_i = c(X_i, \beta) + \varepsilon_i$ (13)

where the logarithm of costs c_i, depend upon a vector of variables, X, and parameters, β, and a composite error term $\varepsilon_i = u_i + v_i$ which is made up of a non-negative random variable, u_i, which represents inefficiency, and a random error term, v_i. Battese and Coelli *(1988)* define technical efficiency of a unit sample, i, as the ratio of its average cost, given the level of inefficiency, to the corresponding average cost if the inefficiency level were zero. Using this definition, technical efficiency, TE_i is:

$TE_i = E(c^*_i \mid u_i, x_i) / E(c^*_i \mid u_i = 0, x_i)$ (14)

where c^*_i is the value of cost (in original units) for the i^{th} unit sample. If it is assumed that the cost function (1) is expressed in logarithmic form, then the inefficiency term will be:

$TE_i = \exp(-u_i)$ (15)

The stochastic frontier approach implies the application of the Maximum Likelihood (ML) estimator that considers the residuals as random variables. To isolate the random component from the inefficiency component, the stochastic frontier method uses a symmetric distribution for the random component, usually a standard normal distribution, while the inefficiency component is modelled according to an asymmetric function, usually a half-normal, because it considers the fact that the inefficiency term cannot be negative and thus they need to be linked to a truncated distribution. The inefficiency and the error terms need to be both orthogonals to the input and output variables or to the other variables used in the specification.

From a methodological point of view, the ML approach allows to fine tune and extend the efficiency analysis to a fixed model, even if it introduces the risk of distortions of the estimates due to the correlation between individual terms and the regressors. It is possible to include in the production/cost frontier time invariant variables, thus obtaining estimates of the individual inefficiency isolated from the effects linked to heterogeneity of the environment where the different institutions operate. Furthermore, the ML methodology developed by Battese and Coelli *(1995)* allows the estimation of the impact of certain important explicative variables.

In this context one question that arises from inspection of (13) is how one models the effects of exogenous factors on costs. These factors, which are beyond the control of the unit under observation, at least in the short-run, are often called exogenous, background or environmental factors. In this sense it is possible rewrite (13) as:

$c_i = c(X_i, Z_i, \beta) + u_i(E_i, \Delta) + v_i$ (13)*

where Z_i and E_i represent background variables affecting costs directly through the cost frontier and indirectly through inefficiency, respectively and Δ is a vector of parameters to be estimated *(Stevens, 2004:4,* Pierleoni with the research support of *Gori, Pierleoni, 2012:58-66).*

There is an extensive literature on the question of what determines inefficiency which was raised soon after stochastic frontiers were developed *(Pitt and Lee (1981), Kalirajan (1981), Kumbhakar, Gosh and McGulkin (1991) and Reifschneider and Stevenson (1991) and Huang and Liu (1994)).* Battese and Coelli *(1995)* extended this analysis to allow for panel data estimation. More recently, Coelli, Perelman and Romano *(1999)*, examined the effect of including the background factors either all in Z or all in E (discussed by Pierleoni with the research support of *Gori, Pierleoni, 2012:60-63, Gori 2013b:16-17)*. There is also extensive literature identifying SFA(ML) as the best technique among other parametric and non parametric methods since it is crucial to isolate the term of efficiency by environmental variables *(Filippini et al., 2003)*, it is more efficient than GLS Random Effect *(Sena, 2003)*, it is better than DEA when the panel data *(Badukenko et al., 2011)*.

3.2.2 The non parametric

The single stage Data Envelopment Analysis (DEA) is a non-parametric method which employs mathematical programming *(Coelli et al. 1998)*. It is based on the work by Farrell *(1957)* further elaborated by Charnes et al. *(1978)* and Banker et al. *(1984)*. There are several advantages of using this method: it takes into account multiple inputs and outputs for calculating, it comes with a single scalar value as a measure of efficiency and does not require any specification of functional forms as is required under parametric methods *(Tripathy, 2009:4)*. With DEA, relative efficiencies of a set of decision-making units (DMUs) are calculated, for each of them the highest possible efficiency score is assigned considering the inputs and outputs. It constructs an efficient frontier composed of the more efficient firms (in terms of either output maximization for a set of inputs or input minimization for a given output) while those firms below the efficient frontier are inefficient. For every inefficient DMU, it identifies a set of benchmark efficient units *(Coelli et al. 1998, Tripathy et al., 2009:5, Seiford & Thrall, 1990).*

DEA compares each producer with only the "best" producer. The production process for each producer is to take a set of inputs and produce a set of outputs. Each producer has a varying level of inputs and gives a varying level of outputs. The objective is to find the "best" virtual producer and to compare it to the producer under consideration. It assigns an efficiency score less than one to (relatively) inefficient units, thus a linear combination of other units from the sample could produce the same vector of outputs using a smaller vector of inputs. There are three assumptions when the production frontier is created: 1) every observed production plan belongs to the production set; 2) any unobserved production plan that is weakly dominated by another production plan is also part of the production set; and 3) the production plan will determine the returns to *scale (Gori & Pierleoni, 2013).*

The CRS has two strong assumptions a proportionate increase in the inputs results in the same proportionate increase in the output and that the optimal mix of inputs and outputs is independent of the firm's scale of operation *(Coelli, 1996 and Tripathy et al, 2009:9-11).*

The CCR model can be expressed as:

$$\min_{\theta, \lambda} \theta \qquad (16)$$

subject to $-y_i + Y\lambda \geq 0$,

$\theta x_i - X\lambda \geq 0$,

$\lambda \geq 0$,

where θ is the technical efficiency value; λ is the intensity weight; x_i is the K×1 input vector of the ith DMU; y_i is the M×1 output vector of the ith DMU; X is the K×N input matrix; and Y is the M×N output matrix.

The CCR model assumption of constant returns to scale (CRS) is appropriate when all DMUs are operating at an optimal scale. This is not always the case since several exogenous and endogenous factors may prevent a DMU from operating at an optimal scale *(Gori, 2013a)*. To address this proposed Banker et al. (1984) proposed an extension of the CRS DEA model to account for variable returns to scale (VRS) *(Cooper et al. 2007:114-115).*

This was done by adding a convexity constraint, $N1'\lambda = 1$,

$$\min_{\theta, \lambda} \theta \qquad (17)$$

subject to $-y_i + Y\lambda \geq 0$,

$\theta x_i - X\lambda \geq 0$,

$N1'\lambda = 1, \quad \lambda \geq 0$,

where N1 is an N×1 vector made up of ones.

The BCC model forms a convex hull of efficiency frontier which envelops the data points more tightly than the CCR model *(Hu & Chu, 2008: 228-229).*

DEA can be a powerful tool when used wisely: it can handle multiple input and multiple output models, it doesn't require an assumption of a functional form relating inputs to outputs, exogenous variables can be taken into account, DMU_s are directly compared against peers, inputs and outputs can have very different units. However, since DEA is an extreme point technique, noise can cause significant problems, it is good at estimating "relative" efficiency but it converges very slowly to absolute efficiency, and statistical hypothesis tests are difficult and since a standard formulation of DEA creates a separate linear program for each DMU, it is computationally complex *(Gori & Pierleoni, 2013).*

3.2.3. The semi parametric

The two-stage DEA is a semi-parametric method which entails that in the first stage efficiency scores are obtained considering only endogenous variables. In the second stage, the efficiency scores are used as a dependent variable, which is then regressed on the exogenous factors to isolate the impact of these variables on the efficiency levels.

This method is useful when exogenous factors need to be taken into account in the efficiency analysis. According to Ray *(Ray 1988,* further developed in *1991*) non-discretionary factors, such as environmental factors, should not be included in a DEA assessment. Since the DEA should include only controllable factors, it is possible to include these non-discretionary variables in a second phase, in a regression to estimate the part of efficiency that is explained by uncontrollable factors. In the past two decades new two stage models have been developed and it is more appropriate to consider in the second stage Tobit models rather than traditional regression models to account for the fact that the dependent variable is bounded between 0 and 1 *(Despic et al., 2008, 349-350).* Simar and Wilson *(2007)* present an exhaustive list of 47 published papers that have used these models which can be used not only to adjust efficiency scores by incorporating non-discretionary factors (following the intuition from *Ray, 1991*) but even more to explain differences in efficiency between the units under investigation *(Despic et al., 2008, 405).* As we have seen above an important decision in the first step when using the DEA is to decide between the two possible models the CCR (from *Charnes, Cooper, Rhodes, 1978*) model based on the assumption that constant return to scale (CRS) the BCC model (from *Banker, Charnes, Cooper, 1984*) which assumes variable returns to scale (VRS) frontiers (*Cooper et al. 2007:114-115*). The VRS is more appropriate than CRS for regulated sectors with high fixed or recurrent costs, where the cost structure is heavily impacted by the scale of production *(Gori et al., 2006).* However, a proper approach could be that of deciding on the basis of the database of peer companies. If the set of operators under scrutiny are of similar size then the choice of which method is used becomes less relevant and there is extensive literature justifying one case or the other. However, if the set of peers under scrutiny are not very homogenous it comes forward that the VRS could perform better *(Amornkitvikai & Harvie, 2010).*

In the second stage, the efficiency scores obtained from the first stage (the DEA efficiency score lies in the interval 0 and 1) are regressed on external environmental factors using a censored regression model called Tobit *(Tobin, 1958, Maddala, 1983).* The reason for the use of the Tobit is that it is applicable in cases where the dependent variable is constrained in some way and because it uses all observations, both those at the limit and those above it to estimate a regression line *(McDonald & Moffitt, 1980:318).*

The Tobit model may be defined as:

y^* ; $0 <= y^* <=1$ (17)

$y = 0$; $y^* < 0$;

1 ; $1< y^*$

$y^* = \beta xi + \varepsilon t$

where y is the DEA VRS score. $\varepsilon t \sim i\ e\ N(0, \sigma 2)$

y^* is a latent (unobservable) variable.

β is the vector of unknown parameters which determines the relationship between the independent variables and the latent variable, xi is the vector of explanatory variables (*Tripathy et al, 2009:12* discussed in *Gori, 2013a:2-5*). A two stage Data Envelopment Analysis (DEA) with Tobit regression is a methodology which could overcome the weakness of DEA in case of panel data (presented by *Badukenko et al., 2011* and thoroughly discussed by *Pierleoni 2012:87*).

Several papers have used the two-stage DEA to assess the efficiency of local institutions, national and international companies. The selection of papers presented below was carried out with the objective of identifying cross country empirical studies or concerning sectors highly labor intensive or institutions impacted by the demographic structure.

For the banking sector, Pasiouras *(2007:1),* used a sample of 715 banks from 95 countries and two-stage DEA to provide international evidence on the impact on efficiency of exogenous variables such as regulation and supervision. He uses DEA to estimate technical and scale efficiency and the Tobit regression to investigate the regulation’s impact. Delis and Papanikolaou *(2009)* used a two-stage methodology to analyze the efficiency of ten banks operating in the European Union. From their research it emerges that size, industry concentration and the investment environment had a positive impact on the efficiency levels. At the national level, Fethi and Jackson *(2000)* investigated the performance of Turkish commercial banks using DEA and then they used a Tobit model to identify the variables explaining the efficiency of some banks (e.g. size, profitability, ownership). They found that larger and more profitable banks are more likely to operate at higher levels of technical efficiency. Similarly Tripathy et al, 2009 analyzed the levels and determinants of efficiency of firms in the Indian pharmaceutical industry by using firm-level data and a two-stage DEA. In the first stage, they carry out a technical efficiency analysis of 90 sample firms with on output, sales of the sample firms from 2001-02 to 2007-08, and three inputs; raw material cost; cost of salaries and wages; and cost of advertising and marketing. In the second stage, the efficiency scores obtained from the first stage are regressed using a Tobit model on environmental factors such as the age of the firms, export of goods, import of capital goods, profit rate, R&D intensity, ownership, patent regime and foreign direct investment *(Tripathy et al, 2009:1).*

Another study in the health sector conducted by Luoma, et al. *(1998)* applied DEA and Tobit to evaluate how various economic, structural and demographic factors effect inefficiency of Finnish health centers. Their results indicate that a higher level of central government grants to the health sector and a higher income per inhabitant are predictors of inefficiency. While Hsuan Lien & Chia-Yu apply the two-stage DEA to measure Taiwanese public hospital efficiency. They use the Tobit regression to investigate the effects of the national program Strategy Hospital Alliance (SHA) on hospital efficiency after controlling for other factors effecting hospital efficiency (e.g. size, degree of competition) *(Hsuan Lien & Chia-Yu, 1998:7-8).*

Concerning the automotive industry, Leachman, et al. *(2005)* adopted a two-stage DEA to examine data from eight automobile manufacturers. They reached a strong conclusion that a strong R&D commitment and the ability to compress production time generate differences in manufacturing performance. Moving to the software industry, Hwang and Oh *(2008)* measured the performance of Korean software firms. In the first step they measured efficiency by using DEA, and then they applied a Tobit regression to investigate the impact of Intellectual Property Rights (IPR) on efficiency. Their results indicated that the average efficiency of firms possessing IPR was higher than those without them. Fethi, et al. *(2000)* used instead the two-stage DEA application to assess the efficiency of European airlines. Their findings confirm that industry concentration and subsidy policies have a negative impact on the efficiency of firms *(Tripathy, 2009:6-7).*

This technique has also been used to scrutinize the efficiency and productivity of telecom firms. A two-stage method was applied in 2008 to examine the efficiency scores of twenty-four telecom firms in APEC member economies during the 1999-2004 period. The DEA was used in the first stage to measure the technical efficiency of these companies. In the second stage efficiency scores are regressed on the environmental variables using the Tobit regression *(Hu J.L. & Chu W. K, 2008:223).*

Loikkanenand Susiluoto, use the two-stage DEA modelling to assess the economic performance of Finnish regions. Their data consists of regional input and output variables and other characteristics concerning 83 regions in Finland during the period 1988-1999 *(Loikkanen & Susiluoto, 2002:3).* The two-stage method has been utilised also in the agricultural field, where the efficiency scores might be influenced by external environmental factors not controlled by farmers *(Kapfer et al., 2012).*

3.2.4 Spearman rank correlation coefficient

Using several econometric methods to carry out cross country analysis generates a ranking among the countries/operators being benchmarked. It is important to assess if these rankings are consistent across the different methods used, parametric, non-parametric and semi-parametric. To do so the appropriate method is the Spearman rank correlation coefficient *(Spearman, 1904* reprinted in *2010).* It is a non-parametric measure of statistical dependence between two variables when they are ordinal numeric. It assesses the relationship between two variables using a monotonic function. A perfect Spearman correlation of +1 or −1 occurs when each of the variables is a perfect monotone function of the other. The main objection to using ranks is that it is based on the assumption that all subjects differ from one another by the same amount whereas in reality this could be unrealistic. However, there are advantages in using ranks and trying to find a correlation between different ranks *(Gori, 2013a).* First of all, by means of rank a series presenting the normal curve can be compared with another series presenting a different curve. Secondly, it has a useful property of allowing any two series to be combined into a third composite one *(Spearman, 2010:1141).*

4. Way forward and future research

To be able to implement this comprehensive framework to real court cases involving network industries there is one precondition which needs to be taken into account and some technical improvements to the models need to be scrutinized further.

The precondition is that the quality of data is crucial for a proper assessment of the performance of the different players in a market. For example a full investigation of Postal operators' efficiency requires the availability of detailed data on the cost structure of postal operators which are very difficult to collect mainly because of the reluctance of operators to disclose sensitive information in an already competitive environment. For example, the thorough analysis on mail demand from

Harding *(2004, 2006)* was possible because he had access to not public data. Only from this type of collaboration it is possible to extend the analysis by type of mail, by origin and destination of mail and the cost structure linked to each phase. This would allow a more detailed analysis for each type of postal submarket and sub-phase also. Furthermore, even if data is available there is an issue of consistency of data in case of benchmarking exercises. A possible solution lies in cooperation between the Association of European Postal Operator (PostEurop), the Association of National Regulators formally the European Conference of Postal and Telecommunications Administrators (CERP) and the European Union's institutions. This cooperation should lead to the definition of a methodology in gathering data, a clear taxonomy of the different sub-phases and a flexible and on-going process to be able to provide a consistent updated statistical database. This principle can be than extended also to national regulatory bodies to develop a national database on the cost structure for each sub-phase and where applicable data on access conditions.

Furthermore, for both competitive and non-competitive segments of the postal market, it is important to use complete and reliable datasets which incorporate regulatory variables (for example mandatory quality levels and price caps) and extend the use of exogenous variables to understand the operational impact on postal operators of the environmental and geographical characteristics of a country. Mountains, islands, extreme weather conditions, overseas territories and other characteristics impacting the operations are variables worthwhile to take into account in an objective and justifiable way.

Concerning the improvement of the techniques utilized in this paper there are two topics which could need further scrutiny:

1) For competitive markets, from the analysis of Gori (2013b) it has emerged that it is appropriate to use the Survivor Technique. The survivor technique type diagrams have emerged as a very valid instrument to capture the trends but for an appropriate utilization of this technique it is important to monitor two types of curve and their relation, the long run average cost curve (LRAC) and the short run average cost (SRAC). The LRAC curve depicts the cost per unit of output in the long run and is created as an envelope of an infinite number of short-run average total cost curves, each based on a particular mix of inputs. All points on the line represent an optimal combination and is typically U-shaped. This can turn out to be very difficult in a multiproduct industry like postal services, hence the creation of these curves needs careful scrutiny and further research.

2) Concerning the End to End less competitive segments of the industry it is more appropriate to use econometric methods. First of all the Principle Component analysis helps to identify the determinants of demand which will become the inputs of the cost function which will be put into scrutiny. After that a series of parametric, non-parametric and semi-parametric methods need to be

applied and the use of the Spearman method turns out to be a useful tool to assess the consistency of the efficiency analysis. From the research work it emerges that semi-parametric such as Two stage DEA with Tobit could be useful to incorporate the advantages of both non-parametric and parametric methods. More specifically, the DEA appears to be useful to carry out in the efficiency analysis in the first stage, and then in the second one the efficiency scores obtained from the first stage (the DEA efficiency score lies in the interval 0 and 1) are regressed on external environmental factors using a Tobit censored regression model.

3) To go beyond Gori *(2013a)* and to improve the two stage DEA approach other techniques, not analyzed in this paper, can be applied. The first one is the outlier detection approach alongside a bootstrapping analysis of a classical dataset (this approach applied to the postal sector was suggested by Dr. Maria Rita Pierleoni, based on *Johnson & Mcginnis, 2008, Fox, 2002, Guan, 2003)*. More specifically, Johnson and Kuosmanen propose the use of a bootstrap method to correct for the small sample bias and serial correlation of the DEA efficiency estimates (*Johnson & Kuosmanen, 2011* discussed in *Gori, 2013a*).

The first use of the bootstrap in frontier models dates to Simar (*1992*) and the application of this method nonparametric envelopment estimators was developed by Simar and Wilson (*1998, 2000*). While the analysis of the theoretical properties of the bootstrap with DEA estimators is thoroughly analysed in Kneip et al. *(2003)*. The essence of this approach *(Efron, 1979, 1982; Efron and Tishirani, 1993)* is to approximate the sampling distribution by simulating the Data Generating Process *(Simar and Wilson, 2008:445)*. It is the intention of Dr. Maria Rita Pierleoni, who initially suggested of the semi-parametric approach, to further scrutinize the possibility of applying the bootstrap method to the postal sector *(Pierleoni, 2013f).*

For future research, it might be also interesting to take on the suggestion of Dr. Meloria Meschi to compare the results obtained through a two stage DEA-Tobit to new models emerging in empirical efficiency studies applied to regulated industries such as the True Fixed Effect (TFE) and the True Random effects (TRE) discussed in Greene *(2005).* These models produce separate estimates of time invariant unobserved heterogeneity, random error and efficiency. Unfortunately, the TFE model is impossible or difficult to use if there are time invariant drivers or if cost drivers show minimal variations *(Gori, 2013a).*

5. Conclusion

The objective of this paper was to identify innovative ways of addressing the issue of efficiency in network industries and more specifically in the postal sector and present original approaches for this sector. The whole Research project was put in place based on the belief that through the use of different appropriate econometric techniques a benchmarking exercise on network industries, more specifically in the postal sector, could be carried out while taking account exogenous variables identifying national characteristics. Thus, comparisons based on reasonable causal chains with exogenous variables (e.g. factors independent from provider's responsibility) and modeled by specific econometric techniques allow the building of a benchmarking framework that is able to account for efficiency differences *(Gori & ViscoComandini, 2009:9-10).*

Special emphasis has been placed along the research project on the different methodological approaches to allow assessing whether a postal undertaking beneficiary of public subsidies in fulfilling its obligations operates at an efficient level or not.

More specifically, this paper attempted to define the analytical economic framework needed to address the issue of state aid while it presents a comprehensive method. In a very recent decision *(European Commission, 2012)* on state aid applied to the postal sector the European Commission took a very clear position. It highly criticized the use of the two pillars of cost analysis in the postal sector in the past decade (*Nera 2004,* and *Cohen et al., 2002,* and *2004*) and it suggests to develop a more robust and comprehensive methodology. This paper responds to this suggestion by offering a comprehensive three step approach. This method can be applied not only by economists, but also by law scholars, national regulatory authorities and European and national courts from other industrialized countries.

7. References

Aigner D.J., Lovell C.A.K. & Schmidt P., 1977, *Formulation and Estimation of Stochastic Frontier Production Function Models*, in Journal of Econometrics, 6, 21-37.

Ali H. M. M., 2008, *Development of Arab Water Sustainability Index Using Principal Component Analysis*, National Water Research Center, Water Quality Unit, Ministry of Water Resources and Irrigation, Egypt presented during The 3rd International Conference on Water Resources and Arid Environments and the 1st Arab Water Forum.

Amornkitvikai, Y. & Harvie, C., 2010, *Identifying and measuring factors of technical inefficiency: evidence from unbalanced panel data of Thai listed manufacturing enterprises*. Korea and the World Economy, IX (pp. 1-32). Korea: The Association of Korean Economic Studies, University of Incheon

Asteriou D. & Hall S. G., 2007, *Applied Econometrics,* Palgrave Macmillan.

Badunenko O., Henderson D.J. &Kumbhakar S.C., 2011, *When, where and how to perform efficiency estimation,* MPRPA.

Bain J. S., 1956, *Barriers to New Competition,* Harvard University Press.

Bain J. S., 1959, *Industrial Organization*, New York, Wiley.

Bain J.S., 1969, *Survival ability as a test of efficiency*, American Economic Review, Papers and Proceedings, 59: 99–104.

Bank for International Settlements,

Banker R.D., Charnes A. & Cooper W.W., 1984, *Some models of estimating technical and scale inefficiencies in data envelopment analysis*. Management Science, Vol. 9:1078-1092.

Bays C. W., 1986, *The determinants of hospital size: a survivor analysis,* Applied Econometrics, 18: 359-377.

Battese G.E. &Coelli T.J., 1988, *Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data,* Journal of Econometrics 38:87-99.

Battese G.E. &Coelli T.J., 1995,*A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data,* in Empirical Economics, 20:325-332.

Blair R.D. & Vogel R.J. 1978, *A survivor analysis of commercial health insurers,* Journal of Business 51:521–529.

Carnero, A., Koopman, S.J. & Ooms, M., 2003, *Periodic heteroskedastic RegARFIMA models for daily electricity spot prices*, Discussion Paper 03-071/4, Tinbergen Institute, Amsterdam.
Charnes A., Cooper W.W. & Rhodes E., 1978, *Measuring the efficiency of decision making units.* European Journal of Operational Research, Vol. 2:429-444.
Chaterjee, S., Hadi, A.S. & Price, B., 1999, *Regression Analysis by Example*, third ed., New York: Wiley.
Christensen L.R., Jorgenson D.W. & Lau L.J., 1971, *Conjugate Duality and the Transcendental Logarithmic Production Function*, in Econometrica, 39:255-256.
Christensen L.R., Jorgenson D.W. & Lau L.J., 1973, *Transcendental Logarithmic Production Frontiers*, in The Review of Economics and Statistics, 55.
Coelli T., Perelman, S. & Romano E., 1999, *Accounting for Environmental Influences in Stochastic Frontier Models: With Application to International Airlines,* in Journal of Productivity Analysis, 11:251-73.
Coelli T.J., Rao D.S.P., & Battese G.E., 1999, *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers.
Coelli T.J., Rao, D.S.P., O'Donnell, C.J. & **Battese, G.E., 1998,** *An Introduction to Productivity and Efficiency Analysis*, Springer Science.
Coelli, T., 1996, *A Guide to DEAP Program Version 2.1: A Data Envelopment Analysis (Computer) Program,* Working Paper 96/08, Centre for Efficiency and Productivity Analysis, Department of Econometrics, University of New England, Armidale, Australia.
Cohen R., Pace C., Robinson M., Scarfiglieri G., Scocchera R., Visco Comandini V., Waller J. and Xenakis S., 2002, *A comparison of the burden of Universal Service in Italy and the United States* in M. Crew e P. Kleindorfer, Postal and delivery services: Pricing, productivity, regulation and Strategy, Kluwer Academic Publishers.
Cohen R. H., Robinson M., Scarfiglieri G., Sheehy R., Visco Comandini V., Walzer J., Xenakis S. S., 2004, *The role of scale economies in the cost behavior of posts*, Proceedings of the Wissenschaftliches Institut für Kommunikationsdienste (WIK), 8th Köenigswinter Seminar on "Regulating Postal markets- Harmonised Versus Country Specific Approaches", February 16-18.

Cooper W. W., Seifurd L. M. & Tone K., 2007, *Data Envelopment Analysis,* Second Edition, Springer.
Crews C., Deduytsche L., Ducasse J.P., Gori S., 2006, *Scenario Planning applied to the postal sector: In a decade what type of cake will the European Postal Industry turn out to be: a pan cake, a mille-feuille or a Black Forrest,* presented at the 14th Conference on Postal and Delivery Economics, in Bern, on June 2nd.
Delis M.D. &Papanikolaou, N.I., 2009. *Determinants of bank efficiency: evidence from a semi-parametric methodology,* Managerial Finance, 35:260-275.
Despic, O., Portela M. C. S. &Thanassoulis E., 2008, *Data Envelopment Analysis: The Mathematical Programming Approach to Efficiency Analysis,* in Fried H.O., Knox Lovell C.A., Schmidt S. S., The Measurement of Productive Efficiency and Productivity Growth, OxfordUniversity Press:251-420.
Dranove D., 2010, *Fixed Effects Models, material for Empirical Methods in Strategy Course,* Northwestern University.
Dunteman G. H., 1989, *Principal Component Analysis,* Sage Publications.
Efron B., 1979, *Bootstrap methods: another look at the jackknife,* Annals of Statistics 7:7-16.
Efron B., 1982, *the Jackknife, the Bootstrap and the Other Resampling Plans*, CBMS-NSF Regional Conference Series in Applied Mathematics, no. 38, Philadelphia: Society for Industrial and Applied Mathematics.
Efron B. &Tibshirani R. J., 1993, *An Introduction to the Bootstrap,* Chapman and Hall.
Elzinga K.G. (1990), *The beer industry. In The Structure of American Industry*, Adams W (ed.). Macmillan: New York.
European Commission, 2000, *Communication from the Commission on Services of General Interest in Europe*, COM/2000/0580 final.
Farrell M.J., 1957, *Measurement of Productive Efficiency*, Journal of the Royal Statistical Society, Series A:253-281.
Fethi, M.D., &Jackson P.M. 2000,*Evaluating the technical efficiency of Turkish commercial banks: An Application of DEA and Tobit Analysis*, In: International DEA Symposium, University of Queensland, Brisbane, Australia.
Field A., 2005, *Factor Analysis using SPSS,* material for Research Methods II Class.
Filippini M., Fazioli R. & Kunzle M., 2003,*Valutazione dell'efficienza delle compagnie di bus italiane e svizzere*, Bank of Italy.
Fox J, 2002, *Bootstrapping Regression Models,* Appendix to An R and s-Plus Companion to Applied Regression.
Frech H.E. & Ginsburg P.B., 1974, *Optimal scale in medical practice: a survivor analysis*, Journal of Business 47:23–46.
Garrett-Mayer E., 2006, *Statistics in Psychosocial Research*, Lecture 8 on Factor Analysis, John Hopkins Bloomberg School of Public Health.
Geddes R., 1999, Public utilities, Encyclopedia of Law and Economics, 1162-1205.
Giordano J.N., 1995, *Deregulation without apology: a truncated survivor analysis of long-run efficiency gains in the U.S. trucking industry*, Review of Industrial Organization 10: 635–650.

Giordano J.N., 1997, *Returns to scale and market concentration among the largest survivors of deregulation in the US trucking industry*, Applied Economics 29: 101–110.
Giordano J.N., 2003, *Using the survivor technique to estimate returns to scale and optimum firm size,* Topics in Economic Analysis and Policy 3: 1–21.
Giordano J. N., 2008, *Economies of Scale after Deregulation in LTL Trucking: A test case for the survivor technique*, Managerial and Decision Economics 29:357-370.
Gori S., 2013 a, *Using a semi-parametric method to carry out efficiency analysis in the postal sector,* presented at the 21st Conference on Postal and Delivery Economics, Portmarnock, County Dublin , Ireland , May 30th.
Gori S., 2013 b, *Applying the survivor technique to the postal sector,* presented at the 21st Conference on Postal and Delivery Economics, Portmarnock, County Dublin , Ireland, June 1st, 2013.
Gori, S., Lettieri C., Marè M. and Visco Comandini V., 2002, *The myth of the uniform price*, paper presented at the International Conference on Public Finance, Helsinki, August 29–31.
Gori S. & Marcone M., 2007, *How to reduce cream skimming through zonal pricing geared to costs in a postal market with mandated access,* presented at the SIDE 3rd Annual Conference, Universita' Bocconi 9-10th November.
Gori S., Piccinin E., Romito S. & Scarfiglieri G., 2006, *On the Use of Cost Functions in the Assessment of the Impact of Liberalization on Postal Universal Service Burden,* in Progress toward Liberalization of the Postal and Delivery Sector, edited by Crew and Kleindorfer, Springer Publishing.
Gori S. & Pierleoni M. R., 2013,*Efficiency analysis postal operators: a comparison between United States and Europe,* in Reforming the Postal sector in the face of Electronic competition, edited by Crew and Kleindorfer, Edward Elgar Publishing Ltd.
Gori S. & ViscoComandini V., 2009,*Antitrust law and public services performances with reference to the postal industry: a discussion paper*, Working Paper of the Institute of Governmental Studies, University of California, Berkeley.
Greene W., 2005, *Fixed and random Effects in Stochastic Frontier Models,* Journal of Productivity Analysis, Vol. 23:7-32.
Guan W., 2003, *From the help desk: Bootstrapped standard errors,* The Stata Journal, 3, Number 1:71-80.
Hansen M., Zuehlke S. & Van Ysendyck A. 2003, *Altmark Trans: European Court of Justice outlines conditions for public service compensation*, Practical Law Company, 7 August.
Harding M. C., 2006, *A Panel Data Study of Mail Demand in Advanced Economies,* Prepared for the 14th Conference on Postal and Delivery Economics, May.
Harding M. C., Jimenez L. & Lintell M., 2008, *The impact of Economic and Demographic Dispersion on aggregate mail volumes,* in Crew M. & Kleindorfer, Competition and Regulation in the postal and delivery sector, Edward Elgar Publishing limited.
Himmelweit S., Simonetti R. & Trigg A, 2001, *Microeconomics: Neoclassical and Institutional Perspectives on Economic Behaviour* , Thomson Learning.
Hotelling, 1933, *Analysis of a complex of statistical variables into principal components,* Journal of Educational Psychology, 24417-441, 498-520.
Hsuan Lien C. & Chia-Yu C., 1998,*The effects of StrategyHospitalAlliance on Hospital efficiency. Evidence from Public Hospitals in Taiwan,* Working paper National Taipei University.
Hu J.L. &Chu W. K, 2008,*Efficiency and Productivity of Major Asia-Pacific Telecom Firms*, Chang Gung Journal of Humanities and Social Sciences:223-245.
Huang C.J. & Liu J.T., 1994, *Estimation of a Non-Neutral Stochastic Frontier Production Function*, in Journal of Productivity Analysis, 5:171-80.
Hutchenson, G., & Sofroniou N., 1999, *The Multivariate Social Scientist: Introductory Statistics Using Generalized Linear Models,* Thousand Oaks, CA: Sage Publications.
Hwang D.S. & Oh, D., 2008. Do Software Intellectual Property Rights Affect the Performance of Firms? Case Study of South Korea. In: The Third International Conference on Software Engineering Advances, Sliema, Malta.
Jacobs, R., Smith P.C. & Street A., 2006, *Measuring Efficiency in Health Care: Analytic Techniques and Health Policy*, Cambridge, Cambridge University Press.
Johnson A. L. & Kuosmanen T., 2011, *One-Stage and Two-Stage DEA Estimation of the Effects of Contextual Variables,* Working Paper. Reprinted in 2012, *One-Stage and Two-Stage DEA Estimation of the Effects of Contextual Variables*, European Journal of Operational Research, Elsevier B.V.:1-12.
Johnson A. L. & McGinnis L. F., 2008, *Outlier detection in two-stage semiparametric DEA models*, European Journal of Operational Research, Volume 187, Issue 2: 629-635.
Kaiser H. F., 1958, *The varimax criterion for analytic rotation in factor analysis,* Psychometrika 23.
Kaiser H. F., 1970, *A second generation Little Jiffy*, Psychometrika, 35:401-415.
Kaiser H. F. ,1974, *An index of factorial simplicity,* Psychometrika, 39: 31-36.
Kalirajan K.P., 1981,*An Econometric Analysis of Yield Variability in Paddy Production*, in Canadian Journal of Agricultural Economics, 29:283-94.

Kapfer M., Kantelhardt J., Eckstein K. &Hübner R., 2012, *Spatially Explicit Efficiency of Agricultural Land Use, A Two-Stage DEA Approach,*presentation at the International Association of Agricultural Economists (IAAE) Triennial Conference.

Kim S. G., Koizumi A., Koo J. Y., Shim M. H. and Yu M. J., 2005, *Estimating regional water demand in Seoul, South Korea, using principal component and cluster analysis,* Water Science and Technology: Water supply, Vol.5 No.1:1-7.

Kneip A., Simar L. & Wilson P. W., 2003,*Asymptotics for DEA estimators in non-parametricfrontiermodels,* Discussion Paper n. 0317, Institut de Statistique, Université Catholique de Louvain.

Kumbhakar S.C., Ghosh, S. &McGuckin J.T., 1991,*A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms,* Journal of Business and Economic Statistics, 9:249-86.

Leachman C., Pegels, C.C. & Shin, S.K., 2005, Manufacturing performance: evaluation and determinants, International Journal of Operations & Production Management, 25:851-874.

Li Jinchao, 2011, *Electric power supply and demand early warning based on PCA and SVM method,* paper presented on the second International Conference on Intelligent Control and Information Processing, 25-28 July.

Loikkanen H. A. &Susiluoto I., 2002, An evaluation of Economic efficiency of Finnish regions by DEA and Tobit models, paper presented at the 42st Congress of the European Regional Science Association.

Lovell C.A.K. & Schmidt S.S., 1993,*The Measurement of Productive Efficiency: techniques and Applications*, Oxford University Press.

Luoma K., Jarvio, M., Suoniemi, I. &Hjerppe, R.T., 1998, Financial incentives and productive efficiency in Finnish health centres, Health Economics, 5:435-445.

Maddala G.S., 1983, Limited-Dependent and Qualitative Variables in Econometrics. Econometric Society Monographs No. 3. Cambridge University Press.

Maddala G. S., 2005, *Introduction to Econometrics,* John Wiley & Sons.

Mastromarco C., 2008, *Stochastic Frontier Models,* Department of Economics and Mathematics-Statistics, CIdE.

McDonald J. F. & Moffitt R. A., 1980,*The Uses of Tobit Analysis,* The Review of Economics and Statistics Vol. 62, No.2, May, The MIT Press: 318-321.

Meeusen W. & Van Den Broeck J., 1977, *Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error,* International Economic Review, 18:435-444.

Mill J. S., 1929, *Principals of Political Economy,* New York: Longmans, Green. Reprinted in 1948 *Principles of Political Economy,* John W. Parker: London. Reprinted in 1987. Augustus M. Kelley.

Mundlak Y. 1978, *On the pooling of Time Series and Cross Section Data,* in Econometrica, 46, 1, 69-85.

Norton, S. W. and Norton, W. Jr., 1986, 'Economies of Scale and the New Technology of Daily Newspapers: A Survivor Analysis', Faculty Publications, College of Journalism & Mass Communications. Paper 6.

NERA Economic Consulting, 2004, *Economics of Postal services: Final Report,*a report to the European Commission, July. http://ec.europa.eu/internal_market/post/doc/studies/2004-nerafinal-postal-report_en.pdf.

Pasiouras F., 2007, *International evidence on the impact of regulations and supervision on banks' technical efficiency: an application of two-stage data envelopment analysis,*University of Bath Working Paper Serie:1-37 republished in 2008.

Pearson K., 1901, *On lines and planes of closest fit to systems of points in space*, Philosophical Magazine 2:559-572.

Pierleoni M. R., 2013f, *Parametric and semi-parametric techniques to estimate the effects of exogenous variables on efficiency levels of the main postal operators,* forthcoming.

Pierleoni M. R. with the research support of Gori S., 2012, *Liberalization, privatization and competition in the postal services: the state of the art,* CRRI Research Project.

Pitt M.M. & Lee L.F., 1981,*Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry*, in Journal of Development Economics, 9:43-64.

Ray S. C., 1988,*Data envelopment analysis, Nondiscretionary inputs and efficiency: An alternative interpretation,* Socio-Economic Planning Sciences 22/4:167-176.

Ray S. C., 1991, *Resource-Use Efficincy in Public SchoolsA Study of Connecticut Data,* Management Science 37/12:1620-1628.

Ramanathan, R., Engle, R., Granger, C.W.J., Vahid-Araghi, F. & Brace C., 1997, *Short-run forecasts of electricity loads and peaks,* International Journal of Forecasting, 13, pp. 161-174.

Rees, R.D., 1973, *Optimum Plant Size in United Kingdom Industries: Some Survivor Estimates*, Econometrica, 40: 394-440.

Reifschnieder D. & Stevenson R., 1991, *Systematic Departures from the Frontier: A Framework for the Analysis of Firm Inefficiency*, in International Economic Review, 32:715-23.

Saving T.R., 1961, *Estimation of Optimum Size of Plant by the Survivor Technique,* The Quarterly Journal of Economics, Vol. 75, No. 4, Oxford University Press: 569-607.

Schmidt P., 1976, On the Statistical Estimation of Parametric Frontier Production Functions, Review of Economics and Statistics 58:2 (May): 238-39.

Seiford L.M. and Thrall R.M., 1990, *Recent developments in DEA, the mathematical programming approach to frontier analysis,* Journal of Econometrics:7-38.
Sena V., 2003, *The frontier approach to the measurement of productivity and technical efficiency,* University of Leeds.
Shepherd, W. G., 1967, *What does the Survivor Technique show about Economies of Scale?,* Southern Economic Journal, vol. 34: 113-22.
Simar L., 1992, *Estimating efficiencies from frontier models with panel data: A comparison of parametric, non-parametric and semi-parametric methods with bootstrapping,* Journal of Productivity Analysis 3:167-203.
Simar L. & Wilson P. W., 1998,*Sensitivity analysis of efficiency scores: How to bootstrap in non parametric frontier models,* Management Science 44:49-61.
Simar L. & Wilson P. W., 2000, Statistical *inference in nonparametric frontier models: The state of the art*, Journal of Productivity Analysis 13:49-78.
Simar L. & Wilson P. W., 2007,*Estimation and inference in two-stage, semi parametric models of production processes,* Journal of Econometrics 136:31-64.
Simar L. & Wilson P. W., 2008,*Statistical Inference in Nonparametric Frontier Models: Recent Developments and Perspectives,* in Fried H. O. Knox Lovell C. A. Schmidt S. S. The Measurement of Productive Efficiency and Productivity Growth, OxfordUniversity Press:421-521.
Spearmani C., 2010, *The Proof and Measurement of Association between two things,* reprints and reflections, International Journal of Epidemiology 2010;39:1137–1150. . Reprinted with permission Spearman C., 1904, The Proof and Measurement of Association between two things. The American Journal of Psychology, 1904;15:72–101.
Stigler G. J., 1958, *The Economies of Scale*, The Journal of Law and Economics, Vol. I.
Stevens P.A., 2004,*Accounting for background variables in stochastic frontier analysis,* National Institute of Economic and Social Research.
Taylor J. W., De Menezes L. M. & McSharry P. E., 2006, A Comparison of Univariate Methods for Forecasting Electricity Demand Up to a Day Ahead, International Journal of Forecasting, Vol. 22, pp. 1-16.
Thorp W.J., 1924, *The Integration of Industrial Operation*, Census Monograph No. 3, Washington: U.S. Government Printing Office.
Tobin J., 1958,*Estimation of Relationships for Limited Dependent Variables,*Econometrica Vol. 26:24-36.
Torres Reyna O., 2010, *Getting Started in Factor Analysis,* Princeton University Training, http://dss.princeton.edu/training/Factor.pdf
Tripathy I. G., Yadav S. S. and Sharma S. 2009, *Measuring the Efficiency of Pharmaceutical Firms in India: An Application of Data Envelopment Analysis and Tobit Estimation,* Working paper of the Research Unit for Statistical and empirical analysis in social sciences.
Uğur A., 2004, *Technical Efficiency in Irish Manufacturing Industry, 1991-1999, in Trinity Economic Papers* – 4/2004, Trinity College Dublin.
Vartanian T. P. & Buck P.W., 2005, Childhood and Adolescent Neighborhood Effects on Adult Income: Using Siblings to Examine Differences in OLS and Fixed Effect Models, in *Social Service Review,* 78 (1), March, 60-94.
Weiss, L. R., 1964, *The Survival Technique and the Extent of Suboptimal Capacity*, Journal of Political Economy, LXXII, No.3: 246-261.

[i] Any views or opinions presented in this paper are solely those of the author and do not represent those of the affiliated institution. Detailed results of the tests will be provided on request. Special thanks to: Prof. Don Webber and Prof. Andrew Mearman for their constant advice during the research project at Bristol Business School, Bruna Ceci for helping me in gathering data, Dr. Maria Rita Pierleoni for her intuition to apply a two-stage DEA with Tobit to the postal sector and for her technical support during the whole research project at Bristol Business School, Dr. Meloria Meschi for the discussion session in Portmarnock's conference in 2013 on DEA and for the advice to consider using TFE and TRE in future research, Marzia Padrone for helping me in building a solid bibliography, Gennaro Scarfiglieri for the discussion at a conference in Brighton, UK in 2012 and the idea to use the Spearrman method to benchmark efficiency rankings emerging from the use of econometric methods, Jean Philippe Ducasse for the discussion session in Portmarnock's conference in 2013 on the Survivor technique, Dr. Emiliano Piccinin for the technical support during the research on the Principal Component Analysis, Prof. Vincenzo Visco Comandini for the joint research especially on the

impact of the Altmark ruling on the postal sector and Mark Fardelli, Darius Anderson, Ethan Rarick, Marisol Lopez, Erik Maurice for their comments and their support during the research project in the past few years.