ABSTRACT
By using stochastic frontier analysis, single- and multiple-output methods, this paper empirically investigates the efficiency of Italian regions in providing public primary and secondary education for the period 2011–2018. The analysis detects strong interregional differences, outlining a clear North–South geographical pattern. To single out the reasons for the poor performance of Southern regions, context variables (per capita gross domestic product, poverty, institutional quality, adult education) are considered and shown to be highly relevant in shaping regional efficiency. Finally, when interregional disparities in socio-economic factors are accounted for, no residual geographical pattern in regional efficiency emerges.

KEYWORDS
efficiency; education; stochastic frontier analysis; Italy

INTRODUCTION
In the last few years, the role of public support in the formation and accumulation of human capital, and the importance of an efficient use of resources devoted to education, have been increasingly recognized in the political and academic debate. The focus of research has been primarily concentrated on the performance of single schools and universities, but also the relative efficiency of entire educational systems has been evaluated through comparisons among countries, or even regions and districts within a single country.

Interregional comparisons are especially motivated in the case of strongly heterogeneous outcomes across different subnational areas (Rodriguez-Pose & Tsakelos, 2011). In this respect, Italy is certainly a case in point for at least two reasons. First, in Italy interregional differences in students’ ability scores are among the largest in the developed world (Agasisti & Cordero-Ferrera, 2013), with some regions performing as well as countries at the top of the Organisation for Economic Co-operation and Development’s (OECD) ranking and others doing much worse. The variability of students’ performance seems to match the diversity in social and economic conditions between the more affluent Northern regions and the relatively poor South of the country, highlighting the link between context variables and educational outcome, recognized at least since the Coleman et al. (1966) report. However, while the geographical socio-economic divide is certainly a reason for the observed heterogeneity in student skills, it still remains unclear whether other factors connected to differences in teacher ability and commitment, and managerial and political choices contribute to account for (part of) interregional variability.

Second, given that per capita public expenditure on schools in Southern regions is in general not lower (and sometimes higher) than in the North, a dualism in terms of efficiency in education also emerges. This contributes to nurturing a long-standing discussion about the different costs of public goods and services across regional jurisdictions in Italy, and the distribution of administrative functions across levels of government, an issue which has been renewed by the recent request of additional competencies by some Northern regions, aiming at achieving greater efficiency and more fiscal autonomy. Again, a key unresolved question concerns the motivation behind the...
geographical efficiency divide, in particular whether lower efficiency in Southern regions is mainly caused by socio-economic conditions involving poorer abilities, learning capacity, etc. of children, and therefore inferior student performance, or it is due to worse teachers' quality, organizational setting and managerial skills.1

Despite the interest of the issue, the literature on the performance of the Italian educational system is rather limited, and especially the efficiency of public spending on education in Italian regions is a relatively little explored topic. This paper, aimed at partially filling this gap, carries out an evaluation of the efficiency of Italian regions in providing public primary and secondary education over the years 2011–18. In particular, the specific targets of the paper are: (1) assessing the width of the efficiency gap between Northern and Southern regions of Italy; (2) evaluating the weight of the socio-economic context in determining worst students' performance and lower efficiency in the South; and (3) deriving implications for educational policies suitable to deal with deep regional diversity.

To perform this task, the approach of efficiency frontiers is adopted. This approach is based on the idea that the best-performing schools maximize their output (students' knowledge, skills, abilities) for any given amount of expenditure (on teachers' salaries, books, computers, infrastructure, etc.) – or alternatively minimize costs for a given output – while others are inefficient, producing less than the largest output feasible with available resources. Inefficiency may be determined by the behaviour and choices of teachers, directors and managers (discretionary factors), as well as by economic, cultural and social context factors defining students' abilities, over and above efforts and commitment of the agents directly involved in the teaching/learning process. By exploiting information from a rich dataset provided by territorial public accounts (CPT), the Italian Institute for Evaluation of the Educational System (INVALSI) and other sources mentioned below, regional efficiency (or inefficiency) is first assessed through a stochastic frontier analysis (SFA), in a one-input one-output framework by using the true fixed effect (TFE) model (Greene, 2005a, 2005b). We then allow for multiple outputs (student scores at primary, lower secondary and upper secondary level, or student scores in mathematics and reading, and university enrolment rates) and inputs by adopting the approach of distance functions and using the stochastic ray production function (SRPF) model (Löthgren, 1997, 2000). For further robustness, the analysis is replicated by employing a semi-parametric version (Simar & Wilson, 2007) of data envelopment analysis (DEA).

The analysis documents the presence of large and non-decreasing variability in students’ performance and efficiency of public spending on education across Italian regions. A clear geographical pattern emerges, with Southern regions characterized by substantially lower outcomes in terms of both student performance and school efficiency. The socio-economic context (and in particular variables such as per capita income, education rate, family poverty and institutional quality) comes out to play a decisive role in determining regional differences. In particular, per capita gross domestic product (GDP), as a proxy of the level of regional development, proves to exert a significant impact on the efficiency of educational systems. However, social and cultural variables (poverty and adult education rates, institutional quality) turn out to be the most important determinants of school efficiency, emphasizing that the cultural conditions of students living in a deteriorated environment play a major role in shaping the relative inefficiency of Southern schools. Finally, the econometric exercise points out that once context determinants are properly accounted for, no residual inefficiency can be attributed to school geographical location (i.e., being in a Southern region) so that teachers’ and school managers’ choices and behaviour seem to play a minor role.

The paper is organized as follows. The next section is devoted to a short survey of related literature. Afterwards, some descriptive evidence on public expenditure and student performance in Italian regions during the period 2011–18 is supplied. The following sections describe the empirical investigation, presenting the methods and data, the inputs and outputs employed in regressions, and the results. The most important conclusions are finally summarized. Appendix A in the supplemental data online collects tables comparing student scores across countries and Italian regions.

RELATED LITERATURE

There is an extensive literature on efficiency in education and related issues, reviewed by several comprehensive surveys, such as De Witte and López-Torres (2017) and Johnes (2015). Based on different theoretical frameworks and methods to compute or estimate efficiency, most of the papers share the common idea that education can be viewed as the output of a production process employing variables related to student, family, school and community features as production inputs.

Since defining the output in the production of education is difficult and somehow arbitrary, many different indicators have been considered in the literature. Some studies (Alexander et al., 2010) focus on attendance, passing and dropout rates; some (Agasisti & Johnes, 2015) on the number of graduates; others (Nazarko & Šaparuskas, 2014) look at research activities and publications. Proxies of job market success, for example, graduate employability, starting salary (Johnes, 2013) and student satisfaction (Mainardes et al., 2014), are also taken into account. Overall, the variable most frequently used to represent the educational output is a proxy of student ability, usually the score that students obtain in standardized tests (according to De Witte & López-Torres, 2017, student score is used in 127 published papers between 1977 and 2015).

The choice of the variables to use as inputs is important and delicate as well. De Witte and López-Torres (2017) single out four types of input, related to student, family, institution and community characteristics. At the student level,
psychological and demographic features as well as personal abilities are considered, for example, by Cordero-Ferrera et al. (2017) and Podinovski et al. (2014). The socio-economic status, parental education and family structure are included among family-related inputs by Kirjavainen (2012) and Mancebón et al. (2012). Educational resources and school expenditure are taken into account as institution-related determinants of the educational output (Crespo-Cebada et al., 2014), while adult education rate and competition (represented, for example, by the number of educational institutions in the area) are common indicators of community characteristics (Nazarko & Šaparauskas, 2014).

Another taxonomy (partially overlapping the previous one) distinguishes between discretionary and context (or environment) variables. Discretionary variables include all the costly resources (i.e., expenditure on education for books, reading materials, computers, places to study, infrastructure, teachers, etc.) used in the school production process to optimize the educational outcome, and amenable to decisions and control by school directors and managers. Context variables not dependent on teachers' and managers' behaviour are connected to structural, institutional and socio-economic factors, and affect efficiency as well, by conditioning students' abilities and performance over and above the efforts and commitment of teachers and directors. The literature has considered a large number of these factors, which are both macro-variables, such as GDP (Cordero-Ferrera et al., 2017; Zoghbi et al., 2013), poverty (Houck et al., 2010; Millimet & Collier, 2008) and adult education rates (De Witte & Kortelainen, 2013; Zoghbi et al., 2013), and micro-variables, such as students’ ethnicity (Crespo-Cebada et al., 2014; De Witte & Kortelainen, 2013) and the family’s socio-economic status (Zoghbi et al., 2013), confirming their importance in determining the educational outcome.

The debate on the Italian case is characterized by the striking gap between Southern and Central-Northern regions observed in both student performance and school efficiency, which closely parallels the historical North–South socio-economic dualism. Checchi (2004) is among the first to document significant regional disparities in student scores and pass/fail rates. He shows that differences are strongly correlated to socio-economic and cultural indicators, and little connected to resource availability and organizational factors. While a similar result is obtained by Quintano et al. (2009), arguing that territorial differences in student scores are mainly determined by socio-economic variables, Bratti et al. (2007) find a significant positive correlation between mathematical literacy, the expenditure in infrastructure (equipment and buildings), and the quality of school management across Italian provinces. In the same vein, Agasisti and Vittadini (2012) recognize the weight of individual, school-level and regional factors, and conclude that together with socio-economic conditions and the cultural climate, school-level factors (teaching quality, leadership, availability of instructional materials) play an important role in explaining the geographical variability in educational results.

Regarding more specifically efficiency, the literature on the Italian educational system is relatively scant, as recalled above. Nonetheless, some contributions document a remarkable interregional heterogeneity, with Northern regions largely outperforming the Southern ones. The investigation of Sibiano and Agasisti (2013) detects noteworthy territorial diversity, explaining it mainly on the basis of differences in per capita regional GDP. Di Giamo and Pennisi (2015) confirm the existence of an efficiency gap between schools located in Centre–North and South, for both primary and lower-secondary education, and ascribe differences primarily to teacher quality and experience, and school organizational characteristics. More specific aspects are investigated by: Barbetta and Turati (2003), dealing with the impact of schools’ ownership structure (public, private for-profit and non-profit) on efficiency; Agasisti (2013), studying the effect of competition on the efficiency of Italian high schools; Agasisti et al. (2014), distinguishing between the impact of external factors (socio-economic background and location in urban or rural areas) and school management, and arguing that the former is more important in shaping overall efficiency. Finally, more recently Ferraro et al. (2020) focus on the efficiency of Italian local governments in producing education ancillary services (mainly meals and school transportation).

Summarizing, while a clear North–South pattern in student competence and school efficiency emerges, matching the socio-economic Italian geographical divide, there is no unanimity on the educational dualism, whose motivations remain to be better clarified, in particular with regard to the importance of the socio-economic context and the nature of variables ultimately defining the Southern disadvantage. A purpose of this paper is to shed some more light on this issue.

PUBLIC EXPENDITURE ON EDUCATION AND STUDENT PERFORMANCE IN ITALIAN REGIONS

The Italian educational system is strongly based on public schools and universities. For primary and secondary education, public funds account for almost 95% of total spending on education (OECD, 2020), which between 2015 and 2018 amounted to around €50 billion per year (ISTAT, 2020s). In comparison with other advanced countries, the latter figure is relatively low; as a percentage of GDP, public expenditure on primary and secondary education is close to 2.8%, versus 3.2% of the OECD average and more than 3.5% of most European countries, that is, Belgium, Denmark, Finland, France, Iceland, Norway, Portugal, Sweden and the UK (OECD, 2020). In addition, outlays are mainly destined to pay teachers’ salaries (almost 90% of total costs), with only a minor share of resources devoted to investment, infrastructure and personnel training.

Regarding the regional distribution of resources, per capita public expenditure on primary and secondary education” is rather uniform throughout the country. Looking
at the period considered in this paper, in 2011 the North-East benefited from the highest levels of per capita expenditure (around €5600), the South and Islands were at €5300, and the Centre and North-West at about €5100. By 2015, expenditure rose to about €5500 in the South and Islands, and declined slightly below €5000 in the rest of the country.

The overall performance of the Italian educational system looks rather unsatisfactory. A comparison of PISA mean scores of Italian students with those obtained in other countries reveals that the former largely underperform in both reading and mathematical abilities. Table A1 in the supplemental data online shows that between 2000 and 2018, Italian students recorded mean scores for reading abilities between 469 and 490, against OECD average scores ranging between 493 and 501. Similarly, the mean scores in mathematics skills for Italian students were between 462 and 490, whereas OECD averages varied between 491 and 499. More generally, the evidence reported in Table A1 highlights that the Italian primary and secondary schools rank in last places among the 31 countries considered, showing a striking gap with respect to leading Asiatic countries (Japan, Korea and Hong Kong) and noteworthy differences with Northern European countries.

Educational attainments are also characterized by a large interregional heterogeneity. Although common to many other OECD countries, regional differentiation is particularly evident in Italy, paralleling the deep North–South economic and social dualism of the country. Exploiting the subnational detail of the PISA dataset, Table A2 in the supplemental data online supplies information on macro-regional mean scores. The comparison of data in Tables A1 and A2 reveals that in Northern Italy regional mean scores are close to those of most European countries (France, Germany, the Netherlands and Sweden), while scores recorded in the South and Islands are lower than those of other European peripheral areas (Greece, Spain and Portugal).

Data supplied by INVALSI convey further information consistent with the PISA assessment. Table A3 in the supplemental data online shows the mean ability scores for Italian language and mathematics obtained by students of primary, lower secondary and upper secondary schools (precisely by students attending the 5th, 8th and 10th grades) by regions in the last available wave (2018), as well as the regional ranking for each single evaluation. An inspection of Table A3 corroborates the hypothesis that students’ performance follows a geographical pattern, with Northern regions showing in most cases the highest scores, especially for secondary school. A very similar pattern emerges when examining data of the previous waves (data not shown, but which are available from the authors upon request).

Summing up, the data analysed in this section highlight steadily worse educational results for Southern Italy, together with values of per capita expenditure not lower than the Centre–North and increasing from 2014. This evidence points to a likely problem of relative inefficiency of schools located in Southern regions. The following econometric exercise is devoted to investigating this issue and singling out the possible determinants of inefficiency of the Southern educational system.

**ASSESSING EFFICIENCY: METHODS**

Efficiency of educational institutions and systems is frequently investigated by adopting the approach of efficiency frontiers, which are either derived from non-parametric mathematical optimization models, such as DEA or free disposal hull (FDH), or estimated through parametric methods, such as the SFA, based on the assumption that the relationship between input(s) and output(s) has a given functional form (Aigner et al., 1977). By employing the latter methodology, we evaluate the efficiency (or inefficiency) of regional educational systems in Italy first through a one-input one-output framework by using the TFE model (Greene, 2005a, 2005b). To allow for multiple inputs and outputs (student scores at primary, lower secondary and upper secondary level; in mathematics and reading), we then adopt the approach of distance functions using the SRPF model (Löthgren, 1997, 2000). For further robustness, the analysis is also replicated by employing a semi-parametric version (Simar & Wilson, 2007) of DEA. It is worth noting that in all cases our decision-making units (DMUs) are not individual schools but aggregate administrative entities (Regions). This choice is obligatory owing to the lack of data on costs at the school level for the whole period under investigation. We are aware that it involves the risk of omitting substantial information on school heterogeneity within regions.

The TFE model simultaneously estimates the stochastic frontier and the inefficiency model to explain the effects of context determinants on inefficiency. By taking advantage of the panel dimension of the dataset, the model can account for time-varying inefficiency and unobserved regional specific time-invariant heterogeneity. The stochastic frontier is specified starting from the equation:

\[ y_{rt} = \alpha_r + \beta x_{rt} + \varepsilon_{rt} \]

where \( y_{rt} \) is the regional educational output at time \( r \); \( \alpha_r \) is the fixed unobservable effect of region \( r \); \( x_{rt} \) is a vector of discretionary inputs; and \( \beta \) the vector of coefficients to be estimated. The error term \( \varepsilon_{rt} \) is specified as:

\[ \varepsilon_{rt} = \psi_{rt} - u_{rt} \]

\[ \psi_{rt} \sim N(0, \sigma^2_{\psi}) \]

\[ u_{rt} \sim N^\perp[\mu_{rt}, \sigma^2_{u}] \]

where \( \psi_{rt} \) is the truly idiosyncratic error term (statistical noise); \( u_{rt} \) is the technical inefficiency term to be estimated; \( \mu_{rt} \) is a vector of context variables (including a constant term) affecting mean inefficiency; and \( \gamma \) a vector of unknown coefficients. Parameters \( \beta \) and \( \gamma \) are simultaneously estimated by applying a simulated maximum likelihood (ML) estimator; in particular, by parameterizing the mean of the pre-truncated inefficiency
distribution, the model estimates \( u_{rt} \) as a truncated (at zero) normal random variable with mean \( \gamma z_{rt} \) and variance \( \sigma_u^2 \) (Battese & Coelli, 1995).

The distance function is the most common method to estimate a production function with multiple outputs and inputs. The SRPF proposed by Lõhberg (1997, 2000) can be seen as a particular specification of the output distance function, allowing for the possibility that some output quantities are zero (Henningsen et al., 2015, 2017). Starting from the classical Shephard output distance function:

\[
D_0(x, y) = \min \{ \lambda > 0 \mid (y/\lambda, x) \in T \}
\]

where \( x = (x_1, x_2 \ldots x_N) \) is the vector of \( N \) inputs; \( y = (y_1, y_2 \ldots y_M) \) is the vector of \( M \) outputs; and \( T \) is the technology set, the SRPF is based on a polar-coordinate representation of the output vector by which the output \( y \) is expressed through its magnitude \( ||y|| \) and its direction \( \rho(\theta) \), so that:

\[
y = ||y|| \cdot \rho(\theta), \quad ||y|| = \left( \sum_{m=1}^{M} y_m^2 \right)^{1/2};
\]

\[
p_m(\theta) = \cos \theta_m \prod_{j=0}^{m-1} \sin \theta_j \quad \text{for} \quad m = 1, \ldots, M
\]

where \( \gamma ||y|| \) represents the Euclidean distance of the vector \( y \); and \( p_m(\theta) \) is a vector of directional measures depending on the polar-coordinates \( \theta = (\theta_1, \theta_2, \ldots, \theta_{m-1}) \). Rewriting the Shephard output distance function by replacing \( y \) with \( ||y|| \cdot \rho(\theta) \), using the linear homogeneity property and rearranging, we obtain:

\[
D_0(x, y) = D(x, \ ||y|| \cdot \rho(\theta)) = ||y|| \cdot D(x, \rho(\theta))
\]

By taking logarithms of both sides of (5), defining the inefficiency term as \( u = -\ln D(x, y) \), the SRPF as \( f(x, \theta) = -\ln D(x, \rho(\theta)) \), and adding a random error term \( v \), the final formulation becomes:

\[
\ln ||y|| = \ln f(x, \theta) - u + v
\]

Following Lõhberg (1997), a translog functional form can be assumed for estimation (6), in order to estimate the panel data SRPF model:

\[
\ln ||y_{rt}|| = \beta_0 + \sum_{m=1}^{M-1} \beta_m x_{mrt} + \frac{1}{2} \sum_{m=1}^{M-1} \sum_{j=1}^{M-1} \beta_{mj} x_{mrt} x_{jrt} + \sum_{m=1}^{M} \beta_m \ln(x_{mrt}) + \frac{1}{2} \sum_{m=1}^{M} \sum_{j=1}^{M} \beta_{mj} \ln(x_{mrt}) \ln(x_{jrt}) + \sum_{m=1}^{M-1} \sum_{n=1}^{N-1} \beta_{mn} \theta_m \ln(x_{mrt}) - u_{rt} + v_{rt}
\]

where \( v_{rt} \sim N(0, \sigma_v^2); u_{rt} \sim N(0, \sigma_u^2); \quad \mu_u = \gamma z_{rt}; \quad \sigma_u \) is the standard deviation of the error term.

ASSESSING EFFICIENCY: INPUTS AND OUTPUTS

The variable \( EXP \) (real regional per capita public expenditure on education in constant 2010 prices) is one of the discretionary inputs (or the only input) considered as determinants of educational performance and efficiency.
in all estimations presented below. When more than one input is allowed for (i.e., when a distance function model is adopted), the variable STRU, accounting for the infrastructure endowment, is also included among the inputs of the educational production function.

Concerning the choice of context variables, we consider first of all the real regional per capita GDP in constant 2010 prices (GDP), which we expect to affect positively educational output and efficiency, as found by many authors (e.g., Cordero–Ferrera et al., 2017; Zoghbi et al., 2013). There are several reasons to believe that GDP is important: on the one hand, more productive and richer regions are likely to enjoy better educational facilities, a larger supply of private and public cultural services and in general a livelier cultural environment; on the other, being attracted by better environments, more capable managers and teachers preferring to work in richer regions and better endowed schools. As an alternative, we consider the possibility that school efficiency be affected not (only) by income but the more comprehensive regional socio-economic and cultural context. To account for the latter, we resort to indicators of regional poverty, and regional institutional quality, as evaluated by the Charron et al. (2013, 2019) European index of quality of government (EQI). The impact of POVE on efficiency is expected to be negative, consistent, for example, with Houck et al. (2010) and Millimet and Collier (2008), since affluent families are willing to invest more in child education and can afford higher educational private expenses. Concerning EQI, a higher institutional quality is likely to have multiple positive effects on school efficiency: for example, good institutions encourage investment in human capital (Eicher et al., 2006; Niño et al., 2017) and thus stimulate students’ and parents’ efforts; also, institutional quality involves broader merit recognition of students’ commitment and teachers’ competence, whence better educational results and more efficiency are engendered; finally, high levels of EQI imply for themselves better functioning of government operations so as to impact positively on the capability of the public sector to produce educational services efficiently.

Moreover, as documented by an abundant literature (De Witte & Kortelainen, 2013; Zoghbi et al., 2013), student performance and school efficiency are likely to be correlated to the educational level of the adult population EDU (measured by the share of adult population with a university degree or higher qualification), because higher educated parents are more aware of the importance of education; they are careful to motivate their children and inclined to demand better educational services. The variable STUD, here measured by the ratio of the number of enrolled students to the number of available classrooms, is also frequently considered in the literature (Agasisti, 2014; De Witte & Kortelainen, 2013). A priori, it has an ambiguous effect on efficiency because, on the one hand, ‘crowded’ classrooms can involve less attention of teachers to students (De Witte & Hudrlíková, 2013), and therefore imply a negative impact on students’ performances, while, on the other, larger classes lead to substantial cost saving (fewer teachers and facilities per student), and even positive peer effects from gender and race diversity (Hoxby, 2000).

To account for the outputs of the educational process, we consider the students’ ability, as measured by INVALS1 ability scores, and the regional university enrolment rates ENROL, defined as the ratio of people enrolled at university to the regional population aged 19–25 years. INVALS1 regional ability score is considered first as a unique measure calculated as the yearly regional weighted average of scores recorded in mathematics and reading at primary, lower secondary and upper secondary schools. Then, when allowing for multiple outputs, three distinct regional average scores for primary, lower secondary and upper secondary schools, or alternatively two different average scores for mathematics and reading, are employed. The regional university enrolment rate is also considered as an output, since the ability of school graduates to attend tertiary education courses is deemed to be an indicator of the good quality of the school educational process.

Table 1 reports data sources and main statistics of the variables used in the empirical investigation. Information about regional public expenditure on primary and secondary education and students’ ability scores is retrieved, respectively, from the CPT and INVALS1. Concerning infrastructure and context variables, data are drawn from ISTAT’s (2020b) database. The indicator of regional institutional quality is taken from Charron et al. (2019).

**RESULTS: TFE ESTIMATION**

Regional efficiency is first assessed through a TFE model, assuming a one-input (real regional per capita public expenditure on education EXP), one-output (INVALS1 regional average student scores) production function. Estimates are shown in Table 2. The suitability of this approach is tested by calculating for each estimated specification the ratios $\lambda = \sigma_b / \sigma_u$ (i.e., inefficiency standard deviation over error standard deviation), which are always significantly different from 0, and $\psi = \sigma_b^2 / (\sigma_u^2 + \sigma_b^2)$ (i.e., inefficiency variance over total variance), which are always very high (> 0.9). Diagnostics support the hypothesis that a relatively large amount of total variability is explained by inefficiency, and therefore justify the use of the TFE model. In addition, the appropriateness of this model is checked versus an alternative random-effect model by performing a likelihood ratio test (statistics not shown); since likelihood ratio (LR) $\chi^2$ is always highly significant, the fixed effect model is confirmed to be the preferred one. Finally, to take into account possible changes in technology, time effects are included among regressors. Since diagnostics indicates that the coefficients of year dummies are always insignificant, they are not reported in Table 2.

The upper panel of Table 2 concerns the efficiency frontier estimation, evaluating the productivity of the discretionary input, that is, $\beta$ of equation (1). The estimated value of $\beta$
Table 1. Variables employed: data sources and descriptive statistics.

| Source            | Mean   | SD    | Minimum | Maximum | Observations |
|-------------------|--------|-------|---------|---------|--------------|
| **Production function variables** |        |       |         |         |              |
| INVALSI score     | 3.000  | 0.840 | 0.861   | 4.185   | 160          |
| EXP – real per capita expenditure (thousands of euros at 2010 prices) | 5.120  | 0.810 | 3.828   | 7.968   | 160          |
| STRU – infrastructure equipment rate (%) | 58.59  | 6.23  | 42.60   | 73.46   | 160          |
| ENROL – university enrolment rate (%) | 40.11  | 6.69  | 22.75   | 55.23   | 160          |
| **Context variables** |        |       |         |         |              |
| GDP – real per capita GDP (thousands of euros at 2010 prices) | 25.29  | 6.49  | 15.31   | 39.31   | 160          |
| POVE – family poverty index (%) | 12.12  | 7.79  | 2.90    | 35.30   | 160          |
| EQI – European quality of government index | −1.01  | 0.75  | −2.37   | 0.70    | 160          |
| EDU – adult education rate (%) | 10.96  | 1.85  | 7.60    | 17.40   | 160          |
| STUD – student per classroom ratio | 20.48  | 1.34  | 17.14   | 22.81   | 160          |

Note: INVALSI, Italian Institute for Evaluation of the Educational System.

Table 2. Estimating the determinants of regional inefficiency: true fixed effect (TFE) estimates (Greene, 2005a, 2005b).

| Variables | (1) | (2) | (3) | (4) | (5) |
|-----------|-----|-----|-----|-----|-----|
| **Efficiency frontier model** |       |     |     |     |     |
| EXP       | 0.1123** | 0.1356** | 0.1223*** | 0.1312*** | 0.1405*** |
|           | (0.0479) | (0.0539) | (0.0503) | (0.0511) | (0.0501) |
| **Inefficiency model** |       |     |     |     |     |
| GDP       | −0.1051*** | 0.0257 |     |     | 0.0242 |
|           | (0.0309) | (0.0348) |     |     | (0.0362) |
| POVE      | 0.0354*  | 0.0266* | 0.0337** | 0.0384** |
|           | (0.0190) | (0.0140) | (0.0160) | (0.0173) |
| EQI       | −0.4176** | −0.3214** | −0.4092* | −0.4564* |
|           | (0.2023) | (0.1532) | (0.2380) | (0.2388) |
| EDU       | −0.1849*** | −0.1695*** | −0.2046*** | −0.2105*** |
|           | (0.0610) | (0.0559) | (0.0650) | (0.0648) |
| STUD      | −0.1367*  | −0.1496* | −0.1839** | −0.1664* |
|           | (0.0771) | (0.0778) | (0.0842) | (0.0865) |
| Regional status dummy | 0.7024*** | 0.0830 | 0.1334 |     |     |
|           | (0.2602) | (0.2732) | (0.2790) |     |     |
| Centre dummy |     |     |     | 0.0408 | 0.1102 |
|           |     |     |     | (0.4455) | (0.4683) |
| South and Islands dummy |     |     | −0.3025 | −0.1376 |     |
|           |     |     |     | (0.5300) | (0.5640) |
| Constant  | 2.2007*** | 3.3136 | 4.2344** | 5.2687*** | 4.2131* |
|           | (0.5162) | (2.1973) | (1.9395) | (2.0431) | (2.5023) |
| $\lambda = \sigma_\alpha / \sigma_\nu$ | 3.405*** | 3.313*** | 3.122*** | 3.141*** | 3.219*** |
|           | (0.076) | (0.052) | (0.050) | (0.055) | (0.057) |
| $\psi = \sigma_\alpha^2 / \sigma_\nu^2 + \sigma_\nu^2$ | 0.921  | 0.916  | 0.907  | 0.908  | 0.912 |
| Observations | 160 | 160 | 160 | 160 | 160 |

Note: Standard errors are shown in parentheses; significance levels: ***p < 0.01, **p < 0.05, *p < 0.1. Since Italian Institute for Evaluation of the Educational System (INVALSI) tests are administered in the spring, one-year lagged values of EXP are used.
is stable and always significantly positive, showing that an increase in regional per capita expenditure on education positively impacts on the INVALSI score achieved by students resident in the region. More importantly, in the lower panel of Table 2 (inefficiency model), TFE estimates highlight the relevance of context variables to explaining regional inefficiency.\(^{17}\) The benchmark parsimonious specification of column 1 considers only two context variables: regional per capita GDP and regional constitutional status. Consistent with the extant literature, GDP turns out to affect inefficiency negatively at a high level of statistical significance: not surprisingly, richer regions tend to perform more efficiently. The constitutional status of regions is considered by introducing a dichotomous variable taking a unit value when the region has a special (rather than ordinary) status,\(^{18}\) and is therefore entitled to enjoy broader spending powers and retain a larger share of its own fiscal revenues. Table 2, column 1, shows that the constitutional status dummy has a positive, strongly significant impact on inefficiency, probably because of the relatively high levels of per capita expenditure allowed by larger fiscal autonomy,\(^{19}\) which may correspond to students’ performance not so good as to offset larger outlays.

The relationship between regional per capita GDP and school efficiency shown in column (1) might be partly due to possible correlation between GDP and other variables which in the parsimonious benchmark specification are omitted. Therefore, exploring the conjecture that beside purely economic variables, even social and distributional factors can be crucial in determining students’ performances and school efficiency, an alternative specification is estimated with some additional regressors (column 2). In particular, indicators of regional poverty (POVE), measured by the share of families living in conditions of relative poverty, regional institutional quality (EQI), evaluated by the EQI, and adult education (EDU), that is, the share of adult population with a university degree or higher qualification, are considered. Moreover, the specification of column (2) is also enriched by the ratio STUD, that is, the number of enrolled students divided by the number of available classrooms, measuring class size.

Results reported in column (2) confirm that when social and distributional factors are included among explanatory variables, per capita GDP becomes irrelevant to determining school efficiency. Instead, poverty, bad quality of institutions and poor adult education do matter, contributing to deteriorate student performance and school efficiency. The coefficient of the variable STUD takes a negative sign (i.e., efficiency increases with class size), even if with little statistical significance.

Very similar results are obtained when GDP is excluded from the set of right-hand-side variables (columns 3 and 4) and geographical dummies are added (columns 3–5). While the estimated impact of EQI is slightly reduced, the significance of the other explanatory variables is unaltered. Interestingly, the regional constitutional status dummy loses significance, while geographical dummies for Southern and Central regions do not seem to exert any impact on the efficiency of regional educational systems, in

\(^{17}\) Valentina Chiariello et al.
both specifications with and without GDP (columns 4 and 5). The latter point deserves attention, since it emphasizes that once properly accounted for context variables such as local institutional quality, poverty and adult education, no residual inefficiency connected to being located in Southern regions is detected. This evidence shows that

### Table 3. Estimating the determinants of regional inefficiency: stochastic ray production function (SRPF) estimates (Löthgren, 1997).

| Variables  | (1)          | (2)          | (3)          | (4)          |
|------------|--------------|--------------|--------------|--------------|
| Inefficiency model |              |              |              |              |
| POVE       | 0.0037***    | 0.0038***    | 0.0027***    | 0.0029***    |
|            | (0.0012)     | (0.0014)     | (0.0012)     | (0.0014)     |
| EQI        | −0.0267**    | −0.0257**    | −0.0299***   | −0.0314***   |
|            | (0.0112)     | (0.0111)     | (0.0083)     | (0.0081)     |
| EDU        | −0.0096***   | −0.0118***   | −0.0110***   | −0.0112***   |
|            | (0.0031)     | (0.0033)     | (0.0023)     | (0.0024)     |
| STUD       | 0.0017       | 0.0020       | 0.0035***    | 0.0035***    |
|            | (0.0014)     | (0.0016)     | (0.0011)     | (0.0012)     |
|            |              |              |              |              |
| Efficiency frontier model |              |              |              |              |
| EXP        | 4.3075**     | 4.2913***    | 2.3843*      | 2.8938**     |
|            | (2.0024)     | (2.0283)     | (1.3690)     | (1.4582)     |
| STRU       | 9.4430***    | 3.1824*      | 2.3843*      | 2.8938**     |
|            | (2.9801)     | (1.8581)     | (1.3690)     | (1.4582)     |
| θ₁         | 36.973       | 28.923       | 5.788        | −0.892       |
|            | (30.093)     | (30.574)     | (14.125)     | (15.434)     |
| θ₂         | 43.694       | 60.331**     | 7.705        | 23.408       |
|            | (28.479)     | (29.186)     | (48.618)     | (47.390)     |
| θ₃         | −16.131      | −15.775      |              |              |
|            | (25.478)     | (25.316)     |              |              |
| STRU*STRU  | 0.166        | 0.433*       |              |              |
|            | (0.236)      | (0.225)      |              |              |
| θ₂*θ₂      | −33.135***   | −44.350***   | −25.103      | −34.293      |
|            | (11.156)     | (11.760)     | (47.689)     | (45.850)     |
| EXP*STRU   | 0.021        | 0.119*       |              |              |
|            | (0.081)      | (0.065)      |              |              |
| EXP*θ₁     | −2.134***    | −2.174***    | −1.248***    | −1.527***    |
|            | (0.748)      | (0.818)      | (0.439)      | (0.431)      |
| EXP*θ₃     | −0.961       | −2.092***    |              |              |
|            | (0.754)      | (0.802)      |              |              |
| STRU*θ₁    | −3.880***    | −1.514***    |              |              |
|            | (1.152)      | (0.676)      |              |              |
| STRU*θ₂    | −4.049***    | −3.776***    |              |              |
|            | (1.140)      | (1.854)      |              |              |
| Constant   | −39.097      | −60.473      | −3.288       | −11.566      |
|            | (42.951)     | (42.908)     | (28.250)     | (27.392)     |
| λ = σₓ/σᵧ  | 3.592***     | 3.313***     | 3.590***     | 3.464***     |
|            | (0.073)      | (0.052)      | (0.101)      | (0.056)      |
| ψ = σₓ²/σᵧ² | 0.928        | 0.916        | 0.928        | 0.923        |
| LLF        | 446.376      | 454.689      | 442.504      | 447.190      |
| Observations | 160          | 160          | 160          | 160          |

Note: Standard errors are shown in parentheses; significance levels: ***p < 0.01, **p < 0.05, *p < 0.1. Iterative maximum likelihood (ML) estimations are carried out by the R package ‘Frontier: Stochastic Frontier Analysis’ version 1.1-8 (Coelli & Henningsen, 2020). The dependent variable is logged. Estimated coefficients of second-order terms EXP*EXP, θ₁*θ₁, θ₂*θ₂, EXP*θ₁, STRU*STRU, θ₁*θ₂, θ₁*θ₃ and θ₂*θ₃ are statistically insignificant and therefore not reported. LLF, log-likelihood function.
(at least at an aggregated level) socio-economic conditions play a key role in shaping regional efficiency, while disciplinary factors such as teacher and managers’ quality and behaviour cannot be deemed a source of heterogeneity in the efficiency of regional educational systems.

Figure 1 displays the dynamic path of regional efficiency scores as estimated by TFE throughout the period 2011–18.20 It shows an evident and persistent geographical divide, with most of Northern and Central regions being highly efficient and Southern regions performing much worse. In the light of the results reported in Table 2, this evidence should be mainly ascribed to the much worse socio-economic environment in the South which negatively conditions students’ abilities and performances.

### RESULTS: SRPF ESTIMATION

A possibly serious drawback of estimates presented above is that the TFE model is designed to handle single-output production functions. In order to enrich the analysis by considering multiple outputs, the approach of distance functions is here adopted by implementing the SRPF model. In this case, instead of using an aggregate regional mean score as unique indicator of the production of educational services, we first employ three distinct regional average INVALSI scores, respectively, for primary, lower secondary and upper secondary schools; then, we consider as distinct outputs the average scores obtained in mathematics and reading. Moreover, as an additional output, the regional university enrolment rate is considered, which is a measure of the ability of schools to train people suitable to access tertiary education. Likewise, the number of inputs is increased as well, by including the infrastructure endowment $STRU$ beside the real regional per capita public expenditure on education $EXP$.

Table 3 collects the results of the ML estimation of four different specifications of equation (7). Estimations are carried out by using the R package ‘Frontier: Stochastic Frontier Analysis’ version 1.1-8 (Coelli & Henningsen, 2020). Column (1) refers to the specification considering as outputs three regional average scores (i.e., scores for primary, lower secondary and upper secondary schools) and the university enrolment rate $ENROL$. In column (2), the indicator of infrastructure endowment $STRU$ is also included among inputs. Estimations reported in columns (3) and (4) include two different average scores for mathematics and reading (respectively, without and with the input $STRU$). The statistical significance of some second-order terms of inputs and polar-coordinate angles (accounting for the output mix) suggests focusing on the unrestricted translog functional form of the production function. As in the case of the TFE model, the statistics $\lambda = \sigma_u/\sigma_c$ and $\psi = \sigma_u^2/(\sigma_u^2 + \sigma_c^2)$ testify that the inefficiency variance is a large share of total variance, confirming the validity of the approach adopted.

As shown in Table 3, the evidence supplied by SRPF estimates is fairly consistent with the results obtained by the TFE model. The regional per capita public expenditure on education and the infrastructure endowment are always positive and statistically significant, confirming
that they participate importantly in the production of the educational output. Concerning context variables, the role of poverty, quality of institutions and adult education in steering inefficiency is fully confirmed, stating once again that the environment decisively affects student performance and school efficiency. Unlike the TFE estimation, in this case the impact of class size (STUD) on inefficiency is positive (and statistically significant in columns 3 and 4); the partial inconsistency of TFE and SRPF estimation regarding STUD can be justified in the light of the ambiguous role attributed to this variable by the previous literature.

Finally, including GDP and constitutional and geographical dummies does not alter estimates, as the coefficients of these variables are always statistically insignificant. In conclusion, once again, if regression equations are specified to take into account the socio-economic disadvantage of Southern regions, no residual inefficiency emerges for Southern regions as a result of possibly omitted discretionary variables.

**RESULTS: SIMAR AND WILSON ESTIMATION**

As a further robustness check, the efficiency of educational systems in Italian regions is also evaluated by implementing the DEA approach. Once obtained DEA first-stage efficiency scores, the latter are regressed on a set of context variables (the same ones employed with the TFE model) by using the bootstrapped truncated estimation proposed by Simar and Wilson (2007). As shown in Table 4, results are quite consistent with the outcome of parametric methods used in previous sections. In particular, once again, in the first parsimonious specification, regional per capita GDP and regional constitutional status appear to be highly significant, with school efficiency increasing in GDP and lower in case of special status regions. Like in the case of TFE estimation, when social and distributional factors are considered as context variables (columns 2–5), GDP loses significance, while the indicators of regional institutional quality (EQI) and alternatively poverty (POVE) in specifications (2) and (3), or adult education rate (EDU) in specifications (4) and (5), turn out to be decisive in shaping school efficiency. The role of region constitutional status is confirmed even in specifications (2) and (3). More importantly, once again, the geographical location dummy for the South and the Islands remains insignificant (columns 4 and 5), confirming that efficiency in Southern regions is not lower than elsewhere, once the worse socio-economic conditions of those regions are controlled for.

The latter remark is corroborated by the evidence shown in Figure 2, where DEA first-stage efficiency scores, based on only the discretionary input (per capita expenditure), and Simar and Wilson predicted efficiency scores, taking into account context variables, are depicted. In line with Figure 1, efficiency scores describe a clear and persistent geographical pattern, with Northern regions
located either on the frontier or very close to full efficiency, Southern regions much less efficient, and Central regions at intermediate levels. Comparing DEA first-stage and Simar and Wilson predicted efficiency scores in Southern regions (especially Basilicata, Campania and Molise), the latter are often lower, showing that in those regions the socio-economic conditions predict a level of efficiency even lower than the one actually recorded by DEA scores.25

CONCLUSIONS

This paper carries out an empirical investigation on the efficiency of the Italian education system in the period 2011–18. As the first result, in accordance with the extant literature, a remarkable heterogeneity in educational outcomes and efficiency scores across Italian regions is found, with a clear geographical North–South pattern, matching the longstanding dualism of the Italian economy. Consistent with this evidence, the efficiency analysis reveals strong interregional differences, with Northern regions (especially Lombardy, Veneto and Emilia–Romagna) being either on the frontier or very close to full efficiency, Southern regions (in particular Calabria, Campania and Sicily) much less efficient, and Central regions at intermediate levels (slightly higher for Tuscany and Umbria and lower for Latium). Over time, the observed North–South duality looks to be remarkably persistent, with only minor fluctuations in mean scores.

To investigate the determinants of the poor efficiency performance of Southern regions, we consider a number of context variables accounting for differences in per capita GDP, family poverty, institutional quality and adult education. All these factors turn out to be important to explain diversity across regions. When interregional differences in context variables are appropriately accounted for, no longer geographical pattern in regional efficiency emerges, so that (at least at the regional level) other possibly relevant discretionary inputs of the education production function, such as teacher and school manager quality and organizational features, seem to play a minor role.

A sound conclusion of the analysis is that disparities in students’ abilities and school efficiency across Italian regions go hand in hand with the socio-economic North–South divide. The poor performance of the Southern educational system may result in human capital loss, impairment of future growth, and aggravation of the geographical socio-economic dualism. Breaking this vicious circle should be treated as a major task for policymakers, who may use education as a main driver to promote the recovery of lagging regions. Indeed, overcoming the low growth – poor education trap needs to reject any fatalistic vision for which better educational results cannot be achieved before changing the context conditions; on the contrary, development policies can profitably start from the educational system, by granting it the resources suitable to offset the influence of a penalizing social and cultural environment, and designing incentives to favour a strong commitment of all the actors involved in the educational process.

ACKNOWLEDGEMENTS

We thank Mariarosaria Agostino and two anonymous referees for their comments. We also thank the participants in the XLI annual Scientific Conference of the Italian Association of Regional Sciences (online conference, 2–5 September 2020) and the Workshop of the cMET05 University Centre for Applied Economic Studies (Prato, Italy, 9–10 September 2020) for discussing an earlier version of this paper.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

NOTES

1. At school level there are, of course, many other factors possibly leading to different educational outcomes, for example, school ownership, that is, private or public (Barbetta & Turati, 2003), degree of competition (Agasisti, 2013), teachers’ age (Bryson et al., 2020) and status, that is, tenured or not (Di Giacomo & Pennisi, 2015), share of disabled students (Agasisti & Vittadini, 2012), etc. Since these variables take on similar values in all Italian regions, they are not explicitly considered in the following analysis.

2. Per capita public expenditure on education is calculated as the ratio of public expenditure to the population aged 3–18 years.

3. The reference to macro-regions is common in the literature on the Italian geographical divide. Macro-regions are North–West (including Piedmont, Aosta Valley, Lombardy, Liguria), North-East (Veneto, Trentino–Alto Adige, Friuli Venezia Giulia, Emilia–Romagna), Centre (Tuscany, Umbria, Marche, Latium), South (Abruzzi, Molise, Campania, Apulia, Basilicata and Calabria) and Islands (Sicily and Sardinia).

4. PISA is the OECD Program for International Student Assessment launched in 2000. The PISA test measures and compares the skills in reading, mathematics and science of 15-year-old students. In 2018, mean scores in reading ranged from 340 (Philippines) to 555 (China), while mean scores in mathematics were included between 325 (Dominican Republic) and 591 (China).

5. The Istituto Nazionale per la Valutazione del Sistema educativo di Istruzione e di formazione (INVALSI) is the government agency charged with the assessment of the Italian educational system. Every year in spring, the mathematics and Italian language skills of all students are evaluated through a standardized test including multiple-choice questions and open-response items. INVALSI data refer to school
and class average marks (for multiple-choice questions) and overall ability scores.
6. Including the 2018 wave allows attenuating concerns of data reliability connected to the possibility of score manipulation due to student and teacher cheating, which can occur mainly through dishonest transcription of students’ responses on machine-readable answer sheets (Bertoni et al., 2013). Indeed, from 2018 answers from secondary schools’ students are collected and processed through an automatized and centralized process, ruling out any kind of influence by teachers and local administrators (INVALSI, 2018). Comparing regional scores in 2018 and previous years, we find that rankings remain basically unaltered and data variability across regions is even reduced, which seems to exclude the possibility that dishonest behaviour was significantly more widespread in the South than other areas.
7. The polar coordinates θ are recursively obtained by:
\[
\theta_m(y) = \cos^{-1}\left(\frac{y_m}{y} \prod_{j=0}^{m-1} \sin \theta_j\right)
\]
for \(m = 1, \ldots, M\).
8. For a single-output technology, \(p(\theta) = 1\), and the ray production function, \(f(x, \theta)\) simplifies to the single-output production function \(f(x)\).
9. The polar coordinates θ are not in logarithmic form (Henningsen et al., 2015).
10. A model with \(m\) outputs includes \(m - 1\) polar coordinates. For example, considering the output vector \(y = (y_1, y_2, y_3)\), the polar coordinates are \(\theta_1\) when measuring the angle between \(y_1\) and the plane spanned by \(y_2\) and \(y_3\), and \(\theta_2\) when measuring the angle between \(y_2\) and \(y_3\).
11. This amounts to considering the largest output that educational institutions are able to achieve with available resources as a benchmark, and deriving a measure of the inefficiency of those producing less.
12. The key idea is that disturbance terms \(e_{it}\) in equation (9) are assumed to be \(\sim\) conditional on \(z_{it}\) — independently truncated normally distributed, with a two-sided truncated normal distribution (left-truncation at \(-\gamma z_{it}\) and right truncation at \(1 - \gamma z_{it}\)). Since this distributional assumption on \(e_{it}\) is not appropriate for panel data, equation (9) is estimated on the pooled dataset. We are aware that a residual problem of separability (Badin et al., 2014) may exist, even if in our case the context variables may not impact on the teaching/learning technology.
13. The regional public expenditure on education is divided by population ages 3–18 years, that is, the share of the population that possibly benefits from that expenditure. Since INVALSI tests are administered in spring, we use one-year lagged values of EXP because scores are likely to be affected by expenditure of time \(t-1\) rather than current time. Data on public expenditure on education at a regional level are collected by the Italian Regional Public Accounts System (CPT, 2020).
14. In particular, the infrastructure endowment \(STRU\) is proxied by the share of schools providing specific services to disabled pupils (e.g., access ramps, elevators, standard stairs, etc.). Schools supplying better instruments to ease access to education are assumed to improve students’ results and school efficiency.
15. As the level of difficulty of INVALSI tests is not constant over time, we standardize annual scores by subtracting the national mean and dividing by the standard deviation. As a result, a part of our data assume negative values, which induce us to employ the SRPF model when allowing for multiple outputs.
16. The EQI proposed by Charron et al. (2013) is based on a large survey of about 34,000 European citizens living in 172 NUTS-1 and NUTS-2 regions within 18 European countries. Respondents are asked about their own experiences about and perceptions of the quality, impartiality and corruption of public services. We employ EQIs for 2010, 2013 and 2017 published by Charron et al. (2019); missing values are imputed by interpolation of the closest observations.
17. When evaluating the impact of context variables, the TFE model, as well as the distance function model, estimates the effects of regressors on inefficiency (not on efficiency). This implies that a positive (negative) sign of the estimated coefficient means that a variable negatively (positively) impacts the efficiency of the regional educational system.
18. The special status is acknowledged to islands (Sardinia and Sicily) and Northern border regions with relatively large linguistic minorities (i.e., Aosta Valley, Trentino-Alto Adige and Friuli Venezia Giulia).
19. According to CPT (2020), for the period 2011–18, regional per student real expenditure on primary and secondary schools was higher than the national average by 1.65 times in Trentino-Alto Adige, 1.19 times in Aosta Valley, 1.17 times in Sardinia and 1.14 times in Friuli Venezia Giulia.
20. For simplicity, and to facilitate comparison with Figure 2, here TFE estimated inefficiency \(\hat{\mu}_t \geq 0\) is expressed through a logarithmic transformation as an efficiency index \(1/e^{\hat{\mu}_t}\) included between 0 and 1.
21. These additional estimations are not shown in Table 3 and are available from the authors upon request.
22. Alternatively, a Tobit regression is also used, yielding results similar to those obtained by the Simar and Wilson method. In addition, to take possible changes in technology into account, we also include time effects; however, the hypothesis that the coefficients of year dummies are jointly equal to 0 is never rejected. Both Tobit and time-effect regressions are not reported, but are available from the authors upon request.
23. In Table 4, the estimated values of coefficients measure the impact of regressors on efficiency, rather than inefficiency. Thus, here the expected signs of the coefficients are opposite to those of Tables 2 and 3.
24. Robustness is confirmed by inspection of the kernel density of efficiency scores obtained by Simar and Wilson and TFE models (data not shown, but they are available from the authors upon request). The shape of probability distributions of efficiency derived from the two estimation methods are pretty similar, with a large part of the scores
being concentrated in the interval 0.75–0.95 and, especially in the Simar and Wilson case, very few scores < 0.4.

25. We also regressed Simar and Wilson residuals on regional dummies to test whether geographical determinants affect residual inefficiency (i.e., inefficiency not due to the context variables employed in estimation). The results (available from the authors upon request) corroborate the hypothesis that when the socio-economic peculiarities of Southern regions are appropriately accounted for, the North–South geographical pattern of inefficiency vanishes or is strongly attenuated.

**ORCID**

Valentina Chiariello Ⓞ http://orcid.org/0000-0001-8742-2510
Domenico Scaleri Ⓞ http://orcid.org/0000-0001-5582-6924

**REFERENCES**

Agasisti, T. (2013). The efficiency of Italian secondary schools and the potential role of competition: A data envelopment analysis using OECD-PISA2006 data. *Education Economics, 21*(5), 520–544. https://doi.org/10.1080/09645292.2010.511840

Agasisti, T. (2014). The efficiency of public spending on education: An empirical comparison of EU countries. *European Journal of Education, 49*(4), 543–557. https://doi.org/10.1111/ejed.12089

Agasisti, T., Bonomi, F., & Sibiano, P. (2014). Measuring the ‘managerial’ efficiency of public schools: A case study in Italy. *International Journal of Educational Management, 28*(2), 120–140. https://doi.org/10.1108/IJEM-02-2013-0032

Agasisti, T., & Cordero-Ferrera, J. M. (2013). Educational disparities across regions: A multilevel analysis for Italy and Spain. *Journal of Policy Modeling, 35*(6), 1079–1102. https://doi.org/10.1016/j.jpolmod.2013.07.002

Agasisti, T., & Johnes, G. (2015). Efficiency, costs, rankings and heterogeneity: The case of US higher education. *Studies in Higher Education, 40*(1), 60–82. https://doi.org/10.1080/03075079.2013.818644

Agasisti, T., & Vittadini, G. (2012). Regional economic disparities as determinants of pupils’ achievement in Italy. *Research in Applied Economics, 4*(3), 33–53. https://dx.doi.org/10.5296/rae.v4i2.1336

Aigner, D. J., Lovell, C. A. K., & Schmidt, P. (1977). Formulation accounted for, the North

Bertoni, M., Brunello, G., & Rocco, L. (2013). When the cat is near, the mice won’t play: The effect of external examiners in Italian schools. *Journal of Public Economics, 104*, 65–77. https://doi.org/10.1016/j.jpubeco.2013.04.010

Bratti, M., Checchi, D., & Filippini, A. (2007). Geographical differences in Italian students’ mathematical competencies: Evidence from PISA 2003. *Giornale degli Economisti e Annali di Economia, 66*(3), 299–333. http://www.jstor.org/stable/23248253

Bryson, A., Corsini, L., & Martelli, I. (2020). Teacher allocation and school performance in Italy (Discussion Paper No. 13669). IZA.

Charonn, N., Dijkstra, L., & Lapuente, V. (2013). Regional governance matters: Quality of government within European Union member states. *Regional Studies, 48*(1), 68–90. https://doi.org/10.1080/00343404.2013.770141

Charron, N., Lapuente, V., & Annoni, P. (2019). Measuring quality of government in EU regions across space and time. *Papers in Regional Science, 98*(5), 1925–1953. https://doi.org/10.1111/pirs.12437

Checchi, D. (2004). Da dove vengono le competenze scolastiche? L’indagine PISA 2000 in Italia. *Stato e Mercato, 72*, 413–453. doi:10.1425/18796

Coelli, T., & Henningsen, A. (2020). Frontier: Stochastic frontier analysis. R Package version 1.1-8. https://CRAN.R-Project.org/package=frontier

Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *The Coleman report. Equality of educational opportunity*. Washington, DC.

Cordero-Ferrera, J. M., Santín, D., & Simancas, R. (2017). Assessing European primary school performance through a conditional nonparametric model. *Journal of the Operational Research Society, 68*(4), 364–376. https://doi.org/10.1057/jors.2015.42

CPT. (2020). *Sistema dei Conti Politici Territoriali*. Agenzia per la Coesione Territoriale. http://old2018.agenziacoesione.gov.it/it/cpt/07_english_version/index.html

Crespo-Cebada, E., Pedraja-Chaparro, F., & Santín, D. (2014). *Does school ownership matter? An unbiased efficiency estimation with continuous and discrete environmental variables*. *Journal of the Operational Research Society, 65*(1), 543–557. https://doi.org/10.1080/02637758.2013.770141

Crespo-Cebada, E., Pedraja-Chaparro, F., & Santín, D. (2014). Does school ownership matter? An unbiased efficiency comparison for regions of Spain. *Journal of Productivity Analysis, 41*(1), 153–172. https://doi.org/10.1007/s11123-013-0338-y

Dario, C., & Simar, L. (2007). Conditional nonparametric frontier models for convex and nonconvex technologies: A unifying approach. *Journal of Productivity Analysis, 28*(1), 13–32. https://doi.org/10.1007/s11123-007-0049-3

De Witte, K., & Huddrikova, L. (2013). What about excellence in teaching? A benevolent rank of universities. *Scientometrics, 96*(1), 337–364. https://doi.org/10.1007/s11192-013-0971-2

De Witte, K., & Kortelainen, M. (2013). What explains the performance of students in a heterogeneous environment? Conditional efficiency estimation with continuous and discrete environmental variables. *Applied Economics, 45*(17), 2401–2412. https://doi.org/10.1080/00036846.2012.665602

Di Giacomo, G., & Pennisi, A. (2015). Assessing primary and lower secondary school efficiency within Northern, Central and Southern Italy. *Italian Economic Journal, 1*(2), 287–311. https://doi.org/10.1015/j.ijor.2015.92

Eicher, T., García-Peñalosa, C., & Teksoz, U. (2006). How do institutions lead some countries to produce so much more output per worker than others? In T. Eicher & C. García-Peñalosa (Eds.), *Institutions, development and growth* (pp. 65–80). MIT Press.
Ferraro, S., Agasisti, T., Porcelli, F., & Soncin, M. (2020). Local governments' efficiency and educational results: empirical evidence from Italian primary schools (Working Paper No. 90). Società Italiana di Economia Pubblica.

Greene, W. (2005a). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics*, 126(2), 269–303. https://doi.org/10.1016/j.jeconom.2004.05.003

Greene, W. (2005b). Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis*, 23(1), 7–32. https://doi.org/10.1007/s11123-004-8545-1

Henningsen, A., Kaymak, M., & Henningsen, G. (2017). New insights into the stochastic ray production frontier. *Economics Letters*, 156, 18–21. https://doi.org/10.1016/j.econlet.2017.04.006

Henningsen, G., Henningsen, A., & Jensen, U. (2015). A Monte Carlo study on multiple output stochastic frontiers: A comparison of two approaches. *Journal of Productivity Analysis*, 44(3), 309–320. https://doi.org/10.1007/s11123-014-0416-9

Houck, E. A., Rolle, R. A., & He, J. (2010). Examining school district efficiency in Georgia. *Journal of Education Finance*, 35(4), 331–357. https://doi.org/10.1353/jef.0.0022

Hoiny, C. (2000). Peer effects in the classroom: Learning from gender and race variation (Working Paper No. 7867). National Bureau of Economic Research (NBER).

INVALSI. (2018). *Rapporto prove INVALSI 2018*. https://wwwINVALSI.it/invalsi/doc_evidenza/2018/Rapporto_prove_INVALSI_1_2018.pdf

ISTAT. (2020a). *Spesa per istruzione*. http://dati-capumano.istat.it/Index.aspx?DataSetCode=CU_SPEISTROCSE

ISTAT. (2020b). *I.Stat*. http://dati.istat.it/Index.aspx

Johnes, G. (2013). Efficiency in English higher education institutions revisited: A network approach. *Economics Bulletin*, 33(4), 2698–2706.

Johnes, J. (2015). Operational research in education. *European Journal of Operational Research*, 243(3), 683–696. https://doi.org/10.1016/j.ejor.2014.10.043

Kirjavainen, T. (2012). Efficiency of Finnish general upper secondary schools: An application of stochastic frontier analysis with panel data. *Education Economics*, 20(4), 343–364. https://doi.org/10.1080/09645292.2010.510862

Löthgren, M. (1997). Generalized stochastic frontier production models. *Economics Letters*, 57(3), 255–259. https://doi.org/10.1016/S0165-1765(97)00246-2

Löthgren, M. (2000). Specification and estimation of stochastic multiple-output production and technical inefficiency. *Applied Economics*, 32(12), 1533–1540. https://doi.org/10.1080/000368400418943

Mainarde, E., Alves, H., & Raposo, M. (2014). Using expectations and satisfaction to measure the frontiers of efficiency in public universities. *Tertiary Education and Management*, 20(4), 339–353. https://doi.org/10.1080/13583883.2014.978358

Marchsb, M. J., Calero, J., Choi, A., & Ximénez-de-Embún, D. P. (2012). The efficiency of public and publicly subsidized high schools in Spain: Evidence from PISA–2006. *Journal of the Operational Research Society*, 63(11), 1516–1533. https://doi.org/10.1057/jors.2011.156

Millimet, D. L., & Collier, T. (2009). Efficiency in public schools: Does competition matter? *Journal of Econometrics*, 145(1–2), 134–157. https://doi.org/10.1016/j.jeconom.2008.05.001

Nazarko, J., & Šapar aukas, J. (2014). Application of DEA method in efficiency evaluation of public higher education institutions. *Technological and Economic Development of Economy*, 20(1), 25–44. https://doi.org/10.3846/20294913.2014.837116

Nifo, A., Scara, D., & Vecchie, G. (2017). The rule of law and educational choices: Evidence from Italian regions. *Regional Studies*, 51(7), 1048–1062. https://doi.org/10.1080/00344304.2016.1262945

Organization for Economic Co-operation and Development (OECD). (2020). *Education at a glance 2020: OECD indicators*. OECD Publ. https://read.oecd-ilibrary.org/education/education-at-a-glance-2020_6909673-en#page305

Podinovski, V. V., Iismai, I., Bouzidine-Chameeva, T., & Zhang, W. (2014). Combining the assumptions of variable and constant returns to scale in the efficiency evaluation of secondary schools. *European Journal of Operational Research*, 239(2), 504–513. https://doi.org/10.1016/j.ejor.2014.05.016

Quintano, C., Castellano, R., & Longobardi, S. (2009). L'influenza dei fattori socio economici sulle competenze degli studenti italiani. Un’analisi multilevel dei dati PISA 2006. *Rivista di Economia e Statistica del Territorio*, 2(2), 109–149. https://doi.org/10.3280/REST2009-002004

Rodriguez-Pose, A., & Tselios, V. (2011). Mapping the European regional educational distribution. *European Urban and Regional Studies*, 18(4), 358–374. https://doi.org/10.1177/0969776411399345

Sibiano, P., & Agasisti, T. (2013). Efficiency and heterogeneity of public spending in education among Italian regions. *Journal of Public Affairs*, 13(1), 12–22. https://doi.org/10.1002/pa.1404

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage semi-parametric models of production processes. *Journal of Econometrics*, 136(1), 31–64. https://doi.org/10.1016/j.jeconom.2005.07.009

Zoghi, A. C., Rocha, F., & Mattos, E. (2013). Education production efficiency: Evidence from Brazilian universities. *Economic Modelling*, 31, 94–103. https://doi.org/10.1016/j.econmod.2012.11.018