

RESEARCH

Open Access



Spatial variations of school-level determinants of reading achievement in Italy

Chiara Sacco*  and Patrizia Falzetti

*Correspondence:
chiara.sacco@invalsi.it
Italian National Institute
for the Evaluation
of the Educational System,
Via Ippolito Nievo, 35,
00153 Rome, Italy

Abstract

The study of the territorial difference in educational achievement is a widely debated topic, in particular in Italy for the presence of the well-known strong regional disparities. National and international large scale assessments confirmed that the main characteristic of the Italian school system is the geographical cleavage between North and South. Policymakers have pressing needs to find solutions to reduce geographical disparities. In this study, we investigate the spatial disparities of academic achievement from a new perspective, assuming that the relationship between academic achievement and predictors varies across Italy. Our aim is to examine the extent of the spatial disparities in the relationship between academic achievement and some school-level factors related to inequalities in educational outcomes, moving beyond the regional administrative confines, in order to identify new spatial patterns. We exploited the reading standardized tests administered by INVALSI in 2018–2019 focusing on the 8th-grade students. Crucial to our contribution is the use of the geographically weighted regression and the k-mean clustering, which allows studying the spatial variability of the impact of the school-level factors on academic achievement and to gather schools in new spatial clusters. The findings of this paper demonstrate the necessity to design a more specific education policy and support the identification of the main critical factors for different geographical areas.

Keywords: Geographically Weighted Regression, Large-scale Assessment, Student achievement, Local model, Inequality

Introduction

A well-educated and well-trained population is crucial for the social and economic well-being of a country. Improving students' achievement can contribute to both economic growth and social development. In recent years, the UN Member States and the European Union have emphasized the necessity to strengthen the educational system promoting efficiency and equity (European Union 2009). In particular, Goal 4 of the 2030 Agenda for Sustainable Development launched by UN Member States in 2015 aims to “ensure an inclusive and equitable quality education and promote lifelong learning opportunities for all”. Equity in education means that the school system is able to offer to each student the same opportunities to achieve the best possible outcome, in terms of academic skills, regardless of gender, ethnic origin, socioeconomic

status and where one lives. To reach this goal, monitoring the educational outcome and understanding the factors behind students' and schools' results have become paramount, accentuating the importance assumed by the large-scale assessments, able to guarantee comparable and reliable data on students' achievement. In the last decade, large-scale national and international assessments have been widely used to measure what students know and can do in order to provide an image of the school system's status and are considered as a benchmark for policymakers for planning reforms to improve the performance of the school system (Breakspear 2012; Fullan 2009).

Student's achievement is a complex phenomenon and, as it has been widely demonstrated in the literature. Although student performances are primarily influenced by individual characteristics, the impact of the socioeconomic and cultural environment is also fundamental (Bakker et al. 2007; Marks 2006; Teodorovic 2012). The socioeconomic condition of the territory is central to understand the students' outcome since it affects the family background but, on the other side, this effect is reciprocal: the educational outcome of a country influences the development of the country. Several researchers demonstrated that socioeconomic inequality tends to increase in presence of inequality in educational outcomes (Checchi and Peragine 2010). Rodriguez-Posè and Tselios (2009) demonstrated that education can impact on economic inequality at regional levels. Since the human capital plays a central role for the socio-economic development of countries (Hanushek and Kimko 2000; Hanushek and Woessman 2009; Hanushek and Woessmann 2011), it is important to understand which factors contribute to the divergence in the educational outcome among the regions.

It is well known that one of the main characteristics of the Italian school system is the geographical cleavage between North and South. National and international large scale assessments reveal that North of Italy outperforms students in both Central and South of Italy, highlighting above all the worrying situation of students in South Italy starting from the lower secondary school (INVALSI 2018, 2019; OECD 2016). Whereas the Northern regions are characterized by small variability in performance, in the South the variability within and between region is high, suggesting a complex and manifold reality (Benadusi et al. 2010; Bratti et al. 2007). Benadusi et al. (2010) demonstrated that one of the main factors for inequality in student scholastic performance is the gap among Italian macro-areas (North-West, North-East, Centre, South-West-Islands and South-East). Panichella and Triventi (2014) confirmed the crucial role played by the geographical factors in the exacerbation of educational inequalities focusing on educational careers. Several studies analysed the strong variations across macro-geographical areas in the levels of students' performance. Agasisti and Cordero-Ferrara (2013) proposed a multilevel analysis applied to PISA 2006 data to study educational disparities across regions in Italy and Spain. Hippe et al. (2018) exploited the regional distribution of skills in Italy and Spain, studying the extent of the regional inequalities in PISA 2015 using descriptive statistics and estimating several regression models. Costanzo and Desimoni (2017) explored inequalities in education using a quantile regression approach applied to primary school data from INVALSI large-scale assessments; mathematics and reading scores were regressed on students' characteristics and geographical variables.

The presence of the well-known regional inequalities in education in Italy points out the need to better understand the regional disparities and the connected policy implications.

Previous studies analyse the geographical disparities in Italy assuming that only the student outcome depends on the geographical areas but that the effect on the student outcome of the contextual variables is the same across Italy. However, the effect of the contextual variables on the student performance may be different across the country (Agasisti et al. 2017) and, so, different critical factors could be identified in different geographical areas. To better understand the determinants of the mechanism of inequality it is important to verify if and how educational predictors have a different effect on the student performance according to the geographical location of the schools.

The aim of this study is to investigate how the impact of the contextual variables on academic performance varies by geographic areas in Italy, exploiting data aggregated at school-level. This work exploits the reading comprehension standardized test administered in Italy by the National Evaluation Institute for the School System (INVALSI) for the school years 2018–2019, focusing on the whole population of the students at the last year of the lower secondary school, involving more than 500 000 students in more than 5000 schools in Italy. For the first time, to analyse the inequality of the reading achievement in Italy, the georeferenced data of the schools are used. Our main concern is to examine the extent of the spatial disparities in the relationship between the academic achievement and some school-level factors related to inequalities in educational outcomes, moving beyond the regional administrative confines, in order to identify new spatial patterns. Geographically weighted regression (GWR) is utilized to explore the spatial variations in the above relationship. K-means clustering method has been applied to classify schools into homogeneous regions based on local regression patterns. In terms of policy implications, it is crucial to identify which and where local factors have the strongest impacts to plan different effective policy strategies. The findings of this paper demonstrate the necessity to design more specific education policy and support with the identification of the main critical factors for different geographical areas.

Educational system in Italy

The Italian schooling system has been centrally managed and financed since the unification of the country in the nineteenth century. The educational system in Italy is articulated in three main cycles: primary school (grades 1 to 5), lower secondary school (grades 6 to 8) and upper secondary school (grade 9 to 13). The total number of students is equal to about 7.6 million, attending 31 000 schools of which only 10% is private but periodically accredited by the Ministry of Education (Italian Ministry of Education 2012). The Ministry of Education in Italy plays a central role in the regulation of the schools' activities since the public schools are characterized by a low degree of autonomy. Although during the nineties a slow process for assigning more controls to the single schools in organising their own teaching activities began (laws n. 537/1993 and n. 59/1997), the school autonomy is still constrained by the inability to choose the teachers and manage the budget for the human resources. The Ministry of Education has the responsibility to recruit and to allocate teachers to schools and to determine their wages or to fire them. Regarding the funding, according to the official statistics of the Italian

Ministry of Education (2018), the schools in Italy are mainly financed by the Central government, particularly more than 90% of the educational expenses are covered by the central government. Only the schools belonging to the Autonomous Provinces of Trento and Bolzano are financed by their own regional funds.

Hence, overall in Italy the school system is characterized by a low degree of autonomy and is conceived as a unified system aiming to create a national identity and to ensure a high degree of homogeneity in student education in each part of the national territory. Despite the centralization of the school system, strong differences in student performance can be found by territorial area.

Background

A topic of great interest in educational research in order to deal with inequality in the school system is understanding the factors that influence the learning outcomes. Understanding the factors that influence student achievement could help educators and policymakers in setting up a support system to allow all students to perform better and to develop equality in the educational system. This study investigates the inequality of the educational system at a geographical level. We aim to identify in which schools/geographical areas it is necessary planning an intervention to flatten the effects of some factors associated in the literature with the inequality in education in order to guarantee greater equality. We focused on the composition of the student body from two different points of view: the students previous learning levels and socio-demographic characteristics. Moreover, we analysed the effect of the school size on the school mean performance to examine if the geographical position of the school influenced the role of school size on the improvement of the achievements.

A fair analysis of the mean school performance takes into account the initial educational attainment of the students. The analysis of the gain between previous and current learning is widespread in particular in the contest of the evaluation of the school effectiveness (Heck 2000; Grilli and Rampichini 2009). The effect of the previous learning on the current learning helps to identify schools that produce a greater or lesser impact on student academic improvement and the omission of the previous achievements could lead to an inflation of the coefficients of the others variables since, it is well-known, that it is one of the most important predictors of the student achievement.

Many studies in the field of education have recognized the importance of the economic, social and cultural status as a determinant of the learning process (Schuetz et al. 2008; Giambona and Porcu 2015; OECD 2016; Gursakal et al. 2016; Benadusi 2016; INVALSI 2019). Agasisti and Vittadini (2012) demonstrated that high variability in the student outcome can be explained by contextual and regional factors; in particular, the composition of the school student body, in terms of socio-economic and cultural background (ESCS) matters more than school's resources, however, this effect is mitigated when considering the effect of macro-areas of the country. Several researchers recognized the impact of the ESCS on the performance varies across regions (Longobardi et al. 2009; Benadusi et al. 2010; Matteucci and Mignani 2014), suggesting the necessity to analyse more in-depth the relationship between the ESCS and the territorial conditions.

One of the major concern at the global policy level is reducing the gender inequality in education; this aim has been recently ratified as one of the seventeen Sustainable Development Goals of the UN. Researchers depict two different realities in the analysis of the gender as a determinant of the learning: boys outperform girls in the STEM (Science Technology Engineering and Mathematics) disciplines (Robinson and Lubiensky 2011; OECD 2016; Wang and Degol 2017; Contini et al. 2017), but, focusing on the reading comprehension test scores, the traditional gender gap is reversed in favour of girls (Department of Education and Skills 2007; Legewie and Di Prete 2012; INVALSI 2019). From a policy perspective, it is important to understand when the gap first shows up and identify the geographical critical areas. The results of the quantile regression analysis on mathematics and reading performance in Italy performed by Costanzo and Desimoni (2017) suggest the necessity analyse the role of the gender in a more complex framework than the traditional regression model. Another variable widely associated with inequalities in educational outcomes is the immigrant status (Azzolini et al. 2012; Schnell and Azzolini 2015). Azzolini and Barone (2013) showed that immigrants' scholastic adaptation in Italy follows heterogeneous paths, suggesting a not complete integration of the immigrant into society. Agasisti and Vittadini (2012) showed that a high concentration of immigrant students within the schools is associated with lower performance, suggesting a negative peer effect.

Instead, the literature is not unified on the directions of the effects of the school size on the students' outcome. Some studies from American evaluations found that student performance decrease as school size grows (Lee and Loeb 2000; Wasley et al. 2000); studies, based on the analysis of other countries, claimed a positive effect of larger school (Luyten, 2014; Scheerens et al. 2014). However, the heterogeneity of the effects of small schools on student performance and engagement for students' subgroups is demonstrated in several studies (Leithwood and Jantzi 2009; Weiss et al. 2010; Schwartz et al. 2013).

Data

The National Evaluation Institute for the School System (INVALSI), annually, carries out large-scale survey assessments in Italy to monitor students' achievement in Reading (reading comprehension and grammatical knowledge), Mathematics and English Language (reading and listening comprehension). In particular, INVALSI standardized tests are administrated to students attending the 2nd and the 5th grade of the primary school, the 8th grade of the lower secondary school, the 10th and the 13th grade of the upper secondary school, involving about 3 000 000 students within approximately 15 000 schools. In 2018, INVALSI shifted from paper-and-pencil to computer-based tests for the assessments of the competence of students in lower and upper secondary school. Moreover, from 2018 INVALSI standard test is compulsory at the end of the lower secondary school for all students to be admitted to the state final exam. Computer-based administration and the test compulsoriness provided a consistent and cleaner dataset and allowed to minimize cheating and, thus, to avoid cheating correction procedure (Quintano et al. 2009, Longobardi et al. 2018).

In addition to the standardized test, INVALSI requires each student to compile a questionnaire after the test to collect socio-demographic variables regarding the student.

School' secretarial offices provide to INVALSI further information about classes' and schools' characteristics (e.g. the number of classes within the school, the number of buildings of school and the school address).

In this study, we focused on the reading standardized tests administered by the INVALSI at the end of lower secondary school in the school year 2018–2019. In 2018–2019, the INVALSI test has been administered to a population of 542 689 students in 5761 schools. In this study, we considered the data aggregated at school-level and we focused on schools with at least 20 students, a percentage of students' participation higher than 80%, geospatial location available. The number of schools analysed in this study is 5520.

The students' performances at Italian INVALSI test are measured using the Weighted Likelihood Estimation (WLE) estimated by the Rasch model (for more information a technical report is available on the official website: <http://INVALSI-areaprove.cineca.it/>).

As predictors of the students' performance, we selected a set of common variables to represent to the socio-demographic background of the school, one variable to evaluate the impact of the school students' previous knowledge on the school mean performance and two variables to evaluate the effect of the school size.

The socio-demographic variables at school-level included in the analysis are those usually associated with the inequalities in the educational outcome, namely the percentage of immigrant students, the percentage of females, the percentage of late-enrolled students (i.e. students enrolled at least 1 year after the of 6 or repeated one or more years) and socio-economic background. The index of economic, social and cultural status (ESCS) is based on the following set of variables related to students' family background: parents' occupation, parents' education and the number of books at home that is a proxy of the family educational resources. The index is derived by INVALSI following an OECD's standard (OECD 2005). Campodifiori et al. (2008) detail the methodological aspects of INVALSI procedure to estimate the ESCS.

The indicator of the students' previous achievement is the students' WLE score at reading INVALSI test at the end of primary school in the s.y. 2015–2016 (called WLE in G05) aggregated at school level. Indeed, using the INVALSI student identifier, it is possible to match individual students' test records from year to year, to map the student's academic achievement over time.

The last group of variables are included in the analysis to represent the school size: the number of classes and the number of buildings. This information is derived from the data provided annually by the school's secretarial office.

All the variables included in the models have been standardized with mean 0 and variance 1. Table 1 shows the descriptive statistics of all the standardized variables included in the model as outcome and predictors by geographical area (North West, North East, Centre, South, South and Islands).

Methods

To assess the geographic dimension of the association between schools' performance and predictor variables, this study followed a two steps procedure. First, both linear regression model and geographically regression model were estimated to study the association at global and local level, respectively.

Table 1 Descriptive statistics, mean and (standard deviations), on the outcome and the standardized predictors by geographical area (North West, North East, Centre, South, South and Islands)

	North West	North East	Centre	South	South and Islands
Outcome					
WLE G08	0.44 (0.85)	0.38 (0.70)	0.21 (0.82)	− 0.43 (0.96)	− 0.79 (1.03)
Predictors					
WLE G05	0.29 (0.79)	0.20 (0.72)	0.22 (0.90)	− 0.27 (1.1)	− 0.58 (1.18)
ESCS	0.21 (0.96)	0.24 (0.8)	0.31 (0.89)	− 0.34 (1.01)	− 0.52 (1.02)
Number buildings	− 0.07 (0.86)	− 0.04 (0.91)	− 0.11 (0.89)	− 0.03 (1.00)	0.29 (1.30)
Number classes	− 0.11 (0.91)	− 0.05 (0.88)	− 0.08 (1.01)	0.27 (1.2)	0.01 (0.95)
Female	0.02 (1.00)	− 0.01 (0.98)	− 0.01 (1.06)	0.01 (0.97)	− 0.02 (0.99)
Immigrant	0.41 (1.13)	0.53 (1.00)	0.19 (0.87)	− 0.65 (0.47)	− 0.66 (0.47)
Late-enrolled	0.21 (1.11)	0.19 (0.96)	0.04 (0.93)	− 0.40 (0.83)	− 0.12 (0.95)

An OLS model can be expressed as.

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \varepsilon_i \quad (1)$$

where y_i is the dependent variable for i -th observation, β_0 is the intercept, β_k is the coefficient for the predictor x_k , x_{ik} is the k -th predictor for i -th observation and ε_i is the error term.

GWR allows the relationship between the dependent variable and the predictors to vary geographically considering locally weighted regression coefficients. In other words, the regression coefficients are variable since they depend on the geographical coordinates of the observations. In the framework of the GWR model (Brunsdon et al. 1998), the relationship between y_i and the predictors can be written as.

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (2)$$

where (u_i, v_i) denotes the geographical coordinates (latitude and longitude) of the i -th observation. Thus, $\beta_0(u_i, v_i)$ represents the intercept of the observation i with spatial coordinates (u_i, v_i) and $\beta_k(u_i, v_i)$ denotes the coefficient for the predictor x_k of the observation i with spatial coordinates (u_i, v_i) . Instead of estimating one single regression, the GWR model generates separate regressions, one for each observation. In this way, GWR allows to examine the spatial variation in the relationship between the dependent variable and the predictors. The GWR model calibrates separate regressions for each observation assuming that observations are weighted on the base of their proximity to i -th observation. The estimation of the coefficients is performed using weighted least squares, assuming closer observations have a greater influence on the estimation of regression parameters of i -th observation than remote observations. The weighting is controlled by the weight matrix that is a diagonal matrix in which each diagonal element w_{ij} is a function of the location of the observation. In particular, a kernel density function determinates the weight w_{ij} . In this study, we exploited the Gaussian weighting function as kernel density function, thus w_{ij} can be written as.

$$w_{ij} = \exp\left(-\frac{d_{ij}^2}{h^2}\right) \quad (3)$$

where d_{ij} is the distance between the i -th observation with spatial coordinates (u_i, v_i) and the j -th observation with spatial coordinates (u_j, v_j) and h is the bandwidth. The choice of a weighting function involves choosing the bandwidth h and the choice of this parameter is crucial since it has the strongest influence on the results. The bandwidth is the number of observations to which the kernel assigns a non-zero weight, i.e. the bandwidth represents the distance from the observation i of interest beyond which the weight of the observations is equal to the value 0. We assumed a Gaussian adaptive kernel, that allows the bandwidth varying on the base of the density of observation points. The bandwidth has been computed through a calibration process based on the minimization of the Akaike Information Criteria (Fotheringham et al. 2003). The spatial analyses have been performed using the R package *spdep* (Bivand and Wong 2018; Bivand et al. 2013) e *GWmodel* (Gollini et al. 2015; Binbin et al. 2014).

In the second step, the k-means clustering method has been applied to the matrix of the spatial coefficients estimated using the GWR to divide schools into areas where performance are homogeneously affected by the analysed predictors, based on local regression patterns. K-means clustering is one of the most popular unsupervised machine learning algorithms, with the objective to group similar data points together and discover underlying patterns. To achieve this objective, k-means looks for a fixed number (k) of clusters in a dataset. Determining the optimal number k of clusters in a data set is a fundamental issue and, in literature, a wide variety of indices have been proposed. For the identification of the number of clusters the package *NbClust* package (Charrad et al. 2014) has been used. *NbClust* package provides 30 indices for determining the optimal number of clusters and proposes to user the best clustering scheme from different results obtained by varying all combinations of number of clusters.

Results

To examine the geographical patterns in the relationship between schools' performance and ESCS, WLE in G05, number of classes, number of school buildings, percentage of immigrant students, percentage of late-enrolled students and percentage of females, the OLS and GWR models were estimated.

Table 2 summarize the results of the OLS model that pictures the relationships through the entire country. All the predictors included in the model are significant; only the number of buildings and the percentage of late-enrolled students have a negative impact on the mean school performance. We checked the presence of multicollinearity between the predictors using the variance inflation factor (VIF), that measure how much the variance of the estimated regression coefficient is inflated by its correlation with the other predictors. As general rule, when the $0 < VIF < 10$ there is no multicollinearity. All the VIF measures of the predictor variables range from 1.01 to 1.78, suggesting that the predictors in the model are not correlated with the other variables.

Table 3 reports different measures of goodness of fit for the OLS model and GWR. The GWR fits better the data as it has higher adjusted R^2 and lower AICc. In correspondence of the OLS model, the adjusted R^2 is equal to 0.60, which indicates that about the 60% of

Table 2 Results of OLS regression model

	Coefficient	Std. Error
Intercept	0.000	0.008***
WLE in G05	0.326	0.010***
ESCS	0.521	0.011**
Number of classes	0.023	0.009**
Number of buildings	− 0.027	0.009***
Female (%)	0.059	0.009***
Immigrant (%)	0.072	0.010***
Late-enrolled (%)	− 0.068	0.011***

Table 3 Goodness-of-fit model comparison

	OLS	GWR
R-squared	0.60	0.71
Adjusted R-squared	0.60	0.70
AIC	10 593.93	8975.92
AICc	10 593.96	9255.74

the total variance of the dependent variable is explained by the model, whereas for the GWR the adjusted R^2 is equal to 0.70.

Figure 1a shows the residuals of the OLS model suggesting the presence of spatial correlation in the residuals: we can observe high residuals in North Italy and low residuals in South Italy. The Moran's I has been used as a measure of the spatial correlation for the residuals: positive and negative values of the Moran's I suggest positive and negative autocorrelation, respectively; under the null hypothesis of no spatial autocorrelation the Moran's I is expected to be close to zero. The Moran's I results statistically different from zero ($I=0.085$, $p<2.2E-16$) and confirms the spatial autocorrelation among model residuals of the OLS model. The GWR residuals exhibit a reduction of the spatial pattern (Fig. 1b); the Moran I ($I=0.007$, $p<5.39E-07$) suggests a substantial reduction although not total elimination of the spatial autocorrelation among model residuals.

Another measure of goodness of fit of the GWR model is provided by the local R^2 that ranges from 0.25 to 0.95, with the 75% of schools having a value of 0.63 or higher (Table 4).

Table 4 shows the estimated coefficients of the GWR model. The WLE obtained in G05 and the ESCS result to be consistently positively associated with the mean school performance, whereas the other indicators have both positive and negative values in different locations. Small interquartile range is observed in correspondence with the percentage of females, the number of buildings, the number of classes and the percentage of late-enrolled students. These variables result to be not significant in the majority of the locations. For the WLE in G05, the ESCS, the percentage of immigrant students and the number of buildings, the interquartile range of the GWR coefficients falls below the magnitude of the OLS coefficients.

Figure 2 shows the spatial patterns of the intercept and the coefficients of the 7 predictors estimated exploiting the GWR. Only the significant coefficients are represented in

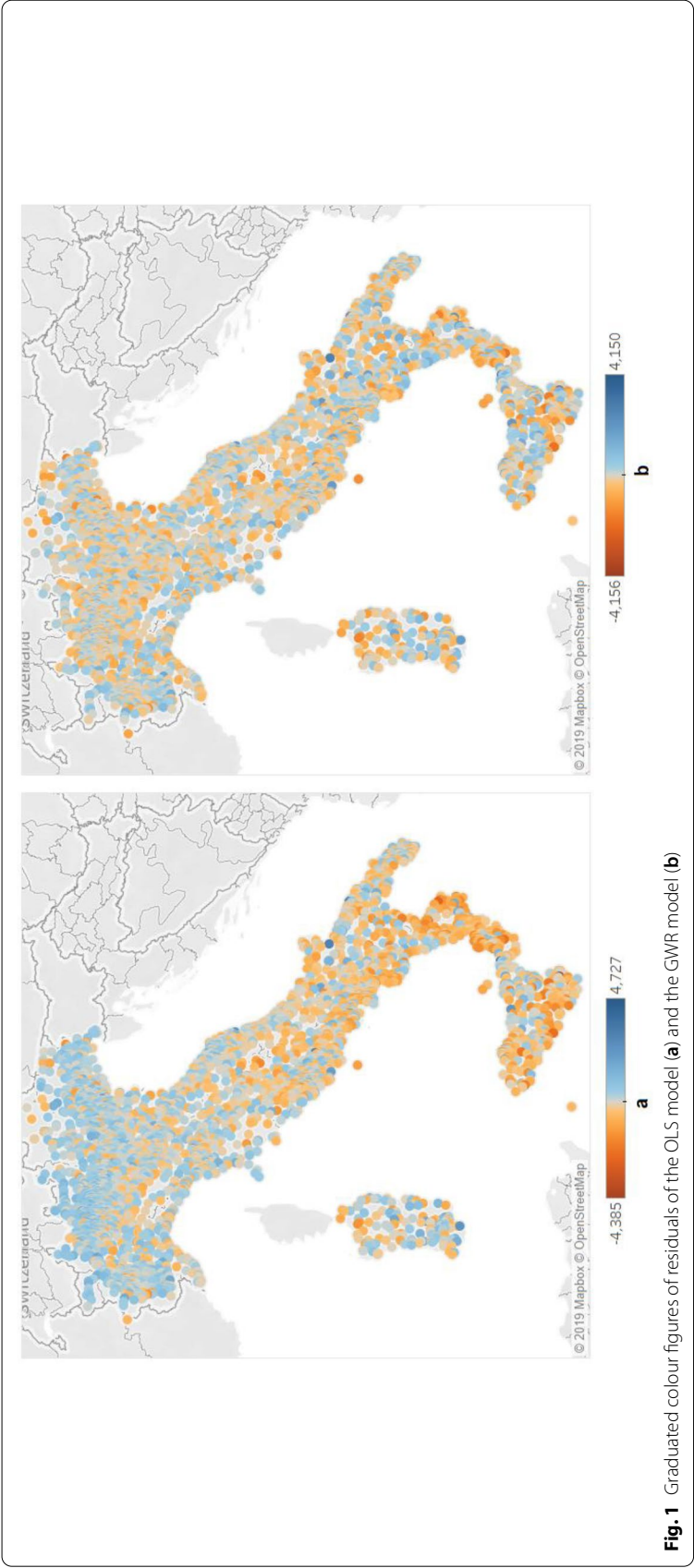


Table 4 Results of the geographically weighted regression (GWR) model

	Coefficient Range						% Significant coefficients
	Min	1st Quantile	Mean	Median	3rd Quantile	Max	
Intercept	− 1.024	− 0.264	0.008	0.136	0.320	0.684	85.52
WLE in G05	0.073	0.192	0.248	0.257	0.305	0.471	99.66
ESCS	0.171	0.295	0.404	0.402	0.511	0.627	100.00
Number of classes	− 0.104	− 0.029	0.008	0.016	0.053	0.187	13.59
Number of buildings	− 0.180	− 0.032	0.016	0.014	0.048	0.175	17.57
Female (%)	− 0.070	0.033	0.057	0.055	0.084	0.208	34.84
Immigrant (%)	− 0.463	− 0.177	− 0.126	− 0.125	− 0.062	0.397	49.69
Late-enrolled (%)	− 0.313	− 0.133	− 0.090	− 0.089	− 0.037	0.112	38.95
Local R-Squared	0.253	0.626	0.686	0.703	0.768	0.947	

the map. The WLE obtained in 5th grade has a low impact on the mean performance in South Italy, in the north of Marche, in Modena and in Brescia province. On the other hand, a high positive effect of prior achievements is observed in North Italy, in the metropolitan area of Rome and in the coastal area of Abruzzo and Molise. The socio-economic indicator has a high positive impact in South Italy and in few areas of Central and North Italy. Only in a few locations, the effect of the number of classes results to be significantly different from zero. In the area of Rome, in the Venice province and in Umbria the number of classes has a positive impact on school performance, whereas in the northern areas of Campania, Apulia and Calabria and in the eastern provinces of Sicily the number of classes is negatively associated with school performance. Only in Sardinia, the number of buildings has a consistent significant negative effect on the performance. The percentage of females results to be positively associated with school performance, only in the area of Rome the mean performances of a few schools are negatively affected by the percentage of females. In the entire Calabria and in several areas of North Italy the presence of immigrant students affects the mean school performance negatively; in particular, in Calabria the magnitude of the coefficients is higher than in the rest of Italy. The schools with high values of the late-enrolled students' coefficients are located in the province of Lecce, Siracusa, in the metropolitan area of Rome and in Tuscany.

To summarize the results of the GWR and to identify clusters of schools with performances homogeneously affected by the analysed factors, the k-means clustering has been applied to the regression coefficients matrix. The comparison between 30 different indices leads to the identification of seven as the optimal number of clusters. Table 5 shows the summary statistics for each cluster, in particular, we reported the mean and the percentage of significant coefficients in correspondence of each variable. It can be noticed that the WLE in G05 and ESCS play a predominant role in the interpretation of the clusters. The use of the k-means has facilitated the reading of the results since it is now possible to identify easily the important factors of a specific geographic area (cluster).

Figure 3 reports the maps and radar plot of the identified seven school clusters. The two clusters, illustrated in the Fig. 3a, include a particular section of the national territory, Alps, Friuli, Veneto, the Po valley, the Apennines and Sardinia. These two

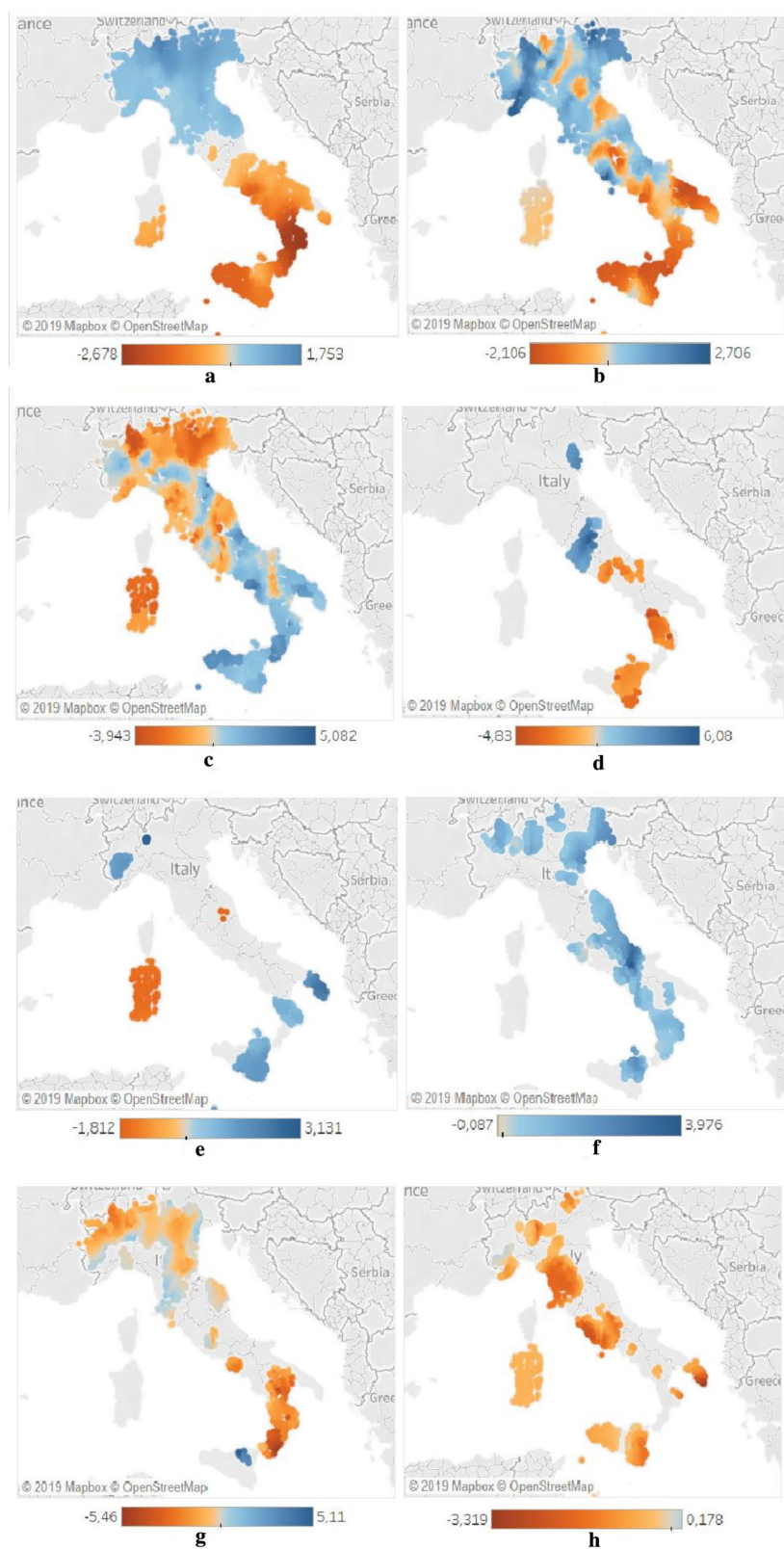
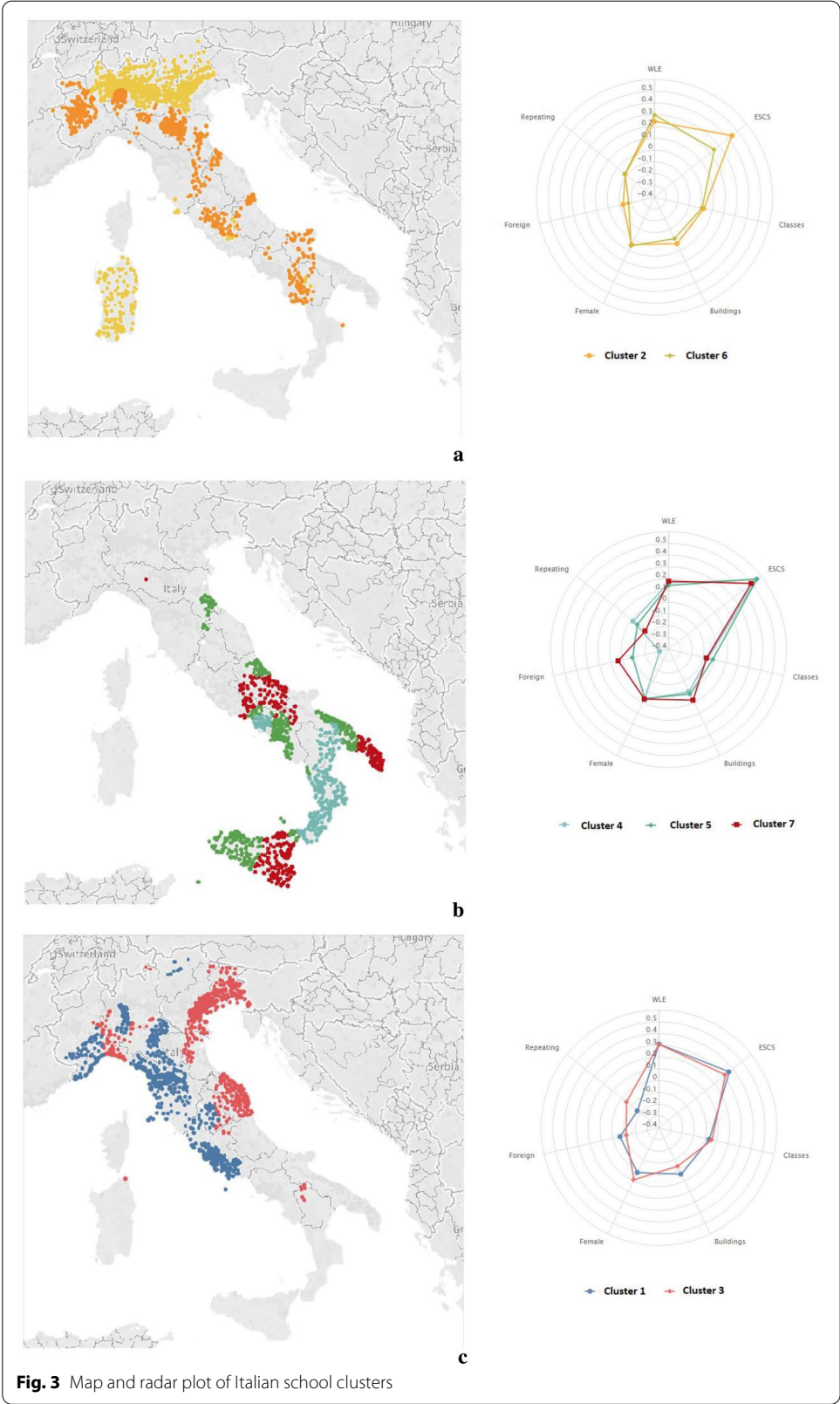


Fig. 2 Graduated colour figures of local coefficient estimates: intercept (a), WLE in G05 (b), ESCS (c), number of classes (d), number of buildings (e), percentage of females (f), percentage of immigrant students (g), percentage of late-enrolled students (h)

Table 5 Mean and percentage of significant coefficients for each cluster and for each predictor

Cluster	Statistics	WLE	ESCS	Number buildings	Number classes	Female (%)	Immigrant (%)	Late-enrolled (%)
1	Mean	0.311	0.362	0.037	0.038	0.022	− 0.062	− 0.168
	Sign Coeff (%)	100.00%	100.00%	2.01%	30.92%	10.54%	23.19%	90.76%
2	Mean	0.245	0.439	0.038	0.030	0.046	− 0.128	− 0.083
	Sign Coeff (%)	99.40%	100.00%	7.35%	5.84%	34.14%	63.54%	31.02%
3	Mean	0.309	0.320	− 0.038	0.062	0.090	− 0.117	− 0.048
	Sign Coeff (%)	99.85%	100.00%	0.77%	17.62%	69.09%	57.81%	0.77%
4	Mean	0.166	0.550	− 0.001	− 0.058	0.074	− 0.323	− 0.017
	Sign Coeff (%)	100.00%	100.00%	17.45%	32.38%	54.19%	94.97%	1.68%
5	Mean	0.142	0.559	0.022	− 0.016	0.064	− 0.088	− 0.062
	Sign Coeff (%)	98.96%	100.00%	3.55%	5.62%	23.37%	1.48%	31.51%
6	Mean	0.298	0.249	− 0.011	0.013	0.058	− 0.176	− 0.081
	Sign Coeff (%)	99.73%	100.00%	14.97%	0.00%	31.28%	81.55%	31.02%
7	Mean	0.179	0.497	0.078	− 0.069	0.069	0.038	− 0.49
	Sign Coeff (%)	99.59%	100.00%	72.65%	52.86%	40.82%	3.27%	73.88%

clusters are quite similar considering the mean of the effect of the previous competencies (0.245 and 0.298 for Cluster 2 and Cluster 6, respectively) and the negative impact of the presence of immigrant students (− 0.128 for Cluster 2 and − 0.176 for Cluster 6), which results to be around the national average for both variables (Table 5). On the other hand, an opposite behaviour of the family background is observed: for schools in Cluster 2 the ESCS effect is greater than the national average, whereas in Cluster 6 the ESCS effect results to be less than the national average. Cluster 2 and Cluster 6 included 18.73% and 21.66% of the analysed schools, respectively. The three clusters, reported in Fig. 3b, represent almost completely the schools located in South Italy, covering 25.85% of Italian schools (Cluster 4: 6.25%; Cluster 5: 12.08%; Cluster 7: 7.52%). The radar plot highlights that focusing on the WLE in G05 and the ESCS these clusters result to be very similar. In all three clusters, the ESCS effect ranges only from 0.497 to 0.550 and its value is greater than the national average (0.40). On the other hand, the effect of the student previous knowledge, which varies between 0.142 and 0.179, is less than the national one (0.248). This means that in South is more important the family background instead of the previous competencies of the students. In the red cluster (Cluster 7) the 73.88% of the schools is significantly influenced by the presence of late-enrolled students, that has on average a greater negative effect on the performance (− 0.149) than the national average (− 0.09). The presence of immigrant students has a significant and negative effect on the performance for 94.97% of the schools located in Cluster 4, which covers Calabria and Basilicata. For these schools the coefficient results to be on average (− 0.323) three times higher than the national average (− 0.126).



The behaviour of the schools belonging to Cluster 1 and Cluster 3 (Fig. 3c) is very similar; both have a positive effect of the previous competencies greater than the national average (0.31 towards 0.25), whereas the socio-economic background effect is less than the national average. Cluster 1 and Cluster 3 included 20.40% and 13.36% of the analysed lower secondary schools, respectively. In contrast to the reality described for South, in these territories it is very important for the school what the student knows at the beginning of the lower secondary school and not his socio-economic status. Another important thing, for the blue cluster that covers part of the Tyrrhenian coast, is the negative effect, greater than the national average, of the presence of immigrant students in the school (-0.168 towards -0.09).

Concluding remarks

The aim of the study is to investigate how the impact of the contextual variables on academic performance varies by geographic areas in Italy, examining the extent of the spatial disparities moving beyond the regional administrative confines in order to understand the factors that most influence the educational outcome of each school. We aim to identify in which schools/geographical areas it is necessary planning an intervention to flatten the effects of some factors associated in the literature with inequality in education (the school mean of the students previous learning levels, some socio-demographic characteristics of the student body as the school mean of socio-economic and cultural indicator, the school percentage of immigrant students, of late-enrolled students and of females, and the school size defined as the number of classes and the number of school buildings) in order to guarantee greater equality. We exploited the reading standardized tests administered by INVALSI in 2017–2018 focusing on 8th-grade students, analysing at the school-level the data of 5520 out of 5761 Italian lower secondary schools.

In this paper, for the first time, the performances of all Italian schools in INVALSI large-scale assessment survey are analysed using the geo-reference data. The well-known geographical cleavage of Italy is analysed from a new perspective: the geographical disparities relies not only on an unequal distribution of the school performance in terms of academic achievement in Italy but on a different effect of the contextual variables on the school performance. Whereas the OLS model provides an overall picture of the relation between academic performance and contextual variables, the GWR models this relation at local level studying the spatial variation of the regression coefficients. The GWR local modelling outperforms the OLS model, explaining the 10% more of the outcome's variations. This analysis sheds light on how the predictor variables and their effects are related to geography, showing a clear spatial pattern. Applying the k-means clustering on the local regression coefficients we identified seven school clusters that are homogeneous with respect to the factors' effect on school performance. Each cluster has been characterized geographically and in relation to the intensity of predictors statistically significant in the area, in order to identify for each area the critical local factors.

The finding of this work revealed strong variations across the schools of the impact of the contextual variables analysed on the educational outcome, highlighting the presence of strong inequality in the Italian educational system. As it has been shown, the school performance in South Italy results strongly affected by the socio-economic status. This means that low socio-economic conditions impede the student's educational

development entailing low academic performance. On the other hand, the schools that belong to part of the Tyrrhenian and Adriatic coast are able to moderate the socio-economic differences of the students and school's mean performances result to be affected mostly by students' competences at the beginning of the secondary school. However, in the schools of the Tyrrhenian coast, as in some schools in South Italy, particularly in Sicily and Puglia, the strong negative impact of the percentage of late-enrolled students is worrying. The low effect of the previous knowledge of the students in all South is a warning signal of a not efficient educational system, where variables associated with inequality are too relevant. Another warning signal that highlights mechanisms of inequality in the educational system is the high impact of the immigrant status on the school performance in North-Est and Calabria.

In Italy, there is a single school system but the findings of this work outlined a fragmented reality, corroborating the hypothesis to improve decentralization in the educational field. Several researches claimed that the well-known regional differences in human capital are the results of a long history, and the central government should take care of such differences as they turn impact on territorial socio-economic differences that are driving elements in understanding the inequalities in the educational outcome.

In terms of policy implications, the use of the geo-referenced data has the benefit to study the reality at local levels and it could be an important help for policymakers, local administration and schools to design effective strategies to improve students learning based on the evaluation and the identification of critical local factors. The finding of the analyses suggests the need to use different policy strategies depending on the territorial necessity. In the areas where the school outcome is strongly affected by the socio-economic background, it is important to plan actions directly on the students to compensate the family lack in terms of cultural background and to sustain families, organizing, for example, out-of-school care and recreation programs. In the South where the impact of the previous knowledge is particularly high, suggesting that one of the school problems is the efficiency, one possible strategy is to improve the learning opportunities of all the students promoting the educational services starting from the early school years. In the territories where the school outcome is strongly and negatively affected by the percentage of the immigrants, should be planned actions to promote cultural integration and to support students and families. The presence of migrant in schools is no longer a deviant case but is new normality which schools have to address. For new immigrants the first obstacle is the language, one possible strategy is the organization of intensive course, in which language learning is the central effort, and the assistance after regular classes with the instalment in schools of learning and homework centres. The integration of cultural items from countries of emigration in the school life could help the self-esteem of the immigrant students. The high impact of the percentage of the late-enrolled students on the school outcome suggests that the children who are retained learn less than they should have and some schools are not able to integrate the retaining students limiting the learning opportunities of the whole class. For these schools too, one possible strategy could be to plan measures regarding teaching, to increase the elements of support in the teacher's role and to introduce a teacher assistant for help underachievers.

In spite of the many promising aspects of GWR, there are some limitations in this study. Multicollinearity issue in the GWR is still a debated topic. Despite GWR has

been demonstrated to be sufficiently robust to withstand the multicollinearity effects (Fotherigan and Oshan, 2016), the collinearity issues restrain the number of the predictors that can be included into the model. Secondly, while this study supports the use of GWR over nonspatial OLS methods for the prediction of school academic performance, the GWR model is not able to control for the entire spatial autocorrelation as shown from the Moran's I on the residuals. To address this limitation better diagnostic tools should be exploited in future investigations, although approaches to calculate the goodness of fit are still studied. Finally, in this study for the estimation of the GWR model, the data has been aggregated at school level and the student-level risk factors were masked. Multilevel models overcome this problem by combining an individual-level model with a macro-level model. On the other hand, the multilevel model is limited in the modelling of spatial processes by the necessity of an a priori definition of a discrete set of spatial units at each level of hierarchy (Fotherigan et al. 2003). This assumption implies that the outcome is modified in exactly the same way throughout a particular spatial unit, e.g. a region, but the process is modified in a different way outside the boundaries of the spatial unit. This assumption is often unrealistic since the effects of the space are continuous. For this reason, the application of the multilevel model for the spatial analysis is limited, whereas the geographically weighted regression has the advantage to estimate local model and analyse the spatial variation of the predictors of the academic achievement at school-level.

In summary, this work allows to visualize the Italian situation going beyond the classic territorial definition and offers an in-depth analysis of the Italian education system. The identification of new spatial clusters is a useful tool to differentiate supports for schools on the basis of their unique specific needs and it is the first step to understand how to address a more contextualized policy response to educational needs in Italy. To have a complete overview of the Italian situation, future research aims to extend this work to the school performance at mathematics INVALSI tests.

Acknowledgements

Not applicable.

Authors' contributions

PF, CS conceived of the presented idea. CS analysed the data and was a major contributor in writing the manuscript. PF contributed to the interpretation of the results and provided revisions to scientific content of manuscript. All the authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

The datasets analysed during the current study are not publicly available because they contain information that could compromise research participant privacy, but are available from the corresponding author on reasonable request.

Declarations**Ethics approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 5 March 2021 Accepted: 3 June 2021

Published online: 09 June 2021

References

- Agasisti, T., & Vittadini, G. (2012). Regional economic disparities as determinants of students' achievement in Italy. *Research in Applied Economics*, 4, 33–54.
- 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.
- Agasisti, T., Ieva, F., & Paganoni, A. M. (2017). Heterogeneity, school-effects and the North/South achievement gap in Italian secondary education: evidence from a three-level mixed model. *Statistical Methods & Applications*, 26(1), 157–180.
- Azzolini, D., Schnell, P., & Palmer, J. R. (2012). Educational achievement gaps between immigrant and native students in two "new" immigration countries: Italy and Spain in comparison. *The Annals of the American Academy of Political and Social Science*, 643(1), 46–77.
- Azzolini, D., & Barone, C. (2013). Do they progress or do they lag behind? Educational attainment of immigrants' children in Italy: The role played by generational status, country of origin and social class. *Research in Social Stratification and Mobility*, 31, 82–96.
- Bakker, J., Denessen, E., & Brus-Laeven, M. (2007). Socio-economic background, parental involvement and teacher perceptions of these in relation to pupil achievement. *Educational Studies*, 33, 177–192.
- Benadusi, L., Fornari, R., & Giancola, O. (2010). La questione dell'equità scolastica in Italia. Analisi delle performance in scienze dei quindicenni della scuola secondaria superiore, Fondazione Giovanni Agnelli, http://www.fga.it/uploads/media/L_Benadusi_R_Fornari_O_Giancola_La_questione_dell_equita_scolastica_in_Italia_-_FGA_WP26.pdf
- Benadusi, L. (2016). Le malattie della scuola italiana: molto serie ma non drammatiche. *Scienze e Ricerche*, 23, 12–14.
- Binbin, L., Paul, H., Martin, C., & Brunsdon, C. (2014). The GWmodel R package: further topics for exploring spatial heterogeneity using geographically weighted models. *Geo-spatial Information Science*, 17(2), 85–101.
- Bivand, R., Pebesma, E., & Gomez-Rubio, V. (2013). *Applied spatial data analysis with R* (2nd ed.). Springer.
- Bivand, R., & Wong, D. W. S. (2018). Comparing implementations of global and local indicators of spatial association. *Test*, 27(3), 716–748. <https://doi.org/10.1007/s11749-018-0599-x>
- Bratti, M., Checchi, D., & Filippin, A. (2007). *Da dove vengono le competenze degli studenti? I divari territoriali nell'indagine OCSE PISA 2003*. Il mulino.
- Breakspear, S. (2012). The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance. OECD Education Working Papers, No. 71. *OECD Publishing (NJ)*
- Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(3), 431–443.
- Campodifiori, E., Figura E., Papini, M., Ricci, R. (2010). Un indicatore di status socio-economico-culturale degli allievi della quinta primaria in Italia. Working paper n. 2, INVALSI. (http://www.invalsi.it/download/wp/wp02_Ricci.pdf)
- Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A., & Charrad, M. M. (2014). Package 'nbclust'. *Journal of statistical software*, 61, 1–36.
- Checchi, D., & Peragine, V. (2010). Inequality of opportunity in Italy. *The Journal of Economic Inequality*, 8(4), 429–450.
- Contini, D., Tommaso, M. L. D., & Mendolia, S. (2017). The gender gap in mathematics achievement: Evidence from Italian data. *Economics of Education Review*, 58, 32–42.
- Costanzo, A., & Desimoni, M. (2017). Beyond the mean estimate: a quantile regression analysis of inequalities in educational outcomes using INVALSI survey data. *Large-scale Assessments in Education*, 5(1), 14.
- Department for Education and Skills(2007)Gender and education: the evidence on pupils in England,<http://webarchive.nationalarchives.gov.uk/20090108131525/http://dcsf.gov.uk/research/data/uploadfiles/rtp01-07.pdf>, 2007Department for Education and Skills (2007) Gender and education: the evidence on pupils in England, <http://webarchive.nationalarchives.gov.uk/20090108131525/http://dcsf.gov.uk/research/data/uploadfiles/rtp01-07.pdf>
- European Union (2009) Council conclusions of 12 May 2009 on a strategic framework for European Cooperation in education and training ('ET 2020') *Official Journal of the European Union*, 52, C–119/2–10.
- Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2003). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Wiley.
- Fotheringham, A. S., & Oshan, T. M. (2016). Geographically weighted regression and multicollinearity: dispelling the myth. *Journal of Geographical Systems*, 18(4), 303–329.
- Fullan, M. (2009). Large-scale reform comes of age. *Journal of educational change*, 10(2–3), 101–113.
- Gennaioli, N., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2012). Human capital and regional development. *The Quarterly journal of economics*, 128(1), 105–164.
- Giambona, F., & Porcu, M. (2015). Student background determinants of reading achievement in Italy. A quantile regression analysis. *International Journal of Educational Development*, 44, 95–107.
- Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. (2015). GWmodel: An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models. *Journal of Statistical Software*, 63(17), 1–50.
- Grilli, L., & Rampichini, C. (2009). Multilevel models for the evaluation of educational institutions: a review. In P. Monari, M. Bini, D. Piccolo, & L. Salmaso (Eds.), *Statistical methods for the evaluation of educational services and quality of products* (pp. 61–80). Physica.
- Gursakal, S., Murat, D., & Gursakal, N. (2016). Assessment of PISA 2012 results with quantile regression analysis within the context of inequality in educational opportunity. *Alphanumeric Journal*. <https://doi.org/10.17093/aj.2016.4.2.5000186603>
- Hanushek, E. A., & Kimko, D. (2000). Schooling, labor force quality, and economic growth. *American Economic Review*, 90(5), 1184–1208.
- Hanushek, E. A., & Woessman, L. (2009). Do better schools lead to more growth? Cognitive skills, economic outcomes and causation. *NBER Working Paper No. 14633*.
- Hanushek, E. A., & Woessmann, L. (2011). The Economics of International Differences in Educational Achievement. *Handbook of the Economics of Education* (Vol. 2, pp. 89–200). North Holland.
- Heck, R. H. (2000). Examining the impact of school quality on school outcomes and improvement: A value-added approach. *Educational administration quarterly*, 36(4), 513–552.

- Hippe, R., Jakubowski, M., & Araújo, L. (2018). *Regional inequalities in PISA: the case of Italy and Spain*. Publications Office of the European Union.
- INVALSI (2018) I risultati delle prove INVALSI 2018 a colpo d'occhio. https://invalsi-areaprove.cineca.it/docs/2018/Uno_sguardo_generale_sui_risultati_delle_prove_INVALSI_2018.pdf
- INVALSI (2019) I risultati delle prove INVALSI 2019 a colpo d'occhio. https://invalsi-areaprove.cineca.it/docs/2019/Uno_sguardo_generale_sui_risultati_delle_prove_INVALSI_2019.pdf
- Italian Ministry of Education (2012) (*La scuola in cifre 2011*) *Numbers about schools 2011*. Rome, Italy: The Ministry of Education.
- Italian Ministry of Education (2018) Focus "Principali dati della scuola scuola – Avvio Anno Scolastico 2018/2019. Rome, Italy: The Ministry of Education. <https://www.miur.gov.it/documents/20182/0/Principali+dati+della+scuola+--+avvio+anno+scolastico+2018-2019.pdf/fb3e7b10-e2bc-49aa-a114-c41ef53cac9?version=1.0>
- Lee, V. E., & Loeb, S. (2000). School size in Chicago elementary schools: Effects on teachers' attitudes and students' achievement. *American Educational Research Journal*, 37, 3–31.
- Legewie, J., & DiPrete, T. A. (2012). School Context and the Gender Gap in Educational Achievement. *American Sociological Review*, 77(3), 463–85.
- Leithwood, K., & Jantzi, D. (2009). A review of empirical evidence about school size effects: A policy perspective. *Review of educational research*, 79(1), 464–490.
- Longobardi, S., Quintano, C., & Castellano, R. (2009). The influence of social and economic factors on the competencies of Italian students. A multilevel analysis of data OECD-PISA 2006. *Rivista di Economia e Statistica del Territorio*, 2, 109–149.
- Longobardi, S., Falzetti, P., & Pagliuca, M. M. (2018). Quis custodiet ipsos custodes? How to detect and correct teacher cheating in Italian student data. *Statistical Methods & Applications*, 27(3), 515–543. <https://doi.org/10.1007/s10260-018-0426-2>
- Luyten, H. (2014). Quantitative summary of research findings. In H. Luyten, M. Hendriks, & J. Scheerens (Eds.), *School size effects revisited: A qualitative and quantitative review of the research evidence in primary and secondary education* (pp. 177–218). Springer.
- Marks, G. N. (2006). Are between-and-within school differences in student performance largely due to socio-economic background? Evidence from 30 countries. *Educational Research*, 48, 21–40.
- Matteucci, M., & Mignani, S. (2014). Exploring regional differences in the reading competencies of Italian students. *Evaluation review*, 38(3), 251–290.
- OECD. (2005). *PISA 2003 Technical Report*. OECD Publishing.
- OECD. (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. OECD Publishing.
- Panichella, N., & Triventi, M. (2014). Social inequalities in the choice of secondary school: Long-term trends during educational expansion and reforms in Italy. *European Societies*, 16(5), 666–693.
- Quintano, C., Castellano, R., & Longobardi, S. (2009). A fuzzy clustering approach to improve the accuracy of Italian student data. *Statistica e Applicazioni*, 7(2), 149–171.
- Robinson, J. P., & Lubiensky, S. T. (2011). The Development of Gender Achievement Gaps in Mathematics and Reading During Elementary and Middle School Examining Direct Cognitive Assessments and Teacher Ratings. *American Educational Resources Journal*, 48, 2268–302.
- Rodriguez-Posé, A., & Tselios, V. (2009). Education and income inequality in the regions of the European union. *Journal of Regional Science*, 49, 411–437.
- Schwartz, A. E., Stiefel, L., & Wiswall, M. (2013). Do small schools improve performance in large, urban districts? Causal evidence from New York City. *Journal of Urban Economics*, 77, 27–40.
- Scheerens, J., Hendriks, M., & Luyten, H. (2014). School size effects: Review and conceptual analysis. In H. Luyten, M. Hendriks, & J. Scheerens (Eds.), *School size effects revisited: A qualitative and quantitative review of the research evidence in primary and secondary education* (pp. 7–40). Springer.
- Schnell, P., & Azzolini, D. (2015). The academic achievements of immigrant youths in new destination countries: Evidence from southern Europe. *Migration Studies*, 3(2), 217–240.
- Schuetz, G., Ursprung, H., & Woessmann, L. (2008). Education policy and equality of opportunity. *Kyklos*, 61(2), 279–308.
- Teodorovic, J. (2012). Student background factors influencing student achievement in Serbia. *Educational Studies*, 38, 89–110.
- Wang, M. T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational psychology review*, 29(1), 119–140.
- Wasley, P. A., Fine, M., Gladden, M., Holland, N. E., King, S. P., Mosak, E., & Powell, L. (2000) *Small schools: Great strides*. New York, NY: Bank Street College of Education.
- Weiss, C. C., Carolan, B. V., & Baker-Smith, E. C. (2010). Big school, small school:(Re) testing assumptions about high school size, school engagement and mathematics achievement. *Journal of Youth and Adolescence*, 39(2), 163–176.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.