WORKSHOP ON EFFICIENCY IN EDUCATION

Lancaster University Management School
The Work Foundation

London, September 19 – 20, 2014

"The teacher effect: an efficiency analysis from a natural experiment in Spanish primary schools"

Daniel Santín
dsantin@ccee.ucm.es
Gabriela Sicilia
gabriels@ucm.es
Complutense University of Madrid (Spain)

Outline

- Introduction: Education and efficiency
- Methodology
  - School choice, endogeneity and education
  - Identification strategy
- Database and relevant variables
- Results
- Conclusions

Introduction

- Mediocre results of Spain in international studies (TIMSS-PIRLS, PISA, PIAAC) compared to educational expenditure effort.
- School dropout rate (26.5%) is far above the EU-27 (13.5%)
- Public sector is the main provider and producer of education.
- In the current context of economic recession and financial crisis the public expenditure devoted to education requires new priorities and clear targets.
- Lack of consensus among stakeholders on how to improve education. The new Educational Law (2013) claims for more evaluation and introduces standarized tests in schools.

Measuring efficiency in Education

- A myriad of papers have estimated technical efficiency for high schools (Worthington, 2001)
- There are some international empirical evidence measuring efficiency in primary schools (Grosskopf et al. 2001; Banker et al.2004; Blackburn et al. 2013; Mancebón and Mar-Molinero 2000; Mizala et al. 2002; Thanassoulis 2002).
- No previous evidence for Spain in primary education (lack of databases).

The Educational Production Funcion

- The EPF for a group of school was proposed by Levin (1974) and Hanushek (1979, 2012). Assuming inefficient behaviors the EPF is:

$$A_i = f(B_i, S_i) * exp(u_i) \qquad \text{being } u_i \leq 1$$

- This model implicitly assumes:
  - B_i, S_i are exogenously determined
  - u_i, the efficiency term, is independent of inputs
- When these assumptions do not hold the problem of endogeneity arises. (Orme and Smith, 1996; Bifulco and Bretschneider, 2001; Ruggiero, 2004; Coelli and Peyrache, 2009; Grosskopf et al., 2014).

School choice and endogeneity (I/III)

- Endogeneity is a common issue in education (Schlotter et al. 2011). Students are not randomly assigned to schools.
- It is well-known that better schools attract relatively more advantaged students and more motivated parents self-select in best schools.
- Parents motivation; $\mu \sim N(0; \sigma_\mu^2)$(unobserved) is positively correlated with SES. $corr(B; \mu) > 0$
- Best pupils (and thus the schools they attend) will tend to obtain better academic results for two reasons:
  1. High SE level which is an essential input to produce education.
  2. High motivated students are more prone to be efficient.

School choice and endogeneity (II/III)

- In practice, this means that the efficiency term u_i that it is measured in efficiency analysis is composed by two terms:
  - The managerial technical efficiency θ_i
  - The omitted in the model average parents' motivation at each school μ_i ; $\mu \sim N(0; \sigma_\mu^2)$

$$A_i = f(B_i, S_i) * \underbrace{exp(\mu_i) * exp(\theta_i)}_{exp(u_i)}$$

- Two possible situations:

1) If students are randomly assigned to schools then

$E(\mu_i) = 0 \ \forall i$ and $exp(u_i) \cong exp(\theta_i)$

2) If students <u>are not</u> randomly assigned to schools then

$E(\mu_1) \neq E(\mu_2) \neq \ldots \neq E(\mu_i) \ \forall i$; $corr(B; \mu) > 0 \rightarrow corr(B; u) > 0$

School choice and endogeneity (III/III)

- Difficult to disentangle what part of u_i is due to school efficiency and the non-observed average motivation.
- Kousmanen and Johnson, (2010, p.152) demonstrate in their work that the DEA problem can be interpreted as a nonparametric least-squares model under the assumption that $\varepsilon_i \leq 0$.
- We can derive straightforward that the same problems of bias caused by the presence of endogeneity in econometrics can also arise within the DEA approach.
- In a recent working paper Cordero-Ferrera et al. (2013) show that when there is a high positive correlation between one input and the efficiency term them DEA obtains flawed estimations.

Identification Strategy (I/III)

- We need a procedure to isolate and explain technical efficiency θ_i from parents' motivation μ_i
- Our strategy employs 'causal inference' frequently used in econometrics to deal with endogeneity in education (Schlotter et al. 2011).
- We use the 'Educational General Diagnostic Database for Primary Education'. A survey for Spanish 4th grade students that also collets information about results, parents, teachers and schools (2009). Two advantages:
  1. It is possible to observe two classrooms inside the same school.
  2. A question in this survey ask the principal: 'how students are assigned to classrooms?'

Identification Strategy (II/III)

- **Randomization** in assigning students to schools (alphabetical order, boys and girls equilibrium, heterogeneity) **guarantees that parents' motivation is not significantly different in both classrooms within each school**.
- The lack of randomization in assigning students to classrooms, for example grouping by ability, language at home, homogeneity within the class, etc. possibly leads to endogeneity.
- Half of our schools use random methods to assign students to classrooms.

Identification Strategy (III/III)

1. Compute u_{iG} at classroom level using DEA
2. Compute efficiency differences for two groups belonging the same school

$$exp(u_{i1}) - exp(u_{i2}) = (exp(\theta_{i1}) - exp(\mu_{i1})) - (exp(\theta_{i2}) - exp(\mu_{i2}))$$

$$exp(u_{i1}) - exp(u_{i2}) = (exp(\theta_{i1}) - exp(\theta_{i2})) - (exp(\mu_{i1}) - exp(\mu_{i2}))$$

- Randomization guarantees

$$E(exp(\mu_{i1}) - exp(\mu_{i2})) = 0 \ \forall i$$

- (Assignation strategies within schools produce)

$$E(exp(\mu_{i1}) - exp(\mu_{i2})) \neq 0 \ \forall i$$

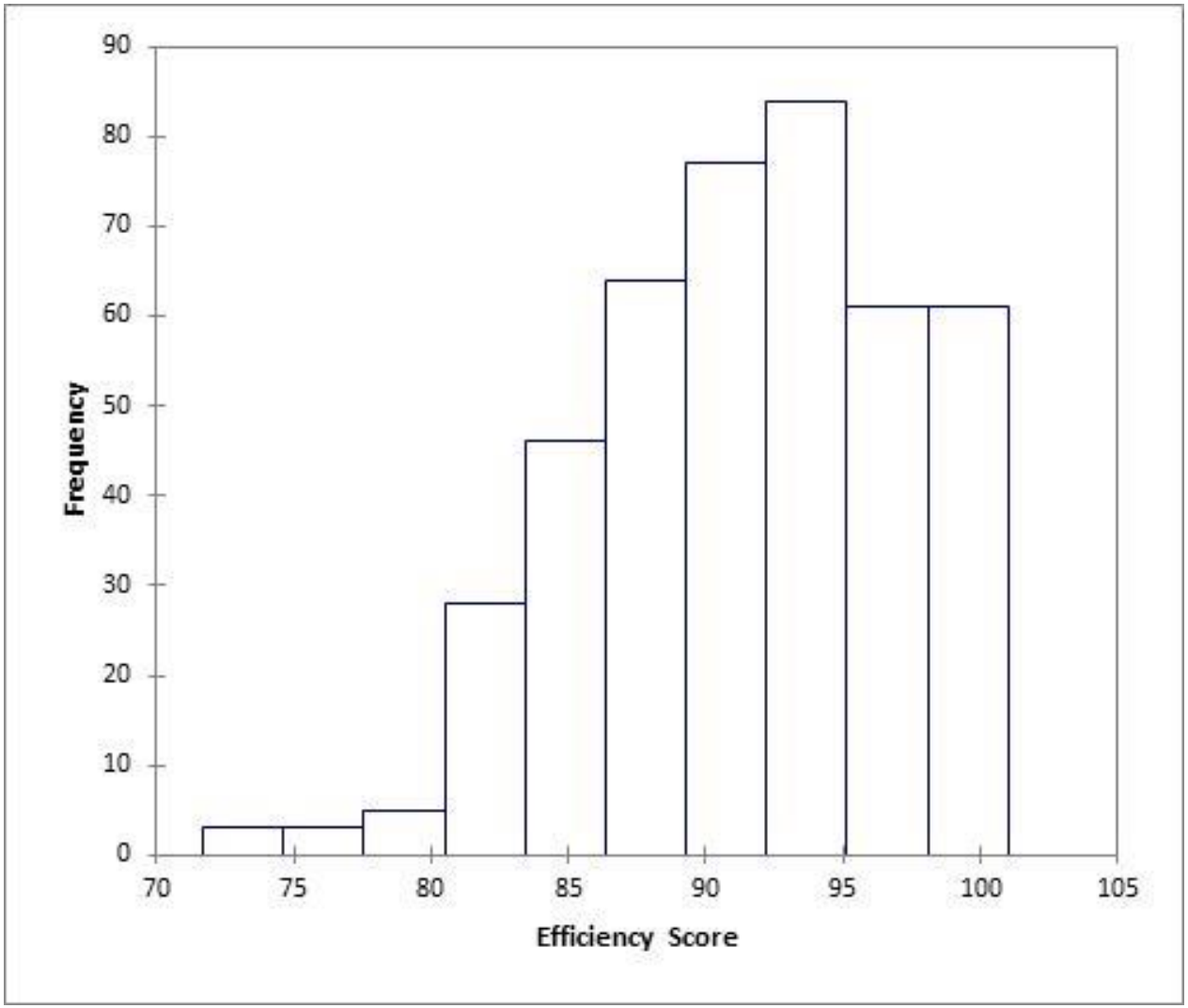
Identification Strategy (III/III)

- Normally we will observe that a classroom is more efficient than the other one, basically for two effects.
  - The teacher applies different 'observable' educational techniques or has some 'observable' characteristics.
  - The existence of an unobservable 'fix teacher effect'.
- What makes an efficient classroom? For every school compute $\Delta u_i = u_{iT} - u_{iC} > 0$ being u_{iT} the most efficient one.
- In this 'natural experiment' we 'treat' (T) some groups of students with the most efficient teachers to analyze what variables characterize (explain) the best performers with respect the non-treated or 'control group' (C).
- After this, regress $\Delta u_i = Z(\Delta z_i)$ with respect to a vector of explanatory of efficiency variables in group differences.

Inputs and Outputs

- Outputs:
  - Classroom average results in mathematics tests.
  - Classroom average results in reading tests.
- Inputs:
  - Index of SES
  - Percentage of native students
  - Percentage of students without learning difficulties
  - School index of educational resources quality
- 432 groups in 216 schools; two groups by school.

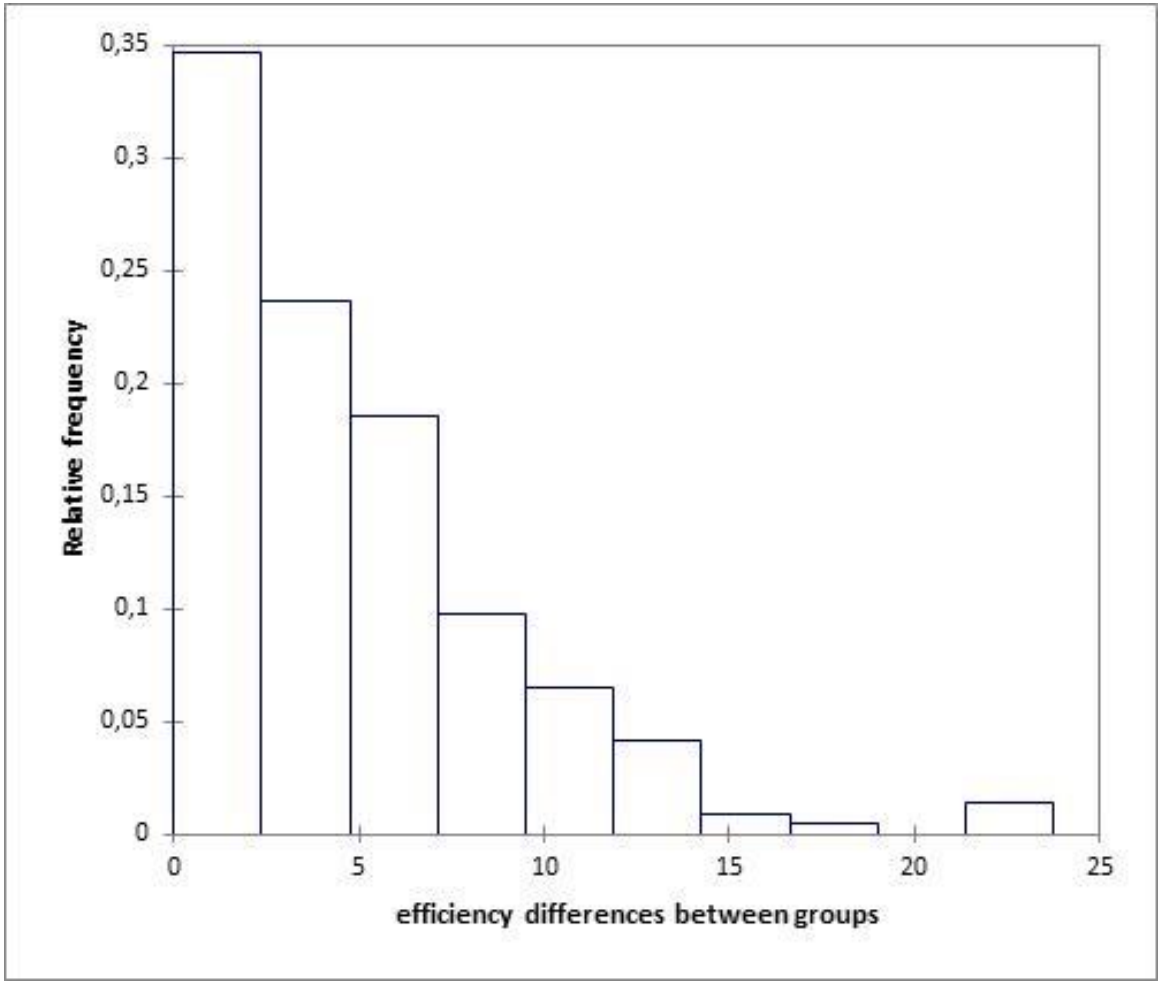
Efficiency distribution



- Mean: 91.51
- StD: 5.67

- …But this u_{iG} estimations are possibly biased (μ_{iG}, θ_{iG})

Efficiency distribution in differences



- In 60% of schools differences are less than 5 efficiency points.
- In 30% of schools differences are between 5 and 10 efficiency points.
- In 10% of schools differences are more than 10 efficiency points.

What explain these efficiency differences?

	B	t	p-value
Intercept	3.820	8.718	**0.001**
dif_early schooling	1.472	1.702	0.177
dif_monoparental family	2.600	0.921	0.354
dif_repiters	-4.098	-1.046	0.387
dif_teacher explain most of time	-2.141	-1.548	0.115
dif_doing exercises at class	2.263	0.946	0.371
dif_individual work at class	1.333	0.842	0.389
School ownership	0.955	1.771	**0.068**

OLS results after 1,000 bootstrap samples

- There is a 'teacher effect'. Efficiency channels are difficult to find out.
- It seems than in public schools the efficiency gap between classrooms tends to be higher.
- The reason is possibly due to the fact that most of teachers in public schools are civil servants that cannot be fired because of poor results.
- Results are robust when only diff_eff > 5 points schools are considered.

Conclusions

- Endogeneity is a well-known problem in Education Economics.
- This work tries to alert DEA and efficiency practitioners that if the endogeneity problem is present then efficiency analysis could be biased.
- We suggest to explore the use of causal inference in non-parametric analysis in order to overcome endogeneity problems through identification strategies.
- The "teacher effect" exists, but it is not clear what observables are behind this effect. (Hanushek et al. 2005)
- Schools efficiency differences must be analyzed in depth to take rational decisions.

WORKSHOP ON EFFICIENCY IN EDUCATION

Lancaster University Management School
The Work Foundation

London, September 19 – 20, 2014

"The teacher effect: an efficiency analysis from a natural experiment in Spanish primary schools"

Daniel Santín
dsantin@ccee.ucm.es
Gabriela Sicilia
gabriels@ucm.es
Complutense University of Madrid (Spain)