



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

## **WORKING PAPER N. 35/2018**

---

**Un esercizio di calcolo del valore aggiunto delle scuole italiane sulla base dei risultati della prova  
INVALSI di Matematica 2017 del grado 8**

**Andrea Bendinelli - INVALSI**

**Angela Martini - INVALSI**

**Collana: Working Papers INVALSI**

**ISSN: 2611 - 5719**

-----  
*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

La valutazione della qualità dell'istruzione fornita dalle scuole è al centro delle riforme educative degli ultimi decenni. Per comune riconoscimento, perché il confronto tra le scuole possa essere equo, è necessario valutare il contributo che ognuna di esse dà all'apprendimento dei suoi studenti al netto dei fattori estranei al suo operato ma che pure incidono sui risultati scolastici: le caratteristiche socio-demografiche degli alunni e il loro livello di preparazione all'ingresso nella scuola. È cioè necessario valutare le scuole non tanto sulla base dei risultati raggiunti dai loro studenti in termini assoluti, bensì in termini di indicatori di valore aggiunto. Dal 2016 l'INVALSI ha iniziato a restituire alle scuole gli esiti delle prove standardizzate di Italiano e Matematica cui sottopone annualmente gli studenti italiani, oltre che come punteggi grezzi, anche come punteggi di valore aggiunto. È tuttavia un tema meritevole di riflessione quali siano le variabili da prendere in considerazione al fine di depurare il risultato delle prove dal peso dei fattori esogeni che, pur influenzando sui livelli di apprendimento, sfuggono al controllo delle scuole. In questo lavoro, utilizzando i dati campionari relativi alla prova di Matematica di terza secondaria di primo grado del 2017, si pongono a confronto due modelli di stima del valore aggiunto delle scuole italiane: nel primo non si tiene conto della regione in cui l'istituto è ubicato mentre nel secondo il valore aggiunto è calcolato al netto anche di questa variabile.

*Parole chiave:* valore aggiunto; rendicontazione; regressione multilivello; effetto di contesto.

*Keywords:* value added; accountability; multilevel regression; contextual effect.

## Introduzione

Da alcuni decenni, in particolare nei Paesi anglosassoni, hanno preso piede sistemi cosiddetti di “rendicontazione” (*accountability*) delle scuole i quali ne valutano l’operato sulla base dei risultati raggiunti dai loro studenti, di cui esse sono tenute direttamente responsabili. La prima a muoversi su questa via è stata l’Inghilterra, che con l’*Education Reform Act* del 1988 ha introdotto un sistema di *accountability* di tutte le scuole finanziate con fondi pubblici basato sulla verifica sistematica dei livelli di apprendimento degli studenti, misurati mediante prove standardizzate in alcune discipline chiave (lingua nazionale, matematica, scienze), al termine di ognuno degli stadi in cui il curriculum nazionale è scandito. All’Inghilterra hanno fatto seguito nel 2002 gli Stati Uniti che con il *No Child Left Behind Act* hanno imposto a tutti gli stati dell’Unione che già non l’avessero fatto di dotarsi di un sistema di *accountability*, a livello dei distretti scolastici e delle singole scuole, basato anch’esso sulla misurazione mediante test dei livelli di apprendimento degli alunni.

Senza entrare nei dettagli degli sviluppi che hanno avuto le riforme citate nei due Paesi cui si è fatto cenno, è importante qui sottolineare che fin dalla loro introduzione il mondo della ricerca ha preso posizione contro la pretesa di valutare l’efficacia pedagogico-didattica di una scuola rispetto a un’altra senza tener conto della diversità delle popolazioni di studenti da esse reclutate. Il principale problema che dev’essere affrontato quando ci si prefigge di valutare la qualità dell’istruzione fornita da una scuola è, infatti, quello di separare l’effetto che è ad essa propriamente attribuibile da tutte le altre influenze che agiscono sull’apprendimento degli alunni, in particolare quelle dovute all’ambiente sociale di provenienza e alle competenze possedute all’atto dell’ingresso in una determinata istituzione scolastica, sulle quali non ha possibilità di controllo. Da ciò l’esigenza di valutare l’effetto di una scuola non tanto in base ai risultati assoluti (o grezzi) ottenuti dai suoi studenti ma in base a indicatori di “valore aggiunto”.

## Che cosa è il valore aggiunto e come si determina

Riprendendo una definizione proposta dall’OCSE (OECD, 2008), possiamo definire il valore aggiunto come il contributo dato dalla scuola all’apprendimento dei propri studenti “al netto” dei fattori esterni alla sua azione che incidono su di esso. Più precisamente, il valore aggiunto di una scuola è la differenza nel livello di apprendimento conseguito in una certa disciplina dai suoi alunni e quello mediamente conseguito da studenti con caratteristiche comparabili che abbiano frequentato altre scuole.

Per il calcolo dell’effetto della scuola è possibile utilizzare differenti modelli statistici. A questo proposito l’OCSE distingue fra “modelli di apprendimento contestualizzato”, che stimano il contributo della scuola all’apprendimento dei suoi alunni misurandolo in un solo punto del tempo al netto delle caratteristiche degli studenti, e “modelli di valore aggiunto” vero e proprio, il cui principale tratto distintivo consiste nel rilevare l’apprendimento degli stessi studenti in almeno due momenti successivi del tempo, all’inizio e alla fine di un periodo più o meno lungo di frequenza della scuola. I modelli di valore aggiunto si suddividono poi a loro volta in due categorie: modelli basati su due misurazioni, una in ingresso e una in uscita, dei livelli di apprendimento degli studenti e sull’aggiustamento delle covariate per tenere sotto controllo i fattori esogeni che agiscono su di essi, e modelli multivariati che si basano su più misurazioni ripetute nel tempo del progresso cognitivo degli alunni (curve di crescita), senza aggiustamento delle covariate sulla base del presupposto che in tal caso (Sanders e altri, 1997) ogni alunno funge da controllo di sé stesso.

Un elemento di complicazione nel calcolo del valore aggiunto è dato dal fatto che dalla ricerca in campo educativo in generale e, da quella sull'effetto della scuola in particolare, è emerso che il livello di apprendimento di uno studente in un dato momento non è influenzato solo dalle sue caratteristiche personali (genere, condizione sociale, ecc.) e dal suo grado di capacità e preparazione all'inizio di un ciclo d'istruzione, ma anche dalle caratteristiche e dal livello medio delle competenze possedute dagli alunni della sua classe e della sua scuola. In altri termini, sussiste un "effetto di contesto" o di composizione del gruppo di cui un alunno fa parte che condiziona in una qualche misura il suo stesso apprendimento. La presenza di un effetto di contesto comporta, per la misurazione del valore aggiunto, che essa deve fondarsi sulla comparazione di alunni non solo simili dal punto di vista delle variabili individuali rilevanti ma che anche frequentano scuole simili sotto il profilo della composizione della popolazione scolastica reclutata.

### **La regressione multilivello per il calcolo del valore aggiunto**

Per far fronte al problema degli effetti dell'appartenenza a un certo gruppo su una variabile dipendente misurata a livello individuale sono nati i modelli di regressione multilivello, che a differenza della regressione ordinaria, basata sul presupposto dell'indipendenza delle osservazioni, si fanno carico della struttura gerarchica intrinseca ai dati di molti fenomeni sociali, tra cui quelli nel campo dell'educazione costituiscono un caso esemplare: gli alunni, infatti, non sono normalmente isolati gli uni dagli altri ma si trovano riuniti in classi e queste a loro volta nelle scuole, le scuole nelle province, e così via. È altamente probabile che gli alunni di una stessa classe o scuola, a causa sia di processi di autoselezione o di selezione all'ingresso sia della condivisione di un ambiente educativo comune, siano più simili tra loro rispetto agli alunni di altre classi o scuole. L'uso di un modello *multilevel* consente di tenere sotto controllo l'effetto di ciascuna delle variabili esplicative prese in considerazione nel modello al netto delle altre, ma anche l'effetto dovuto all'appartenenza allo stesso gruppo. Una volta stimato il peso esercitato dalle caratteristiche degli alunni, individuali e aggregate, sul livello di apprendimento in una certa materia, è possibile calcolare il valore aggiunto basandosi sui residui della regressione. Per maggiore chiarezza, diamo di seguito le equazioni che definiscono un modello *multilevel* a due livelli, studente e scuola.

Nella formulazione *multilevel*, la relazione tra una variabile-risposta  $Y$  (nel nostro caso, il risultato in una prova di profitto d'una certa materia), osservata sull'individuo  $i$ , appartenente alla scuola  $j$ , e una variabile esplicativa individuale  $X$  (ad esempio, la condizione sociale dello studente) è rappresentata nel modo seguente (Snijders e Bosker, 1999):

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \quad (1)$$

dove i pedici  $i$  e  $j$  indicano rispettivamente le unità di primo livello (gli studenti) e le unità di secondo livello (le scuole). La prestazione di uno studente  $Y_{ij}$  è funzione della prestazione media della scuola ( $\beta_{0j}$ ) cui lo studente appartiene, dell'effetto ( $\beta_{1j}$ ) della variabile  $X_{ij}$  all'interno della scuola, e di una componente di errore ( $e_{ij}$ ), con media nulla e varianza fissa pari a  $\sigma^2$ .

Il principio fondamentale del modello *multilevel* è che i coefficienti  $\beta_{0j}$  e  $\beta_{1j}$  non assumono necessariamente lo stesso valore in ogni gruppo (scuola) ma possono variare da gruppo a gruppo. La variabilità dei coefficienti può essere espressa, a sua volta, in funzione di un livello medio generale ( $\gamma_{00}$  e  $\gamma_{10}$ ), di una qualche

caratteristica di gruppo  $Z_j$  (ad esempio, lo status sociale medio degli alunni della scuola) e di un termine di errore casuale  $u_{0j}$  e  $u_{1j}$ :

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} \quad (2)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j} \quad (3)$$

Gli errori  $u_{0j}$  e  $u_{1j}$  hanno media nulla e varianza pari a  $\text{var}(u_{0j}) = \sigma_{u_0}^2$  e  $\text{var}(u_{1j}) = \sigma_{u_1}^2$ .

Sostituendo nella 1 i termini  $\beta_{0j}$  e  $\beta_{1j}$  si ottiene la formulazione definitiva del modello *multilevel* completamente specificato:

$$Y_{ij} = \gamma_{00} + \gamma_{01}Z_j + \gamma_{10}X_{ij} + \gamma_{11}Z_j X_{ij} + u_{0j} + u_{1j} X_{ij} + e_{ij} \quad (4)$$

## I dati analizzati

I dati su cui il nostro esercizio di calcolo del valore aggiunto è stato condotto sono i punteggi, espressi su scala Rasch centrata alla media nazionale posta uguale a 200 e deviazione standard pari a 40 e corretti per il *cheating*<sup>1</sup> (Quintano et al., 2009), che gli alunni del campione estratto sugli studenti di terza secondaria inferiore hanno ottenuto nella prova INVALSI di Matematica del 2017. Sebbene le prove INVALSI siano censuarie, vengano cioè sostenute da tutti gli studenti delle classi oggetto di rilevazione, tra di essi viene comunque estratto un campione<sup>2</sup> con metodo a due stadi: nel primo stadio sono selezionate le scuole e nel secondo, di norma, una o due classi intere per ognuna delle scuole selezionate al primo stadio. Ai fini di questo lavoro l'analisi dei dati è stata estesa alla totalità degli studenti delle scuole campionate. La tabella che segue riporta il numero di studenti e di scuole, dell'Italia e delle cinque macro-aree<sup>3</sup> in cui il Paese è suddiviso, sul quale le analisi di cui si illustrano qui i risultati sono state eseguite.

**Tabella 1: Numero di scuole e studenti del campione analizzato**

| Area        | Numero scuole | Numero studenti |
|-------------|---------------|-----------------|
| Nord Ovest  | 287           | 31.761          |
| Nord Est    | 274           | 30.237          |
| Centro      | 280           | 33.612          |
| Sud         | 297           | 38.731          |
| Sud e Isole | 265           | 27.900          |
| ITALIA      | 1.403         | 162.241         |

<sup>1</sup> I risultati delle prove INVALSI sono normalmente sottoposti a una procedura statistica per la correzione delle distorsioni dei punteggi introdotte da eventuali comportamenti opportunistici (copiare, suggerire le risposte, ecc.) da parte di studenti o insegnanti.

<sup>2</sup> Lo scopo del campione è di garantire l'attendibilità dei dati raccolti in quanto nelle classi campione è presente un osservatore esterno con il compito di assicurare il regolare svolgimento delle prove e di trasmettere le risposte degli studenti all'Istituto di valutazione. Nella terza secondaria di primo grado tale compito è affidato al presidente della commissione di esami di licenza media, tra le cui prove, fino all'anno 2016-17 erano incluse anche le prove INVALSI.

<sup>3</sup> Le macro-aree sono: Nord Ovest (Valle d'Aosta, Piemonte, Lombardia Liguria), Nord Est (Provincia aut. di Bolzano, Provincia aut. di Trento, Veneto, Friuli-Venezia Giulia, Emilia-Romagna), Centro (Toscana, Umbria, Marche, Lazio), Sud (Abruzzo, Molise, Campania, Puglia), Sud e Isole (Basilicata, Calabria, Sicilia, Sardegna).

Si è scelto di focalizzare l'attenzione sui dati relativi alla prova di Matematica perché l'apprendimento di questa disciplina è meno influenzato dalle condizioni socio-economiche del contesto (soprattutto familiare) rispetto all'Italiano (Checchi, 2010) e dunque si presta meglio a valutare l'efficacia dell'insegnamento impartito. Ai fini del calcolo del valore aggiunto si sono stimati due modelli di regressione a due livelli, studente e scuola<sup>1</sup>. Le variabili esplicative prese in considerazione al primo livello (studente) sono, in entrambi i modelli, quelle riferite alle caratteristiche socio-demografiche degli alunni e al loro livello di competenza all'ingresso nella scuola secondaria di primo grado, mentre al secondo livello sono state considerate, in una prima formulazione, le variabili, relative alla composizione del corpo studentesco della scuola, alle quali, in una seconda formulazione, è stata aggiunta la regione di localizzazione dell'istituto. I valori mancanti di due delle variabili, l'indice di status socio-economico e culturale (ESCS) e il punteggio di matematica in V Primaria, che avevano un'elevata percentuale di *missing* (16,8% e 19,7% rispettivamente) sono stati imputati con la media della classe o, in assenza, con quella della scuola o, in assenza anche di questa, con la media generale. Inoltre, per tenere sotto controllo l'effetto dovuto all'eventuale mancanza dell'informazione relativa a queste due variabili, sono state create e inserite nei modelli, al primo livello, due variabili dicotomiche (*dummies*) che assumono valore 0 se il dato è presente, valore 1 se è mancante<sup>2</sup>. La tabella che segue mostra tutte le variabili inserite nei modelli di regressione. Per le variabili di tipo categoriale (ad es. il genere), sono indicate tra parentesi le categorie considerate, di cui in corsivo quella assunta come riferimento.

**Tabella 2: Variabili inserite nei modelli di regressione multilivello**

| <i>Variabili di livello 1</i>                            | <i>Variabili di livello 2</i>                              |
|--|--|
| Genere ( <i>maschio/femmina</i> )                        | % stranieri (/10)  |
| Regolarità <sup>3</sup> ( <i>regolare / in ritardo</i> ) | % alunni in ritardo  |
| Cittadinanza <sup>4</sup> ( <i>italiano/straniero</i> )  | % studenti assenti alla prova (/10)                        |
| Status socio-economico e culturale (ESCS) <sup>5</sup>   | Indice di status ESCS medio                                |
| Punteggio di matematica in V Primaria                    | Punteggio medio di matematica in V Primaria                |
| Indice di status mancante ( <i>no/sì</i> )               | Localizzazione della scuola ( <i>Lazio/altre regioni</i> ) |
| Punteggio in V Primaria mancante ( <i>no/sì</i> )        |  |

<sup>1</sup> I modelli stimati sono a intercetta casuale e pendenza fissa. Le variabili quantitative di livello 1 sono centrate sulla media di gruppo (scuola), mentre le variabili categoriali non sono centrate. Le variabili quantitative di livello 2 (Escs medio e punteggio medio di Matematica all'ingresso) sono centrate sulla media generale.

<sup>2</sup> Ciò perché - come anche nel nostro caso - spesso i valori mancanti non sono randomizzati.

<sup>3</sup> In questo lavoro sono considerati come alunni in regola sia gli alunni che frequentano una classe corrispondente alla loro età anagrafica secondo la normativa vigente, sia gli alunni anticipatori che non hanno mai ripetuto.

<sup>4</sup> Sono considerati di cittadinanza italiana gli studenti nati in Italia da genitori entrambi italiani, mentre sono considerati stranieri, senza distinzione qui tra prima e seconda generazione, gli studenti nati in Italia da genitori stranieri, o gli studenti immigrati dall'estero.

<sup>5</sup> Sulla base delle risposte al questionario-studente, l'Invalsi calcola per ogni alunno un indice di status socio-economico e culturale (ESCS: *Economic, Social and Cultural Status*), che integra tre variabili: il livello d'istruzione dei genitori, la loro professione, e le risorse educative e culturali di cui l'alunno può fruire a casa. L'indice è standardizzato con media uguale a 0 e deviazione standard uguale a 1 (Campodifiori et al., 2010). Per gli alunni di terza secondaria di primo grado, cui il questionario non è fatto compilare, facendo le prove Invalsi parte dell'esame di licenza fino al 2017, l'indice è calcolato in base alle risposte date al questionario dagli stessi alunni in quinta primaria

## Gli effetti netti delle variabili considerate sul risultato della prova INVALSI di matematica 2017

La tabella che segue riporta i coefficienti della regressione a due livelli sul punteggio della prova INVALSI di Matematica somministrata nel 2017 al termine della scuola secondaria di primo grado.

**Tabella 3: Effetti netti sul punteggio della prova INVALSI 2017 di Matematica di III SPG**

|                                       | Modello 0 | Modello 1 | Modello 2 | Modello 3 |
|---------------------------------------|-----------|-----------|-----------|-----------|
| Intercetta                            | 197,4     | 203,6     | 201,1     | 203,1     |
| <b>Componenti fisse</b>               |           |           |           |           |
| <b>Livello 1 (studente)</b>           |           |           |           |           |
| Femmina                               |           | -4,2**    | -4,1**    | -4,2**    |
| In ritardo                            |           | -13,4**   | -13,5**   | -13,4**   |
| Straniero                             |           | -4,3**    | -4,6**    | -4,5**    |
| ESCS individuale                      |           | 5,1**     | 5,1**     | 5,1**     |
| Punt. Matematica in V Primaria        |           | 18,6**    | 18,6**    | 18,6**    |
| ESCS mancante                         |           | -9,9**    | -9,6**    | -10,1**   |
| Punteggio in V Primaria mancante      |           | -4,6**    | -4,5**    | -4,4**    |
| <b>Livello 2 (scuola)</b>             |           |           |           |           |
| % stranieri (/10)                     |           |           | 2,8**     | -1,0*     |
| % alunni in ritardo                   |           |           | 0,2*      | 0,1       |
| % assenti alla prova (/10)            |           |           | -1,8*     | -2,3**    |
| ESCS medio di scuola                  |           |           | 11,5**    | 8,4**     |
| Punteggio medio in V Primaria         |           |           | 11,6**    | 4,6**     |
| Valle D'Aosta                         |           |           |           | 21,5**    |
| Piemonte                              |           |           |           | 4,9**     |
| Liguria                               |           |           |           | 6,3**     |
| Lombardia                             |           |           |           | 8,5**     |
| Veneto                                |           |           |           | 10,9**    |
| Friuli-Venezia Giulia                 |           |           |           | 10,5**    |
| Emilia-Romagna                        |           |           |           | 8,8**     |
| Toscana                               |           |           |           | 7,5**     |
| Umbria                                |           |           |           | 4,5**     |
| Marche                                |           |           |           | 7,1**     |
| Abruzzo                               |           |           |           | -1,0      |
| Molise                                |           |           |           | -0,8      |
| Campania                              |           |           |           | -9,1**    |
| Puglia                                |           |           |           | -0,9      |
| Basilicata                            |           |           |           | -4,0*     |
| Calabria                              |           |           |           | -13,2**   |
| Sicilia                               |           |           |           | -7,5**    |
| Sardegna                              |           |           |           | -5,0**    |
| Prov. Aut. di Bolzano (lingua it.)    |           |           |           | 7,9**     |
| Prov. Aut. di Trento                  |           |           |           | 12,3**    |
| <b>Componenti casuali</b>             |           |           |           |           |
| Varianza livello 1 ( <i>within</i> )  | 1463,3    | 1089,5    | 1089,7    | 1089,8    |
| Varianza livello 2 ( <i>between</i> ) | 171,2     | 184,3     | 114,1     | 79,9      |
| % varianza <i>between</i> (I.C.C.)    | 10,5      |           |           |           |
| Riduzione varianza <i>within</i> (%)  |           | 25,5      | 25,5      | 25,5      |
| Riduzione varianza <i>between</i> (%) |           | -7,7      | 33,4      | 53,3      |

\**p*-value < 0,05; \*\**p*-value < 0,01

Il modello 0, detto anche modello “nullo” o “vuoto”, non contiene variabili esplicative né di primo né di secondo livello ed ha il solo scopo di ripartire la varianza totale dei punteggi di Matematica in terza secondaria di primo grado in due componenti: quella dovuta a differenze tra gli alunni all’interno della scuola (varianza *within*) e quella dovuta a differenze tra le scuole (varianza *between*). Come si può vedere, la varianza dovuta a differenze tra le scuole, ovvero il coefficiente di correlazione intraclasse (I.C.C.) corrisponde a poco più del 10% della varianza totale, una quota in sé non particolarmente elevata<sup>1</sup>.

Nel modello 1 sono introdotte le variabili di primo livello riferite alle caratteristiche personali degli alunni. L’intercetta rappresenta il punteggio medio di un alunno “tipo” con le seguenti caratteristiche: maschio, italiano, in regola con il percorso degli studi, con un indice ESCS eguale alla media della sua scuola e un punteggio in V Primaria anch’esso eguale alla media degli alunni della sua scuola.

A parità delle altre condizioni, l’essere di genere femminile comporta una diminuzione del punteggio, rispetto all’intercetta, di 4 punti. Nello stesso modo si leggono le variazioni del punteggio connesse all’essere in ritardo e all’essere di origine immigrata. Per quanto riguarda l’effetto delle variabili quantitative, l’aumento di una unità di deviazione standard dell’indice di status dello studente rispetto alla media della sua scuola comporta l’aumento di 5 punti del punteggio di Matematica in terza secondaria inferiore, mentre l’aumento di una deviazione standard del punteggio in V Primaria rispetto alla media della scuola comporta una sostanziosa crescita di più di 18 punti della variabile dipendente. L’assenza del dato sulle due ultime variabili esplicative ha invece su di essa un effetto negativo.

Nel modello 2 sono inserite le variabili di secondo livello relative alla composizione del corpo studentesco della scuola<sup>2</sup>. L’intercetta corrisponde qui al punteggio di un alunno medio che, oltre ad avere le caratteristiche sopra descritte, frequenta una scuola dove non vi sono alunni stranieri, né in ritardo, né assenti alla prova.

Inoltre, tale scuola è frequentata da alunni con un indice medio di status e un punteggio medio di Matematica in V primaria pari alla media generale. Come già al primo livello, le due variabili che hanno il più forte effetto sul risultato in Matematica al termine della scuola secondaria inferiore sono l’ESCS medio e il punteggio medio degli alunni in V Primaria degli alunni dell’istituto: l’aumento di una deviazione standard su queste due variabili rispetto alla media di tutte le scuole comporta in entrambi i casi un aumento di più di 11 punti. Essendo il modello additivo, tale aumento va ad aggiungersi alla crescita del punteggio dello studente dovuta a una variazione unitaria delle due variabili misurate a livello individuale.

In ogni caso, al di là della loro entità, tutti i coefficienti sono significativi alla soglia di probabilità d’errore dell’1% (tolti due casi per cui la probabilità d’errore è al 5%). Da notare che, per quanto piccolo, risulta significativo anche il coefficiente corrispondente all’aumento di 10 punti percentuali degli alunni stranieri nella scuola, effetto che probabilmente è legato alla presenza nelle scuole del Nord-Italia, i cui studenti hanno in genere punteggi più alti, di una quota di stranieri più elevata che non al Sud.

Le variabili esplicative considerate spiegano il 25,5% per cento della varianza all’interno delle scuole e il 33,4% della varianza tra le scuole.

---

<sup>1</sup> La variabilità tra scuole è diversa tra le macro-aree italiane ed è assai più alta nelle regioni meridionali e insulari rispetto a quelle centrali e settentrionali. Il dato in tabella 3 corrisponde al valore medio a livello nazionale.

<sup>2</sup> Nel modello 2 le variabili di secondo livello sono tutte ottenute per aggregazione dal primo livello. La localizzazione della scuola, inserita tra le variabili di secondo livello del modello 3, è invece una variabile cosiddetta “globale”, non essendo derivata per aggregazione dal livello precedente.



Nel modello 3 è stata introdotta al secondo livello, in aggiunta alle variabili del modello 2, anche la regione di localizzazione della scuola.

I coefficienti corrispondenti alla regione dove la scuola è situata indicano le variazioni del punteggio di Matematica al termine del primo ciclo d'istruzione, a parità di tutte le altre condizioni, rispetto al punteggio di uno studente che frequenta una scuola ubicata nella regione Lazio, assunta come base di riferimento. Come si può vedere, i coefficienti delle regioni settentrionali e centrali sono positivi mentre i coefficienti delle regioni meridionali e insulari sono negativi e tutti, tranne nel caso di tre regioni, sono statisticamente significativi. Se confrontiamo il modello 3 con il modello 2, possiamo inoltre constatare che l'introduzione della localizzazione della scuola, com'era da attendersi, assorbe in parte, riducendoli, gli effetti dovuti all'indice medio di status e al punteggio medio di Matematica in V Primaria degli alunni della scuola, più bassi nelle regioni meridionali e insulari. Da rilevare, inoltre, che ora il coefficiente connesso all'aumento di 10 punti percentuali della quota di stranieri nella popolazione della scuola è divenuto negativo. Infine, la percentuale di varianza *between* spiegata aumenta, passando dal modello 2 al modello 3, dal 33,4% al 53,3%.

### **Gli indicatori di valore aggiunto delle scuole in base ai modelli 2 e 3**

Per determinare il valore aggiunto ci si basa sui residui di secondo livello della regressione, cioè su quella parte della varianza tra scuole della variabile dipendente che non è spiegata dalle variabili esplicative prese in considerazione (nel nostro caso, le caratteristiche, individuali e aggregate, degli studenti e, nel modello 3, la localizzazione dell'istituto), nell'ipotesi che tale variabilità sia da attribuire ai processi, sul piano organizzativo e pedagogico-didattico, messi in atto dalle scuole. Per il calcolo dei residui di livello scuola sono percorribili due strade (Grillo e Rampichini, 2009): la prima consiste nel calcolare la differenza tra il risultato osservato e il risultato predetto, la seconda – che è anche quella seguita in questo lavoro – consiste nel calcolare i cosiddetti *empirical Bayes* (EB) o *shrunk residuals*, aggregando per ogni scuola la componente residuale dell'effetto casuale, ovvero la differenza tra il valore predetto e la componente predetta dovuta alla parte fissa del modello<sup>1</sup>.

Dopo aver effettuato la stima dei residui – o indicatori di valore aggiunto – si è innanzitutto calcolata la correlazione tra i residui generati, rispettivamente, dal modello 2 e dal modello 3, correlazione che risulta molto forte, essendo pari a 0,88<sup>2</sup>.

I residui sono stati quindi ordinati in cinque categorie<sup>3</sup>. Osservando la distribuzione percentuale delle scuole italiane nelle cinque categorie a livello nazionale (tabella 4), si può vedere che non emergono differenze di rilievo sia che si tenga sia che non si tenga conto della localizzazione della scuola. La categoria più popolata è quella intermedia, che comprende circa il 74% delle scuole: ciò significa che la stragrande maggioranza di

---

<sup>1</sup> Per “parte fissa” del modello si intendono gli effetti delle variabili degli studenti e delle scuole inserite nei modelli di regressione a due livelli.

<sup>2</sup> Si fa riferimento al coefficiente di correlazione lineare di Bravais-Pearson

<sup>3</sup> Per la definizione delle cinque categorie di valore aggiunto si utilizza la deviazione standard dei residui, sulla cui base sono stabilite le soglie che le delimitano. I residui delle scuole con valore aggiunto “positivo” sono compresi tra una e due deviazioni standard dalla media (eguale a zero), mentre i residui delle scuole con valore aggiunto “molto positivo” raggiungono o superano le due deviazioni standard dalla media. In maniera simmetrica sono definite le categorie di scuole con valore aggiunto “negativo” e “molto negativo”. Sono considerate, infine, scuole con valore aggiunto non significativamente diverso dalla media quelle il cui residuo è compreso tra -1 e +1 deviazione standard.

esse consegue i risultati che era prevedibile ottenesse in base alle caratteristiche, individuali e aggregate, dei suoi alunni. Le scuole con valore aggiunto “molto positivo” o “positivo” assommano all’incirca al 12%, percentuale analoga a quella riscontrabile riunendo le due rimanenti categorie all’estremità opposta della graduatoria.

**Tabella 4: Percentuali di scuole con valore aggiunto molto positivo, positivo, nella media, negativo e molto negativo - Italia**

| Modello 2 |      |      |     |     | Modello 3 |      |      |     |     |
|-----------|------|------|-----|-----|-----------|------|------|-----|-----|
| ++        | +    | =    | -   | --  | ++        | +    | =    | -   | --  |
| 2,0       | 10,3 | 74,4 | 9,6 | 3,7 | 2,2       | 11,3 | 73,8 | 9,4 | 3,3 |

Il quadro cambia se andiamo a vedere la tabella 5: la distribuzione percentuale delle scuole delle cinque categorie di valore aggiunto, entro ognuna delle macro-aree in cui l’Italia è suddivisa, varia a seconda che si tenga o no conto della localizzazione della scuola. In particolare, passando dal modello 2 al modello 3, le scuole con valore aggiunto “molto positivo” o “positivo” aumentano nel Centro e ancor più nel Sud e nel Sud e Isole, mentre diminuiscono parallelamente le scuole con valore aggiunto “molto negativo” o “negativo”; nello stesso tempo, nelle due macro-aree del Nord diminuiscono le scuole delle prime due categorie e aumentano le scuole delle ultime due.

**Tabella 5: Percentuali di scuole con valore aggiunto molto positivo, positivo, nella media, negativo e molto negativo per macro-area**

| Area        | Modello 2 |      |      |      |      |      | Modello 3 |      |      |      |     |      |
|-------------|-----------|------|------|------|------|------|-----------|------|------|------|-----|------|
|             | ++        | +    | =    | -    | --   | Tot. | ++        | +    | =    | -    | --  | Tot. |
| Nord Ovest  | 3,1       | 12,2 | 79,8 | 3,8  | 1,0  | 100  | 0,3       | 8,4  | 80,8 | 9,1  | 1,4 | 100  |
| Nord Est    | 1,1       | 16,8 | 79,6 | 2,6  | 0,0  | 100  | 0,0       | 7,3  | 85,8 | 6,6  | 0,4 | 100  |
| Centro      | 1,4       | 7,5  | 83,2 | 5,7  | 2,1  | 100  | 1,4       | 10,4 | 78,6 | 8,2  | 1,4 | 100  |
| Sud         | 2,4       | 7,4  | 70,7 | 14,1 | 5,4  | 100  | 4,4       | 14,5 | 64,6 | 10,1 | 6,4 | 100  |
| Sud e Isole | 1,9       | 7,9  | 58,1 | 21,9 | 10,2 | 100  | 4,9       | 15,8 | 59,2 | 13,2 | 6,8 | 100  |

La differenza fra le distribuzioni degli indicatori di valore aggiunto delle scuole prodotti dai due modelli diviene, come ovvio, molto più evidente se si esaminano le percentuali delle scuole di ogni macro-area che si collocano all’interno di ciascuna delle cinque categorie, come illustrato nella tabella che segue.

**Tabella 6: Percentuali di scuole di ogni macro-area all’interno di ciascuna delle categorie di valore aggiunto**

| Area        | Modello 2 |      |      |      |      |      | Modello 3 |      |      |      |      |      |
|-------------|-----------|------|------|------|------|------|-----------|------|------|------|------|------|
|             | ++        | +    | =    | -    | --   | Tot. | ++        | +    | =    | -    | --   | Tot. |
| Nord Ovest  | 32,1      | 24,1 | 21,9 | 8,2  | 5,8  | 3,2  | 15,2      | 22,4 | 19,7 | 8,7  | 32,1 | 24,1 |
| Nord Est    | 10,7      | 31,7 | 20,9 | 5,2  | 0,0  | 0,0  | 12,7      | 22,7 | 13,6 | 2,2  | 10,7 | 31,7 |
| Centro      | 14,3      | 14,5 | 22,3 | 11,9 | 11,5 | 12,9 | 18,4      | 21,2 | 17,4 | 8,7  | 14,3 | 14,5 |
| Sud         | 25,0      | 15,2 | 20,1 | 31,3 | 30,8 | 41,9 | 27,2      | 18,5 | 22,7 | 41,3 | 25,0 | 15,2 |
| Sud e Isole | 17,9      | 14,5 | 14,8 | 43,3 | 51,9 | 41,9 | 26,6      | 15,2 | 26,5 | 39,1 | 17,9 | 14,5 |
| Totali      | 100       | 100  | 100  | 100  | 100  | 100  | 100       | 100  | 100  | 100  | 100  | 100  |

Passando dal modello 2 al modello 3, la percentuale di scuole con valore aggiunto “molto positivo” e “positivo” cresce nel Sud dal 25% al 42% nel primo caso e dal 15% al 27% nel secondo, e nel Sud e Isole dal 18% al 42% e dal 14% al 26% rispettivamente, mentre diminuiscono parallelamente le percentuali di scuole con valore aggiunto “molto negativo” e “negativo”; al contempo, le corrispondenti percentuali nel Centro, nel Nord Ovest e nel Nord Est aumentano e diminuiscono nella direzione opposta.

## Conclusioni

Il primo punto da sottolineare è che la posizione delle scuole muta considerevolmente se, in una ipotetica graduatoria, le si ordina in base al punteggio grezzo o in base al punteggio di valore aggiunto, e ciò sia che lo si calcoli tenendo conto oppure non tenendo conto della localizzazione dell’istituto. A questo proposito si è calcolato il coefficiente di correlazione a ranghi  $\tau$ -b (Tau\_b di Kendall), che consente di misurare il grado di coincidenza della posizione occupata dai valori di due variabili ordinate: esso oscilla da +1, nel caso di una totale convergenza dell’ordine di rango dei valori delle due variabili, a -1 nel caso di una completa divergenza. Nella tabella 7 sono riportati i coefficienti  $\tau$ -b fra i punteggi medi grezzi delle scuole e gli indicatori di valore aggiunto stimati col modello 2 e col modello 3.

**Tabella 7: Correlazione tra i punteggio grezzi delle scuole e gli indicatori di valore aggiunto**

|   | $\tau$ -b di Kendall |
|---|----------------------|
| Modello 2 (senza la regione di localizzazione della scuola) | 0,38                 |
| Modello 3 (con la regione di localizzazione della scuola)   | 0,32                 |

Per quanto riguarda l’opportunità di mettere o no sotto controllo, nel calcolo del valore aggiunto delle scuole, la loro collocazione geografica, va innanzitutto rilevato che le scuole del campione che cambiano posizione a seconda che gli indicatori di valore aggiunto siano stimati con il modello 2 o con il modello 3 corrispondono al 19,2% del totale, mentre la posizione del restante 80,8% rimane la stessa. Delle scuole per le quali si osserva un cambiamento della categoria in cui si collocano, il 10,9% migliora la propria posizione e l’8,3% la peggiora.

Ciò detto, attualmente l’INVALSI restituisce ogni anno alle scuole i loro risultati in termini di punteggi grezzi nelle prove di Italiano e Matematica e, dall’anno 2015-16, anche in termini di indicatori di valore aggiunto, calcolati però senza tener conto della localizzazione della scuola<sup>1</sup>. Si può senz’altro giustificare questa scelta sulla base dell’argomentazione che il sistema scolastico è strutturato nello stesso modo in tutto il territorio nazionale e le procedure amministrative per quanto concerne il reclutamento e l’assegnazione del personale dirigente e docente alle unità scolastiche sono le medesime, così come sono gli stessi i curricoli previsti per le materie d’insegnamento nel primo ciclo d’istruzione.

<sup>1</sup> Bisogna però precisare che gli indicatori di valore aggiunto restituiti alle scuole sono stimati tre volte, a livello nazionale, di macro-area e di regione.



Se la comparazione tra le scuole sulla base non dei risultati grezzi ma del loro valore aggiunto sconta le differenze sociali e di preparazione iniziale esistenti tra gli alunni dei vari istituti, livellando il terreno del confronto, rimane tuttavia la domanda se tra i fattori esogeni che ne condizionano l'operato vada considerato il contesto territoriale in cui esse sono inserite e dunque se il calcolo del loro valore aggiunto debba essere effettuato al netto dell'effetto, oltre che delle variabili, individuali e aggregate, degli alunni che le frequentano, anche dell'effetto legato alla localizzazione. A sostegno di ciò si può portare l'argomento che tra i punteggi dei test standardizzati sui livelli di apprendimento e lo sviluppo economico di un territorio esiste una relazione positiva, come mostrano anche i dati delle indagini internazionali. È pur vero che tenendo conto delle variabili di *background* degli studenti si tiene indirettamente conto del grado di sviluppo del luogo dove è situata la scuola che essi frequentano; tuttavia questo è corretto solo in parte, poiché l'indice di status socio-economico e culturale calcolato dall'INVALSI misura la qualità dell'ambiente familiare dell'alunno e riflette dunque solo in parte la forza dell'economia di un territorio. A questo proposito può essere istruttivo confrontare la correlazione tra l'ESCS medio delle regioni italiane e il punteggio medio nella prova INVALSI di Matematica 2017 di III Secondaria di primo grado<sup>1</sup> e la correlazione tra questo e il PIL pro-capite<sup>2</sup>: mentre la prima è pari a 0,68, la seconda è pari a 0,80, una correlazione elevata che giustifica l'interrogativo di cui sopra e lascia aperta la discussione su quale sia la strada migliore da seguire per una valutazione equa e responsabile dei risultati di apprendimento raggiunti dalle scuole del nostro Paese.

---

<sup>1</sup> I punteggi sono quelli pubblicati nel Rapporto "Rilevazioni nazionali degli apprendimenti 2016-17" edito da INVALSI e reperibile on-line.

<sup>2</sup> I dati su cui il calcolo della correlazione è basato sono desunti da: Istat\_Conti\_regionali\_2015\_tabella\_121216.



## Bibliografia

Campodifiori E., Figura E., Papini M., Ricci R. (2010). Un indicatore di status socio-economico e culturale degli allievi della quinta primaria in Italia. *Working Papers INVALSI N. 02* (disponibile al seguente link: [http://www.provincia.bz.it/servizio-valutazione-italiano/download/escs\\_invalsi.pdf](http://www.provincia.bz.it/servizio-valutazione-italiano/download/escs_invalsi.pdf) )

Checchi D. (2010), Orientamento verso la scuola superiore: cosa conta davvero?, *RicercaAzione*, vol.2 (2), 215-236.

Grilli L., Rampichini C. (2009). Multilevel models for the evaluation of educational institutions: a review, in: M. Bini et al. (eds), *Statistical Methods for the Evaluation of Educational Services and Quality of Products*, cap. 5, Berlin Heidelberg: Springer-Verlag,

Quintano C., Castellano R. e Longobardi S. (2009), A Fuzzy Clustering Approach to Improve the Accuracy of Italian Student Data. An Experimental Procedure to Correct the Impact of Outliers on Assessment Test Scores, *Statistica & Applicazioni* 7 (2), 149-171

OECD (2008). *Mesuring improvements in learning outcomes. Best practices to assess the value-added of schools*, Paris: OECD Publishing

Sanders W. L., Saxton A. M. e Horn S. P. (1997). The Tennessee value-added assessment system: A quantitative outcomes-based approach to educational assessment, in: J. Milliam (ed.), *Grading Teachers, Grading schools: Is Student Achievement a Valid Evaluation Measure?*, Corwin Press, Thousand Oaks (CA), 137-162.

Snijders, T., Bosker, R. (1999). *Multilevel Analysis*, London: Sage