ASOCIACION ARGENTINA
DE ECONOMIA POLITICA

ANALES | ASOCIACION ARGENTINA DE ECONOMIA POLITICA

LIII Reunión Anual

Noviembre de 2018

ISSN 1852-0022
ISBN 978-987-28590-6-0

Efficiency in Higher Education Teaching Activities: A Stochastic Frontier Analysis

Ferro Gustavo
D'Elía Vanesa

Efficiency in Higher Education Teaching Activities: A Stochastic Frontier Analysis

Vanesa D'Elía[1] and Gustavo Ferro[23]

Abstract

We aim to study technical efficiency of undergraduate teaching activity in national universities in Argentina. We surveyed the literature seeking the theoretical and applied consensus, and we estimate a production frontier for the sector. We employ a stochastic frontier analysis (SFA) with a panel of 37 national universities over a 9 year period. We compare conventional models that do not take into account heterogeneity with heterogeneity-extended stochastic frontier models.

We find 20% to 25% of inefficiency on average in terms of lost outcomes (graduates). Besides, as expected, those models that take into account unobserved heterogeneity present higher values of technical efficiency while the random effect model presents the lowest value. This result indicates that there is evidence of heterogeneity that overestimates inefficiency. Our results are robust to different model specifications.

We do not address extension activities of the universities, which are usually difficult to measure but use resources of the system. The results could be useful to devise some remedial measures to improve efficiency. To our knowledge is the first study using this method to analyze universities' efficiency in Argentina.

Resumen

Nuestra meta es estudiar la eficiencia técnica en la actividad docente universitaria de grado en las universidades nacionales de Argentina. Efectuamos una revisión de la literatura buscando los consensos teóricos y aplicados y estimamos una función de producción para el sector. Luego empleamos un modelo de análisis de frontera estocástica (SFA) con un panel de 37 universidades nacionales sobre un período de 9 años. Comparamos modelos que no tienen en cuenta heterogeneidad con aquellos que extienden los modelos estocásticos de frontera para contemplarla.

Encontramos en promedio valores de entre 20 y 25% de ineficiencia en términos de producción (graduados) perdida. Los modelos que toman en cuenta la heterogeneidad no observada presentan valores mayores de eficiencia técnica que los modelos de efectos aleatorios, como podría esperarse. Tal resultado indica evidencia de heterogeneidad no observada que sobreestima la ineficiencia. Nuestros resultados son robustos a diversas especificaciones del modelo.

No contemplamos actividades de extensión como producto de las universidades, que son normalmente muy difíciles de medir, pero que usan recursos de aquellas. Los resultados pueden ser útiles para el diseño de medidas que busquen aumentar la eficiencia. A nuestro conocimiento, es el primer estudio usando este método para analizar eficiencia de universidades argentinas.

Keywords: Argentina, Public Universities, Efficiency, Stochastic Frontier Analysis, Heterogeneity

JEL Codes: I23, C23

[1] Universidad del CEMA (UCEMA). vvd04@ucema.edu.ar

[2] Universidad del CEMA (UCEMA) and CONICET. Córdoba 374, (1054) Buenos Aires, Argentina. Telephone: +54-11-63143000, extension 780, email: gaf37@ucema.edu.ar and gferro05@yahoo.com.ar. Corresponding Author.

[3] The views and opinions expressed in this publication are those of the authors and do not necessarily represent those of UCEMA.

Efficiency in Universities Teaching Activities: A Stochastic Frontier Analysis

1. Introduction

Higher education is increasing in recent years. The average gross enrollment rate in higher education (within population ages 18-24) in Latin America and the Caribbean grew from 21 to 43 percent between 2000 and 2013 (54 percent in Argentina in 2015). The expansion happened at a time of economic growth, abundance of fiscal resources and a middle class rise, and was concentrated in the low- and middle-income segments. Completion rates in the Latin American region are closed to 50 percent (compared to around 65 percent in the United States). Lower income and lower ability students are more likely to drop out. Time to degree in Latin America is 136% of the stipulated time (the former is similar in the United States, but the latter is typically longer in Latin America than in the United States). Student/faculty ratio is similar in Latin America and comparable countries of Eastern Europe or Eastern Asia, but salaries represent a higher proportion with respect to total expenses in Latin America than in comparable countries (Ferreyra et al.., (2017).

There are some market failures in higher education which call for intervention to put the results closer to social optimum (maximizing students' potential and meeting the skills society demands):

1) Externalities since knowledge creation and dissemination spill over.
2) Liquidity constraints where talented but poor students can be affected.
3) Information asymmetries, being difficult to assess the quality and variety of offerings, with the potential danger of making wrong or suboptimal decisions. The cost of a wrong decision increases with the difficulty of moving across different programs.
4) Differentiated products (in location, programs, student peer ability, focus, rigor and expectations) and a certain degree of market power.
5) Individual decisions have a certain level of irreversibility, once years, efforts and resources had been sunk in a bad choice.

Externalities justify some government subsidies; liquidity constraints call for subsidies or for student loans; information-related problems call for information provision and consumer protection; imperfect competition calls for monitoring and regulating choice and mobility of students (Ferreyra et al.., 2017).

In Argentina, Law 24521 of superior education established a National State responsibility in financial supporting national universities, and does so by funding them from federal revenue. National universities have economic and financial autonomy and they have to manage their resources and approve their budget, with considerable freedom in their activities. In the country there are no standardized exams to finish high school, to access to the university or to conclude university degrees; no tuitions are applied in national universities, which are completely free (Coria, 2008). There were 57 national universities, 49 private ones (all non-for profit), 4 provincial and 1 foreign in 2016. There is consistent information on the former. In 2015 there were 1.9 million undergraduate students in Argentine universities, of which 1.49 million in public universities. The same year, there were 458,000 new entrants to the system and 125,000 degreed. Between 2006 and 2015, students increase in 19 percent, new entrants in 27 percent and degreed in 47 percent. Only 27 percent of the degreed completed their studies in the established time in public universities, while 37 percent dropped-out their studies during the first year. All data from Argentina is originated in the University Information Departament (DIU) of the Secretary of University Policies (SPU) of the Ministry of Education.

The system is characterized by high university dropout rates, long duration of studies, low graduation rates and (in some disciplines) difficulty in finding a suitable job for graduates in accordance with the level of education achieved. Colomé (1996) highlights the scarce

academic complementary of institutions within the system, being the student mobility more an administrative matter tan an academic issue. Among national universities situated at short distances, there is overlapping in degrees offered.

In the whole world, and with different intensities, universities face resources' constrain, because of the explosion in enrollment rates, and are being forced to consider raising revenues and increasing productivity, the latter as a long-term response which demands knowing more on the outcome process and efficiency measurement (Gates and Stone, 1997).

Thus, our objective is twofold: first we aim to explore the literature in order to model reasonably the outcome process and second to estimate comparative efficiency for undergraduate teaching activity in national universities in Argentina. The results could be useful to devise some remedial measures to improve efficiency.

The paper organizes as follows: Section 2 accounts for the literature review on universities' efficiency; the Section 3 explains the methodology employed; the Section 4 shows the different models that were estimated, the results and inferences; and finally, the Section 5 presents the conclusions of the study.

2. Literature Review

According Avkiran (2001) universities key characteristics are "the lack of profit motive, goal diversity and uncertainty, diffuse decision making, and poorly understood production technology". Productivity improvement is thence a poorly understood concept in the higher education context, and an issue viewed with hostility by insiders. It is often associated with quality-insensitive cost cutting or attempts to increase administrative efficiency (Gates and Stone, 1997). Information technology is another promise of improving productivity, with dubious results to date.

The outcomes (or outputs) produced in a university can be split in teaching (knowledge dissemination), research (knowledge production; basic and/or applied, where basic research implies learning the principles which explain phenomena, without developing an immediate application), and extension / transfer / public or "third mission" services (in the form of externalities and public goods directed to varied audiences beyond campuses) (Cohn and Cooper, 2004, Johnes and Johnes, 2009).

In general, choices for empirical work are conditioned by data availability and by the possibility of measuring the phenomena under study. Most analysis omits extension activities for the difficulty in quantifying meaningfully its outcomes. In general, externalities are difficult to measure, thus also some teaching or research outcomes cannot be properly measured (as learning by interaction in peer groups by the students, or discoveries not subject to intellectual property rights). Nearly all efficiency studies of higher education institutions focus only in teaching, research of both products of universities, leaving aside extension services (Salerno, 2003).

The so called Humboldt model of universities is characterized by the coexistence of teaching and research, on the one hand, and of many disciplines within the same institutional umbrella, on the other hand (Daraio et al., 2015). The coexistence of activities can be explained by economies of scope, due to the presence of non-specialized and indivisible inputs, and thence susceptible of being shared. In turn, different production technologies exist among academic disciplines, with varying degrees of labor/capital intensity but administrative and supporting expenses can be shared.

Thus, connected activities are produced with the same inputs: human and non-human resources. There are also complex interactions of substitution and complementarity between teaching and research, because once recognizing the potential synergies, both teaching and research are time consuming and the short term and long term rewards are different between

them (Kuo and Ho, 2008). If research activities are not included, implicitly, it is assumed that no scope exists for economies between teaching and research activities (Horne and Ho, 2008). The other side of scope economies is economies of specialization, which arise when the cost of producing a specific good by a specialized decision unit is lower than the cost of the same good made by a decision unit which produces together two or more goods (Daraio et al., 2015).

Productivity is a measure of output(s) per unit of input(s). In the higher education context, productivity can be thought as how much outcomes individuals and society are getting from the sector, given the resources they put in. Some features of the sector, shared with other service industry, difficult productivity measurement, including the identification of the output unit, determining the value added, isolating the consumer own contribution to the outcome (for example, personal effort devoted to study), and accounting for quality respects. The standard practice in defining output is to focus only in those that are used outside the organization under consideration (Gates and Stone, 1997, Salerno, 2003).

Teaching is a complex process, whose technology is yet poorly understood. It is subject to economies of scale since expanding the number of students attending to a lecture increases the output while keeping constant the input (teacher effort). Teaching also requires one-to-one interaction with students, such as examinations and tutoring, for which costs are roughly proportional to the output. The exact combination between these two opposite forces is responsible for the overall effect and jointly determines the optimum scale of production. Research is an even less understood production process, for which the arguments for economies of scale are mostly linked to indivisibilities in cognitive capital (minimum scale of research teams), and above all in physical capital (laboratories) when it applies (Daraio et al., 2015).

The absence of profit motivation, combined with the diversity of goals pursued, makes the measurement of efficiency in higher education particularly problematic (Johnes, 2006). The major problems found in educational efficiency studies are: to obtaining data to specify adequate input measures, specify outputs, and communicating the results (Worthington, 2001). Some characteristics of universities complicate the evaluation of efficiency. They have multiple objectives and outcomes, some of them with externalities or public good characteristics. There are often conflicting opinions regarding the goals, and their relative importance. Many of the inputs of an educational organization cannot be unambiguously measured or quantified. Still other educational outcomes such as socialization defy parametrization (Worthington, 2001).

Productivity measures are rank-free indicators of the rate at which inputs are translated into outputs. Efficiency can be seen as an index to rank the different productivity values. Technical efficiency captures the extent to which physical inputs are efficiently allocated. Technical efficiency is defined as the ability to minimize input use for a given output. Unlike cost efficiency, technical efficiency does not involve the imposition of behavioral assumptions such as cost minimization (Horne and Hu, 2008). Allocative or price efficiency measures the extent to which inefficiency occurs because a university is using suboptimal combination of inputs given what they cost to purchase. Economic or overall efficiency, jointly considers technical and allocative efficiencies. If the university is allocative and technically efficient, hence it is overall efficient. Which inputs are considered depends on what type of efficiency is being assessed: technical efficiency estimates routinely employ physical input units while cost efficiency uses expenditure-based units (Salerno, 2003).

Universities are subject to two phenomena which affect their productivity and efficiency, labeled in the literature as academic ratchet and administrative lattice (accretion). Academic ratchet describes the tendency of faculty (over time) to shift effort away from student teaching, advising and counseling, towards administrative governance tasks, research and personal income opportunities. Administrative lattice or accretion describes the tendency for the number of administrative staff to increase faster than faculty. This faster rate

of growth of administrative personnel is seen partly as a reaction to academic ratchet, but also because of the self-perpetuating growth tendency of bureaucracies (Mensah and Werner, 2003).

In the university activities the quality of the outcomes is closely related with the quantity and intensity of human effort devoted to that task. The resulting impediment to labor-saving changes (shared with many other services' industries) condemns the costs of those effort-intensive services to keep rising year after year, in comparison with goods whose growing automation declines unit costs. The most recent source of change in production activity in the universities is the increased use of information technology and e-learning. In teaching, they may allow class sizes to increase, but an increasing ratio of students to staff may have an adverse effect on technical efficiency (Johnes, 2007). As Wolff et al. (2014) state, the productivity cannot be enhanced while maintaining quality, by putting students on an assembly line or substituting robots for teachers, or making class sizes larger and larger. It is not clear if technology-based approaches will yield educational results that match those of current educational methods. It is also not clear if online learning will deliver permanent cost reductions. It is not yet clear how online instruction affects the learning and career opportunities of those student participants who are not physically connected with a sponsoring university (Wolff et al., 2014)

All the existent studies vary in the precise definitions of the variables used to reflect inputs and outputs. Most conclude that inputs can be grouped as student inputs, staff inputs and capital inputs, while outputs can be divided into teaching and research output. Some studies have focused on the efficiency of universities at producing either teaching only or research only, while others have attempted to measure efficiency in the joint production of the two outputs (Johnes and Yu, 2008). The Table 1 presents the outputs of universities grouped in teaching, research and extension.

Teaching is the delivering of educational services, which implies human capital accumulation, including both knowledge, competencies and skills. It is proxied as the number of degrees completed (which is the more precise measure even when it may underestimate the outcomes, because of drop outs), results in standardized exams (recalling that student´s grade is a complex function of the student´s entry-level ability, the marking standards of the university, as well as teaching and supervision quality), students (which is a better indication of an input, the "raw material" of the process), courses/hours taught, job or remuneration attainment by degreed, admission to graduate studies and byproducts (such as training to non-degreed students).

Research is the new knowledge development and accumulation. Research output can be proxied by published products, citation indexes, Ph.D. awarded, patents and other intellectual property issues. Some research outcomes are not ex-ante observable or ex-post measurable. A large unobserved research effort may well lead to no results, and conversely, given that "serendipity and luck may yield huge returns at little cost", thus it is very difficult to infer and reward effort from the results obtained (De Fraja and Valbonesi (2012). Number of publications, often interpreted as a measure of research activity, suffers from the problem of different practices across disciplines. The number of citations attracted by an article has its drawbacks as well; for example, articles in mathematics have a longer shelf life than those in pharmacy (Avkiran, 2001). The means for estimating the value of research output of universities has proved to be a controversial topic. Not only is it necessary to capture the quantity of output, which can be quite varied, and to attribute some weight, but also the quality of the work must be accounted for. An alternative to using some weighted-index of publications is to use external research finance attracted by a university as a proxy for research output. However, grants are spent not only on research assistance but also on other facilities which are input into production. To include grants to measure both, outputs and inputs, is double counting. Another proxy for research output is research income spent during the year if money is granted on the basis of the research output records. In that case, there is correlation between expenditure and research output (Abbott and Douccouliagos,

2003). Other problems are that the reputation of a whole university can benefit from a strong and very public activity of just one small group within that university (externality effects); and the measurement may be based on past or accumulated activity and hence may not be an adequate reflection of current research outcome (Johnes, and Yu, 2008).

Extension (also called Transfer / Public Services or Third Mission), consists in the generation of public goods and externalities (with possible but difficult to measure rewards in terms of publicity and prestige, in tuition values or supporting fundraising activities). Universities have also the responsibility of storage and preservation of knowledge. It includes cultural and sports activities, non-formal education for the elder and other collectives, informed opinion and advice on social or community issues and the even more difficult to measure, building of (desirable) social values and citizenship.

Table 2 characterizes the resources (inputs) of university education, which can be classified in human and non-human. The former includes labor and “raw materials”, and the latter encompasses hardware. Human resources are academic and non-academic staff, while “raw materials” of the process are the students to be taught. To address the possible substitution between teaching and research activity, it can be calculated the ratio of research teachers (or research workload) on total faculty. Alternatively, salaries paid to different categories of staff, is another form of recognizing the human inputs in efficiency estimates. Non-human resources include capital goods and materials, which can be measured in physical units (square meters of laboratories or classrooms, classroom seats, computers, books in libraries, etcetera) or money expenditure in hardware.

Table 3 presents quality variables used in university efficiency studies. Quality can be assessed either in outcomes or in resources, by different ratios or dummy variables. They are normally indications of completion, achievements and recognition (given duration, structure and contents of the programs, time dedication and qualification of the staff). It can be also applied to expenditures. The process can be addressed by the technology in use, for example by establishing a ratio between on line to off line students. The quality of the staff input is reflected by faculty proportion with professor status, and/or full time on part time ratio. The premise underlying this variable is that the promoted /tenured faculty is more productive than their colleagues. Nevertheless, it may were promoted when research demands were less than nowadays, or, once promoted, these faculty have less motivation to continue to be that productive, thus the net effect is a priori uncertain (Johnes and Yu, 2008).

Table 1: Characterization of outputs/outcomes of university education

Outputs/Outcomes	Interpretation	Variable	Measurement
Teaching	Delivering of educational services, which implies human capital accumulation (knowledge, competencies and skills)	Degrees completed (discriminating undergraduate from graduate)	Degreed/new students in t-(theoretical duration of the program)
		Results in standardized tests	Possible if this tests are taken
		Head count of enrolled students	Standardized by full time equivalent Actually an input
		Hours / Courses /Credits taught	Proxy to the value added to student's knowledge basis
		Job attainment once graduated	Students employment potential
		Remuneration attainment once graduated	Students earnings potential Rate of return of the degree
		Admission to graduate studies	Head count
		Byproducts	Students which did not complete degrees, yet incorporated some qualifications
Research	New knowledge development and accumulation	Published products	Some weighted sum of articles, books, conference papers, etc. Problems of heterogeneity and ageing of the contributions
		Citation indexes	Impact of only part of the research outcomes
		Ph.D. awarded	Head count
		Patents granted	Number of patents
		Other intellectual property obtained / Awards	Number of registers. Difficult to weigh.
		Grants / Projects awarded	Money achieved. Actually an input
		Research project expenditures	Money spent. Actually an input
		Contract money / Partnership with business	Money awarded. Actually an input.
Extension (Transfer / Public / Community Services or Third Mission)	Generation of public goods and externalities. Storage and preservation of knowledge.	Citizenship development / Behavioral changes in students, including values and interpersonal skills	Hard to measure
		Cultural, sport and recreational activities open to the community	Number of events. Difficult to weigh.
		Technical / qualified / informed public opinion / advice in community issues or social debate	Number of events. Difficult to weigh.
		Non-formal education to social groups (such as the elder) / Disadvantaged collectives	Can be proxied by the same type of variables than in regular teaching

Source: own elaboration from Coria (2008), Salerno (2003), Worthington (2001), Laureti et al. (2014), Cantele et al. (2016), Johnes (2006), Katharaki and Katharakis (2010), Kuah and Wong (2011), Avkiran (2001), Zoghbi te al (2013), Cohn and Cooper (2004), Johnes and Jones (2009), Ferreyra et al. (2017), De Fraja and Valbonesi (2012), Kao and Hung (2008), Worthington and Lee (2008).

Table 2: Characterization of resources (inputs) of university education

Resources (Inputs)	Interpretation	Variable	Measurement
Human resources	Labor and "raw materials"	Academic staff	Full time equivalent faculty
		Researchers / Doctoral students ratio	Head count / Workload on total
		Students to be taught	Full time equivalent students who are submitted to a cultural transformation process
		Non-academic staff	Head count
		Salaries paid to each category of staff	Money spent
Non-human resources	Hardware: capital and other tangible items	Facilities	Lecture halls square meters Labs square meters Materials Bibliography items Computers Classroom seats
		Non-human expenditure	Money spent

Source: own elaboration from Johnes (1996), Worthington (2001), Coria (2008), Laureti et al. (2014), Avkiran (2001), Worthington and Lee (2003), Kuah and Wong (2011), Johnes and Yu (2008), Kao and Hung (2008), Johnes (1996), Cantele et al. (2016), Cao and Hung (2008).

Table 3: Quality variables in efficiency studies of university education

Quality	Interpretation	Variable	Measurement
Could be formed as ratios or dummies	Indications of completion, achievements and recognition	Drop-out rates	Proportion of drop outs on cohort
		Low-progress rate	Proportion of lagged students on cohort
		Student / Faculty ratio	Proportion on total
		Staff expenditure / Total expenditure	Proportion on total
		Professorship / Tenured academics ratio	Proportion of professors over certain status / seniority
		Index of full time faculty / Research, Teaching and Administrative Workload	Full time on total faculty
		On line / Off line students' ratio	Proportion

Source: own elaboration from Zogbi et al. (2013), Coria (2008), Ferreyra et al. (2017), Johnes and Yu (2008), Kuo and Ho (2003), Sav (2012) and Wolff et al. (2014).

Table 4 refers to environmental or contextual variables, which allow fair comparisons between institutions and permit address for observable heterogeneity. They are used for separating the effect of uncontrollable inputs. At least three groups of environmental variables can be distinguished: those referred to students' intellectual, economic, and social background, those referred to place of origin (poor or rich regions, both in term of GDP and human capital), ethnicity and gender, and those referred to the type of university (big or small, old or new, private or public, profit or non-for profit, laic or religious, specialized of generalist, mixed specialization among labor intensive or capital intensive disciplines).

Table 4: Environmental variables in efficiency studies of university education

Environmental (contextual) variables	Interpretation	Variable	Measurement
Used to separate the effect of uncontrollable inputs	Context (student and place characteristics, type of university)	Students' intellectual background	High school completion grades. Selection process exams results. Individual effort, innate ability, academic readiness, and peer externalities (learning by group interaction) are difficult to measure
		Parental socioeconomic index	Family per capita GDP
		Parental qualification index	Years of parents' educational training (or degrees attained)
		Index of full time students	Full time on total students Working students on total
		Gender composition of students	Proportion of female students
		Ethnic composition of students	Student minority proportion
		Internationalization of students	Foreign /out of region to domestic /regional origin rate
		Age of students	Average age of students /Students' age under certain value
		Region GDP	GDP of the region with respect to national average
		Region Human Capital	Years of average educational training of region population with respect to national average
		Size of the university	With respect to local context (history matters)
		Ownership of the university/System governance	Public or private If private, non-profit or for-profit Laic or religious
		Degree of specialization	Mixture of disciplines taught with different hardware intensity / Share of science, engineering and medicine on total.
		Age of the university	Older or newer with respect to local context

Source: own elaboration from Ferreyra et al. (2017), Laureti et al. (2014), Worthington (2001), Zoghbi et al. (2013), Sav (2012), Johnes (2006), Worthington (2001), Cantele et al. (2016), Daraio et al. (2015), Millot (2015), Cohn and Cooper (2004), McGukin and Winkler (2015), Johnes and Johnes (2009), Horne and Ho (2008).

3. Methodology and Data

There are two families of techniques for measuring efficiency: parametric or regression based and non-parametric, or mathematically programming estimators. Because each possesses its own strengths and limitations, neither is generally regarded to be superior to the other.

The non-parametric approach, being the most common the Data Envelopment Analysis DEA, characterize the set of efficient producers (those on the frontier) and then derive estimates of efficiency for inefficient observations based on how far they deviate from the most efficient ones. They are Total Factor Productivity measures which require finding some way to fairly assign weights, or importance, to the various inputs and outputs included. In DEA there is no need to draw assumptions about efficiency a priory or even to objective functions of the units under analysis (Salerno, 2003). The DEA-oriented studies tend to group institutions by type to address for observed heterogeneity while in the SFA studies all tend to account for it through the use of dummy variables. One of the main advantages of the SFA approach compared to the non-parametric techniques is that it easily accommodates the various environmental factors (Laureti et al., 2014). Meanwhile, the Stochastic Frontier Analysis SFA (or error-inclusive estimator) is very similar to traditional parametric regressions. In the former, the traditional random error term is divided into two components: a normally distributed random error term of pure randomness (with zero mean and positive variance) and an error term that capture inefficiency (strictly positive in the case of production, or strictly negative in the case of costs, with different possible statistical distributions). Once a distribution form is selected for the efficiency residual term, it is a relatively straightforward exercise to derive all of the estimates found in a traditional production / cost function and also determines for each observation the deviations from production / cost efficient (frontier) behavior (Salerno, 2003).

Greene (2008) distinguishes between observed and unobserved heterogeneity: the former is reflected in measured variables represented by various factors related to the organizational aspects of educational processes, students' characteristics and the socio-economic context in which the faculties are located.

Two categories of models in which observed heterogeneity enters in the form of parameterized functions of exogenous variables, are those which directly include the observed factors in the production function structure (production model) and those that consider them in the inefficiency term (inefficiency model). The issue of heterogeneity in the production model is due to different types of technology which may simultaneously coexist as the various faculties carry out their production in different environments. A frequently used method for solving this problem is to add exogenous shifts, environmental variables and technical changes to the production function. The more recent one-step approach by which exogenous effects are estimated, simultaneously estimates it with other parameters of the model (Laureti et al., 2014).

Unobserved heterogeneity is related to environmental factors that are unobserved but constant for each unit, and it enters into the SFA frontier model in the form of "effects" and is usually viewed as an issue of panel data. A SFA model already includes a random error term that captures the idiosyncratic heterogeneity among different production units. In the context of panel data in which a single unit is observed several times, the unit-specific unobserved variations can also be taken into account though fixed or random effects. Various models have been introduced for separating the two effects, including true fixed and true random effects, which only treat the time invariant components as unobserved heterogeneity and random coefficient models in which cross-unit parameter variation is viewed as heterogeneity in the technology being employed (Laureti et al., 2014).

Choosing a functional from for the production function is not straightforward in higher education as there is no rule upon which to base the decision. The Cobb-Douglas production function is appealing in that it involves the estimation of relatively few parameters thus

allowing an easy interpretation of results. The Trans-Logarithmic functional form may be preferred to the Cobb-Douglas form due to the restrictive elasticity of substitution and the scale properties of the latter. However, the quadratic and interaction terms in the Trans-Logarithmic are difficult to interpret (Laureti et al., 2014).

3.1 Stochastic Frontier Analysis

Prior to the stochastic frontier analysis, we examine the most appropriate functional form that relates the number of university graduates (Y_{it}) of university *i* at year *t* with the selected inputs (included in the vector D_i for dummy variables, X_{it} otherwise). A temporary variable τ is also included to capture the technological change over time. Assuming a Cobb Douglas form for the production function, the general specification of the model has the following form:

$$Y_{it}^{\theta} = \alpha + \beta X_{it}^{\lambda} + \delta D_i + \theta\tau + \varepsilon_{it} \quad (1)$$

where $Y_{it}^{\theta} = \frac{Y_{it}^{\theta}-1}{\theta}$ and $X_{it}^{\lambda} = \frac{X_{it}^{\lambda}-1}{\lambda}$ if $\theta, \lambda \neq 0$, respectively

and $Y_{it}^{\theta} = \log(Y_{it})$ if $\theta = 0, X_{it}^{\lambda} = \log(X_{it})$ if $\lambda = 0$.

Then, we compare the following pooled Ordinary Least Squared (OLS) specifications: linear (θ=λ=1), log-lin (θ=0, λ=1), lin-log (θ=1, λ=0) and log-log (θ=λ=0).

To select within the set of candidate models, we consider that the best model is the one with the minimum Bayesian Information Criteria (BIC) value. The BIC is defined as: BIC = − 2logL + log(N)K, where L is the likelihood of the model, and K is the number of parameters to be estimated in the model. The BIC rewards goodness of fit (as assessed by the likelihood function), and penalizes the number of estimated parameters.

With the selected specification, we estimate de stochastic production frontier of Eq. (1), where the residual ε_{it} is the difference of the normally distributed term v_{it} which represents the classical random error ($v_{it} \sim iid\ N(0, \sigma_v^2)$), and a one-sided disturbance u_{it} which represents inefficiency:

$$\varepsilon_{it} = v_{it} - u_{it} \quad (2)$$

It is also assumed that v_{it} and u_{it} are independent of each other and of the regressors, and identically distributed across observations. Finally, when estimating the SFA model via maximum likelihood (ML) or simulated maximum likelihood (SML) techniques, it is assumed the distribution of u_{it} is usually half normal, truncated normal, exponential or normal gamma.[4]

The main purpose of the SFA is the estimation of the technical efficiency (TE). If TE is defined as the ratio of observed output to the stochastic frontier output, then TE is: $ET_{it} = \exp(-u_{it})$.[5] However, the econometric procedure allows to estimate ε_{it}, so, we use the strategy proposed by Jondrow, Lovell, Materov and Schmidt, JLMS, (1982) to decompose

[4] In general, SFA is based in two sequential steps: in the first, estimates of the model parameters are obtained by maximizing the log-likelihood function. In the second step, point estimates of inefficiency can be obtained through the mean (or the mode) of the conditional distribution$f = (u_i \setminus \widehat{\varepsilon_{it}})$. For details, see Belotti, et al.(2012).

[5] $ET_{it} = \frac{f(X_{it}\beta)\exp(v_{it}-u_{it})}{f(X_{it}\beta)\exp(v_{it})} = \frac{1}{\exp(u_{it})} = \exp(-u_{it})$.

the inefficiency term from the compounded error. A point estimate of the inefficiencies can be obtained using the mean $E(u_{it} \setminus \varepsilon_{it})$:

$$E(u_{it} \setminus \varepsilon_{it}) = \frac{\sigma_* \, \phi(\frac{\mu_{*it}}{\sigma_*})}{\Phi(\frac{\mu_{*it}}{\sigma_*})} + \mu_{*it} \quad (3)$$

where $\mu_{*it} = \frac{-\sigma_u^2 \, \varepsilon_{it}}{\sigma_v^2 + \sigma_u^2}$ and $\sigma_*^2 = \frac{\sigma_u^2 \, \sigma_u^2}{\sigma_v^2 + \sigma_u^2}$

An issue that is worth considering is that the different universities within the country could have own specificities that are beyond the universities' authorities control and that affect their production. For example, different academic characteristics, regional differences in socio-economic characteristics of the students and teachers, and geographical features that prioritizes some fields over others (for example, Agronomics in regions whose productive activity is more linked to agriculture). If this heterogeneity is not taken into account, the efficiency estimates could be significantly biased. Then, we concentrate in models that take into account individual specific observed and unobserved heterogeneity.

3.1.1 Observed heterogeneity

Random Effect models (RE and REH)

We perform the random effect (RE) model proposed by Pitt and Lee (1981), which is specified as:

$$Y_{it}^{\theta} = \alpha + \beta X_{it}^{\lambda} + \delta D_i + \theta \tau + \varepsilon_{it} \qquad \text{i = 1,…,N, t=1…T} \quad (4)$$

$$\varepsilon_{it} = v_{it} - u_i$$

$$v_{it} \sim \mathrm{N}(0, \sigma_v^2)$$

$$u_i \sim \mathrm{N}^{+}(0, \sigma_u^2)$$

Or:

$$Y_{it}^{\theta} = \alpha_i + \beta X_{it}^{\lambda} + \delta D_i + \theta \tau + v_{it} \qquad \text{i =1,…,N, t=1…T} \quad (5)$$

where $\alpha_i \equiv \alpha - u_i$

This model, fits by ML, assumes that technical inefficiency is half normal, and that, in proportional terms, it is constant over time. As can be seen, individual heterogeneity cannot be distinguished from inefficiency. Then, the whole time invariant heterogeneity is confounded into inefficiency, and therefore, $\widehat{u_i}$ might be picking up heterogeneity in addition to, or even instead of, inefficiency (Greene, 2005b). Besides, Kumbhakar and Hjalmarsson (1993) argued that the time varying component v_{it} would inappropriately capture time varying inefficiency.

This model assumes that inefficiency is not correlated with the regressors, and this is a problem when, in the presence of heterogeneity, the inefficiency term is not independent of the frontier inputs. One way to deal with this issue is to introduce the exogenous variables in the location of the distribution of the inefficiency. The most common approach is to parameterize the mean of the pre-truncated inefficiency distribution.[6]

This model specification, called here random effect with observed heterogeneity (REH) model can be written as:

$$Y_{it}^{\theta} = \alpha + \beta X_{it}^{\lambda} + \delta D_i + \theta\tau + \varepsilon_{it} \qquad \text{i=1,…,N, t=1…T} \qquad (6)$$

$$\varepsilon_{it} = v_{it} - u_i$$
$$v_{it} \sim \mathrm{N}(0, \sigma_v^2)$$
$$u_i \sim \mathrm{N}^{+}(\mu_i, \sigma_u^2)$$
$$\mu_i = z_0 + z_1 h_i$$

It is important to note that unobserved heterogeneity still remains in the inefficiency term. To overcome these limitations, Greene (2005a) proposes the following models which deals with unobserved heterogeneity.

3.1.2 Unobserved heterogeneity

True Fixed Effects Model

To overcome the problem of the previous models where inefficiency cannot be separated from time-invariant individual effects, Greene (2005a) proposes the following model, called the true fixed effects model (TFE), where the university specific constant term is included in the stochastic frontier:

$$Y_{it}^{\theta} = \alpha_i + \beta X_{it}^{\lambda} + \delta D_i + \theta\tau + \varepsilon_{it} \qquad \text{i=1,…,N, t=1…T} \qquad (7)$$

$$\varepsilon_{it} = v_{it} - u_{it}$$
$$v_{it} \sim \mathrm{N}(0, \sigma_v^2)$$
$$u_{it} \sim \mathrm{N}^{+}(0, \sigma_u^2)$$

This model allows controlling for constant individual unobserved heterogeneity and assumes that inefficiency is time varying, meaning that there is not structural inefficiency that persists over the period of analysis. The model is fit by ML and the normal-half normal model is applied to the stochastic part of the regression.

True Random Effects Model

[6] An alternative approach to analyzing the effect of exogenous variables on inefficiency is to rescale its distribution allowing observable variation in σ_{ui}^2. See for example, Caudill and Ford (1993), Caudill, Ford and Gropper (1995), and Hadri (1999).

The true random effect model (TRE) proposed by Greene (2005a) is an extension of the previous RE model and has the following structure:

$$Y_{it}^{\theta} = \alpha + w_i + \beta X_{it}^{\lambda} + \delta D_i + \theta\tau + \varepsilon_{it} \qquad \text{i=1,…,N, t=1…T} \qquad (8)$$

$$\varepsilon_{it} = v_{it} - u_{it}$$
$$v_{it} \sim \mathrm{N}(0, \sigma_v^2)$$
$$u_{it} \sim \mathrm{N}^{+}(0, \sigma_u^2)$$

As can be seen, it is a model with a traditional random effect w_i but with the additional feature that the time varying disturbance is not normally distributed. The model is estimated by maximum simulated likelihood. It is assumed that $w_i \sim \mathrm{N}(0, \theta^2)$, where θ represents the standard deviation of the unobserved heterogeneity, and $E(w_i \setminus X_{jit}) = 0$, where j is the number of independent variables.

Compared with the Pitt and Lee (1981) model where u_i is the inefficiency, in the TRE, u_{it} is the inefficiency, and it is time varying. Moreover, whereas in the RE model the inefficiency term all also contains all other time invariant unmeasured sources of heterogeneity, in the TRE model these effects appear in w_i and u_{it} picks up the inefficiency.

3.1.3 Observed and unobserved heterogeneity combined

To take into account observed and unobserved heterogeneity, it is possible to combine the previous RE and TRE models as follows:

$$Y_{it}^{\theta} = \alpha + w_i + \beta X_{it}^{\lambda} + \delta D_i + \theta\tau + \varepsilon_{it} \qquad \text{i=1,…,N, t=1…T} \qquad (9)$$

$$\varepsilon_{it} = v_{it} - u_{it}$$
$$v_{it} \sim \mathrm{N}(0, \sigma_v^2)$$
$$u_i \sim \mathrm{N}^{+}(\mu_i, \sigma_u^2)$$
$$\mu_i = z_0 + z_1 h_i$$
$$w_i \sim \mathrm{N}(0, \sigma_w^2)$$

With this specification, observed heterogeneity is included in the model by parametrizing the mean of the inefficiency distribution, and the unobserved heterogeneity (which still remains in the inefficiency term in the RE model), is taken into account by introducing the specific random effects and by assuming that the inefficiency component varies across individuals and time.

3.2 Data

We work with the "*Anuarios de Estadísticas Universitarias"* of the Ministry of Education of Argentina. We use a balanced panel of 37 National Universities in Argentina with information over the period 2005-2013. The year of foundation was the main criterion followed to choose the Universities of the panel, together with the availability of consistent and continuous information (see Table 5).[7] New Universities were not chosen since the time

[7] The National Universities that were not considered in the panel are: Instituto Universitario del Arte

to “generate” graduates is not long enough, and in a couple of cases the information was discontinuous and the university had been dropped. There is available information prior to 2005, but it is not completely compatible in some variables with that produced after that date. For those reasons, we arrive to the current composition of the database.

The output variable is measured as the number of graduates. This variable is defined as the number of degreed students who completed all courses and other academic requirements. Graduates of undergraduate and pre-graduate courses are included.

The inputs that we consider are students, human resources (faculty), financial resources, and environmental variables. We include the number of students defined as all the people who sign up for a career and who were assigned a number of file. Pre-graduate and undergraduate students are included. As human resources, we incorporate the number of teachers, taking into account their role (professor or assistant) and their teaching time dedication (exclusive, semi-exclusive or simple). We calculate a professor variable (Professor_w) and an assistant variable (Assistants_w) that weights the role with the time dedication. Following the definition of the Ministry of Education for the conversion of the teaching roles into the equivalent exclusive position, the weights are 1 for exclusive teachers, 0.5 for semi-exclusive teachers and 0.25 for simple teachers. Then, the teaching roles (professor or assistant) are converted into equivalent exclusive time dedication roles.

Concerning financial resources, we consider total annual expenditures (i.e. costs of personnel, goods and services and monetary transfers) in constant terms (base 2013=100). [8] Personnel expenses refer to the remuneration of personal services, employer contributions, family allowances, extraordinary services and social benefits received by the agents, while the rest of the costs are the operating expenditures.

As the local context influences the development of the educational process in universities, we include dummy variables for the country’s five regions defined by the *Instituto Nacional de Estadísticas y Censos, INDEC* (Gran Buenos Aires –GBA-, Noroeste Argentino –NOA-, Nordeste Argentino -NEA-, Cuyo, Pampeana and Patagonia) to capture this effect. Finally, we include a measure of efficiency in the efficiency equation, defined as the ratio number of students of university *i* in period *t* divided into the maximum number of students of university *i* over the period 2005-2013 (rprodst).

The Table 5 presents a description of these variables by university. On average, the set of universities considered had 1.29 million students, and graduated 66,371 per year, with 25,791 full-time equivalent professors and 28,393 full time equivalent assistants. Considering a nominal length of five years on average for each degree, the rate of degreed students is 25.7 percent (66,371/258,233) on average in the years of the sample. Each year, the financial resources employed by the sample of universities averaged 42,633 million of 2013 pesos (AR$, equivalent to 7,909 million dollars at July 1st, 2013 price[9]). Thence on average each degreed demanded AR$ 642,000 (or US$ 119,000) of financial resources per year.

(1996), Noroeste de la Provincia de Buenos Aires (2002), Chaco Austral (2007), Río Negro (2008), José C. Paz (2009), Moreno (2009), Oeste (2009), Villa Mercedes (2009), Arturo Jauretche (2010), Avellaneda (2010) and Tierra del Fuego (2010).

[8] To deflate the expenditure series, we use the Cost of Construction Index calculated by the Instituto Nacional de Estadística y Censos (INDEC). We do not use the Consumer Price Index, since for the period analyzed, the information was not reliable (in fact, by Decree 55/2016, it was declared the statistic emergency). In the years prior to the unreliable data (before 2007), Cost of Construction Index exhibited a very high correlation with Consumer Price Index.

[9] The Price was AR$ 4.93 = US$ 1 in January 2nd, AR$ 5.39 in July 1st and AR$ 6.53 in December 30th. The July 1st is a reasonable mean of the whole year average Price.

Table 5: Resources and outcomes. Average per university (2003-2013)

University	Graduates	Students	Exclusive	S-exclusive	Simple	Professors_w	Assistants_w	Expenditures	rprodst
Buenos Aires *(1821), GBA*	16,809	322,535	2,392	2,871	18,786	3,025	5,499	13,500	0.90
Catamarca *(1972), NOA*	388	12,379	306	402	385	443	160	811	0.94
Centro PBA *(1974), Pampeana*	585	12,425	563	386	854	491	478	1,050	0.90
Chilecito *(2004), NOA*	48	3,066	44	118	161	99	44	166	0.52
Comahue *(1971), Patagonia*	957	26,997	587	748	1,348	597	701	1,440	0.87
Cuyo *(1939), Cuyo*	2,305	31,322	661	1,942	1,695	1,177	879	2,300	0.99
Córdoba *(1613), Pampeana*	6,875	106,413	1,371	3,275	4,334	1,973	2,118	4,340	0.96
Entre Ríos *(1973), Pampeana*	913	13,043	190	976	1,054	504	438	467	0.88
Formosa *(1988), NEA*	563	11,691	97	331	730	280	165	249	0.92
Gral. Sarmiento *(1993), GBA*	196	5,145	234	88	0	137	141	212	0.72
Jujuy *(1972), NOA*	184	12,864	208	499	403	275	283	510	0.87
La Matanza *(1989), GBA*	1,308	31,184	448	738	444	387	541	811	0.86
La Pampa *(1958), Pampeana*	410	8,953	261	345	1,078	334	369	435	0.97
La Plata *(1897), Pampeana*	4,953	99,699	1,507	2,138	8,911	1,833	2,971	2,760	0.89
La Rioja *(1994), NOA*	603	24,034	54	908	727	405	285	367	0.75
Lanús *(1995), GBA*	472	10,355	38	159	256	144	38	207	0.81
Litoral *(1919), Pampeana*	1,664	39,184	649	1,083	1,878	767	894	1,050	0.89
Lomas de Zamora *(1972), GBA*	2,699	34,735	104	138	2,711	504	346	408	0.93
Luján *(1972), GBA*	895	16,889	286	418	840	326	380	429	0.92
Mar del Plata *(1975), Pampeana*	1,168	23,335	673	642	2,701	706	963	769	0.92
Misiones *(1973), NEA*	686	20,733	290	457	748	379	327	553	0.89
Nordeste *(1956), NEA*	2,912	50,934	536	199	3,424	623	869	920	0.95
Patagonia Austral *(1994), Patagonia*	119	6,959	158	293	491	224	203	297	0.77
Patagonia S. J. Bosco *(1980), Patagonia*	417	13,300	227	566	1,560	472	428	556	0.92
Quilmes *(1989), GBA*	985	15,020	233	116	262	220	136	168	0.67
Rosario *(1968), Pampeana*	5,686	73,227	1,056	2,061	4,991	1,412	1,922	1,230	0.98
Río Cuarto *(1971), Pampeana*	906	16,643	764	576	330	551	584	366	0.87
Salta *(1972), NOA*	468	24,087	429	767	410	388	527	394	0.86
San Juan *(1973), Cuyo*	582	19,458	868	1,038	1,143	1,141	532	493	0.94
San Luis *(1973), Cuyo*	566	12,703	757	444	154	515	502	408	0.93
San Martín *(1992), GBA*	754	10,862	96	225	733	281	111	551	0.84
Santiago del Estero *(1973), NEA*	708	14,001	331	251	203	285	223	181	0.86
Sur *(1956), Patagonia*	1,032	19,493	614	275	1,334	519	566	387	0.94
Tecnológica Nacional *(1948), Pampeana*	4,071	73,144	542	819	17,216	3,009	2,247	2,520	0.88
Tres de Febrero *(1995), GBA*	237	8,936	90	136	145	157	37	247	0.72
Tucumán *(1912), NOA*	2,096	61,418	1,315	2,191	691	1,254	1,329	1,010	0.97
Villa María *(1996), Pampeana*	152	4,004	114	285	143	135	157	71	0.66
Total	66,371	1,291,169	19,096	28,902	83,273	25,971	28,393	42,633	

Source: Own elaboration on “Anuarios de Estadísticas Universitarias”

In the Table 6 we derive some partial productivity and average cost indicators. The degree rate is presented, showing high dispersion into universities, as well as classroom average size, and average expenditures.

Table 6: Partial productivity and average costs ratios

University	Graduates / (0.2*Students)	Students / (Professors_w + Assistants)	Graduates / (Professors_w + Assistants)	Expenditures / Graduates (Million AR$)	Expenditures / Students (Million AR$)
Buenos Aires (1821), GBA	0 .26	37 .84	1 .97	0 .80	0 .04
Catamarca (1972), NOA	0 .16	20 .53	0 .64	2 .09	0 .07
Centro PBA (1974), Pampeana	0 .24	12 .82	0 .60	1 .79	0 .08
Chilecito (2004), NOA	0 .08	21 .45	0 .34	3 .43	0 .05
Comahue (1971), Patagonia	0 .18	20 .79	0 .74	1 .51	0 .05
Cuyo (1939), Cuyo	0 .37	15 .24	1 .12	1 .00	0 .07
Córdoba (1613), Pampeana	0 .32	26 .01	1 .68	0 .63	0 .04
Entre Ríos (1973), Pampeana	0 .35	13 .86	0 .97	0 .51	0 .04
Formosa (1988), NEA	0 .24	26 .28	1 .27	0 .44	0 .02
Gral. Sarmiento (1993), GBA	0 .19	18 .49	0 .70	1 .08	0 .04
Jujuy (1972), NOA	0 .07	23 .06	0 .33	2 .77	0 .04
La Matanza (1989), GBA	0 .21	33 .60	1 .41	0 .62	0 .03
La Pampa (1958), Pampeana	0 .23	12 .74	0 .58	1 .06	0 .05
La Plata (1897), Pampeana	0 .25	20 .75	1 .03	0 .56	0 .03
La Rioja (1994), NOA	0 .13	34 .83	0 .87	0 .61	0 .02
Lanús (1995), GBA	0 .23	57 .01	2 .60	0 .44	0 .02
Litoral (1919), Pampeana	0 .21	23 .60	1 .00	0 .63	0 .03
Lomas de Zamora (1972), GBA	0 .39	40 .84	3 .17	0 .15	0 .01
Luján (1972), GBA	0 .27	23 .95	1 .27	0 .48	0 .03
Mar del Plata (1975), Pampeana	0 .25	13 .98	0 .70	0 .66	0 .03
Misiones (1973), NEA	0 .17	29 .38	0 .97	0 .81	0 .03
Nordeste (1956), NEA	0 .29	34 .15	1 .95	0 .32	0 .02
Patagonia Austral (1994), Patagonia	0 .09	16 .28	0 .28	2 .49	0 .04
Patagonia S. J. Bosco (1980), Patagonia	0 .16	14 .78	0 .46	1 .33	0 .04
Quilmes (1989), GBA	0 .33	42 .11	2 .76	0 .17	0 .01
Rosario (1968), Pampeana	0 .39	21 .96	1 .71	0 .22	0 .02
Río Cuarto (1971), Pampeana	0 .27	14 .67	0 .80	0 .40	0 .02
Salta (1972), NOA	0 .10	26 .33	0 .51	0 .84	0 .02
San Juan (1973), Cuyo	0 .15	11 .63	0 .35	0 .85	0 .03
San Luis (1973), Cuyo	0 .22	12 .48	0 .56	0 .72	0 .03
San Martín (1992), GBA	0 .35	27 .72	1 .92	0 .73	0 .05
Santiago del Estero (1973), NEA	0 .25	27 .58	1 .39	0 .26	0 .01
Sur (1956), Patagonia	0 .26	17 .97	0 .95	0 .38	0 .02
Tecnológica Nacional (1948), Pampeana	0 .28	13 .92	0 .77	0 .62	0 .03
Tres de Febrero (1995), GBA	0 .13	45 .98	1 .22	1 .04	0 .03
Tucumán (1912), NOA	0 .17	23 .78	0 .81	0 .48	0 .02
Villa María (1996), Pampeana	0 .19	13 .68	0 .52	0 .47	0 .02
Total	0 .26	23 .75	1 .22	0 .64	0 .03

Source: Own elaboration on Anuarios de Estadísticas Universitarias

In the Table 7 we present the descriptive statistics of the sample, by variable. As can be seen from Table 5, many variables varies significantly within the sample (the difference between the maximum value and the minimum value is large).

Table 7: Descriptive statistics of the sample (n=333)

Variable	Mean	Std. Dev.	Min	Max
Graduates	1,794	2,996	24	18,124
Students	34,896	54,356	1,111	358,071
Exclusive	516	503	0	2,478
Semi-exclusive	781	796	44	3,589
Simple	2,251	4,176	0	21,214
Professors	1,473	1,992	115	10,522
Assistants	2,075	3,202	2	18,636
Professors_w	702	715	53	3,224
Assistants_w	767	1,046	1	6,011
Expenditures (millions 2013 AR$)	1,140	2,790	3	30,200
GBA	0.24	0.43	0.00	1.00
NOA	0.16	0.37	0.00	1.00
NEA	0.11	0.31	0.00	1.00
Cuyo	0.08	0.27	0.00	1.00
Pampeana	0.32	0.47	0.00	1.00
Patagonia	0.08	0.27	0.00	1.00
rprodst	0.87	0.15	0.19	1.00

Note: Own elaboration on “Anuarios de Estadísticas Universitarias” (Ministry of Education of Argentina, 2005-2013).

4 Results

To select the model specification, we run four models by pool OLS, considering the 37 universities over the period 2005-2013 using the linear, log-lin, lin-log and log-log specifications respectively. The variables were chosen in line with the empirical literature and the objective to assess efficiency in teaching activity. Besides, at the end of this Section we perform a sensitive analysis to check the robustness of our results.

As can be observed, most of the signs of the significant variables are the expected ones and are robust to all specifications. The major difference is that in the lin-log specification the teaching variables seem not to have any effect in the number of graduates, while the university expenditure has a positive impact. As the information criterion BIC favors the log-log specification, we then take it as our “base model”.

Before performing the SFA models, we test the residuals of the pool OLS model to check for the validity of the stochastic frontier specification. The idea behind the test is that for the production frontier model, with a composed error $v_{it} - u_{it}$, $u_{it} \geq 0$ and v_{it} distributed symmetrically around zero, the residual of the corresponding OLS estimation should skew to the left. The calculated statistic for the skewness is equal to -0.4746, and is statistically significant (p-value=0.0007) indicating that the distribution of the residuals skews to the left (see Figure 1).[10]

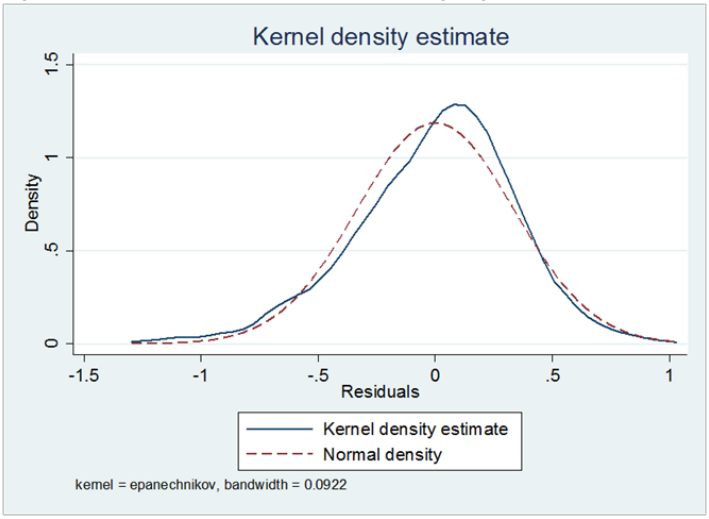
[10] The test statistic performed is that proposed by Schmidt and Lin (1984) and is defined as: $b_1^{1/2} = \frac{m_3}{m_2\, m_2^{1/2}}$, where m_2 and m_3 are the second and third sample moment of the OLS residuals, respectively.

Table 8: Higher education graduates equation under alternative specifications (POLS).

Variables	Model Specification			
	OLS	**log-lin**	**lin-log**	**log-log**
Students	0.0447***	0.0000	1,277.3409***	0.8441***
	(0.0052)	(0.0000)	(332.8796)	(0.0695)
Professors_w	0.3775**	0.0012***	702.3467	0.5311***
	(0.1517)	(0.0003)	(430.4674)	(0.1013)
Assistants_w	0.1555	0.0001	155.7676	-0.1263***
	(0.2645)	(0.0002)	(148.8100)	(0.0433)
Expenditures	0.0000	-0.0000***	472.5505***	0.0025
	(0.0000)	(0.0000)	(146.5827)	(0.0224)
GBA	489.8208***	0.7898***	2,086.4863***	0.7285***
	(83.4449)	(0.1722)	(491.3932)	(0.0900)
NOA	-206.8891***	-0.0775	120.5934	-0.1471*
	(78.9828)	(0.1823)	(363.2633)	(0.0843)
NEA	359.8762***	0.9954***	396.6992	0.6220***
	(85.1771)	(0.1726)	(310.8571)	(0.0960)
Cuyo	142.6727	0.3503	-690.9034**	0.2429**
	(114.5587)	(0.2143)	(333.1239)	(0.1170)
Pampeana	428.5354***	0.6485***	329.8131	0.5954***
	(68.7166)	(0.1553)	(292.8966)	(0.0660)
Trend	4.7253	0.0337**	-140.4987***	0.0026
	(15.3658)	(0.0169)	(54.0648)	(0.0085)
Constant	-480.7804***	5.0811***	-25,410.4323***	-4.6668***
	(96.4947)	(0.1736)	(2,930.9683)	(0.3382)
Observations	323	323	323	323
R-squared	0.9619	0.7036	0.6777	0.9267
BIC	5103,149	720,6561	5792,577	275,1538
Robust standard errors in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

Source: Own elaboration

Figure 1: Distribution of OLS residuals (log-log model)



Then, the SFA models estimated are the following:

- ***RE***

$$lngrad_{it} = \alpha + \beta_1 lnStud_{it}\, \beta_2 lnProf_{it}\, \beta_3 lnAssist_{it}\, \beta_4 lnExp_{it} + \delta_1 GBA_i + \delta_2 NOA_i + \delta_3 NEA_i + \delta_4 Cuyo_i + \delta_5 Pampeana_i + \theta\tau + v_{it} - u_i$$

- ***REH***

$$lngrad_{it} = \alpha + \beta_1 lnStud_{it}\, \beta_2 lnProf_{it}\, \beta_3 lnAssist_{it}\, \beta_4 lnExp_{it} + \delta_1 GBA_i + \delta_2 NOA_i + \delta_3 NEA_i + \delta_4 Cuyo_i + \delta_5 Pampeana_i + \theta\tau + v_{it} - u_{it}$$

$$\mu_i = z_0 + z_1 rstudents_{it}$$

- ***TFE***

$$lngrad_{it} = \alpha_i + \beta_1 lnStud_{it}\, \beta_2 lnProf_{it}\, \beta_3 lnAssist_{it}\, \beta_4 lnExp_{it} + \delta_1 GBA_i + \delta_2 NOA_i + \delta_3 NEA_i + \delta_4 Cuyo_i + \delta_5 Pampeana_i + \theta\tau + v_{it} - u_{it}$$

- ***TRE***

$$lngrad_{it} = \alpha + w_i + \beta_1 lnStud_{it}\, \beta_2 lnProf_{it}\, \beta_3 lnAssist_{it}\, \beta_4 lnExp_{it} + \delta_1 GBA_i + \delta_2 NOA_i + \delta_3 NEA_i + \delta_4 Cuyo_i + \delta_5 Pampeana_i + \theta\tau + v_{it} - u_{it}$$

- ***TRE+REH***

$$lngrad_{it} = \alpha + w_i + \beta_1 lnStud_{it}\, \beta_2 lnProf_{it}\, \beta_3 lnAssist_{it}\, \beta_4 lnExp_{it} + \delta_1 GBA_i + \delta_2 NOA_i + \delta_3 NEA_i + \delta_4 Cuyo_i + \delta_5 Pampeana_i + \theta\tau + v_{it} - u_{it}$$

$$\mu_i = z_0 + z_1 rstudents_{it}$$

We use the ratio number of students of university *i* in period *t* divided on the maximum number of students of university *i* over the period 2005-2013 as a measure of efficiency.

Table 9: Estimation results of the alternative SFA models

Variables	Alternative models				
	RE	**REH**	**TFE**	**TRE**	**TRE+REH**
ln Students	0.7720***	0.8883***	0.8893***	0.8654***	0.8540***
	(0.0757)	(0.0618)	(0.1037)	(0.067)	(0.0661)
ln Professors_w	0.6020***	0.5000***	0.5071***	0.2755***	0.2828***
	(0.1069)	(0.0925)	(0.1283)	(0.0952)	(0.0924)
ln Assistants_w	-0.1582***	-0.1575***	-0.1514***	-0.0851**	-0.0882**
	(0.0410)	(0.0335)	(0.0459)	(0.0361)	(0.0369)
ln Expenditures	0.0374*	-0.0071	-0.0014	0,0226	0.0229
	(0.0201)	(0.0244)	(0.0194)	(0.0170)	(0.0165)
GBA	0.8234***	0.7240***	0.9173	0.6493***	0.6579***
	(0.1402)	(0.0830)	(0.9892)	(0.1455)	(0.1461)
NOA	-0.1771	-0.1469**	-0.2427	-0,1400	-0.1393
	(0.1394)	(0.0747)	(0.8573)	(0.1484)	(0.1486)
NEA	0.6598***	0.6068***	0.8872	0.6712***	0.6794***
	(0.1539)	(0.0876)	(1.0365)	(0.1597)	(0.1551)
Cuyo	0.3692**	0.3723***	0.4542	0.3228**	0.3207**
	(0.1621)	(0.0955)	(1.0355)	(0.1591)	(0.1579)
Pampeana	0.5725***	0.5440***	0.6233	0.5499***	0.5526***
	(0.1297)	(0.0683)	(1.2430)	(0.1385)	(0.1380)
Trend	-0.0035	0.0033	0.0068	0,0058	0.0036
	(0.0069)	(0.0086)	(0.0069)	(0.0056)	(0.0054)
Constant	-4.5656***	-4.2381***		-3.7481***	-3.6659***
	(0.5195)	(0.3597)		(0.4179)	(0.3974)
Observations	323	323	323	323	323
Number of groups	37	37	37	37	37
Log-likelihood	-38,3818	-95,0618	44,727	-17,3357	-15,724
BIC	151,8732	276,7885	164,7627	121,3362	118,1129
Standard errors in parentheses					
*** p<0.01, ** p<0.05, * p<0.1					

As can be seen in Table 9, in terms of signs and significance, results almost do not differ among specifications, except in the case of the TFE model where any of the regional dummy is statistical significant (however, the majority of the specific dummy variables are significant). All the significant coefficients have the expected sign and have similar magnitudes. The number of students has a positive effect on the determination of the production frontier, but the teachers' effects differ: the number of professors has a positive impact while the effect is negative when assistant professors are considered.

Only in the RE model university expenditure explains the production of graduates. Then, it seems that human resources are more important for the definition of the production frontier than the budget allocated to each university. Regarding the regional variables, all the models show that the geographical region explains the number of graduates. Ceteris paribus,

all universities have a positive significant differential in the number graduates regarding the base region (Patagonia).

The time variable is in general positive, but it is never statistically significant, meaning that there is no any technological development over the period under analysis.

As it is shown in Table 10, the student relationship has the expected negative sign but it is not significant for explaining the distribution of the mean of inefficiency in the RE model, however it is significant in the model that considers both observed and unobserved heterogeneity.

Table 10: Error components of the alternative models

Variables	RE	REH	TFE	TRE	TRE+REH
Mu					
rprodst		-1.5813			-170.1476***
		(2.6269)			(0..7588)
Constant		-1.1292		-302.0971***	-106,0048
		(3.1004)		(10.2618)	
sigma_u					
Constant	0.4046***	0.9408	0.2132***	8.0707***	7.3616***
	(0.0604)	(0.7007)	(.0203337)	(0.43057)	(0.0141)
sigma_v					
Constant	0.2410***	0.2056***	0.1057***	0.1143***	0.1121***
	(0.0101)	(0.0227)	(0.0153)	(0.0155)	(0.0121)
Theta				0.2621***	0.2621***
				(0.0355)	(0.0346)

Source: Own elaboration

The estimated inefficiency variance is significant in all models, while presents the highest value in the RE one. This means that some of the inefficiency captured in the RE model is in fact heterogeneity. Finally, the parameter θ is statistically significant, indicating the presence of heterogeneity among universities.

Regarding the estimated efficiency scores, the mean efficiency over the entire period goes from 75% to 82% depending on the model considered (see Table 11). This means that from 20% to 25% of the production (graduates) is lost due to inefficiency. Besides, as expected, those models that take into account unobserved heterogeneity present higher values of TE while the RE model presents the lowest value. This result indicates that there is evidence of heterogeneity that overestimates inefficiency.

Table 11: Descriptive statistics of the estimated inefficiency scores (JLMS technique)

Variable	Mean	Std. Dev.	Min	Max
RE	0,7494	0,1515	0,3707	0,9514
REH	0,7636	0,1432	0,2587	0,9512
TFE	0,8246	0,1408	0,2371	0,9721
TRE	0,8163	0,1413	0,2380	0,9672
TRE+REH	0,8211	0,1372	0,2365	0,9692

Source: Own elaboration

Finally, we present a correlation among the estimated inefficiency scores of the alternative models. As can be observed, the ranks of the RE model are highly correlated with those of the REH model, weakly correlated with the ranks of the TRE+REH specification, and not statistical correlated with the ranks of the TFE and TRE models. This means that the ranking of the TE scores differs depending on the model chosen, more specifically it depends whereas the model takes into account unobserved university-specific differences. In fact, the RE and the REH model rank the universities similarly, while the rankings of the TFE and the TRE models are highly correlated.

Table 12: Correlation of inefficiency estimates

	RE	REH	TFE	TRE	TRE+REH
RE	1				
REH	0,7209	1			
	0,000				
TFE	0,074	0,6573	1		
	0,184	*0,000*			
TRE	0,082	0,6568	0,9835	1	
	0,141	*0,000*	*0,000*		
TRE+REH	0,1373	0,703	0,9886	0,9868	1
	0,014	*0,000*	*0,000*	*0,000*	

P-value in italics.
Source: Own elaboration

As the BIC criteria indicates that the model that best fits the data is the model that takes into account both observed and observed heterogeneity, in the next Table 13 we present the ranking of the public universities according to their estimated efficiency score.

Table 13: Efficiency results by University (JLMS estimates, average 2005-2013)

Nacional University	TE score	Nacional University	TE score
Cuyo	0,8973		
Nordeste	0,8901	Catamarca	0,8478
Córdoba	0,8884	San Juan	0,8392
Sur	0,8863	Villa María	0,8373
Mar del Plata	0,8814	Comahue	0,8253
La Plata	0,8809	Entre Ríos	0,8216
Buenos Aires	0,8801	La Rioja	0,8090
Tecnológica Nacional	0,8795	Rosario	0,7762
Tucumán	0,8791	Santiago del Estero	0,7734
Patagonia San Juan Bosco	0,8754	Jujuy	0,7687
Litoral	0,8709	Patagonia Austral	0,7557
Río Cuarto	0,8705	La Matanza	0,7516
Luján	0,8701	Salta	0,7348
La Pampa	0,8669	General Sarmiento	0,6972
Misiones	0,8664	Tres de Febrero	0,6961
San Martín	0,8608	Quilmes	0,6956
Lomas de Zamora	0,8603	Formosa	0,6763

San Luis	0,8573	Chilecito	0,6594
Centro de la PBA	0,8537	Average	0,8224
Lanús	0,8495	Standard Deviation	0,0700

Source: Own elaboration

As expected, the ranking is very similar among the TFE, the TRE and the TRE_RH models. When analyzing the ranks university by university, we notice that in these models the ranks of the University of Rosario and the University of Entre Ríos, together with the University of Lomas de Zamora decrease compared with the RE and REH rankings, while the position of other big Universities such as the University of Córdoba, University of La Plata and University of Buenos Aires increase (see Table 14).

Table 14: Efficiency results by University (JLMS estimates, average 2005-2013)

RE		REH		TFE		TRE	
Nacional University	TE score	Nacional University	TE score	Nacional University	TE score	Nacional University	TE score
Lomas de Zamora	0,9514	Cuyo	0,9028	Córdoba	0,8906	Cuyo	0,8916
Rosario	0,9442	Lomas de Zamora	0,8990	Cuyo	0,8898	Nordeste	0,8858
Cuyo	0,9324	Rosario	0,8958	Sur	0,8896	Córdoba	0,8836
Entre Ríos	0,9311	Nordeste	0,8873	Tecnológica Nacional	0,8894	Sur	0,8830
Tucumán	0,9285	Tucumán	0,8872	Buenos Aires	0,8887	La Plata	0,8789
Nordeste	0,9284	Entre Ríos	0,8748	Mar del Plata	0,8879	Mar del Plata	0,8780
Comahue	0,9232	Córdoba	0,8722	Misiones	0,8847	Buenos Aires	0,8750
Quilmes	0,9168	Comahue	0,8570	La Plata	0,8845	Tucumán	0,8740
La Rioja	0,8856	San Martín	0,8470	Litoral	0,8843	Tecnológica Nacional	0,8735
San Martín	0,8486	Sur	0,8346	Nordeste	0,8822	Patagonia San Juan Bosco	0,8720
Sur	0,8449	Luján	0,8276	Patagonia San Juan Bosco	0,8813	Litoral	0,8697
Catamarca	0,8329	Río Cuarto	0,8242	La Pampa	0,8751	Río Cuarto	0,8694
San Luis	0,8319	Catamarca	0,8165	Río Cuarto	0,8744	Luján	0,8655
Río Cuarto	0,8298	Patagonia San Juan Bosco	0,8130	Luján	0,8720	Misiones	0,8645
Villa María	0,8297	Buenos Aires	0,8116	Tucumán	0,8716	San Martín	0,8621
Patagonia San Juan Bosco	0,8143	Mar del Plata	0,8086	San Juan	0,8712	La Pampa	0,8620
Santiago del Estero	0,8053	La Plata	0,8052	Centro de la PBA	0,8634	Lomas de Zamora	0,8550
Luján	0,7952	Quilmes	0,8032	Villa María	0,8619	Centro de la PBA	0,8530
Córdoba	0,7727	La Rioja	0,7991	San Martín	0,8518	San Luis	0,8519
La Pampa	0,7459	San Luis	0,7916	San Luis	0,8513	Lanús	0,8509
Salta	0,7408	La Pampa	0,7718	Catamarca	0,8474	Villa María	0,8482
Mar del Plata	0,7402	Santiago del Estero	0,7701	Lanús	0,8462	Catamarca	0,8442
Formosa	0,7108	Litoral	0,7504	Lomas de Zamora	0,8298	San Juan	0,8350
Tecnológica Nacional	0,6981	Centro de la PBA	0,7501	Entre Ríos	0,8276	Comahue	0,8238
Lanús	0,6917	Tecnológica Nacional	0,7389	Comahue	0,8201	Entre Ríos	0,8229
Chilecito	0,6820	Villa María	0,7327	La Rioja	0,8092	La Rioja	0,8167
Centro de la PBA	0,6600	Salta	0,7176	Jujuy	0,7946	Santiago del Estero	0,7782
La Plata	0,6546	La Matanza	0,7170	Patagonia Austral	0,7877	Rosario	0,7724
Litoral	0,6538	Formosa	0,7030	Santiago del Estero	0,7864	Jujuy	0,7684
La Matanza	0,6346	Lanús	0,6939	Rosario	0,7600	Patagonia Austral	0,7617

General Sarmiento	0,6030	General Sarmiento	0,6363	La Matanza	0,7555	La Matanza	0,7560
Jujuy	0,5651	Misiones	0,6221	Salta	0,7360	Salta	0,7364
Patagonia Austral	0,5500	Chilecito	0,6194	Tres de Febrero	0,6997	General Sarmiento	0,7097
Misiones	0,5469	Jujuy	0,5809	General Sarmiento	0,6972	Tres de Febrero	0,7052
Buenos Aires	0,5464	Patagonia Austral	0,5590	Quilmes	0,6818	Quilmes	0,7043
Tres de Febrero	0,3847	San Juan	0,4570	Chilecito	0,6711	Chilecito	0,6771
San Juan	0,3707	Tres de Febrero	0,4322	Formosa	0,6625	Formosa	0,6723

Source: Own elaboration

This result indicates that there are heterogeneity factors that influence inefficiency. One possible explanation could be that in the last group of Universities there is a higher proportion of students who participate in the labor market. We can approximate an analysis using the *Encuesta Permanente de Hogares* (EPH). Over the period under analysis, the EPH shows that Buenos Aires City, Mar del Plata, Greater Buenos Aires and Gran Córdoba are the urban agglomerations with higher proportion of workers among university students. On the other hand, Gran Rosario and Gran Paraná together with Concordia are agglomerates with lower proportion of occupied students than in the previous regions. This fact could negatively affect the production function and then, when controlling for these factors, the inefficiency component decreases when the proportion of students who works is high.

Finally, we undertake a sensitivity to check the robustness of our results. We first separate the total expenditure into its components (i.e. costs of personnel, consumption, goods and services and transfers). We include all the components except those that refer to personnel due to its positive relationship with the teaching variables (M1). Then, we use two alternative specifications to measure the teachers' effects.

In the first alternative (M2), we consider the teachers according to their role (professor or assistant) and we include an environmental variable defined as $exclusive_p = \frac{exclusive}{(exclusive + 0.5\ semi + 0.25\ simple)}$ to take into account the proportion of exclusive teachers.

In the other specification (M3), we consider the number of professors according to their time dedication (exclusive, semi-exclusive or simple). This classification comprised every existing category in the national system, going from full time professors holding a permanent chair, and from full and part-time assistant teachers. We also include an environmental variable defined as $professors_p = \frac{profesor}{(professor + assistant)}$ to take into account the proportion of professors in the regression.

Finally, we include an alternative measure of efficiency in the efficiency equation (rprodnst) defined as the ratio number of enrollees of university *i* in period *t* divided on the maximum number of enrollees of university *i* over the period 2005-2013 (M4). In the Table 15 the different models are presented. The Figure 2, in turn, shows the relationship among the TE scores of the base model and those of the sensitive analysis.

It can be observed that TE scores obtained from the above models are highly positively correlated with those of the base model. This result indicates that the estimated TE in the base regression is robust to the alternative models.

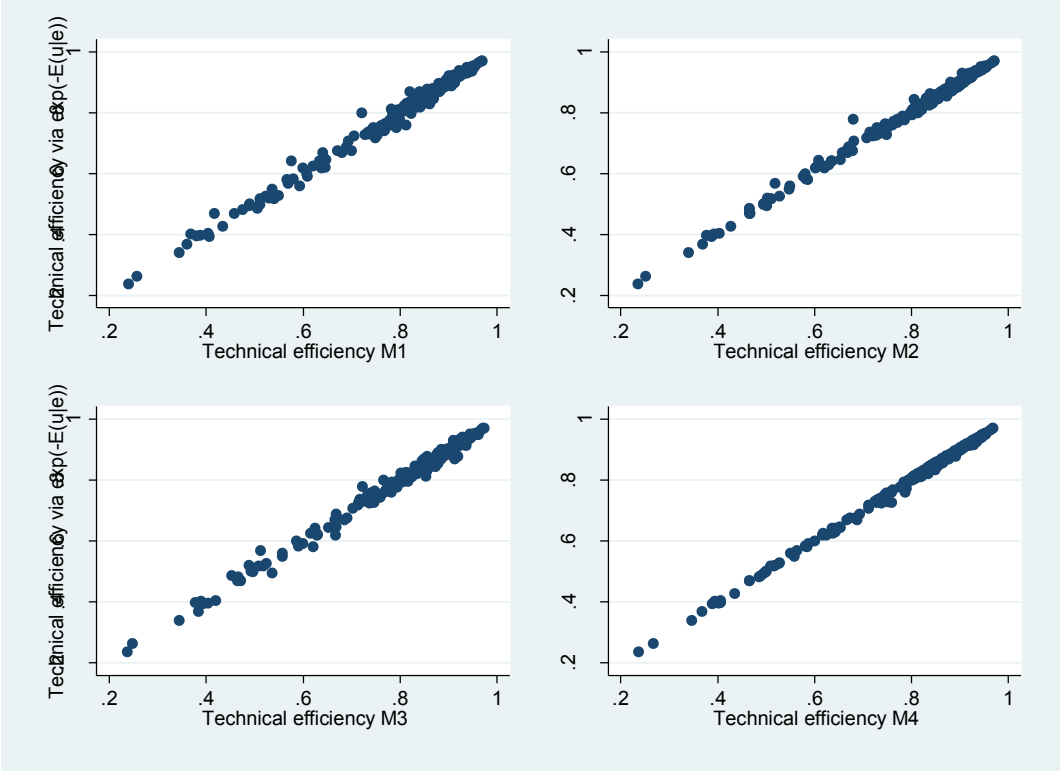
Table 15 Estimation results of the alternative specifications of the TRE_REH model

Variables	Alternative models			
	M1	**M2**	**M3**	**M4**
ln Students	0.8834***	0.8908***	0.8966***	0.8627***
	(0.0054)	(0.0617)	(0.1184)	(0.0666)
ln Professors_w	0.3182***			0.2735***
	(0.0074)			(0.0950)
ln Assistants_w	-0.0910***			-0.0848**
	(0.0097)			(0.0364)
ln Consumption	0.0766***			
	(0.0031)			
ln Services	-0.0675***			
	(0.0039)			
ln Goods	-0.0121***			
	(0.0029)			
ln Transfers	-0.0195***			
	(0.0039)			
ln Professor		0.2279***		
		(0.0801)		
ln Assistants		-0.0948***		
		(0.0356)		
ln Exclusive_p		-0.0183		
		(0.0457)		
ln Expenditures		0.0223	0.0162	0.0229
		(0.0163)	(0.0195)	(0.0168)
ln Exclusive			0.0146	
			(0.1218)	
ln Semi Exclusive			0.0014	
			(0.0621)	
ln Simple			0.0437	
			(0.0375)	
ln Professors_p			0.1433	
			(0.1318)	
GBA	0.7491***	0.5818***	0.7650*	0.6502***
	(0.1535)	(0.2094)	(0.3949)	(0.1455)
NOA	-0.1568**	-0.1388	-0.1025	-0.1432
	(0.0787)	(0.2022)	(0.2172)	(0.1481)
NEA	0.6043***	0.6258***	0.5300	0.6710***
	(0.1933)	(0.2130)	(0.3509)	(0.1591)
Cuyo	0.4151**	0.6173**	0.5965**	0.3227**
	(0.1699)	(0.2551)	(0.2901)	(0.1586)
Pampeana	0.5206***	0.5580***	0.6728**	0.5477***
	(0.1740)	(0.1977)	(0.3134)	(0.1386)
Trend	0.0150***	0.0049	0.0066	0.0047
	(0.0039)	(0.0053)	(0.0085)	(0.0055)
rprodst	-113.0849	-171.3990***	-114.8764	

	(0.0000)	(1.2659)	(200.9621)	
rprodnst				-68.0678***
				(10.3479)
Constant	-3.2610***	-3.7120***	-3.0773	-3.7145***
	(0.0531)	(0.3838)	(0.0000)	(0.4072)
Observations	323	322	311	323
Number of id	37	37	36	37
Log-likelihood	-14,4547	-14,6371	-7,9163	-17,1298
BIC	121,3519	127,4416	107,6693	126,7021
Standard errors in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

Source: Own elaboration

Figure 2: TE score scatterplots



5 Summary and conclusions

This paper presents an estimation of the teaching technical efficiency public universities in Argentina using Stochastic Frontier Analysis. To our best knowledge, this is the first estimation of the university inefficiency in Argentina applying stochastic frontier techniques. In particular, we compare conventional models that do not take into account heterogeneity with heterogeneity-extended stochastic frontier models. More precisely, we applied the basic "random effect model", an extension of the "random effect model that includes observed heterogeneity" to the mean of the inefficiency distribution, a "true fixed effects model" and a "true random effects model" (both models takes into account unobserved heterogeneity), and a combination of the "true random effect model" with the extended random effect model.

We work with statistics from the "*Anuarios de Estadísticas Universitarias"* of the Ministry of Education of Argentina. The data consists of a balanced panel of 37 National Universities in Argentina with information over the period 2005-2013. We find that, in terms of signs and significance, results almost do not differ among specifications, except in the case of the TFE model where any of the regional dummy is statistical significant (however, the majority of the specific dummy variables are significant).

All the significant coefficients have the expected signs and have similar magnitudes. The number of students has a positive effect on the determination of the production frontier, but the teachers' effects differ: the number of professors has a positive impact while the effect is negative when assistant professors are considered.

Only in the RE model university expenditure explains the production of graduates. Then, it seems that human resources are more important for the definition of the production frontier than the budget allocated to each university. Regarding the regional variables, all the models show that the geographical region explains the number of graduates. Ceteris paribus, all universities have a positive significant differential in the number graduates regarding the base region (Patagonia).

The time variable is in general positive, but it is never statistically significant, meaning that there is no any technological progress over the period under analysis.

Regarding the estimated efficiency scores, the mean efficiency over the entire period goes from 75% to 82% depending on the model considered. This means that from 20% to 25% of the production (graduates) is lost due to inefficiency. Besides, as expected, those models that take into account unobserved heterogeneity present higher values of TE while the RE model presents the lowest value. This result indicates that there is evidence of heterogeneity that overestimates inefficiency. Our results are robust to different model specifications.

This study can be considered an important tool for policy makers in Argentina, since the estimation of the technical efficiency of the public universities could help to improve the educational process in order to reach more efficient levels in education.

Nevertheless, this study should be considered preliminary and can be enriched by different ways. As addressed in the literature review, universities produce more than graduates, and variables could be introduced to account for these other outcomes. More attention to characteristics of the fields in which each university is focused can improve the understanding of the variance in costs per graduate and per student. More research is needed for a better knowledge of national universities' efficiency.

References

Abbott, M. and C. Doucouliagos (2003). The Efficiency of Australian Universities: a Data Envelopment Analysis. Economics of Education Review 22, pp. 89-97.

Agasisti, Tomasso and Geraint Johnes (2009). Beyond Frontiers: Comparing the Efficiency of Higher Education Decision-Making Units Across More Than One Country. Education Economics 171:1, pp. 59-79.

Avkiran, Necmi K. (2001). Investigating Technical and Scale Efficiencies of Australian Universities Through Data Envelopment Analysis. Socio-Economic Planning Sciences 35, pp. 57-80.

Belotti F., S. Daidone, G. Ilardi, and V. Atella (2012). Stochastic frontier analysis using Stata. CEIS Tor Vergata, Research Papers Series, 10 (1), No. 251.

Cantele, Silvia, Andrea Guerrini and Bettina Campedelli (2016). Efficiency of Italian Universities: the Effect of Controllable and Non-Controllable environmental and Operational Variables. International Journal of Public Policy, Vol. 12, Nos. 3/4/5/6.

Caudill, S., and J. Ford (1993). Biases in frontier estimation due to heteroscedasticity. Economic Letters 41: 17-20.

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

Chen, Yining, Mary R. Nixon, Ashok Gupta and Leon Hoshower (2010). Research Productivity of Accounting Faculty: An Exploratory Study. American Journal of Business Education, Vol 3, Nº 2, pp. 101-115).

Cohn, Elchanan and Samuel T. Cooper (2004). Multi-Product Cost Functions for Universitites: Economies of Scale and Scope. In International Handbook on the Economics of Education. Geraint Johnes and Jill Johnes (editors). Edgar Elgar Publishing Ltd.

Coria, María Marta (2008). Eficiencia Técnica de las Universidades de Gestión Estatal en Argentina. Pontificia Universidad Católica Argentina, Facultad de Ciencias Sociales y Económicas. Departamento de Economía, documento de Trabajo Nº 27.

Costa, Edward Martins, Francisco de Sousa Ramos and Hermíno Ramos de Souza (2011). The Productive Efficiency of Federal Institutions of Brazilian Higher Education. Encontro ANPEC 2011.

Daraio, Cinzia, Andrea Bonaccorsi and Léopold Simar (2015). Efficiency and Economies of Scale and Specialization in European Universities: A Directional Distance Approach. Journal of Informetrics 9, pp. 430-448.

De Fraja, Gianni and Paola Valbonesi (2012). The Design of the University System. Journal of Public Economics 96, pp. 317-330.

Ferreyra, María Marta, Ciro Avitabile, Javier Botero Álvarez, Francisco Haimovich Paz and Sergio Urzúa (2017). At a Crossroads. Higher Education in Latin America and the Caribbean. Directions in Development. Human Development. World Bank Group.

Gates, Susan and Ann Stone (1997). Understanding Productivity in Higher Education. Prepared for California Education Roundtable. Institute on Education and Training. RAND.

Greene, W. (2005a). Reconsidering heterogeinity in panel data estimators of the stochastic frontier model. *Journal of Econometrics,* 126(2): 269-303.

Greene, W. (2005b). Fixed random effects in stochastic frontier models. *Journal of Productivity Analysis*, 23(1): 7-32.

Greene, W. (2008). The econometric approach to efficiency analysis. In The Measurement of Productive Efficiency and Productivity Growth, edited by Harold O. Fried, C. A. Knox Lovell, Shelton S. Schmidt. Oxford University Press.

Hadri, K. (1999). Estimation of a Doubly Heteroscedastic Stochastic Frontier Cost Function. *Journal of Business Economics and Statistics,* 17(3): 359-363.

Horne, Jocelyn and Baiding Hu (2008). Estimation of Cost Efficiency of Australian Universities. Mathematics and Computing in Simulation 78, pp. 266-275.

Johnes, Geraint and Jill Johnes (2009). Higher Education Institutions' Costs and Efficiency: Taking the Decomposition a Further Step. Economics of Education Review 28, pp. 107-113.

Johnes, Geraint and Jill Johnes (1993). Measuring the Research Performance of UK Economics Departments: An Application of Data Envelopment Analysis. Oxford Economic Papers 45, pp. 332-347.

Johnes, Jill (2006). Data Envelopment Analysis and its Application to the Measurement of Efficiency in Higher Education. Economics of Education Review 25, pp. 273-288.

Johnes, Jill (2004). Efficiency Measurement. In International Handbook on the Economics of Education. Geraint Johnes and Jill Johnes (editors). Edgar Elgar Publishing Ltd.

Johnes, Jill (2008). Efficiency and Productivity Change in the English Higher Education Sector From 1996/97 To 2004/05. Manchester School, vol. 76, issue 6, pages 653-674.

Johnes, Jill (2006). Measuring Teaching Efficiency in Higher Education: An Application of Data Envelopment Analysis to Economics Graduates from UK Universities 1993. European Journal of Operational Research 174, pp. 443-456.

Johnes, Jill and Li Yu (2008). Measuring the Research Performance of Chinese Higher Education Institutions Using Data Envelopment Analysis. China Economic Review 19, pp. 679-696.

Jondrow, J., C. Lovell, I. Materov and P. Schmidt (1982). On estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19(2-3): 233-238.

Kao, Chiang and His-Tai Hung (2008). Efficiency Analysis of University Departments: An Empirical Study. Omega. The International Journal of Management Science 36, pp. 653-664.

Katharaki, Maria and George Katharakis (2010). A Comparative Assessment fo Greek Universities' Efficiency Using Quantitative Analysis. International Journal of Educational Research 49, pp. 115-128.

Kumbhakar, S., and L. Hjalmarsson (1993). Technical efficiency and technical progress in Swedish dairy farms. In H. Fired, K. Lovell and S. Schmidt (eds). The Measurement of Productive Efficiency: Technics and Applications. New York: Oxford University Press.

Kuo, Jenn-Shyong and Yi-Cheng Ho (2008). The Cost Efficiency Impact of the University Operation Fund in Taiwan. Economics of Education Review 27, pp. 603-612.

Kuah, Chuen Tse and Kuan Yew Wong (2011). Efficiency Assessment of Universities Through Data Envelopment Analysis. Procedia Computer Science 3, pp. 499-506.

Laureti, Tiziana, Luca Secondi and Luigi Biggeri (2014). Measuring the Efficiency of Teaching Activities in Italian Universities: An Information Theoretic Approach.

Lee, Boon L. and Andrew C. Worthington (2016). A Network DEA Quantity and Quality-Oriented Production Model: An Application to Australian University Research Services. Omega 60, pp. 26-33.

Marinho, Alexandre, Marcelo Resende and Luís Otávio Façanha (1997). Brazilian Federal Universities: Relative Efficiency Evaluation and Data Envelopment Analysis. Revista Brasileira de Economia 51(4), pp. 489-508.

McGuckin, Robert H. and Donald R. Winkler (1979). University Resources in the Production of Education. The Review of Economics and Statistics, Vol 61, Nº 2, pp 242-248.

Millot, Benoit (2015). International Rankings: Universities vs. Higher Education Systems. International Journal of Educational Development 40, pp. 156-165.

Mensah, Yaw M. and Robert Werner (2003). Cost Efficiency and Financial Flexibility in Institutions of Higher Education. Journal of Accounting and Public Policy 22, pp. 293-323.

Pitt, M., and L. Lee (1981). The measurement and sources of technical inefficiency in Indonesian weaving industry. *Journal of Development Economics*, 9(1):43-64.

Salerno, Carlo (2003). What we know about the efficiency of higher education institutions: the best evidence. [Report].

Siegel, Donald S., David Waldman and Albert Link (2003). Assessing the Impact of Organizational Practices on the Relative Productivity of University Technology Transfer Offices: An Exploratory Study. Research Policy 32, pp. 27-48.

Tyagi, Preeti, Shiv Prasad Yadav and S. P. Singh (2009). Relative Performance of Academic Departments Using DEA With Sensitivity. Evaluation and Program Planning 32, pp. 168-177.

Sav, G. Thomas (2012). Managing Operating Efficiencies of Publicly Owned Universities: American University Stochastic Frontier Estimates Using Panel Data. Advances in Management & Applied Economics, vol. 2, Nº 1, pp. 1-23.

Schmidt, P., and T. Lin (1984). Simple tests of alternative specifications in stochastic frontier models. *Journal of Econometrics*, 24:349-61.

Secretaría de Políticas Universitarias (many years). Anuarios de Estadísticas Universitarias. Ministerio de Educación, Ciencia y Tecnología de Argentina. Available in http://portales.educacion.gov.ar/spu/investigacion-y-estadisticas/anuarios/.

Wolff, Edward, William J. Baumol and Anne Noyes Saini (2014). A Comparative analysis of Education Costs and Outcomes: The United States vs. Other OECD Countries. Economics of Education Review 39, pp. 1-21.

Worthington, Andrew (2001). An Empirical Survey of Frontier Efficiency Measurement Technicques in Education. Education Economics 9(3), pp. 245-268.

Worthington, Andrew and Boon L. Lee (2008). Efficiency, Technology and Productivity Change in Australian Universities, 1998-2003. Economics of Education Review 27, pp. 285-298.

Zhoghbi, Ana Carolina, Fabiana Rocha and Enlinson Mattos (2013). Education Production Efficiency: Evidence from Brazilian Universities. Economic Modelling 31, pp. 94-103.