



**Istituto nazionale per la valutazione del sistema educativo di
istruzione e di formazione**

WORKING PAPER N. 46/2020

Analisi spaziale degli apprendimenti scolastici negli istituti del Sud Italia

Silvia Donno - INVALSI

Cecilia Bagnarol - INVALSI

Michele Marsili - INVALSI

Collana: Working Papers INVALSI

ISSN: 2611 - 5719

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the view and the official policy or position of INVALSI.

Le opinioni espresse nei lavori sono attribuibili esclusivamente agli autori e non impegnano in alcun modo la responsabilità dell'Istituto. Nel citare i temi, non è, pertanto, corretto attribuire le argomentazioni ivi espresse all'INVALSI o ai suoi Vertici

Abstract

Il sistema scolastico italiano è caratterizzato da una grande complessità territoriale che vede storicamente contrapposte le macro-aree Nord e Sud del Paese.

Le Rilevazioni Nazionali INVALSI confermano la presenza di differenze lungo l'asse Nord-Sud anche per gli apprendimenti scolastici, infatti nelle macro-aree del Sud i punteggi medi di Italiano e Matematica risultano significativamente inferiori alla media italiana per tutti i gradi scolastici.

Tale differenze possono, in parte, essere spiegate dalla diversa efficacia educativa dei singoli istituti scolastici (valore aggiunto) ma sono rafforzate dalla differente distribuzione territoriale dell'indice di background socio-economico-culturale (ESCS) e dal diverso background migratorio degli studenti e delle famiglie.

Questo lavoro si pone l'obiettivo di analizzare, per il Meridione, le differenze negli apprendimenti e nelle competenze a livello di sub-area e, in seguito, ove possibile, individuare alcune specifiche realtà scolastiche resilienti rispetto al contesto socio-economico e culturale sfavorevole in cui sono collocate.

Lo studio focalizza l'attenzione sugli esiti delle prove INVALSI di Italiano e Matematica per l'anno scolastico 2018-2019 degli allievi delle classi terze della scuola secondaria di I grado provenienti dalle aree geografiche Sud e Sud e Isole.

La rappresentazione sul territorio degli esiti delle prove INVALSI ha utilizzato, inizialmente, un'aggregazione territoriale di tipo istituzionale (regionale e provinciale). Successivamente, la georeferenziazione delle scuole ha consentito di superare il concetto di unità amministrativa individuando, per comunanza di caratteristiche ed esiti delle prove, alcune aggregazioni funzionali di scuole. Su tali aggregazioni di istituti, infine, è stata condotta un'analisi spaziale univariata e bivariata utilizzando le misure di autocorrelazione spaziale globale (I di Moran) e locale (LISA).

Parole chiave: Analisi spaziale; Scuola secondaria primo grado; Sud Italia; Resilienza

Keywords: Spatial analysis; Lower secondary school; South Italy; Resilience

Introduzione

Esiste un'ampia letteratura che documenta come in Italia le storiche disuguaglianze territoriali nella crescita economica e nello sviluppo lungo l'asse Nord-Sud interessino anche l'efficienza-efficacia del sistema scolastico ed educativo.

Se si tiene conto delle importanti e sostanziali differenze regionali nella capacità delle scuole di assicurare a tutti gli studenti medesimi servizi e medesime condizioni di insegnamento-apprendimento, il sistema scolastico italiano è ancora oggi ben lungi dal rappresentare un continuum equo ed uniforme.

La mancanza di uniformità nell'efficacia educativa dei singoli istituti scolastici (*valore aggiunto*) e delle diverse realtà regionali è un argomento frequentemente trattato a livello internazionale (Indagini OCSE-PISA e TIMMS/PIRLS di IEA) e a livello nazionale (Indagini INVALSI). Ognuna di esse approfondisce uno o più ambiti di studio come la lettura, la matematica, le scienze e il problem-solving, ed evidenzia una forte diversificazione territoriale: risulta infatti che gli studenti meridionali si attestano al di sotto degli standard internazionali e di quelli delle regioni settentrionali, le differenze a livello regionale sono significative e vi è una elevata variabilità dei risultati tra le singole scuole in una stessa area.

I fattori che spiegano tale divario di competenze tra macro-aree geografiche e tra regioni sono molteplici e di diversa natura, ognuno di essi genera una diversa interpretazione del fenomeno (Pavolini et al. 2015; Checchi, 2007; Bratti, Checchi, Filippin, 2007). In primo luogo, è opinione diffusa che la condizione socio-culturale ed economica del *background* degli studenti e delle loro famiglie giochi un ruolo cruciale sui livelli di apprendimento conseguiti e ha su di essi un notevole valore predittivo: chi vive in condizioni di maggiore svantaggio economico, ma anche sociale e culturale, ha minori possibilità di conseguire risultati più soddisfacenti durante il percorso formativo (Brunello e Checchi, 2006; Raimondi, De Luca, Barone, 2013). L'influenza del capitale socio-economico-culturale familiare, inoltre, sembra esercitarsi nella scelta dell'istituto scolastico superiore incidendo sul percorso scolastico dello studente negli anni successivi (Montanaro, 2008; Melchiori, 2013) tenendo presente che il tipo di scuola prescelto (liceo, istituto tecnico o professionale) e le caratteristiche specifiche dell'istituto frequentato sono da considerarsi ulteriori fattori predittivi degli esiti scolastici (MEF, 2012). Il background migratorio di per sé non rappresenta un fattore di insuccesso, ma concorre e influisce sulla precarietà delle condizioni socio-economico-culturali delle famiglie, contribuendo a creare e ad esacerbare il gap territoriale negli apprendimenti scolastici.

Di certo considerare i soli fattori individuali e familiari, seppur rilevanti, non può dar conto dei differenziali territoriali nell'apprendimento: un ruolo importante è attribuito al cosiddetto "effetto di contesto" o effetto di



composizione del gruppo, per cui l'apprendimento di uno studente è influenzato di certo dalle caratteristiche personali ma anche da quelle medie del gruppo-classe-scuola con cui interagisce (Rapporto INVALSI, 2019). In un contesto scolastico come quello italiano in cui permangono differenze territoriali strutturali e radicate, alcuni studi recenti (Agasisti et al., 2015; Longobardi e Agasisti, 2014) hanno rintracciato ed esplorato il caso degli studenti “resilienti”, che superano la propria condizione di svantaggio socio-economico e culturale e ottengono buoni risultati scolastici.

Nella stessa ottica la nostra ricerca ha verificato l'esistenza di “scuole resilienti” nel Sud Italia, realtà in cui la condizione socio-economico-culturale di svantaggio è stata superata e per le quali si registrano buoni livelli di competenze, talvolta superiori alla media nazionale. Inoltre, è nostra opinione che la tradizionale e sempre attuale contrapposizione dualistica Nord-Sud non può e non deve rappresentare l'approccio più adatto a cogliere le caratteristiche territoriali del sistema scolastico italiano, tanto più che, sebbene al Sud si concentrino le maggiori criticità in termini di apprendimenti scolastici, la distribuzione degli stessi è nettamente eterogenea tra gli istituti. Il supporto dell'analisi di Autocorrelazione Spaziale locale (Anselin, 1995) ha concesso l'approccio a nuovi scenari di analisi e di ricerca, fornendo “*dati a un livello di dettaglio geografico più elevato e confini geografici flessibili, individuati a partire dall'influenza che un'area esercita sul territorio circostante [...] e quindi non necessariamente coincidenti con la delimitazione amministrativa o con le caratteristiche morfologiche*” (ISTAT, 2018).

Dati

Il dataset utilizzato per l'analisi ha preso in esame esclusivamente gli esiti della rilevazione in modalità CBT (Computer Based Test) sugli apprendimenti in Italiano e Matematica degli allievi delle classi III della scuola secondaria di I grado (in seguito Grado 8) provenienti dalle aree geografiche Sud (sono comprese le regioni Abruzzo, Molise, Campania e Puglia) e Sud e Isole (sono comprese le regioni Basilicata, Calabria, Sicilia e Sardegna) del Paese per l'anno scolastico 2018-19. Al fine di approfondire l'analisi, alle variabili del dataset sono state associate anche le informazioni fornite dal Questionario compilato dagli studenti, utili al reperimento di una ricca gamma di informazioni su loro stessi e sul loro contesto familiare (ad es. il numero di libri a disposizione in casa, la possibilità di avere un luogo tranquillo in cui studiare, un computer da utilizzare per lo studio con il relativo software, ecc.), per il calcolo dell'indicatore sul background socio-economico-culturale (ESCS) dello studente.

Metodi

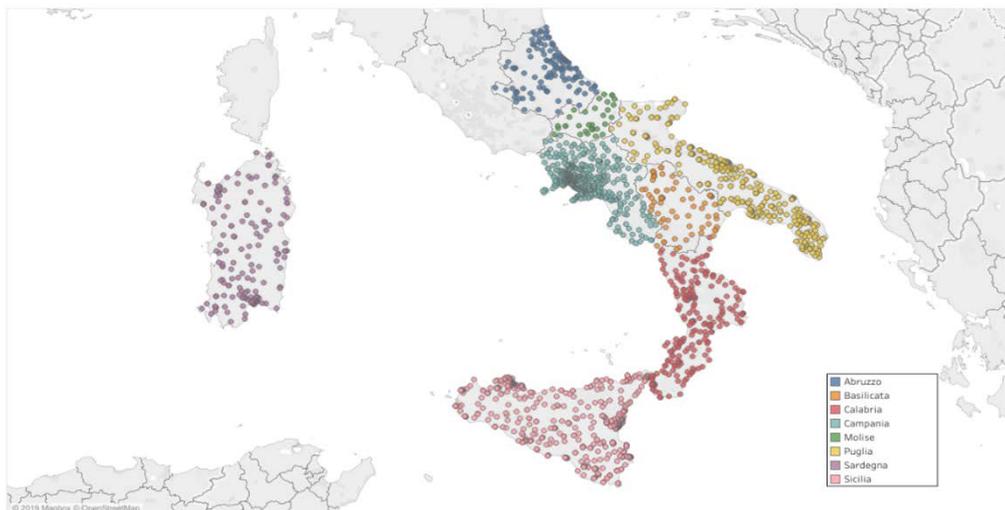
Per la georeferenziazione degli istituti scolastici del Meridione è stato creato un dataset a partire dalle informazioni disponibili sulla localizzazione delle scuole, costituito dalle seguenti variabili:

- codice identificativo dell'istituto scolastico;
- indirizzo;
- nome della città;
- CAP.

Il processo di conversione degli indirizzi in coordinate geografiche (latitudine e longitudine) è stato effettuato tramite le API (Application Programming Interface) di Google Maps.

La preparazione dei dati per le successive analisi spaziali è stata effettuata utilizzando KNIME Analytics Platform (Berthold et al., 2008), una piattaforma open source di analisi dati, reportistica, machine learning e data mining.

Figura 1. Georeferenziazione delle scuole secondarie di primo grado del Sud Italia (a.s. 2018-19)



Cenni sugli Indici di Autocorrelazione spaziale globale e locale univariati e bivariati

Il concetto di autocorrelazione (o associazione) spaziale deriva dall'osservazione che i valori assunti da una variabile non sono distribuiti in maniera indipendente sul territorio ma, al contrario, tendono a concentrarsi in determinate zone (Demarinis et al., 2011). In particolare si parla di:

- *autocorrelazione spaziale positiva* quando valori simili di una variabile tendono a raggrupparsi in prossimità l'uno dell'altro, a formare cioè cluster più o meno grandi;
- *autocorrelazione spaziale negativa* quando valori dissimili di una variabile tendono a raggrupparsi in prossimità l'uno dell'altro;
 - *assenza di autocorrelazione spaziale (o indipendenza spaziale)* quando la distribuzione dei valori nello spazio è casuale.

Tra i metodi che misurano l'autocorrelazione spaziale si distinguono i metodi per la misurazione sull'insieme completo di località oggetto di studio (*misure globali*) e i metodi per la misurazione su un sottoinsieme spazialmente delimitato di località (*misure locali*).

La misura più comunemente utilizzata per testare il grado di *autocorrelazione spaziale globale* è la statistica "I di Moran" (Moran, 1948), formalizzata dalla formula:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

dove:

- N è il numero di osservazioni (cioè delle località o unità geografiche),
- x_i è il valore della variabile considerata nella località i ,
- x_j è il valore della variabile considerata nella località j ,
- \bar{x} è la media campionaria,
- $(x_i - \bar{x})$ rappresenta la deviazione dalla media della variabile di interesse,
- w_{ij} è un peso assegnato alla relazione tra la località i e la località j .

La statistica I di Moran è strutturalmente simile al coefficiente di correlazione e assume valori compresi tra -1 e $+1$.

L'indice I di Moran può essere esteso al caso multivariato: in questo modo l'indice rappresenta l'associazione sistematica tra i valori di una variabile osservata x in una data area di interesse e i valori di un'altra variabile y osservata in aree limitrofe.

Si ottiene in questo modo l'*Indice di Moran bivariato* di x rispetto a y :

$$I_b = \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} z_i v_j}{(\sum_{i=1}^N \sum_{j=1}^N w_{ij})(\sum_{i=1}^N z_i^2) / N}$$

Le statistiche per la misurazione del grado di *autocorrelazione spaziale a livello locale* consentono di individuare il contributo che ogni località fornisce al comportamento globale e quindi permettono di studiare le variazioni dell'autocorrelazione spaziale all'interno del territorio. Focalizzando l'attenzione su ogni singola località, queste tecniche possono quindi essere impiegate per individuare la presenza di cluster spaziali.

L'indicatore locale di associazione spaziale è noto come *Local Indicator of Spatial Association (LISA)*. L'indice LISA più comunemente utilizzato è rappresentato dalla versione locale della statistica I di Moran ed è definito dalla relazione:

$$I_i = N \frac{x_i - \bar{x}}{\sum_i (x_i - \bar{x})^2} \sum_{i=1, i \neq j} w_{ij} (x_j - \bar{x})$$

Risultati

Ai dati INVALSI della Rilevazione Nazionale degli apprendimenti di Italiano e Matematica, per l'a.s. 2018-19 delle classi terze della scuola secondaria di primo grado del Sud Italia sono state connesse le informazioni per il calcolo dell'indice ESCS fornite dal Questionario studente.

In via preliminare è stata svolta una operazione di data cleaning al fine di escludere le scuole non rappresentative (scuole con un numero di studenti inferiore a 20, numerosità media di tutte le classi) e le scuole presenti sulle isole minori per evitare che la loro distanza dalla terraferma potesse avere una forte influenza nel calcolo dell'indice di Moran a cui si farà riferimento in seguito.

Le variabili oggetto di studio sono:

1. il punteggio numerico conseguito dagli studenti rispettivamente nelle prove di Italiano (WLE ITA) e Matematica (WLE MAT), espresso su scala quantitativa (*Rash*);
2. i valori dell'indice di condizione socio-economica ESCS.

L'unità statistica minima presa in esame è la scuola, pertanto i punteggi individuali degli studenti sono stati aggregati prima a livello di scuola (in Italiano si registra un WLE medio di istituto pari a 189,93 con d.s. pari a 37,21 mentre in Matematica un WLE medio pari a 186,22 con d.s. pari a 36,88) e successivamente a livello di provincia.

La lettura congiunta dei punteggi medi a livello provinciale e dei valori provinciali dell'indice ESCS risulta particolarmente significativa: i valori dell'indice socio-economico ESCS (la cui media nazionale è posta pari a 0) sono riportati sull'asse delle ascisse mentre sull'asse delle ordinate è riportato il punteggio medio provinciale, rispettivamente di Italiano e Matematica (Figura 2 e Figura 3 riportate di seguito). In entrambi i grafici le linee tratteggiate rappresentano la media degli istituti scolastici del Sud Italia per i due indicatori rappresentati (il valore dell'indice ESCS nel Sud Italia si registra pari a -0.15). Dai grafici a seguire si possono individuare alcuni particolari cluster di province:

- le realtà provinciali che presentano un valore dell'indice ESCS inferiore alla media nazionale e un WLE medio nelle prove inferiore alla media del Meridione: è il caso di alcune province della Sicilia (Palermo, Agrigento, Trapani, Ragusa, Caltanissetta), e della provincia di Crotona in Calabria, realtà socio-economico-culturali svantaggiate che potrebbero aver penalizzato le performance scolastiche influenzando negativamente sui punteggi. Nella provincia di Crotona si registrano i punteggi più bassi sia in Italiano che in Matematica (rispettivamente pari a 178,80 e 173,05), mentre a Palermo si riscontra il valore dell'indice ESCS più basso, pari a -0,38;
- le realtà provinciali che presentano un valore dell'ESCS superiore alla media del Sud Italia ma WLE ITA e MAT inferiori alla media: ad esempio le province di Vibo Valentia e Reggio Calabria per Italiano a cui si aggiungono Catanzaro, Cosenza e Siracusa per Matematica, le cui scuole, nonostante la situazione socio-economica-culturale sia migliore rispetto a quella del territorio circostante, non conseguono risultati incoraggianti;
- le province che conseguono WLE ITA e MAT superiori alla media del Sud Italia e hanno un valore dell'ESCS superiore alla media del Sud: in queste province la situazione socio-economico e culturale favorevole permetterebbe agli studenti di ottenere risultati migliori. Di questo gruppo fanno parte tutte le province dell'Abruzzo, le due province molisane di Isernia e Campobasso, Lecce, Bari, le province campane di Benevento, Avellino e Salerno, le province di Matera e Potenza. Nella provincia di Pescara si registrano le performance migliori sia in Italiano che in Matematica (WLE ITA= 203,61 e WLE MAT= 203,23) che superano anche la media nazionale, come accade per la provincia di Chieti per entrambe le materie e per L'Aquila, Campobasso e Lecce in Italiano;
- le province "resilienti": benché presentino un valore dell'indice ESCS inferiore alla media, le loro performance scolastiche risultano promettenti. È il caso delle province pugliesi di Barletta-Andria-Trani, Taranto e delle province della Sardegna, e per Italiano delle province siciliane di Catania ed Enna.

Figura 2. Distribuzione delle province del Sud Italia per WLE medio in Italiano e indice ESCS (a.s. 2018-19)

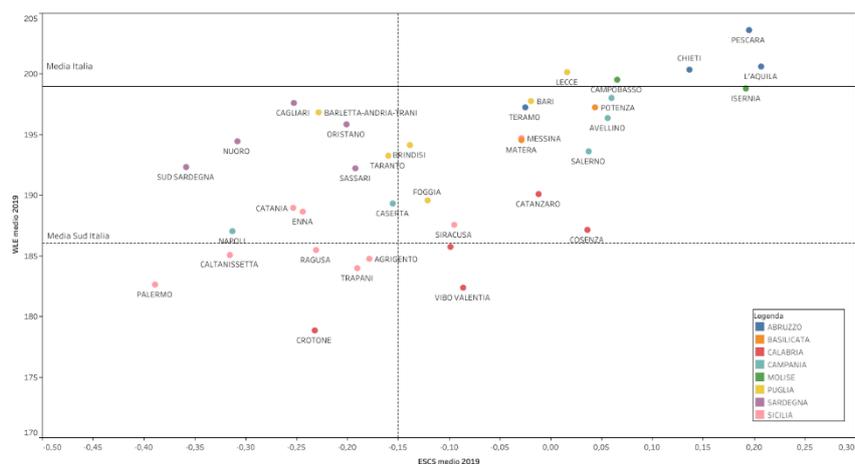
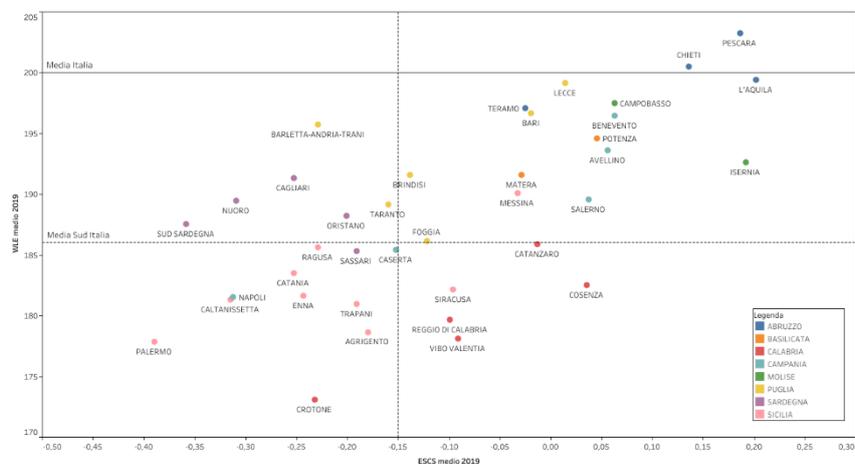


Figura 3. Distribuzione delle province del Sud Italia per WLE medio in Matematica e indice ESCS (a.s. 2018-19)



Data l'ampiezza e la complessità del territorio italiano, l'analisi descrittiva territoriale potrebbe non far emergere delle realtà scolastiche molto diverse tra di loro. Per questo motivo si è ritenuto opportuno approfondire l'indagine mediante l'analisi di autocorrelazione spaziale, utilizzando l'indice I di Moran, lo Scatterplot di Moran (Anselin, 1996) e gli indici di associazione spaziale locale (LISA) (Anselin, 1995), con l'ausilio del software per l'analisi dei dati spaziali GeoDa (Anselin, 2003).

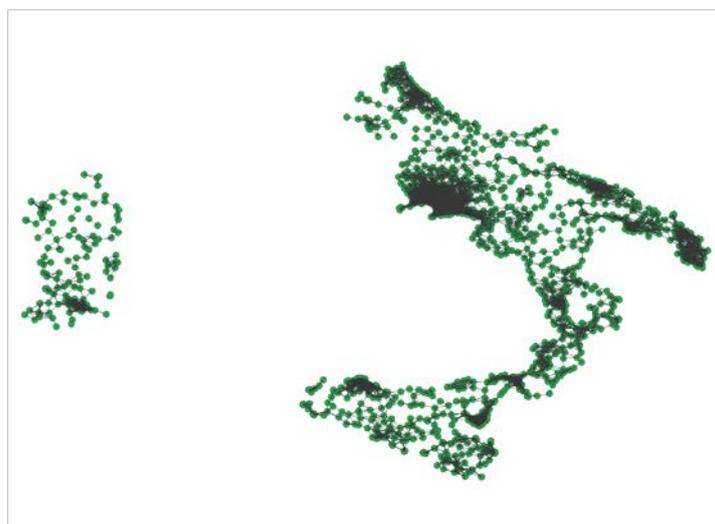
L'indice I di Moran restituisce un'indicazione del grado di associazione lineare tra i valori osservati della variabile in esame e i valori spazialmente ritardati.

Le misure di autocorrelazione spaziale come l'indice I di Moran richiedono la costruzione della matrice dei pesi che definisce un intorno per ciascuna unità geografica. Il valore della variabile di interesse per ciascuna unità è confrontato con il peso medio dei valori delle unità vicine.

Per costruire la matrice dei pesi, basati sulle distanze spaziali tra due punti, quando i punti sono espressi in termini di latitudine e longitudine, si usa la “*great circle distance*” o “*arc-distance*”, che tiene conto della curvatura della terra.

La distanza minima per assicurare che ciascun istituto avesse almeno un “*neighbor*” (*vicino*) è risultata pari a circa 35 chilometri. Il numero medio di “*neighbors*” analizzati è 50,88, la mediana è pari a 19. La Figura 4, di seguito riportata, illustra la connettività tra gli istituti scolastici del Sud d'Italia.

Figura 4. Connectivity graph per gli istituti scolastici del Meridione

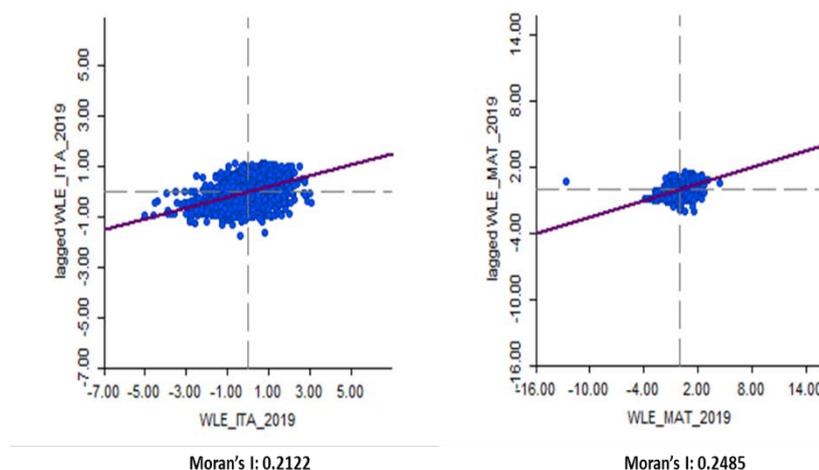


L'indice globale I univariato ha rivelato la presenza di autocorrelazione spaziale per la distribuzione dei punteggi di Italiano e Matematica, rispettivamente $I \approx 0,2122$ e $I \approx 0,2485$, in entrambi i casi l'indice è significativo (lo pseudo p-value è $< 0,001$). Tuttavia, l'indice di Moran non consente di verificare se la dipendenza spaziale generi cluster di scuole per livello di apprendimento, né di individuare confini geografici degli stessi cluster. Al fine di superare questi limiti è stato adottato il Moran Scatterplot e le misure di autocorrelazione locale.

Il Moran Scatterplot (Figura 5) riporta su di un grafico cartesiano i valori della variabile di interesse standardizzati ed espressi in unità di deviazione standard (la media è posta uguale a 0 e la deviazione standard è pari a 1) sull'asse delle ascisse e sull'asse delle ordinate i corrispettivi valori spazialmente ritardati anch'essi standardizzati. L'indice I rappresenta il coefficiente angolare della relazione lineare tra le due variabili

riportate sugli assi: se i punti rappresentati risultano dispersi fra i quattro quadranti questo indicherà assenza di correlazione (il coefficiente angolare è zero), se invece esiste una relazione, lo Scatterplot consentirà di distinguere le diverse tipologie di correlazione spaziale (High-High, Low-Low, High-Low, Low-High).

Figura 5. Moran Scatterplot per WLE_ITA e WLE_MAT (a.s. 2018-19)



I risultati degli Scatterplot precedenti sono stati riportati su di un cartogramma in modo da caratterizzare geograficamente le aree con diverse tipologie di correlazione: tale rappresentazione ha consentito di verificare la contiguità geografica delle scuole accomunate dallo stesso tipo di correlazione e di identificare alcuni cluster omogenei nel Meridione.

Tuttavia, il Moran Scatterplot non fornisce informazioni sulla significatività dei raggruppamenti spaziali rintracciati, dunque è stato calcolato il Local Indicator of Spatial Association (LISA), in modo da misurare per ciascun istituto scolastico l'interdipendenza con gli altri istituti ed indicarne la tipologia (positiva o negativa) e la significatività. I risultati così ottenuti sono stati riportati su di un cartogramma (Figura 6 e Figura 7) su cui è stato possibile individuare 4 agglomerazioni:

- il primo cluster è formato da scuole contigue con punteggi superiori alla media del Sud (High-High, riportati in rosso nella mappa) ed interessa alcune aree situate in Abruzzo, Molise, Puglia (le province di Lecce e Bari), Campania (nelle province di Benevento, Avellino e Salerno) e la provincia di Potenza; inoltre per Italiano le province di Messina, Cagliari e Reggio Calabria e per Matematica la provincia di Palermo;
- il secondo cluster è formato da scuole contigue con punteggi bassi (Low-Low, in blu nella mappa) e interessa principalmente aree situate in Calabria (ad eccezione della provincia di Catanzaro), Sicilia (nelle

province di Palermo, Catania e Agrigento in Matematica a cui si aggiunge Trapani e Ragusa per Italiano) e nella provincia di Napoli;

- le scuole che formano il cluster Low-High (in viola sulla mappa) sono posizionate prevalentemente in Puglia (per Italiano sono concentrate nelle province di Lecce e Bari, mentre per Matematica sono diffuse anche nelle altre province ad eccezione della provincia di Foggia e Taranto), Campania (si può rintracciare un arco geografico tra le province di Benevento, Avellino e Salerno) e Abruzzo (principalmente in Italiano nelle province di Teramo, Pescara e l'Aquila). Si tratta di scuole con punteggi bassi che non sembrano beneficiare della collocazione territoriale in zone in cui le scuole hanno punteggi mediamente più alti;
- l'ultimo cluster (High-Low, in rosa sulla mappa) è formato dalle scuole con punteggi superiori alla media con una bassa contaminazione territoriale, cosiddette "resilienti". Le scuole appartenenti a questo cluster si trovano prevalentemente in Campania (nelle province di Napoli e Caserta), Calabria (sporadiche e sparse in tutte le province) e Sicilia (nelle province di Palermo, nella zona sud orientale e per Matematica nella provincia di Catania);
- i coefficienti di correlazione locale non risultano statisticamente significativi, per entrambe le materie, in Sardegna (ad eccezione di quelli misurati su alcune scuole della provincia di Cagliari), in Basilicata, nelle province più a Sud della Campania, in Molise e nella provincia di Foggia. Per i punteggi di Matematica non sembra esserci significatività nei coefficienti di correlazione locale anche in Sardegna.

Figura 6. LISA cluster map per WLE ITA (a.s. 2018-19)

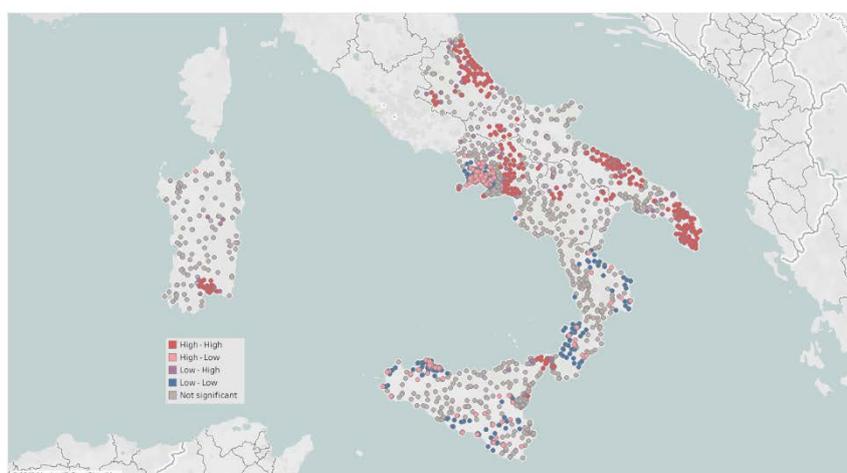
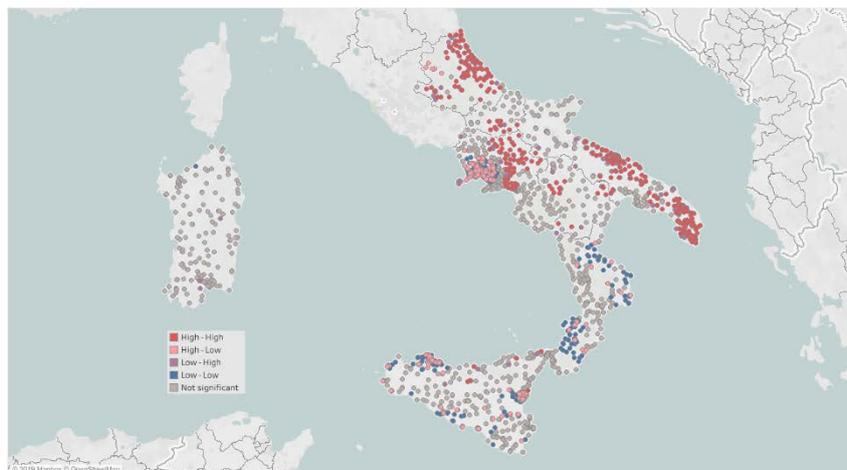


Figura 7. LISA cluster map per WLE MAT (a.s. 2018-19)



Come evidenziato dall'analisi descrittiva territoriale le province dell'Abruzzo hanno punteggi superiori alla media del Meridione per entrambe le materie. I coefficienti di correlazione locale evidenziano dei raggruppamenti di scuole appartenenti al cluster High-High in ognuna di queste province. Tuttavia, nelle province di Teramo, Pescara e Chieti si osservano coefficienti significativi su tutto il territorio, mentre nella provincia di L'Aquila i raggruppamenti di scuole con punteggi più alti della media e significativamente correlati con quelli delle scuole contigue si osservano nella Valle dell'Aterno e nelle zone della Marsica confinanti con il Lazio. Anche nella regione Basilicata, sono rintracciabili due raggruppamenti di scuole High-High, in entrambe le province: un gruppo di scuole della provincia di Potenza, al confine con le province campane di Salerno ed Avellino e un gruppo di scuole della provincia di Matera, al confine con la Puglia, che sembrano definire un continuum con le realtà regionali limitrofe, più che con il resto della propria regione. In Campania, gli istituti con punteggi più alti si registrano nelle province di Benevento, Avellino (al confine con le province molisane e la provincia di Barletta-Andria-Trani) e Salerno. Tra le province di Messina e Reggio Calabria, per Italiano, si riscontra un ulteriore continuum extra-regionale.

Dall'altro lato, l'analisi territoriale ha mostrato punteggi inferiori alla media per tutte le province della Calabria. Raggruppamenti di scuole con coefficienti di correlazione locale negativi e statisticamente significativi sono diffusi nella zona orientale della provincia di Cosenza, Crotona, Vibo Valentia e nella parte settentrionale della provincia di Reggio Calabria. In Sicilia tale situazione riguarda la zona nord-occidentale, nella provincia di Palermo e Trapani, e la zona sud-orientale dell'isola: gli istituti interessati si trovano tra la parte sud-orientale della provincia di Agrigento, le zone più interne della provincia di Ragusa e Siracusa, la

città di Catania e alcuni comuni più interni della sua provincia ed infine alcuni comuni della provincia di Caltanissetta più vicini alla costa.

Un'ulteriore raggruppamento di scuole appartenenti al cluster Low-Low, si osserva in Campania, nelle province di Napoli e Caserta, dove si registrano anche le più alte concentrazioni di scuole appartenenti al cluster High-Low.

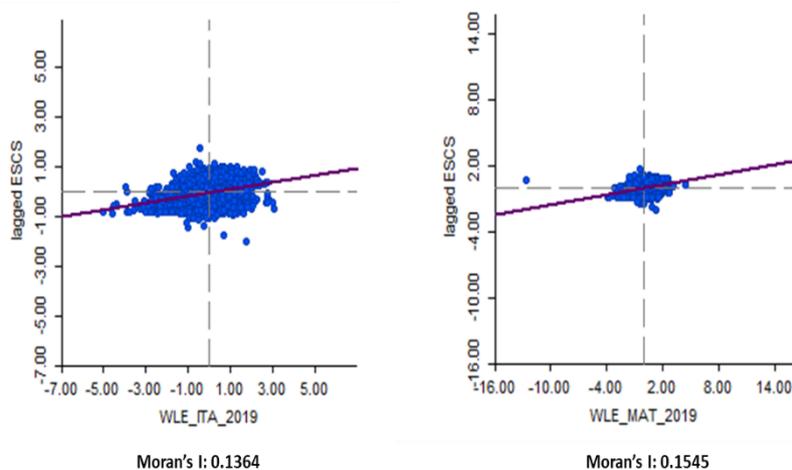
In Puglia, come già evidenziato in fase descrittiva, le province di Lecce e di Bari hanno i punteggi medi più alti. Tuttavia, si evidenzia la presenza di outlier territoriali, ossia la presenza di scuole appartenenti al cluster Low-High. Tali scuole sono presenti principalmente nelle zone periferiche dei comuni capoluogo.

Al fine di approfondire i fattori che influenzano i pattern territoriali rintracciati dall'analisi spaziale univariata, l'indagine è stata estesa all'analisi bivariata, per tener conto del contesto socio-economico circostante.

Come già accennato in precedenza, partendo dall'assunzione per cui i valori di una variabile osservata in una data unità geografica mostrano un'associazione sistematica con un'altra variabile osservata in unità geografiche vicine, l'indice di Moran bivariato è stato impiegato per rintracciare e analizzare la dipendenza spaziale esistente tra il punteggio WLE (in Italiano e Matematica) e l'indicatore di status socio-economico ESCS.

L'indice I globale bivariato ha rivelato la presenza di autocorrelazione spaziale tra la distribuzione di ESCS e i punteggi WLE di Italiano e Matematica, rispettivamente $I \approx 0,1364$ e $I \approx 0,1545$, in entrambi i casi l'indice è significativo (lo pseudo p-value è $< 0,001$) (Figura 8).

Figura 8. Moran Scatterplot bivariato per WLE_ITA e WLE_MAT (a.s. 2018-19)



I risultati dell'indicatore LISA bivariato sono stati riportati sui cartogrammi in Figura 9 e 10, in cui sono stati individuati 4 diversi cluster di istituti scolastici:

- il primo cluster (High-High, riportato in rosso nella mappa) è formato da scuole che conseguono punteggi di Italiano e Matematica superiori alla media del Sud e sono circondate da istituti il cui ESCS è superiore alla media; ne sono interessate Abruzzo, Puglia (per le province di Lecce e Bari), Molise (nella zona più interna), Campania (le province di Avellino, Benevento e Salerno), Basilicata, Calabria (nella parte meridionale della provincia di Cosenza, Catanzaro e Reggio Calabria) per entrambe le materie;
- il secondo cluster (Low-Low, in blu nella mappa) è formato da scuole che registrano punteggi WLE bassi e sono contigue ad istituti scolastici con ESCS inferiore alla media del Sud; ne fanno parte principalmente aree situate in Campania (principalmente nella provincia di Napoli), Sicilia (le province di Palermo, Catania e Caltanissetta) e Sardegna (le province di Nuoro e Cagliari per Italiano e Sassari per Matematica);
- le scuole che formano il cluster Low-High (in viola sulla mappa) sono posizionate prevalentemente in Basilicata, Campania, Abruzzo e Calabria (nella provincia di Reggio Calabria) e per Italiano anche la Puglia. Si tratta di scuole con punteggi WLE bassi che non sembrano beneficiare della collocazione territoriale in zone in cui le scuole hanno ESCS mediamente più alti;
- l'ultimo cluster (High-Low, in rosa sulla mappa) è formato dalle scuole con punteggi WLE promettenti e superiori alla media territoriale circondate da scuole con ESCS inferiore alla media. Le scuole appartenenti a questo cluster, che definiamo “*resilienti*”, si trovano prevalentemente in Campania e Sicilia e in Sardegna in entrambe le materie.

Nello specifico, in Campania le scuole “*resilienti*” si trovano nella provincia di Napoli, in Sicilia nelle province di Palermo, Trapani, Catania per entrambe le materie, alle quali si aggiungono le scuole delle province di Caltanissetta e Agrigento per Matematica. In Sardegna sono interessate le province di Cagliari e Nuoro per entrambe le materie.

Figura 9. LISA bivariato cluster map per WLE ITA e ESCS (a.s. 2018-19)

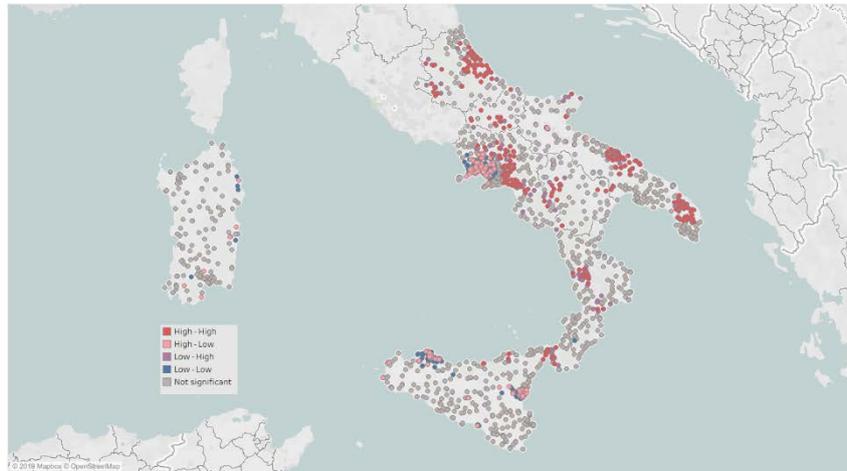
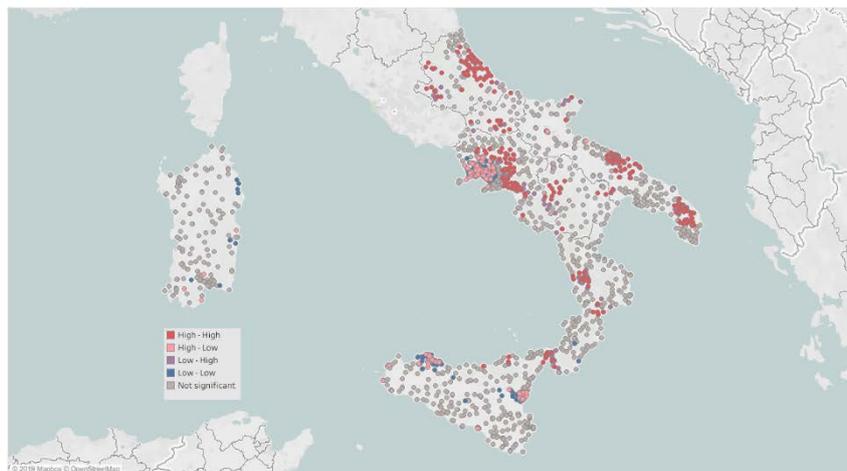


Figura 10. LISA bivariato cluster map per WLE MAT e ESCS (a.s. 2018-19)



5. Conclusioni

La questione delle differenze negli apprendimenti fra studenti del Sud e studenti del Centro-Nord risulta ancora oggi molto delicata poiché è noto che scarti nelle competenze scolastiche odierne potrebbero potenzialmente influenzare differenze future nei livelli di sviluppo socio-economico tra territori.

La georeferenziazione delle scuole sul territorio e l'analisi dei punteggi delle prove INVALSI di Italiano e Matematica di Grado 8 ha permesso di svelare una struttura eterogenea della geografia degli apprendimenti nel Sud d'Italia. Tuttavia le analisi svolte dimostrano che questa forma di eterogeneità spaziale va di pari passo con la presenza di autocorrelazione spaziale a livello sub-regionale e sub-provinciale e spesso consente di rintracciare delle ripartizioni che non tengono conto dei confini istituzionali e amministrativi, ma li superano a vantaggio di "continuum territoriali" nuovi e mutevoli.

Nel tentativo di individuare i fattori che spiegano le differenze territoriali nelle performance degli studenti abbiamo focalizzato l'attenzione sull'indicatore del contesto socio-economico e culturale (ESCS), che coniuga alcuni dei principali fattori coinvolti nel determinare le differenze nei livelli di apprendimento scolastico e di conseguenza nello sviluppo degli interessi culturali: la scolarità genitoriale, il reddito familiare, il contesto sociale. Nonostante i profondi cambiamenti socio-demografici che l'Italia ha vissuto negli ultimi decenni, il contesto familiare rimane una componente decisiva della struttura sociale ed economica del Paese. Ciononostante la nostra ricerca ha verificato l'esistenza di molteplici eccezioni alla relazione già nota tra background socio-economico-culturale e apprendimenti scolastici. Eccezioni rappresentate da tutte quelle realtà scolastiche "resilienti" da noi rintracciate nelle diverse regioni e province meridionali grazie all'analisi spaziale, realtà che combattono ogni giorno una battaglia fondamentale alla ricerca di riscatto da un contesto socio-economico svantaggiato, per dimostrare che anche al Sud conseguire buoni livelli di apprendimento scolastico è possibile.

Dunque riscattarsi da un passato o da un contesto sociale ingombranti, per sovvertire il presente ed il futuro è impresa possibile, faticosa ma possibile.

Rimangono necessari, a nostro avviso, investimenti in campo educativo, in particolare indirizzati a favorire gli studenti più svantaggiati se si vogliono ridurre quanto più possibile gli effetti delle caratteristiche sociali e culturali sulle disuguaglianze nei livelli di apprendimento. I margini di intervento possono essere ampi. Al tempo stesso, resta di cruciale importanza continuare ad indagare in questo ambito, ampliando la ricerca sia dal punto di vista geografico che socio-economico al fine di ottenere una visione più esaustiva delle dinamiche coinvolte nei processi di interazione fra fattori socio-culturali e apprendimento scolastico.

Riferimenti Bibliografici

- Agasisti T., Soncin M., Valenti R. (2015), School factors helping disadvantaged students to succeed: empirical evidence from four Italian cities, *INVALSI Working Paper n.23/2015*.
https://www.invalsi.it/download/wp/wp23_Agasisti_Soncin_Valenti.pdf
- Anselin L. (1995), *Local indicators of spatial association - LISA*, Geographical analysis, Vol. 27, No. 2.
- Berthold M.R. et al. (2008), “KNIME: The Konstanz Information Miner”, in: Preisach C., Burkhardt H., Schmidt-Thieme L., Decker R. (eds), “Data Analysis, Machine Learning and Applications. Studies in Classification”, *Data Analysis, and Knowledge Organization*, Springer, Berlin, Heidelberg.
- Braga M., Checchi D. (2010), Sistemi scolastici regionali e capacità di sviluppo delle competenze. I divari dalle indagini Pirls e Pisa, *La Rivista delle Politiche Sociali*, 3/2010.
- Brunello G., Checchi D. (2006), Does School Tracking Affect Equality of Opportunity? New International Evidence, *IZA Discussion Papers 2348*, Institute for the Study of Labor (IZA).
- Campodifiori E., Figura E., Papini M., Ricci R. (2010), Un indicatore di status socio-economico-culturale degli allievi della quinta primaria in Italia, *INVALSI Working Paper n.02/2010*.
https://www.invalsi.it/download/wp/wp02_Ricci.pdf
- Cecchi D., Peragine V. (2005), Regional Disparities and Inequality of Opportunity: the case of Italy, in *IZA Working Paper 1874*.
- Cecchi D. (2007), *I vincoli del sistema scolastico e la formazione delle competenze*, Il Mulino, 1/2007, testo disponibile sul sito: <http://cecchi.economia.unimi.it/pdf/un37.pdf>, data consultazione: 29 febbraio 2020.
- Cipollone P., Sestito P. (2007), *Quanto imparano gli studenti italiani: i divari Nord-Sud*, Banca d'Italia, mimeo.
- Demarinis G., Iaquina, M., Leogrande, D., Viola D. (2011), Analisi quantitativa della mobilità studentesca negli atenei italiani. Confronto territoriale fra domanda e offerta di formazione universitaria, *Valutazione e qualità degli atenei: Modelli, metodi e indicatori statistici*, 273-303.
- Ferrer-Esteban G. (2011), Beyond the traditional territorial divide in the Italian Education System. Effects of system management factors on performance in lower secondary school, in *FGA Working Paper 43*,
http://www.fondazioneagnelli.it/wp-content/uploads/2017/05/G._Ferrer-Esteban_Beyond_the_traditional_territorial_divide_in_the_Italian_Education_System_FGA_WP43.pdf, data di consultazione: 29 febbraio 2020.
- INVALSI (2017), *Rilevazioni nazionali degli apprendimenti 2016-17. Rapporto risultati*, testo disponibile al sito: https://invalsi-areaprove.cineca.it/docs/file/Rapporto_Prove_INVALSI_2017.pdf, data di consultazione: 29 febbraio 2020.
- INVALSI (2018), *Rapporto prove INVALSI 2018*, testo disponibile al sito: http://www.invalsi.it/invalsi/doc_evidenza/2018/Rapporto_prove_INVALSI_2018.pdf, data di consultazione: 29 febbraio 2020.
- INVALSI (2019), *Rapporto prove INVALSI 2019*, testo disponibile al sito: https://invalsi-areaprove.cineca.it/docs/2019/Rapporto_prove_INVALSI_2019.pdf, data di consultazione: 29 febbraio 2020.
- ISTAT (2018), *Rapporto annuale 2018 La situazione del Paese*, testo disponibile al sito: https://valored.it/wp-content/uploads/2018/06/2018_25_Istat_RapportoAnnuale.pdf, data di consultazione: 29 febbraio 2020.



- Longobardi S., Agasisti T. (2014), Educational institutions, resources, and students' resiliency: An empirical study about OECD countries, *Economics Bulletin*, Vol. 34, No. 2, pp. 1055-1067.
- Melchiori F.M. (2013), La costruzione delle opportunità: il processo di scelta dei percorsi di istruzione e formazione professionale, «*Formazione & Insegnamento*», XI, 2-2013.
- Ministero dell'Economia e delle Finanze, Dipartimento della Ragioneria Generale dello Stato, Servizio Studi (2012), *Analisi dell'efficienza delle scuole italiane rispetto agli apprendimenti degli studenti. Differenze territoriali e possibili determinanti*.
- Montanaro P. (2008), *I divari territoriali nella preparazione degli studenti italiani: evidenze dalle indagini nazionali e internazionali*, Bank of Italy Occasional Paper, 14, testo disponibile al sito: https://www.bancaditalia.it/pubblicazioni/qef/2008-0014/QEF_14_ita.pdf , data di consultazione: 29 febbraio 2020.
- Moran P. (1948), The Interpretation of Statistical Maps, *Journal of the Royal Statistical Society*, 10, 243-251.
- Pavolini E., Argentin G., Barbieri G., Falzetti P., Ricci R. (2015), L'influenza delle scuole e del contesto locale sui divari territoriali delle competenze degli studenti, in F. Asso, E. Pavolini (a cura di), *L'istruzione difficile. I divari nelle competenze tra Nord e Sud*, Donzelli, Roma.
- Raimondi E., De Luca S., Barone C. (2013), Origini sociali, risorse culturali familiari e apprendimenti nelle scuole primarie: un'analisi dei dati Pirls 2006, *Quaderni di Sociologia*, 61/2013, testo disponibile sul sito: <http://journals.openedition.org/qds/481>. data consultazione: 29 febbraio 2020.