

Il 'valore aggiunto' della scuola: problemi di misurazione e risultati

Angela Martini
Esperta INVALSI

Dagli ultimi decenni del secolo scorso l'interesse per i risultati delle scuole e per il contributo che esse danno al progresso cognitivo dei loro studenti è cresciuto nei Paesi sviluppati in concomitanza con l'introduzione di prove standardizzate per la valutazione degli apprendimenti e di sistemi di accountability delle istituzioni educative. La valutazione della qualità dell'istruzione fornita da una scuola richiede di separare l'effetto proprio di questa dalle variabili esogene che influiscono sull'apprendimento degli alunni. Dopo una breve discussione dei principali problemi e limiti nella misurazione del valore aggiunto, l'articolo illustra i risultati delle analisi effettuate dall'Invalsi sulla base di dati della rilevazione 2019 dei livelli di apprendimento in Italiano e Matematica degli studenti della scuola primaria e della scuola secondaria di primo e secondo grado al fine di stimare il valore aggiunto delle scuole italiane.

Since the last decades of the twentieth century interest in schools' performance and in the contribution they make to the cognitive progress of their pupils has grown in developed countries alongside with the introduction of standardized tests for the assessment of learning and school accountability systems. Assessing the quality of education provided by a school requires separating its effect from the exogenous variables that affect pupils' learning. After a brief discussion of the main problems and limitations in measuring schools' value-added, the article illustrates the analysis carried out by Invalsi in order to estimate the value-added of Italian schools on the database from the 2019 survey of learning achievements in Italian and Mathematics acquired by primary and secondary school students.

DOI: 10.1485/2532-8549-202003-2

Citazione

Martini A. (2020), Il 'valore aggiunto' della scuola: problemi di misurazione e risultati, *Sinapsi*, X, n.3, pp.17-31

Parole chiave

Valutazione dei risultati dell'apprendimento
Qualità dell'istruzione
Efficacia della scuola

Key words

Assessment of learnings outcomes
Quality of education
School effectiveness

1. Che cos'è il valore aggiunto

Il valore aggiunto è al centro della ricerca sull'efficacia educativa (*school effectiveness*) o, con altra espressione, sull'effetto della scuola (*school-effect*), ricerca che si è sviluppata soprattutto – ma non solo – nei Paesi anglosassoni a partire dagli ultimi due decenni del secolo scorso, in concomitanza con l'introduzione di sistemi

di valutazione standardizzata degli apprendimenti degli alunni e di *accountability* degli istituti scolastici (Peter-son e West 2003; Maroy 2005; Rosenkvist 2010).

Le domande che la ricerca sull'efficacia si è posta e a cui ha tentato con maggiore o minore successo di rispondere sono molteplici. Ne diamo di seguito un sommario elenco.

- Quale è l'unità di analisi più appropriata per valutare la qualità dell'istruzione? Il sistema scolastico, nel suo insieme e/o nelle sue articolazioni territoriali, la singola scuola, la classe?
- Ci sono scuole migliori o meno buone di altre? In altre parole, esiste un effetto della scuola, e se sì quale ne è la dimensione?
- Le differenze tra scuole più e meno buone sono valutabili obiettivamente e come?
- Se vi sono scuole più efficaci di altre, esse lo sono per tutti gli alunni nello stesso modo e per tutte le materie o i loro effetti sono diversi rispetto ad ogni particolare categoria di alunni e a seconda della materia?
- L'effetto della scuola è stabile nel tempo? La sua dimensione varia attraverso i diversi livelli d'istruzione?
- È possibile identificare delle caratteristiche di scuola significativamente e stabilmente associate a una minore o maggiore efficacia?
- Se sì, è possibile che, intervenendo su questi fattori, si possa ottenere un miglioramento della prestazione degli istituti meno efficaci? In altre parole, le 'buone pratiche' sono esportabili e generalizzabili?

Come si vede, si tratta di domande di rilevante importanza per la progettazione e la valutazione di politiche che si propongano di migliorare i risultati di apprendimento degli alunni migliorando la qualità dell'insegnamento impartito dalle singole scuole, che costituiscono i mattoni di cui il sistema d'istruzione è costruito.

Ma che cosa s'intende, propriamente, per valore aggiunto di una scuola?

Una scuola è un ambiente di apprendimento all'interno del quale gli alunni che la frequentano sviluppano le proprie conoscenze e abilità, pur con risultati diversi dall'uno all'altro. Se guardiamo ai risultati medi conseguiti nella stessa prova da scuole differenti constatiamo anche in questo caso una certa variabilità. A che cosa è dovuta tale variabilità? Sul risultato medio conseguito dagli alunni di una scuola influiscono diverse variabili: le loro caratteristiche socio-demografiche (la famiglia di provenienza, l'origine etnica, il genere, ecc.) e le competenze possedute nella materia oggetto della prova al momento dell'ingresso in quella data istituzione scolastica. Queste variabili, che

possiamo chiamare 'esogene', sono evidentemente estrinseche rispetto all'azione educativa della scuola stessa, che non può modificare l'ambiente sociale di provenienza o la preparazione iniziale degli alunni che in essa si iscrivono. Se dunque vogliamo valutare il contributo specifico che una scuola dà al progresso cognitivo dei propri alunni è necessario 'depurare' i risultati di apprendimento da questi ottenuti da tutte quelle influenze che sfuggono al suo controllo. Il valore aggiunto di una scuola è, infatti, l'effetto che essa ha sui livelli di apprendimento raggiunti in una data materia dai suoi studenti, 'al netto' degli effetti esercitati dalle variabili esogene. In altre parole, per valutare l'efficacia di una scuola è essenziale definire se e in quale misura essa sia riuscita a far sì che i propri alunni apprendessero più di quanto abbiano mediamente appreso alunni in tutto e per tutto 'comparabili' che hanno frequentato nello stesso periodo di tempo altre scuole (Scheerens 2018).

2. Misurare il valore aggiunto

Per stimare l'effetto della scuola sull'apprendimento dei propri alunni sono stati utilizzati diversi approcci e modelli matematici (Lissitz 2005). L'OCSE in una pubblicazione sulle migliori pratiche per valutare il valore aggiunto delle scuole distingue innanzitutto tra due tipi di modelli, quelli che chiama "modelli di apprendimento contestualizzato" e i modelli di valore aggiunto in senso proprio (OECD 2008). La differenza fondamentale tra i due tipi di modello consiste nel fatto che i primi stimano il contributo di una scuola all'apprendimento dei suoi alunni misurando i risultati in una o più prove in un unico punto del tempo e aggiustandoli statisticamente per tener conto degli effetti dovuti al contesto sociale dell'istituto, mentre i secondi misurano i livelli di apprendimento dei medesimi alunni in almeno due momenti successivi del tempo, all'inizio della frequenza di una certa scuola e dopo un certo periodo, mettendo così sotto controllo anche la variabile rappresentata dalle competenze possedute dagli alunni all'ingresso. È importante sottolineare che solo un approccio longitudinale è in grado, come ormai comunemente riconosciuto, di stimare propriamente il valore aggiunto della scuola e di separarne l'effetto dalle altre influenze che condizionano i risultati degli alunni in una data materia, in particolare dall'incidenza del loro grado di preparazione iniziale. Con un modello di apprendimento contestua-

lizzato, anche nel caso in cui le rilevazioni dei livelli di apprendimento siano ripetute nel corso del tempo, poiché queste sono effettuate su alunni ogni volta diversi, non è possibile stabilire se eventuali progressi o regressi nei risultati siano da attribuire all'azione della scuola o a mutamenti nella composizione della popolazione di studenti che la frequentano (Hanushek e Raymond 2003).

I modelli longitudinali di stima del valore aggiunto si suddividono, a loro volta, in due categorie: modelli basati su due misurazioni, una in ingresso e una in uscita, dei livelli di apprendimento degli studenti di una scuola e sull'aggiustamento delle covariate per tenere sotto controllo le variabili esogene che influenzano su di essi, e modelli basati su più misurazioni ripetute nel tempo del progresso cognitivo degli stessi alunni (curve di crescita), senza aggiustamento delle covariate partendo dal presupposto – per altro discusso – che in tal caso ogni alunno funge da controllo di se stesso (Sanders e Horn 1995; Sanders 1998; Collins e Amrein-Beardsley 2014). Questa seconda categoria di modelli, a differenza della prima, richiede che i risultati delle prove cui gli alunni sono sottoposti siano espressi su una medesima scala, o in altri termini che le prove siano ancorate verticalmente.

Ciò detto, tutti i modelli di stima del valore aggiunto hanno in comune, nonostante le differenze metodologiche tra l'uno e l'altro, l'uso di tecniche di regressione statistica. Il valore aggiunto è, infatti, dato dai residui di un'analisi di regressione, vale a dire dalla differenza tra i risultati 'osservati' di una certa scuola e i risultati 'attesi', cioè i risultati che essa teoricamente avrebbe dovuto ottenere tenuto conto delle caratteristiche dei suoi alunni. Il valore aggiunto corrisponde, infatti, a quella parte della variabilità tra i risultati delle scuole in una prova di apprendimento che non è spiegata dalle variabili esogene prese in considerazione nell'analisi di regressione.

Un progresso significativo nella misurazione del valore aggiunto è stato realizzato dall'introduzione dei modelli di regressione multilivello, che, diversamente dai modelli di regressione ordinaria, basati su un assunto di indipendenza delle osservazioni, tengono conto della struttura gerarchica (*nested*) che i dati sull'educazione tipicamente presentano: gli studenti, infatti, sono normalmente riuniti in classi, le classi nelle scuole, queste nei provveditorati, e così via (Goldstein 1995; Snijders e Bosker 1999; Hox 2002;

Raudenbush e Bryk 2002; Grilli e Rampichini 2009). La regressione a più livelli consente di tener conto degli effetti di aggregazione, e dunque della correlazione esistente tra le osservazioni appartenenti a uno stesso gruppo: è molto probabile, infatti, che gli alunni di una stessa classe o di una stessa scuola siano tra loro più simili degli alunni di altre classi o altre scuole, e ciò a causa sia di processi di selezione o autoselezione nella formazione dei gruppi, sia dell'esposizione a fattori comuni una volta che questi si siano costituiti. A tale riguardo la ricerca educativa ha messo in luce la presenza di un 'effetto di contesto', o di 'composizione del gruppo' (*school mix*) di cui un alunno fa parte, e che fa sì che il suo rendimento scolastico sia influenzato non solo dalle sue caratteristiche personali e dal patrimonio di abilità e conoscenze che egli ha in precedenza accumulato, ma anche dalle caratteristiche e dal livello medio delle competenze possedute dagli alunni della sua classe e della sua scuola (Thrupp 1997). La regressione multilivello permette di tener conto dell'effetto di contesto e dunque di stimare l'efficacia di una scuola al netto non solo delle caratteristiche individuali degli alunni, ma anche dell'effetto di composizione della popolazione scolastica che essa recluta.

3. Problemi e limiti nella misurazione del valore aggiunto

La stima del valore aggiunto delle scuole, nonostante i notevoli progressi realizzati, va incontro a una serie di difficoltà e limitazioni, di cui la letteratura offre ampia testimonianza (Van de Grift 2009; Everson 2017).

Una prima difficoltà è rappresentata dalla disponibilità per tutti gli alunni e le scuole coinvolte delle informazioni necessarie per il calcolo, in particolare i livelli di apprendimento progressi degli alunni. In una prospettiva di tipo longitudinale una certa perdita di dati è infatti pressoché inevitabile, cosa che, come ben noto, non è priva di conseguenze sulla robustezza dei risultati delle analisi e sulla tenuta delle inferenze che se ne traggono (Ding 2009).

Secondariamente, la scelta delle covariate da inserire nei modelli di calcolo è esposta al rischio, da un lato, di trascurare variabili degli studenti che pure hanno un peso sull'apprendimento, portando a un'insufficiente specificazione del modello, dall'altro lato di includere un numero eccessivo di variabili complicando il modello senza un apprezzabile guadagno nel grado di precisione delle stime (Papay 2011).

In terzo luogo, i processi sul piano organizzativo e pedagogico-didattico messi in atto da una scuola e da cui dovrebbe ipoteticamente dipendere il suo valore aggiunto non sono senza relazione con la composizione della popolazione di studenti che essa recluta, dalla quale sono invece a loro volta condizionati. Ciò, oltre a mettere in discussione due dei principali presupposti della stessa ricerca sull'efficacia, vale a dire l'indipendenza dei processi dal contesto e l'esistenza di un legame unidirezionale tra processi e risultati, ha una ripercussione sulla possibilità di pervenire a un calcolo preciso della dimensione del valore aggiunto di un istituto, giacché la messa sotto controllo degli effetti contestuali conduce di fatto a sottostimarli, in quanto neutralizza una fonte di variabilità rappresentata dall'interazione tra composizione del reclutamento di una scuola e variabili di processo (Dumay 2004; Dumay e Dupriez 2004; Rothstein 2009; Lockwood e McCaffrey 2009). Solo in una situazione ideale in cui gli alunni e gli insegnanti fossero assegnati in maniera del tutto aleatoria alle scuole e alle classi – cosa che difficilmente trova riscontro nella realtà – sarebbe possibile una misura del valore aggiunto esente da distorsioni.

In quarta istanza, gli indicatori di valore aggiunto si sono rivelati relativamente instabili, tanto più quanto più si allunga l'intervallo temporale tra una misurazione e l'altra dei livelli di apprendimento: una scuola che si sia dimostrata efficace in un certo anno scolastico, non è detto che lo sia anche l'anno successivo, senza contare che una scuola può dimostrarsi efficace nell'insegnamento di una certa materia ma non di un'altra (Kyriakides e Creemers 2008; Ferrão 2012; Agasisti e Minaya 2018).

Infine, non va dimenticato che la stima del valore aggiunto è soggetta, come ogni altra procedura statistica, a margini d'incertezza, così da rendere discutibile ogni operazione di *ranking* delle scuole basata su tale stima (Leckie e Goldstein 2011). In generale, si può dire che due scuole qualsiasi hanno prestazioni significativamente diverse se gli intervalli di fiducia all'interno dei cui limiti si trova il loro risultato 'vero' non si sovrappongono. L'aggiustamento statistico necessario per separare l'effetto della scuola dalle variabili estranee al suo operato 'schiaccia' i risultati delle scuole verso il

centro della distribuzione, cosicché solo quelle collocate ai due estremi si differenziano con sufficiente sicurezza rispetto alle altre e alla media generale.

Nonostante i limiti sopra accennati e altri su cui qui sorvoliamo, la disponibilità, accanto ai risultati di apprendimento osservati o 'grezzi' delle scuole, anche dei loro risultati depurati dal condizionamento su di essi esercitato dalle caratteristiche degli alunni, è essenziale ai fini di una comparazione che voglia ad un tempo essere equa e sensata. La considerazione dei soli risultati osservati, infatti, confonde insieme qualità degli alunni e qualità dell'istruzione e istituisce un ingiusto confronto tra scuole con popolazioni di studenti anche molto diverse tra loro. Una scuola con una larga proporzione di studenti in condizioni di difficoltà potrebbe infatti avere un risultato basso in termini assoluti, ma un buon risultato in termini di valore aggiunto (Martini 2014). Inoltre, sebbene, come prima accennato, le analisi sul valore aggiunto facciano emergere solo una minoranza di scuole suscettibili di distinguersi dalle altre e dalla media per la loro relativa efficacia o inefficacia, ciò non diminuisce in alcun modo l'utilità di tali analisi se, invece che sulla graduatoria come tale, si punta l'attenzione sulle caratteristiche che contraddistinguono gli istituti all'estremità superiore e inferiore della distribuzione: non sfugge infatti ad alcuno la rilevanza delle indicazioni che da queste indagini possono venire, ai fini della progettazione di politiche di intervento, a quanti hanno a vario livello la responsabilità del governo del sistema scolastico, nel suo insieme e nelle unità che lo compongono.

4. La stima del valore aggiunto secondo la metodologia dell'Invalsi

Dal 2016 l'Invalsi restituisce alle scuole i risultati ottenuti nelle prove di Italiano e Matematica di quinta primaria, terza secondaria di primo grado e seconda e quinta secondaria di secondo grado non solo in termini di risultati osservati, ma anche in termini di indicatori di valore aggiunto.

Il modello di stima utilizzato dall'Istituto di Valutazione è un modello di regressione a due livelli, studente e scuola, a intercette casuali e pendenze fisse (*random/fixed effects model*)¹.

1 Nel modello solo le intercette (medie delle scuole) sono lasciate libere di variare, mentre le pendenze sono costrette sul loro valore medio. In pratica, ciò equivale a supporre che le variabili indipendenti considerate nelle analisi di regressione

Tabella 1
Variabili di livello 1 e 2 nel modello di stima Invalsi del valore aggiunto

Nome variabile	Descrizione
Variabili di livello 1: studenti	
Punt. Ing_i	Punteggio d'ingresso in Italiano o Matematica
Femmina	Genere dello studente (<i>Maschio / Femmina</i>)
Straniero	Cittadinanza dello studente (<i>Italiano / Straniero</i>)
ESCS_i	Status socio-economico-culturale
Ritardo	Regolarità nel percorso degli studi (<i>Regolare o in anticipo / In ritardo</i>)
Freq. Sc. infanzia	L'alunno ha frequentato la scuola dell'infanzia (<i>Si / No</i>)
Altra lingua	Lingua parlata in casa dallo studente (<i>Italiano / Altra lingua</i>)
Dialetto	Lo studente parla in casa il dialetto (<i>Si / No</i>)
Variabili di livello 2: scuole	
Punt. Ing._sc	Media dei punteggi in Italiano o in Matematica degli studenti della scuola all'ingresso
ESCS_sc	Status socio-economico-culturale medio degli studenti della scuola
% stranieri	Percentuale di studenti d'origine straniera della scuola
% ritardatari	Percentuale di alunni in ritardo della scuola
% ass. alla prova	Percentuale di assenti alla prova d'Italiano o di Matematica (solo scuola secondaria di secondo grado)

Nota: per le variabili categoriali, la categoria di riferimento è indicata tra parentesi in carattere corsivo.

Fonte: Invalsi 2019

L'Invalsi stima il valore aggiunto delle scuole separatamente per ognuno dei due ambiti disciplinari interessati e per ognuna delle due classi terminali del primo ciclo d'istruzione (grado 5 e 8), mentre per la seconda e quinta classe della scuola secondaria di secondo grado (grado 10 e 13) vengono stimati modelli distinti, oltre che per l'Italiano e la Matematica, anche per ognuno di quattro tipi di scuola: licei scientifici e classici (tipo 1), altri licei (tipo 2), istituti tecnici (tipo 3), istituti professionali (tipo 4). Per l'Italiano i licei classici e scientifici sono posti nella stessa ca-

tegoria (tipo 1), mentre per la Matematica i licei scientifici sono considerati a sé stanti (tipo 1) e i licei classici sono considerati insieme agli altri licei (tipo 2).

Nei modelli di regressione stimati la variabile dipendente è costituita dal punteggio conseguito dall'alunno nelle prove di Italiano e Matematica dei quattro gradi scolari sopra indicati (5, 8, 10, 13)². Nella tabella 1 sono elencate le variabili indipendenti prese in considerazione nelle analisi.

È opportuno precisare che i punteggi d'ingresso sono costituiti dai punteggi che gli studenti hanno conseguito, in Italiano o in Matematica, nella rilevazione del grado precedente a quello i cui risultati rappresentano la variabile dipendente della regressione; quindi, ad esempio, per la quinta primaria il punteggio d'ingresso è il punteggio ottenuto dall'alunno tre anni prima in seconda primaria, per la terza secondaria di primo grado è il punteggio di quinta primaria, e così via.

La procedura seguita dall'Invalsi per la stima del valore aggiunto consiste essenzialmente di tre passi. Al primo passo la variabilità totale dei risultati nelle prove di Italiano e Ma-

tematica dei quattro gradi scolari sopra indicati viene scomposta in due componenti: quella dovuta a differenze tra gli alunni all'interno delle scuole (varianza *within*) e quella dovuta a differenze tra le scuole (varianza *between*); al secondo passo, per ognuna delle variabili di primo e secondo livello considerate nelle analisi, si calcola l'effetto che esse hanno sul punteggio conseguito dallo studente in Italiano o in Matematica; infine, basandosi sui 'residui' a livello 2 della regressione, si stima per ogni istituto un indicatore di valore aggiunto, che, in pratica, corrisponde a quella parte della

abbiano il medesimo effetto in tutte le scuole. Per una descrizione più dettagliata del modello di stima del valore aggiunto si veda: Martini 2018 (a cura di), paragrafi 3 e 4.

2 Il punteggio è standardizzato con media eguale a 0 e deviazione standard eguale a 1.

variabilità dei risultati tra le scuole che non è spiegata dalle caratteristiche, individuali e aggregate, degli studenti o, in altre parole, alla media delle differenze tra i punteggi osservati degli studenti di una scuola e i punteggi che essi avrebbero dovuto raggiungere in base alle loro caratteristiche. Quando il risultato osservato di una scuola è superiore alle attese ciò significa che i livelli di apprendimento dei suoi alunni sono più alti di quelli previsti, mentre quando è inferiore ciò vuol dire che essi sono più bassi di quelli previsti.

È opportuno chiarire, prima di passare a illustrare i risultati delle analisi, che, sebbene si sia fin qui usata, così come si continuerà a usare nei paragrafi successivi, l'espressione 'effetto', con essa si intende il coefficiente associato alle variabili indipendenti prese in considerazione nei modelli di regressione, coefficiente che indica una correlazione tra le suddette variabili e la variabile dipendente, ma non necessariamente l'esistenza di un rapporto di causalità diretta fra le prime e la seconda. L'uso del termine 'effetto' è fatto per semplicità di discorso, ma non implica di per sé un'interpretazione in senso causale della relazione fra le variabili in esame.

5. I risultati della stima del valore aggiunto delle scuole nel 2019

Prima di procedere alla stima del valore aggiunto, la base di dati della rilevazione 2019 è stata sottoposta a un'operazione di 'pulizia' in quattro fasi successive: nella prima fase sono stati eliminati gli alunni con valori mancanti sulla variabile rappresentata dal punteggio d'ingresso; nella seconda fase sono state eliminate le classi dove il numero di alunni per cui era disponibile l'informazione su tale variabile era inferiore al 70%; nella terza fase sono state eliminate le scuole con meno del 50% di classi sopravvissute; infine, nella quarta fase, sono state eliminate le scuole con meno di dieci studenti. L'operazione di pulizia ha avuto come ovvia conseguenza che non per tut-

te le scuole, e per ognuna di esse in entrambe le materie, è stato possibile restituire il dato sul valore aggiunto³.

Le tabelle 2 e 3 danno un quadro complessivo degli effetti sul punteggio di Italiano e Matematica di ognuna delle variabili di livello 1 e 2 inserite nei modelli di regressione, al netto del peso esercitato dalle altre, per ogni grado e tipo di scuola.

Quasi tutte le variabili prese in considerazione nelle analisi hanno un effetto statisticamente significativo sul risultato nelle prove. Tra le variabili socio-demografiche di primo livello, l'essere di genere femminile ha in genere, a parità delle altre condizioni, un effetto positivo sul risultato della prova di Italiano e un effetto negativo su quello della prova di Matematica; nel grado 13, tuttavia, nei licei di ogni tipo l'effetto in Italiano è lievemente negativo e, negli istituti tecnici e professionali, non è statisticamente significativo. Secondo le aspettative, l'aumento dell'indice, denominato Escs, dello status socio-economico-culturale della famiglia dello studente (Campodifiori, Figura, Papini, Ricci 2010) esercita un effetto positivo sui risultati in entrambe le materie, ma tale effetto si riduce passando dal primo ciclo d'istruzione alla scuola secondaria di secondo grado e, negli istituti tecnici e professionali, esso, in Matematica, non è più, nel grado 13, significativo. Hanno per lo più, invece, un effetto negativo l'essere d'origine immigrata e il parlare a casa una lingua diversa dall'italiano o il dialetto, anche se l'effetto di queste variabili nella scuola secondaria superiore non è a volte significativo.

Per quanto concerne le variabili relative alla carriera scolastica degli studenti, l'essere in ritardo sul regolare percorso degli studi ha un effetto negativo, mentre l'aver frequentato la scuola dell'infanzia ha in genere un effetto positivo, che, tuttavia, raggiunge la soglia di significatività nelle scuole del primo ciclo mentre nella scuola secondaria di secondo grado ciò si verifica solo in alcuni casi. L'aumento del punteg-

3 La perdita di informazioni è stata particolarmente consistente per il grado 13, a causa di una più ridotta partecipazione alla rilevazione 2015, anno in cui gli studenti che nel 2018-19 hanno frequentato la quinta secondaria di secondo grado si trovavano in II classe, dovuta a uno sciopero nazionale degli insegnanti indetto nei giorni delle rilevazioni. La limitata partecipazione ha caratterizzato soprattutto alcune regioni del Mezzogiorno e in particolare, come ha dimostrato un'analisi effettuata *ex post*, le scuole che avevano ottenuto bassi risultati nelle prove dell'anno precedente. Tutto ciò ha operato come un fattore di selezione della popolazione di studenti e scuole su cui la stima del valore aggiunto è stata effettuata ed ha avuto come conseguenza una riduzione della variabilità complessiva e delle sue componenti *entro* le scuole e *tra* le scuole. Le stime del valore aggiunto per le scuole di questo grado scolastico vanno dunque considerate con una certa cautela.

Tabella 2
Effetti netti sul punteggio di Italiano

	Gr. 5	Gr. 8	Gr. 10				Gr. 13			
			T.1	T.2	T.3	T.4	T.1	T.2	T.3	T.4
Livello 1										
Punt. Ing_i	0,459	0,558	0,398	0,498	0,533	0,527	0,549	0,570	0,629	0,674
Femmina	0,191	0,138	0,088	0,046	0,105	0,118	-0,039	-0,021	-0,008	-0,011
Straniero	-0,087	-0,081	-0,080	-0,050	-0,033	-0,079	-0,125	-0,118	-0,108	-0,106
ESCS_i	0,145	0,135	0,060	0,042	0,026	0,030	0,033	0,029	0,011	0,023
Ritardo	-0,182	-0,079	-0,172	-0,068	-0,073	-0,028	-0,190	-0,151	-0,122	-0,073
Freq. Sc. inf.	0,092	0,024	0,008	0,010	0,020	0,013	0,049	0,038	0,012	0,060
Altra lingua	-0,180	-0,100	-0,184	-0,173	-0,164	-0,174	0,010	0,000	-0,064	-0,014
Dialetto	-0,069	-0,092	-0,085	-0,044	-0,053	-0,027	-0,069	-0,061	-0,066	-0,018
Livello 2										
Punt.Ing._sc	0,205	0,403	0,859	1,005	1,050	0,983	0,581	0,685	0,758	0,773
ESCS_sc	0,178	0,423	0,315	0,156	0,139	0,188	0,196	0,129	0,177	0,237
% stranieri	-0,001	0,006	0,007	0,001	0,003	0,005	0,009	0,008	0,003	0,001
% ritardatari	-0,005	0,000	-0,006	0,001	0,001	0,000	-0,005	-0,003	0,000	0,002
% ass. prova	-	-	-0,015	-0,010	-0,012	-0,009	-0,020	-0,017	-0,027	-0,015

Nota: i valori in carattere corsivo non sono statisticamente significativi (p-value >0,05).

Fonte: Invalsi 2019

Tabella 3
Effetti netti sul punteggio di Matematica

	Gr. 5	Gr. 8	Gr. 10				Gr. 13			
			T.1	T.2	T.3	T.4	T.1	T.2	T.3	T.4
Livello 1										
Punt. Ing_i	0,493	0,542	0,446	0,514	0,526	0,507	0,457	0,643	0,545	0,673
Femmina	-0,098	-0,040	-0,051	-0,065	-0,054	-0,060	-0,129	-0,077	-0,176	-0,154
Straniero	-0,043	-0,032	-0,029	0,000	-0,008	-0,035	-0,055	-0,047	-0,051	-0,054
ESCS_i	0,136	0,142	0,025	0,044	0,009	0,014	0,014	0,026	-0,002	0,002
Ritardo	-0,218	-0,162	-0,145	-0,133	-0,090	-0,075	-0,249	-0,228	-0,189	-0,157
Freq. Sc. inf.	0,075	0,016	-0,002	0,016	-0,009	-0,009	0,015	0,027	0,007	0,049
Altra lingua	-0,160	-0,089	-0,182	-0,177	-0,145	-0,148	-0,007	0,013	0,024	0,019
Dialetto	-0,069	-0,117	-0,078	-0,051	-0,028	-0,019	-0,069	-0,066	-0,028	-0,002
Livello 2										
Punt.Ing._sc	0,199	0,261	1,019	1,002	1,146	1,056	0,373	0,773	0,969	0,741
ESCS_sc	0,155	0,565	0,098	0,221	0,126	0,283	0,243	0,157	0,129	0,209
% stranieri	-0,002	0,013	0,006	0,006	0,003	0,004	0,013	0,006	0,000	0,002
% ritardatari	-0,005	-0,002	-0,003	0,008	0,001	0,003	-0,005	0,000	0,003	0,003
% ass. prova	-	-	-0,013	-0,014	-0,014	-0,008	-0,029	-0,017	-0,022	-0,015

Nota: i valori in carattere corsivo non sono statisticamente significativi (p-value >0,05).

Fonte: Invalsi 2019

gio d'ingresso di Italiano o di Matematica incide invece positivamente sul punteggio conseguito nel grado successivo in tutti i gradi e ordini di scuola. In generale, è questa la variabile individuale che ha il peso maggiore sulla variabile dipendente.

Al secondo livello, il punteggio medio degli alunni all'ingresso e il loro indice di status medio hanno un effetto positivo in tutti i gradi scolari analizzati e in entrambi gli ambiti disciplinari, ma, come d'altronde a livello individuale, è nettamente maggiore il peso della prima delle due variabili rispetto alla seconda. In generale, in ogni caso, la composizione del corpo studentesco della scuola dal punto di vista socio-economico e del livello di competenza all'ingresso si conferma, come documentato dalla letteratura sul tema⁴, un robusto predittore del risultato degli studenti. L'effetto aggregato del punteggio d'ingresso è, nel primo ciclo, inferiore a quello che si registra a livello individuale, ma generalmente superiore a quest'ultimo nella scuola secondaria di secondo grado, mentre l'effetto aggregato dell'indice di status è maggiore di quello individuale in tutti i gradi e i tipi di scuola. Le altre due variabili relative alla composizione del corpo studentesco – la percentuale di studenti d'origine immigrata e di alunni in ritardo – hanno effetti piccoli, di segno differente a seconda del grado e della materia, e a volte non significativi statisticamente; un effetto sempre significativamente negativo ha, infine, nella scuola secondaria di secondo grado la percentuale di studenti assenti alla prova.

Come ultima notazione, si osservi che, data la natura additiva del modello, le variazioni di punteggio associate alle variabili di livello 2 si sommano alle variazioni associate alle variabili di livello 1. Quindi, ad esempio, dalla tabella 3 si evince che, quando l'indice di status personale di uno studente del grado 8 aumenta di una unità di deviazione standard, il suo risultato in Matematica cresce di 0,142 punti, a cui si

aggiunge un incremento di altri 0,565 punti se l'indice di status medio della sua scuola aumenta della stessa quantità.

Poiché il valore aggiunto, come s'è detto, è dato dai residui di livello 2 della varianza, è il caso innanzitutto di rilevare che la quota di varianza tra le scuole spiegata dalle caratteristiche degli studenti è maggiore della quota di varianza spiegata tra gli alunni, fatta eccezione per la scuola primaria, dove la relazione tra le due è invertita, come si vede dalla tabella 4, dove è riportato per ogni grado e tipo di scuola anche il valore del coefficiente di correlazione intraclassa (ICC). Tale coefficiente è il rapporto percentuale fra la componente di varianza tra le scuole e la varianza totale e fornisce dunque una misura del grado di somiglianza delle osservazioni appartenenti a uno stesso gruppo o, in altre parole, di quanto la variabile risposta, ossia il risultato degli studenti nelle prove dei vari gradi scolastici, è influenzata dal raggruppamento degli alunni nelle scuole.

Dal quadro tracciato dalla tabella 4 emerge, da una parte, che il coefficiente di correlazione intraclassa, basso nella scuola primaria, tende a crescere nel corso dell'itinerario scolastico ed è comunque sempre più alto in Matematica⁵ rispetto all'Italiano. Dall'altra parte, la varianza tra le scuole spiegata dalle caratteristiche degli studenti tende anch'essa ad aumentare da un grado scolastico al successivo, in parallelo col coefficiente ICC, in particolare passando dalla scuola primaria alla scuola secondaria di primo e secondo grado⁶, indicando un rafforzamento del peso della composizione del corpo studentesco della scuola sui suoi risultati, soprattutto al termine del biennio della scuola secondaria di secondo grado.

I risultati delle analisi illustrati nelle tabelle 2, 3 e 4 possono essere interpretati come un segno del fatto che la ripartizione degli alunni tra le scuole è nella scuola primaria relativamente indipendente dalla

4 Ad esempio, dai rapporti sui risultati dell'indagine PISA si evince che se si considerano due studenti della stessa scuola il cui status socio-economico-culturale differisce di mezza unità di deviazione standard e poi due studenti con lo stesso status ma che frequentano scuole diverse, di cui l'una con un indice medio di status inferiore a quello dell'altra sempre di mezza unità di deviazione standard, si può prevedere che la disparità di risultati nei test PISA fra gli studenti della seconda coppia sarà maggiore della disparità di risultati fra gli studenti della prima coppia (OECD 2010).

5 La Matematica è una materia che si apprende soprattutto a scuola mentre le competenze in Italiano sono più dipendenti anche da fattori extra-scolastici.

6 Il fatto che la varianza di livello 2 spiegata dalle caratteristiche degli studenti diminuisca dal grado 10 al grado 13 è probabilmente da mettere in relazione con la riduzione della variabilità dovuta sia a processi di selezione degli studenti nel passaggio dal biennio alla conclusione della scuola secondaria sia a quanto detto nella nota 2.

Tabella 4
Coefficiente intraclasse e varianza di livello 1 e 2 spiegata

	ITALIANO			MATEMATICA		
	ICC (%)	Varianza liv.1 spiegata (%)	Varianza liv.2 spiegata (%)	ICC (%)	Varianza liv.1 spiegata (%)	Varianza liv.2 spiegata (%)
Grado 5	7,1	28,3	20,6	11,8	29,9	10,7
Grado 8	9,3	44,3	59,7	13,3	39,2	47,6
Grado 10	Licei cl. e sc.	21,3	19,9	72,7	29,7	26,3
	Altri licei	18,8	25,1	76,1	25,1	29,2
	Ist. Tecnici	17,6	24,9	67,0	27,6	28,8
	Ist. Profess.	13,6	21,6	50,8	18,7	20,7
Grado 13	Licei cl. e sc.	23,6	21,0	41,0	33,2	24,0
	Altri licei	24,8	25,6	54,9	28,7	29,8
	Ist. Tecnici	23,3	26,4	61,6	37,9	26,0
	Ist. Profess.	20,8	28,0	55,2	24,3	23,0

Fonte: Invalsi 2019

condizione sociale e dal livello di abilità degli alunni, mentre nella scuola secondaria di primo e secondo grado la distribuzione degli alunni appare più diseguale, con una tendenza al raggruppamento nella stessa scuola di alunni simili sotto il profilo sociale e il grado di abilità.

Una seconda osservazione riguarda la scuola primaria: in questo grado scolare, diversamente dagli altri, in entrambe le discipline la varianza spiegata tra le scuole è, come s'è visto, minore di quella tra gli alunni; inoltre, le caratteristiche degli studenti rendono conto di una quota della varianza tra le scuole in Matematica che è la metà di quella spiegata dalle stesse variabili in Italiano. Tutto ciò ci dice che, per render conto della varianza tra scuole, in particolare in Matematica, sarebbe necessario ricorrere ad altre variabili non prese in considerazione in questa analisi, in primo luogo la competenza professionale degli insegnanti.

A questo riguardo è il caso di precisare che, in generale, il valore aggiunto o effetto-scuola stimato dall'Invalsi è di fatto una *black box*: esso infatti comprende tutti gli effetti dovuti alla frequenza di quella determinata scuola senza distinzione alcuna tra gli effetti dovuti alla sua collocazione territoriale, alle infrastrutture e alle dotazioni tecniche di cui dispone, ai processi organizzativi e pedagogico-didattici posti in atto e alla qualità del dirigente e dei docenti che in essa svolgono la propria azione. Penetrare all'interno

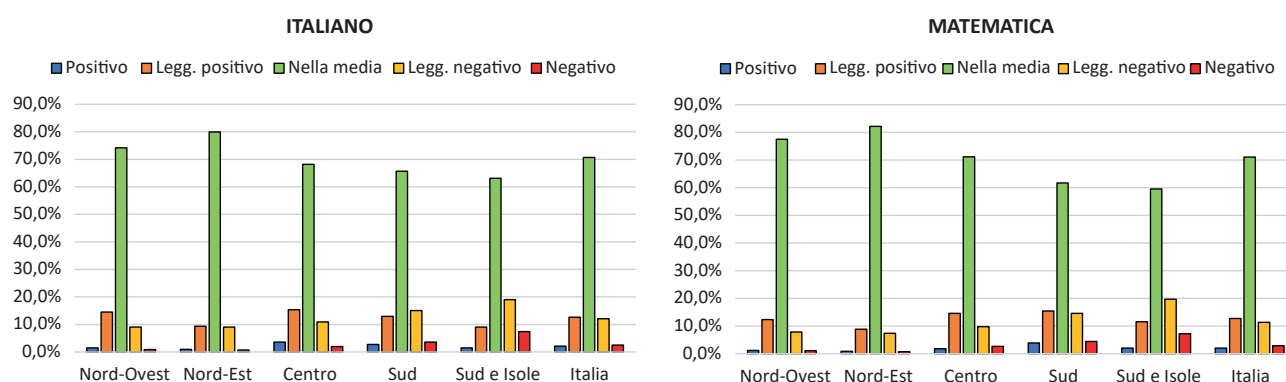
della scatola nera per individuare quali variabili siano responsabili della maggiore o minore efficacia di una scuola rispetto all'altra richiederebbe analisi *ad hoc* e, in particolare, informazioni sugli istituti e sul personale dirigente e docente che vi opera, dati di cui al momento l'Invalsi non dispone o che non ha la possibilità di incrociare con le altre variabili in suo possesso.

6. Il valore aggiunto delle scuole in Italia

Una volta calcolati gli indicatori di valore aggiunto delle scuole di ogni grado e tipo di scuola in base ai residui di livello 2 della regressione, questi, sulla base della deviazione standard dalla media (pari a zero), sono suddivisi in cinque categorie ordinate: la prima categoria (VA positivo) comprende le scuole il cui valore aggiunto raggiunge o supera le due deviazioni standard dalla media, la seconda categoria (VA leggermente positivo) include le scuole il cui valore aggiunto è compreso tra una e due deviazioni standard dalla media, la terza categoria (VA nella media) comprende le scuole il cui valore aggiunto si colloca fra +1 e -1 deviazioni standard dalla media; in modo simmetrico alle prime due categorie, sono definite la quarta e la quinta categoria, quelle delle scuole con valore aggiunto, rispettivamente, leggermente negativo e negativo.

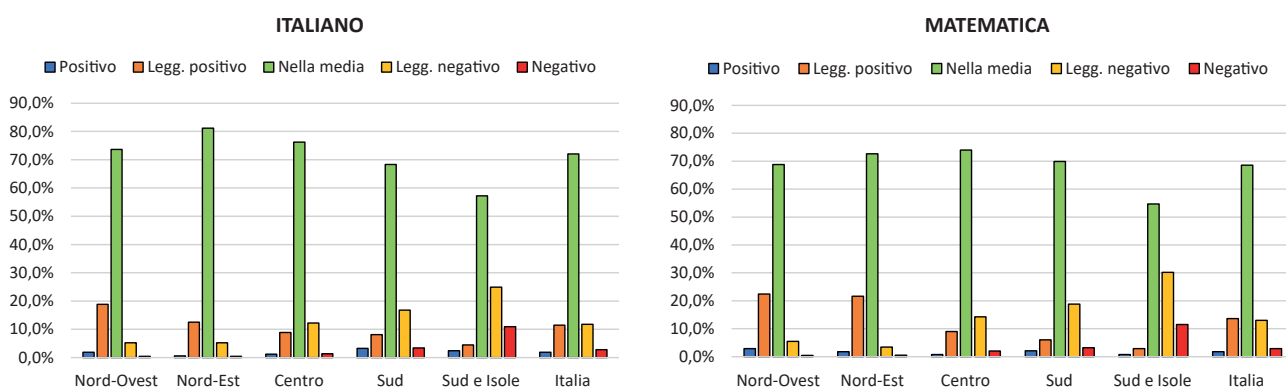
Le quattro figure che seguono rappresentano, per ogni grado scolare, la distribuzione percentuale delle scuole nelle cinque categorie di valore aggiunto in

Figura 1
Distribuzione delle scuole per categoria di valore aggiunto – Grado 5



Fonte: Invalsi 2019

Figura 2
Distribuzione delle scuole per categoria di valore aggiunto – Grado 8



Fonte: Invalsi 2019

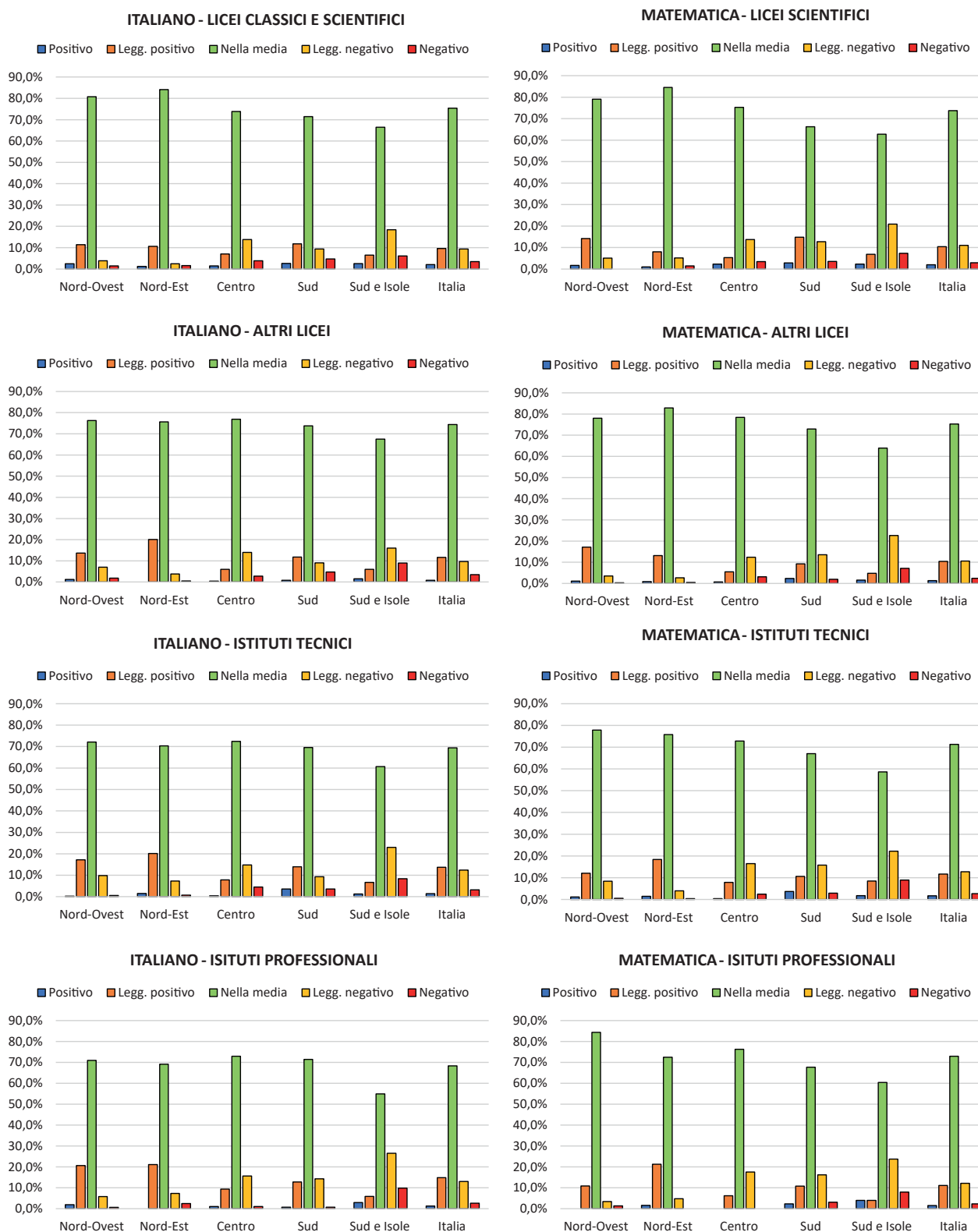
ognuna delle macro-aree in cui è suddiviso il territorio italiano e nell'intero Paese.

Come si può constatare dalle figure, in tutti i gradi e i tipi di scuola, sia in Italiano che in Matematica, la categoria che registra la percentuale più alta di scuole è quella di mezzo, dove si trovano gli istituti il cui valore aggiunto non si differenzia in maniera significativa dalla media, pari a zero; per contro, le due categorie estreme registrano una percentuale bassa di istituti. La grande maggioranza delle scuole, dunque, consegue i risultati che era prevedibile ottenesse tenuto conto delle caratteristiche dei suoi alunni, e solo una minoranza di esse si differenzia in senso positivo o negativo dal valore medio. È anche il caso di ripetere che la categoria di valore aggiunto in cui una scuo-

la si posiziona non è necessariamente la stessa per l'italiano e per la Matematica.

Per evitare che si cada in un possibile malinteso, è opportuno sottolineare che il fatto che la grande maggioranza delle scuole abbia un indicatore di valore aggiunto non significativamente diverso da zero non vuol dire che solo gli alunni delle scuole con valore aggiunto positivo imparano mentre quelli che frequentano una scuola con valore aggiunto nullo o anche negativo non imparano o, peggio, disimparano: in realtà gli studenti di tutte le scuole apprendono, ma solo alcune di queste conseguono risultati che si discostano dalle aspettative formulabili sulla base del background sociale e dei livelli di competenza in ingresso dei propri studenti.

Figura 3
Distribuzione delle scuole per categoria di valore aggiunto – Grado 10



Fonte: Invalsi 2019

Figura 4
Distribuzione delle scuole per categoria di valore aggiunto – Grado 13



Fonte: Invalsi 2019

Per meglio capire questo punto è utile rifarsi a un esempio (Rutter 1983) citato nell'*International Handbook of School Effectiveness Research* (Teddlie e Reynolds 2000, 97). Si immagini che l'unico modo per imparare il sanscrito in una comunità sia attraverso la frequenza della scuola, giacché gli insegnanti sono i soli a conoscerlo e a possedere testi in quella lingua. Si supponga anche che tutti gli insegnanti siano egualmente efficaci nell'insegnare il sanscrito e che gli alunni siano assegnati in maniera completamente casuale alle scuole e alle classi. In questa situazione si riscontrerebbe ancora una variabilità nell'apprendimento del sanscrito tra gli alunni: alcuni infatti lo imparerebbero meglio di altri. Tuttavia, la scuola che gli studenti frequentano non spiegherebbe per nulla tale variabilità. In breve, in queste condizioni, non verrebbe rilevato nessun effetto della scuola sull'apprendimento del sanscrito, nonostante il fatto che esso sia stato evidentemente appreso solo come effetto della scolarizzazione.

Ciò detto, tornando alle figure, si può da queste constatare che la distribuzione delle scuole nelle cinque categorie di valore aggiunto non è uniforme sul territorio nazionale, ma varia in funzione dell'area geografica e del grado scolare.

Nella scuola primaria (figura 1) il numero di scuole con valore aggiunto positivo o leggermente positivo nel Sud e nel Sud e Isole raggiunge o supera in percentuale il numero di scuole della stessa categoria del Nord e del Centro, ma risulta anche un po' più alta la percentuale di scuole con valore aggiunto leggermente negativo o negativo.

Nella scuola secondaria di primo grado (figura 2), il quadro appare in parte mutato: il numero di scuole con valore aggiunto leggermente positivo o positivo nel Nord e nel Centro (in particolare nel Nord-Ovest) supera in percentuale quello delle analoghe scuole del Sud e del Sud e Isole, mentre ad un tempo cresce in queste aree, rispetto alla scuola primaria, la percentuale di scuole con valore aggiunto leggermente negativo o negativo, specie nel Sud e Isole.

Nella scuola secondaria di secondo grado, al grado 10 (figura 3), il quadro delineatosi al grado precedente tende a riprodursi in misura maggiore o minore a seconda del tipo di scuola considerato. Da notare che in questo grado scolare, la percentuale di scuole delle prime due categorie del Nord-Est supera quella del Nord-Ovest, mentre nel Centro essa è simile, e a volte inferiore, a quella che si osserva nel Sud e nel Sud

e Isole. Ciò che differenzia queste due aree dal Centro e ancor più dalle due aree del Nord è la presenza di una percentuale di scuole con valore aggiunto leggermente negativo o negativo maggiore di quella che si riscontra altrove, in particolare negli "Altri licei" e nell'istruzione tecnica e professionale.

Infine, nel grado 13 (figura 4), le scuole con valore aggiunto leggermente positivo o positivo appaiono in genere percentualmente più numerose nel Nord-Ovest rispetto al Nord-Est, mentre nel Sud e in particolare nel Sud e Isole la quota percentuale di scuole con valore aggiunto leggermente negativo o negativo continua a superare quella che si registra nelle altre aree del Paese.

In conclusione, per sintetizzare, si può affermare che ciò che differenzia il sistema scolastico nel Meridione e nelle Isole rispetto al Centro e più ancora al Nord-Italia è una più pronunciata tendenza alla polarizzazione tra istituti più e meno efficaci. In altre parole, nel nostro Mezzogiorno non mancano in ogni grado scolare scuole altrettanto efficaci di quelle del resto d'Italia, ma vi sono nello stesso tempo scuole scarsamente efficaci in misura percentualmente maggiore di quella che si osserva nelle regioni centrali e settentrionali del Paese. Da questo punto di vista gli indicatori di valore aggiunto riconfermano la situazione che emerge anche dai dati grezzi, vale a dire la più alta proporzione di varianza tra scuole nei risultati delle prove Invalsi che si riscontra nelle due macro-aree del Mezzogiorno rispetto al Centro-Nord, fenomeno che si manifesta già a partire dalle scuole del primo ciclo, dove invece dovrebbero essere assicurate a tutti gli studenti la stessa qualità d'istruzione e le stesse condizioni di insegnamento-apprendimento a garanzia dell'eguaglianza di opportunità.

7. La valutazione delle scuole: conclusioni

Al di là dei divari territoriali che le rilevazioni sui livelli di apprendimento nel nostro Paese, internazionali e nazionali, continuano a mettere in luce, sia che si considerino i risultati grezzi sia gli indicatori di valore aggiunto, in ogni caso la comparazione dei risultati delle scuole sulla base di questi ultimi anziché dei punteggi grezzi porta a valutazioni diverse nei due casi. Anche se tra punteggi grezzi e indicatori di valore aggiunto sussiste una correlazione positiva, tuttavia la posizione delle scuole in una ipotetica graduatoria costruita su questi ultimi cambia rispetto a una graduatoria costruita in base ai primi. Una misura di tale variazione è data dal

Tabella 5
Coefficiente di correlazione a ranghi tra punteggi grezzi e indicatori di valore aggiunto

		ITALIANO <i>Tau di Kendall</i>	MATEMATICA <i>Tau di Kendall</i>
Grado 5		0,650	0,733
Grado 8		0,445	0,522
Grado 10	Licei classici e scientifici	0,368	-
	Licei scientifici	-	0,350
	Altri licei	0,355	0,425
Grado 13	Istituti tecnici	0,442	0,374
	Istituti professionali	0,543	0,540
	Licei classici e scientifici	0,643	-
Grado 13	Licei scientifici	-	0,658
	Altri licei	0,603	0,516
	Istituti tecnici	0,562	0,522
	Istituti professionali	0,565	0,632

Fonte: Invalsi 2019

coefficiente di correlazione a ranghi (*Tau* di Kendall). Esso fornisce infatti una misura del grado in cui le posizioni occupate dai valori della distribuzione di due variabili ordinate coincidono fra loro, oscillando in un intervallo tra -1 e +1 a seconda della minore o maggiore difformità o conformità tra gli ordinamenti dell'una e

l'altra variabile. La tabella 5 mostra per ogni grado scolastico e tipo di scuola il valore di tale coefficiente di correlazione tra i punteggi grezzi medi e gli indicatori di valore aggiunto delle scuole in Italiano e in Matematica.

Come si può vedere, tra i punteggi medi grezzi delle scuole e i punteggi di valore aggiunto permane una relativa associazione, maggiore in quinta primaria dove, come s'è visto, la correlazione intraclasse è più bassa rispetto a quella che si registra nei gradi scolari successivi. Tutti i coefficienti di correlazione riportati nella tabella 5 sono comunque inferiori a 1 e questo indica che l'ordinamento delle scuole cambia, in misura più o meno grande, quando si tiene conto dei loro punteggi grezzi o, in alternativa, degli indicatori di valore aggiunto.

Per chiudere, nonostante i limiti segnalati al paragrafo 3, il valore aggiunto rimane dal punto di vista del governo del sistema educativo un indispensabile parametro di giudizio per un'equa valutazione dell'operato delle scuole e della qualità dell'istruzione da esse fornita.

Bibliografia

- Agasisti T., Minaya V. (2018), *Evaluating the stability of school performance estimates for school choice: evidence for Italian primary school*, SIEP WP n.733, Pavia, Società italiana di economia pubblica
- Campodifiori E., Figura E., Papini M., Ricci R. (2010), *Un indicatore di status socio-economico-culturale degli allievi della quinta primaria in Italia*, Roma, INVALSI WP n. 2
- Collins C., Amrein-Beardsley A. (2014), Putting growth and value-added models on the map: a national overview, *Teachers College Record*, 116, n.1, pp.1-32
- Ding C.S. (2009), Measurement issues in designing and implementing longitudinal evaluation studies, *Educational Assessment, Evaluation and Accountability*, 21, n.2, pp.155-171
- Dumay X. (2004), *Effet établissement: effet de composition et/ou effet des pratiques managériales et pédagogiques? Un état du débat*, Les Cahiers de Recherche en Education et Formation n.34, Louvain, GIRSEF
- Dumay X., Dupriez V. (2004), *Effet établissement: effet de processus et/ou effet de composition?*, Les Cahiers de Recherche en Education et Formation n.36, Louvain, GIRSEF
- Everson K.C. (2017), Value-Added modeling and educational: are we answering the real questions?, *Review of Educational Research*, 87, n.1, pp.35-70
- Ferrão M.E. (2012), On the stability of value added indicators, *Quality & Quantity*, n.46, pp.627-637
- Goldstein H. (1995), *Multilevel Models in educational and social research: a revised edition*, London, E. Arnold
- Grilli L., Rampichini C. (2009), Multilevel models for the evaluation of educational institutions: a review, in Monari P., Bini M., Piccolo D., Salmasso L. (eds.), *Statistical Methods for the Evaluation of Educational Services and Quality of Products*, Berlin Heidelberg, Springer-Verlag, cap.5

- Hanushek E.A., Raymond M.E. (2003), Improving educational quality: how best to evaluate our schools, in Kodrzycki Y.K. (ed.), *Education in the 21st century: meeting the challenges of a changing world*, Boston-MA, Federal Reserve Bank of Boston, pp.193-224
- Hox J. (2002), *Multilevel analysis. Techniques and applications*, London-Mahwah New Jersey, Lawrence Erlbaum associates
- Kyriakides L., Creemers B.P.M. (2008), A longitudinal study on the stability over time of school and teacher effects on student outcomes, *Oxford Review of Education*, n.34, pp.521-545
- Leckie G., Goldstein H. (2011), Understanding uncertainty in school league tables, *Fiscal studies*, n.32, pp.207-224
- Lissitz R.W. (2005), *Value Added Models in Education: Theory and Application*, Maple Grove-Minnesota, JAM Press
- Lockwood J.R., McCaffrey D.F. (2009), Exploring student-teacher interactions in longitudinal achievement data, *Education Finance and Policy*, n.4, pp.439-467
- Maroy C. (2005), *Vers une régulation post-burocratique des systèmes d'enseignement en Europe*, Les Cahiers de Recherche en Education et Formation n.49, Louvain, GIRSEF
- Martini A. (2014), *La pubblicità dei risultati della valutazione esterna dei livelli di apprendimento raggiunti dalle scuole: pro e contro, alla luce delle esperienze internazionali*, Fondazione Giovanni Agnelli WP n.52, Torino, Fondazione Agnelli
- Martini A. (a cura di) (2018), *L'effetto scuola (valore aggiunto) nelle prove INVALSI 2018*, Roma, Invalsi <<https://bit.ly/2Jre9GS>>
- Martini A. (a cura di) (2019), *L'effetto scuola (valore aggiunto) nelle prove INVALSI 2019*, Roma, Invalsi (in corso di pubblicazione)
- OECD (2008), *Measuring improvements in learning outcomes. Best practices to assess the value-added of schools*, Paris, OECD Publishing
- OECD (2010), *PISA 2009 Results: Overcoming Social Background*, Vol. II, Paris, OECD Publishing
- Papay, J.P. (2011), Different tests, different answers. The stability of teacher value-added estimates across outcome measures, *American Educational Research Journal*, n.48, pp.163-193
- Peterson P.E., West M.R. (eds.) (2003), *No child left behind? The politics and practice of school accountability*, Washington-D.C., Brookings Institution Press
- Raudenbush S.W., Bryk A.S. (2002), *Hierarchical Linear Models. Applications and data analysis methods* (2nd ed.), Thousand Oaks CA, Sage
- Rosenkvist M.A. (2010), *Using student test results for accountability and improvement. A literature review*, OECD EWP n.54, Paris, OECD Publishing
- Rothstein J. (2009), Student sorting and bias in value-added estimation. Selection on observables and unobservables, *Education Finance and Policy*, n.4, pp.537-571
- Rutter M. (1983), School effects on pupil progress-Findings and policy implications, *Child Development*, 54, n.1, pp.1-29
- Sanders W.L. (1998), Value-added assessment, *School Administrator*, n.55, pp.24-77
- Sanders W.L., Horn S.P. (1995), The Tennessee value-added assessment system (TVAAS): mixed methodology in educational assessment, in Shinkfield A.J., Stufflebeam D.L. (eds.), *Teacher evaluation: guide to effective practice*, Boston, Kluwer Academic Publishers, pp.337-376
- Scheerens J. (2018), *Efficacia e inefficacia educativa. Esame critico della knowledge base*, Milano, Springer Healthcare Italia s.r.l.
- Snijders T., Bosker R. (1999), *Multilevel Analysis*, London, Sage
- Teddlie C., Reynolds D. (2000), *The international handbook of school effectiveness research*, NewYork-London, Routledge-Falmer Press
- Thrupp M. (1997), *The School Mix Effect: How the Social Class Composition of School Intakes Shapes School Processes and Student Achievement*, Chicago, paper presented to the AERA 1997 Annual Meeting
- Van de Grift W. (2009), Reliability and validity in measuring the value added of Schools, *School Effectiveness and School Improvement*, n.20, pp.269-285

Angela Martini

angela.martini44@gmail.com

Laureata a Padova in Filosofia e in Psicologia sperimentale, dopo aver lavorato nella scuola come insegnante e dirigente, dal 1999 si occupa stabilmente di ricerca nell'ambito della valutazione oggettiva degli apprendimenti, della valutazione delle scuole, della comparazione e valutazione dei sistemi scolastici e di analisi dei dati delle indagini internazionali sul rendimento scolastico. Da diversi anni collabora con l'Invalsi alla costruzione delle prove per la rilevazione degli apprendimenti e all'analisi dei dati delle rilevazioni. È autrice di numerosi saggi e articoli pubblicati su riviste italiane e straniere.