



BANCA D'ITALIA
EUROSISTEMA

Questioni di Economia e Finanza

(Occasional Papers)

La qualità dell'istruzione italiana: un confronto tra PISA
e le rilevazioni nazionali dell'Invalsi

di Pasqualino Montanaro e Paolo Sestito

Aprile 2014

Numero

218



BANCA D'ITALIA
EUROSISTEMA

Questioni di Economia e Finanza

(Occasional papers)

La qualità dell'istruzione italiana: un confronto tra PISA e le
rilevazioni nazionali dell'Invalsi

di Pasqualino Montanaro e Paolo Sestito

Numero 218 – Aprile 2014

La serie Questioni di economia e finanza ha la finalità di presentare studi e documentazione su aspetti rilevanti per i compiti istituzionali della Banca d'Italia e dell'Eurosistema. Le Questioni di economia e finanza si affiancano ai Temi di discussione volti a fornire contributi originali per la ricerca economica.

La serie comprende lavori realizzati all'interno della Banca, talvolta in collaborazione con l'Eurosistema o con altre Istituzioni. I lavori pubblicati riflettono esclusivamente le opinioni degli autori, senza impegnare la responsabilità delle Istituzioni di appartenenza.

La serie è disponibile online sul sito www.bancaditalia.it.

ISSN 1972-6627 (stampa)

ISSN 1972-6643 (online)

Stampato presso la Divisione Editoria e stampa della Banca d'Italia

LA QUALITÀ DELL'ISTRUZIONE ITALIANA: UN CONFRONTO TRA PISA E LE RILEVAZIONI NAZIONALI DELL'INVALSI

di Pasqualino Montanaro[♥] e Paolo Sestito[♣]

Sommario

L'obiettivo di questo lavoro è di verificare la coerenza del quadro complessivo che emerge da due diverse rilevazioni sugli apprendimenti degli studenti italiani (PISA e RN-Invalsi) svoltesi nello stesso anno (2012). Specialmente concentrando l'attenzione sulla parte in comune delle popolazioni di riferimento (gli studenti 15enni in II secondaria di secondo grado), entrambe le fonti restituiscono un quadro simile dei differenziali di performance esistenti tra gli studenti, sia in connessione con il diverso background familiare sia con riferimento al pattern territoriale. Simile è anche il quadro relativo alla variabilità complessiva, sebbene criticità da questo punto di vista si riscontrino nelle RN, per via della presenza di una significativa presenza di fenomeni di *cheating* e di una imperfetta correzione dello stesso, che genera una ulteriore fonte di variabilità nei risultati. La coerenza tra le due fonti si estende ai risultati dei singoli studenti e soprattutto delle singole scuole. Tali evidenze costituiscono un buon viatico per un uso – e per un ridisegno – congiunto delle due rilevazioni. In particolare, esse suggeriscono l'opportunità di un disegno *ex ante* di meccanismi di ancoraggio dei risultati delle RN nella metrica fornita da PISA, con il vantaggio di consentire comparazioni nel tempo e nello spazio, rispetto agli altri paesi OCSE.

Classificazione JEL: I20, I21.

Parole chiave: scuola, rilevazioni sugli apprendimenti.

Indice

1. Introduzione	5
2. Una descrizione delle due popolazioni	7
3. Le caratteristiche delle due rilevazioni	9
4. Un confronto tra gli studenti “core” delle due popolazioni	14
5. Il confronto a livello di scuola e individuale.....	18
6. Conclusioni	20
Bibliografia	21
Appendice statistica	22

[♥] Banca d'Italia, Ufficio Analisi e ricerca economica, Sede di Ancona.

[♣] Banca d'Italia, Servizio Struttura Economica.

1 Introduzione¹

PISA è una rilevazione campionaria a frequenza triennale riferita agli studenti 15enni, indipendentemente dal grado scolastico perseguito, laddove le rilevazioni nazionali sugli apprendimenti condotte dall'Invalsi (d'ora in poi RN) sono annualmente rivolte all'universo degli studenti di un certo grado scolastico. Ne discendono differenze anche nella struttura delle due indagini: l'una (PISA) basata su un più ampio set di quesiti, a rotazione proposti ai diversi singoli partecipanti, e svoltasi in condizioni maggiormente controllate e con un maggior numero di domande aperte, grazie alla presenza di somministratori appositamente retribuiti; l'altra (RN) basata su un più ristretto set di quesiti egualmente posti a tutti i singoli partecipanti, con un minor numero di domande aperte e controllori esterni solo in sottoinsieme di classi.

L'obiettivo di questo lavoro è di verificare se il quadro complessivo che emerge dalle due rilevazioni è coerente. Pur attesa la diversa finalità dei due esercizi – l'uno di confronto del sistema Italia col resto del mondo e l'altro di feedback alla singola scuola sulla propria *performance* – è importante capire se le rispettive indicazioni sui differenziali di *performance* all'interno dell'Italia, ad esempio sul piano geografico o su quello dei differenziali tra diverse categorie di studenti, sono tali da consentirne una lettura integrata, che meglio sfrutti le diverse caratteristiche delle due indagini. Se la risposta a tale quesito fosse positiva, allora potrebbe essere vantaggioso, per avere un monitoraggio più fine degli andamenti interni all'Italia, sfruttare la maggiore copertura delle RN e la loro cadenza annuale. Al tempo stesso, sarebbe in tal caso possibile ancorare le RN, che al momento non sono nei loro livelli assoluti confrontabili da un anno all'altro e con il resto del mondo, alla metrica di PISA e delle altre rilevazioni internazionali che si svolgono in altri gradi scolastici (nello specifico, TIMSS e PIRLS), metrica che, per come è costruita, consente confronti con gli altri paesi (e singole regioni di altri paesi) e nel tempo.

Naturalmente, un esplicito ancoraggio delle RN a PISA (o in maniera simile alle altre indagini internazionali) richiede un esplicito intervento sul disegno delle RN, questione che non è oggetto specifico di questo lavoro.² L'esame qui compiuto è tuttavia da ritenere preliminare a tale sforzo, perché

¹ Alcune parti di questo lavoro sono confluite nel *Rapporto nazionale sulle prove PISA 2012* predisposto da Invalsi (2013). Le opinioni qui espresse peraltro impegnano solo gli autori e non riflettono necessariamente quelle dell'Invalsi o dell'istituzione di appartenenza. Gli autori ringraziano, tra gli altri, Roberto Ricci (Invalsi) per gli utili commenti e chiarimenti, Patrizia Falzetti e Paola Giangiacomo (sempre dell'Invalsi) per la predisposizione dei dati, Maria Letizia Cingoli (Banca d'Italia) per l'assistenza prestata nella predisposizione di tavole e figure.

² Le tecniche di ancoraggio di un'indagine all'altra – o nel tempo tra diverse edizioni della stessa rilevazione – passano per la proposizione di un sottoinsieme di domande ripetute a un sottoinsieme di partecipanti a ciascuna delle due iniziative. In PISA ciò avviene mantenendo segrete, e replicandole da un'edizione triennale all'altra, alcune domande. Nell'ancoraggio effettuato da Invalsi all'interno delle RN – al momento attuato per la quinta primaria – si adoperava invece un metodo indiretto, non esistendo per definizione un set di domande segrete (le domande di ciascuna annualità delle RN sono immediatamente rese pubbliche): a un sottoinsieme di alunni che partecipano alla normale edizione della RN si

consente di capire se l'operazione in quanto tale abbia o meno un suo senso e sotto quali condizioni essa possa essere fatta. L'analisi qui effettuata ha anche natura propedeutica rispetto al disegno delle modalità italiane di partecipazione alle indagini internazionali, perché un uso integrato di RN e indagini internazionali potrebbe consentire di risparmiare sui costi, piuttosto elevati, di ampliamento del campione oggetto di queste ultime, ampliamento a suo tempo effettuato proprio per avere informazioni di dettaglio regionale (un grado di dettaglio oggi ormai disponibile a cadenza annuale nelle RN).

Più in particolare, nel presente lavoro il confronto verterà su PISA e sulle RN condotte nella II classe della secondaria di secondo grado nello stesso anno scolastico (2011-12). Quanto a PISA si considereranno solo lettura e matematica, trascurando le scienze (e alcune altre competenze pure testate, su sottoinsiemi della popolazione, nel 2012), in modo da poter puntualmente considerare la corrispondenza con le prove rispettivamente di italiano e matematica contenute nelle RN. Il confronto verrà effettuato sia in termini di *pattern* generale dei risultati e delle sue principali regolarità (divari tra regioni, importanza di varie caratteristiche degli studenti etc.) sia considerando, per un sottoinsieme di casi per cui ciò è stato possibile, la concordanza di risultati a livello di singola scuola (per l'esattezza si tratta del singolo plesso scolastico, potendo più plessi scolastici – anche riferiti a indirizzi di studio diversi – appartenere alla stessa scuola, intesa come unità amministrativa) e di singolo studente. Il *pattern* generale e il confronto a livello di singole scuole e individui terranno naturalmente conto del fatto che le popolazioni di riferimento delle due indagini sono diverse e che solo un sottoinsieme dei 15enni (oggetto di PISA) è rappresentato da studenti della II superiore (studenti che includono anche alunni anticipatori e posticipatori, rispettivamente con meno o più di 15 anni).

Questa diversa composizione delle popolazioni oggetto delle RN e di PISA – e in particolare il fatto che quest'ultima solo in circa tre quarti dei casi si riferisce alla II superiore – è il punto di partenza del confronto qui operato (par. 2). La diversa composizione della popolazione è per molti versi la differenza più facilmente isolabile, perché ci si può concentrare sul sottoinsieme di studenti in comune tra le due rilevazioni (come poi verrà fatto nell'analisi successiva). Nel par. 3 si esporranno invece le differenze nel format delle due indagini, differenze i cui possibili impatti non sono facilmente identificabili ed eliminabili, ma che pure devono essere tenute a mente nell'interpretare i risultati del confronto.

Il confronto vero e proprio è quindi esposto nei paragrafi 4 e 5. Prima si considera il *pattern* generale della variabilità interna alle due indagini (nelle popolazioni proprie di ciascuna indagine e concentrandosi sul *core* di studenti in comune tra le due indagini, ovverosia i 15enni che siano in II superiore): si guarda alla variabilità *tout court* e alla variabilità spiegata da un set di caratteristiche (inclusa

sottopongono anche domande (mantenute segrete) proposte a un simile sottoinsieme di alunni nell'edizione dell'anno precedente e/o di quello successivo della RN.

la regione) sulla base di alcuni semplici modelli di regressione multivariata. Poi si considera il confronto tra scuole e tra singoli studenti che hanno partecipato ad ambedue le rilevazioni, confronto possibile solo però per quelle scuole e quegli studenti che, appartenendo al *core* di entrambe le indagini, si è riusciti ad abbinare puntualmente, sulla base dei rispettivi codici identificativi.

2 Una descrizione delle due popolazioni

In PISA 2012 solo tre 15enni su quattro risultavano frequentare la II classe della scuola secondaria di secondo grado. Tra i restanti, vi sono da un lato i pochi anticipatori (2,6 per cento) che, frequentando già la III secondaria di secondo grado, hanno accumulato un anno in più di scolarità, dall'altro i posticipatori che sono ancora nella I classe della secondaria di secondo grado (14,6 per cento) o addirittura nella secondaria di primo grado (2,1 per cento); infine, il 4,9 per cento frequenta il primo o il secondo anno dei percorsi di formazione professionale. La quota complessiva di 15enni “in ritardo” è del 18,9 per cento circa (tavv. a1 e a3). L'incidenza dei 15enni in ritardo negli studi è più elevata tra i maschi; tra chi frequenta i Tecnici e, soprattutto, gli Istituti professionali; tra gli immigrati (soprattutto se di prima generazione, per i quali il numero di studenti in ritardo supera quello di chi è regolare); nelle regioni del Centro Nord. Gli studenti “anticipatori” sono invece pressoché equamente suddivisi tra maschi e femmine, frequentano in gran parte i Licei, sono quasi tutti nativi e più numerosi nel Mezzogiorno (tav. a1). In Sicilia la quota di 15enni in III secondaria di secondo grado raggiunge l'8,2 per cento, in Campania il 7,5 per cento, in Calabria il 4,3 per cento; essa è invece assai bassa nelle regioni settentrionali (tav. a3).

In linea con le attese, i tre gruppi di studenti (15enni in II, che d'ora in poi chiameremo *core*; in ritardo con gli studi; in anticipo con gli studi) si distinguono nettamente sia in termini di *background* familiare sia in termini di qualità degli apprendimenti. Dal primo punto di vista, l'indicatore sintetico delle condizioni socio-economiche e culturali della famiglia di origine (ESCS) è, infatti, molto più elevato – a parità di altre caratteristiche individuali – per gli studenti in anticipo con gli studi (cioè i 15enni che frequentano la III secondaria di secondo grado) rispetto ai “regolari” e, soprattutto, ai posticipatori (tavv. a1 e a3).

Quanto ai risultati, il *pattern* generale che emerge è quello di una differenza significativa tra gli studenti in anticipo con gli studi e quelli regolari (5-6 punti percentuali, a seconda della materia, lettura o matematica) e tra questi e gli studenti in ritardo (14-15 punti percentuali). Tali differenze, più marcate tra i maschi, maturano però soprattutto nei Licei e tra i nativi; a livello di area geografica, emerge che i differenziali Nord-Sud tra gli anticipatori non sono dissimili da quelli che emergono tra i “regolari” o tra gli studenti in ritardo (tav. a2).

Le evidenze che emergono dalle RN non sono, nel loro complesso, dissimili. Il 76,6 per cento degli studenti testati risulta “regolare”, ossia ha 15 anni di età e frequenta la II secondaria di secondo grado (si tratta quindi del *core* comune a PISA); il 18,7 per cento risulta posticipatario, il 4,7 per cento anticipatario. Anche in questo caso, l’incidenza degli studenti in ritardo negli studi è più elevata tra i maschi; tra chi frequenta i Tecnici e gli Istituti professionali; tra gli immigrati; nelle regioni del Centro Nord. Gli studenti “anticipatari”, invece, frequentano in gran parte i Licei, sono quasi tutti nativi e sono più numerosi nel Mezzogiorno (tav. a4). Nelle RN, la quota di anticipatari raggiunge il 12,8 per cento in Campania, il 9,7 per cento in Sicilia e il 9,3 per cento in Calabria (tav. a5).

A differenza di PISA 2012, i risultati conseguiti dagli studenti anticipatari non sono migliori nel confronto con i “regolari”, anzi sono di poco inferiori, nel complesso nazionale. Ma ciò non sorprende, se si pensa che in RN i primi sono più giovani di un anno rispetto ai secondi, quindi presentano un livello di maturità inferiore, mentre in PISA gli studenti “in anticipo” a parità di età, che si tratta sempre di 15enni, frequentano una classe più avanzata. Detto questo, essi provengono comunque da famiglie avvantaggiate, in termini di condizioni socio-economiche e culturali (ESCS). La *performance* degli studenti anticipatari rispetto ai regolari, nel complesso nazionale, risente tuttavia molto della differente incidenza del fenomeno a livello territoriale. Valutando regione per regione, infatti, la *performance* dei primi rispetto ai secondi appare comunque tendenzialmente migliore, in matematica più che in italiano, anche se il divario è nettamente più contenuto rispetto a quanto emerge in PISA 2012. Il vantaggio derivante dall’anticipare di un anno il percorso scolastico svanisce poi nella media nazionale, poiché tale fenomeno è, come detto, assai più diffuso nel Mezzogiorno, dove i livelli di apprendimento sono comunque più bassi. E per gli anticipatari meridionali, i punteggi riportati sono inferiori rispetto a quelli riportati al Nord dagli studenti regolari; in qualche caso, sono inferiori o comunque non dissimili da quelli che al Nord riportano gli studenti posticipatari (tav. a5). I divari tra regolari e posticipatari sono invece ampi, in tutte le regioni. In termini di punti percentuali, tale differenza è più pronunciata in matematica, ancora di più rispetto a quanto emerge in PISA 2012 per la stessa materia (tavv. a2-a3); in italiano, invece, essa è meno marcata, inferiore a quella che segnala PISA 2012 in *reading literacy*.

Una componente decisiva nello spiegare la variabilità dei risultati tra gli studenti è il *background* familiare (ESCS). Tale variabile è strettamente (e positivamente) correlata con i livelli degli apprendimenti, per ognuno dei tre segmenti della popolazione studentesca considerati (regolari, posticipatari, anticipatari). Tale relazione differisce però di intensità tra le due rilevazioni e a seconda della materia. In particolare, ripartendo in decili la distribuzione degli studenti, i risultati di RN ci restituiscono un quadro nel quale: *i*) per gli studenti “regolari” la relazione diviene decrescente all’aumentare dell’ESCS in italiano, mentre è crescente in matematica; *ii*) all’aumentare dell’ESCS, i divari tra studenti regolari e posticipatari si riducono in italiano e si ampliano in matematica. Il quadro è invece

più univoco in PISA 2012, dove la performance degli studenti cresce in maniera più o meno regolare all'aumentare dell'ESCS, e con eguale intensità, per studenti regolari e studenti in ritardo, senza significative differenze tra le due materie considerate (fig. a1).

3 Le caratteristiche delle due rilevazioni

Le prove PISA 2012 e RN presentano numerose analogie, ma anche sostanziali differenze. Vale la pena soffermarsi sulle une e sulle altre. Gli aspetti sui quali ragionare attengono ai **contenuti** e al **disegno** della rilevazione, all'**assemblaggio** e alla **composizione** dei fascicoli, alla **natura stessa dei punteggi** restituiti, per ogni studente, dalle prove, alla **costruzione** delle singole domande, alle **modalità di conduzione e correzione** dei test.

I contenuti. – Ma andiamo con ordine. Innanzitutto, a differire sono i contenuti. Come riportato nel *Rapporto tecnico sulle rilevazioni nazionali* dell'Invalsi, i Quadri di Riferimento (QdR) Invalsi di italiano e matematica si ispirano ai framework delle prove usate nelle indagini internazionali sugli apprendimenti; tuttavia, vi sono differenze sulle quali è opportuno soffermarsi, la prima delle quali è che se in PISA i QdR sono fondati sulle competenze e slegati, per definizione, dai *curricula* nazionali, nelle RN sono connessi alle Indicazioni nazionali e alle Linee Guida.

PISA verte su tre ambiti disciplinari (comprensione della lettura di testi scritti, matematica e scienze), mentre le RN riguardano solo i primi due ambiti. Ma anche nelle aree di valutazione comuni, i contenuti delle due prove sono in parte diversi. Nelle RN la prova di italiano concerne non solo la comprensione del testo, ma anche le competenze grammaticali (ortografia, morfologia, formazione delle parole, lessico e semantica, sintassi, testualità), con ulteriori differenze nella definizione degli aspetti più specifici che la prova intende misurare. Nella comprensione del testo, ad esempio, le prove Invalsi testano la capacità di comprendere il significato di parole ed espressioni e la capacità di cogliere relazioni di coerenza e coesione testuale, che sono aspetti strettamente legati alla semantica e alle strutture sintattiche e testuali di una particolare lingua (l'italiano, in questo caso), per definizione assenti in una prova internazionale.

Per quanto riguarda la prova di matematica, le aree di contenuto sostanzialmente si sovrappongono (conoscenza, applicazione, ragionamento nella formulazione matematica di situazioni, nell'utilizzazione di concetti, fatti, procedure e ragionamenti matematici, nell'interpretazione, applicazione e valutazione di risultati matematici), con la differenza che gli argomenti specifici sono nelle RN collegati con quelli curriculari.

Il disegno della rilevazione. – Il disegno delle due rilevazioni presenta anch'esso delle sostanziali differenze, legate alle loro diverse finalità. Mentre PISA si pone l'obiettivo di informazioni ricche e

approfondite circa il funzionamento di un sistema scolastico nel suo complesso, le RN aggiungono a questa finalità quella di restituire a ciascuna scuola una serie d'informazioni sul loro funzionamento in termini di apprendimenti prodotti, anche in un'ottica comparativa. Ne discende che PISA segue un disegno campionario (sia a livello di scuola sia a livello di studenti, circa 40 per scuola), mentre le RN sono di natura censuaria. Il fatto che PISA sia una rilevazione campionaria, a livello di scuola e di studenti all'interno delle scuole, implica che i risultati medi di scuola siano per definizione affetti da un errore campionario; inoltre, in PISA vengono messi assieme gli studenti provenienti da più classi della stessa scuola e non è possibile esaminare quanto la variabilità all'interno di una singola scuola si scomponga in una variabilità tra classi e in una variabilità (media) all'interno di ciascuna classe.

La composizione dei fascicoli. – Le differenze nel disegno di rilevazione, unite alla sua periodicità triennale (le RN essendo a cadenza annuale) consentono inoltre di proporre in PISA prove basate sul principio della rotazione delle domande, in modo che, nel complesso, possano essere somministrate agli studenti più domande, senza che ciascuno le debba affrontare tutte. Una rilevazione nazionale, invece, richiede che le prove siano uguali per tutti gli allievi e che siano rese pubbliche immediatamente dopo la loro somministrazione: per questo motivo, esse sono basate necessariamente su un numero minore di quesiti. E qui veniamo agli aspetti specifici riguardanti l'assemblaggio e la composizione dei fascicoli delle prove, che sono poi quelli che a noi più interessano ai fini del nostro lavoro.

La maggiore differenza tra PISA e RN riguarda, infatti, più che i contenuti, la procedura di costruzione dei fascicoli di prova, anch'essa connessa con il diverso carattere, campionario in un caso e censuario nell'altro, delle rilevazioni. I fascicoli di prova delle indagini internazionali sugli apprendimenti sono costruiti partendo da insiemi di item che coprono l'intero spettro di contenuti e processi da misurare. Tali item vengono poi suddivisi in blocchi o sottogruppi, equivalenti per gamma di difficoltà delle domande. I blocchi vengono quindi casualmente combinati tra loro in modo da ottenere diversi fascicoli di prova (o *booklets*) con lo stesso numero totale di domande (solitamente 13 fascicoli in PISA).³ In questo modo, i fascicoli di prova utilizzati in ogni tornata delle indagini differiscono da studente a studente, anche se il livello di difficoltà complessivo di ogni fascicolo è analogo. Inoltre, ciascun blocco di domande all'interno dei fascicoli viene assegnato a uno stesso numero di studenti. I fascicoli delle RN sono invece unici per tutti gli studenti e con una struttura organizzativa che si riproduce in maniera uniforme da una rilevazione all'altra per lo stesso livello scolastico.⁴

In estrema sintesi, PISA ha più fascicoli, mentre le RN ne hanno solo uno: in PISA possono esserci domande di "eguale difficoltà" che poste a ragazzi di "eguale abilità" portano a punteggi diversi,

³ Non è dunque l'assegnazione delle domande a ciascuno studente che è *casuale*, bensì quella dei blocchi.

⁴ I singoli quesiti vengono disposti nel fascicolo non in ordine (crescente o decrescente) di difficoltà, per evitare che gli effetti della "fatica" di rispondere da parte degli alunni si concentrino solo su determinati quesiti e testi.

perché le domande sono diverse, laddove nelle RN questo non può avvenire *per definizione*. D'altro canto, con un solo fascicolo, le RN descrivono un certo livello di abilità con minore precisione: dovendo fare affidamento a poche domande, per identificare un certo livello di abilità, i risultati potrebbero risultare poco precisi, laddove PISA su quel livello di abilità (in fascicoli diversi e quindi su ragazzi diversi) utilizza più domande. In teoria, un ragazzo di una certa abilità nelle RN potrebbe avere dunque risultati PISA anche molto diversi.

I Plausible Values. – Il fatto che in PISA i fascicoli di prova differiscano da studente a studente induce anche ad utilizzare, per ciascuno studente, non un solo punteggio complessivo, bensì un insieme di 5 *Plausible Values* (PV). I PV rappresentano “estrazioni casuali” da una distribuzione empirica (derivata dal test) dei livelli di *proficiency*, condizionati ai valori osservati nelle risposte ai quesiti (Wu and Adams, 2002). In altre parole, i PV sono stime del livello di abilità: invece di stimare *direttamente* il livello di abilità dello studente θ , si stima una distribuzione di probabilità di θ (*posterior distribution*; Wu, 2004), ossia un range di suoi possibili valori, ognuno dei quali con un associato livello di verosimiglianza.

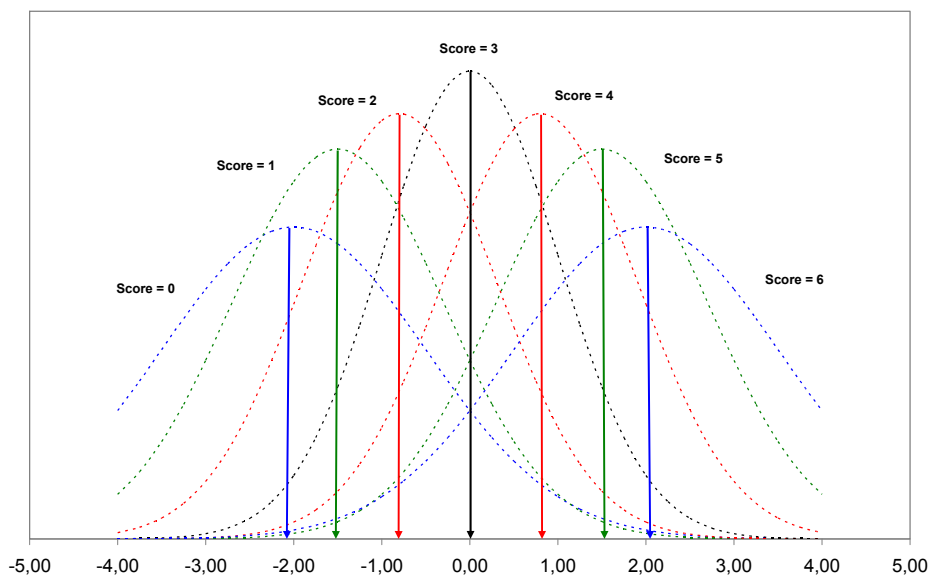
I PV per uno studente con *item response pattern* x sono allora “estrazioni casuali” (*random draws*) da una distribuzione di probabilità con densità $h(\theta|x)$: essi forniscono così non solo informazioni sull'abilità stimata di uno studente, ma anche l'incertezza associata a questa stima, incertezza che a sua volta discende dal fatto che gli studenti rispondono a domande diverse.

Come gestire i PV al fine di ottenere una stima delle caratteristiche della popolazione di una certa abilità Θ ? Consideriamo due possibili stime del parametro $\hat{\theta}_n$ di ciascuno studente: una *Maximum Likelihood Estimate* (MLE) e una stima *Expected A-Posteriori* (EAP), quest'ultima non essendo altro, in definitiva, che la media aritmetica dei 5 PV, soluzione peraltro raccomandata dall'OCSE per l'elaborazione dei risultati PISA (OCSE, 2009). Si può dimostrare (Wu, 2004) che se i $\hat{\theta}_n$ sono MLE, allora la loro media è una stima non distorta della media della popolazione μ , ma la loro varianza è una *sovrastima* di σ^2 , ossia la varianza della popolazione. Di contro, se i $\hat{\theta}_n$ sono EAP, la loro media è anch'essa una stima non distorta della media della popolazione μ , ma la loro varianza è una *sottostima* di σ^2 . Questi *bias* si attenuano non con incrementi della numerosità campionaria, ma solo con un aumento del numero di domande sottoposte agli studenti.

Facciamo un semplice esempio che aiuti a capire meglio questo punto. Poniamo che vi sia un test con 7 domande, nel quale uno studente può conseguire un punteggio che va da 0 a 6. In termini grafici, tale situazione può bene essere espressa dalla fig. 1, nella quale sono riportate le distribuzioni di probabilità (pesate) ex post $h(\theta|x)$ (*posterior distributions*) e le corrispondenti stime EAP (linee verticali),

che altro non sono che i valori medi dei PV, *casualmente estratti* da ognuna delle sette (teoriche) distribuzioni riportate in figura.

Figura 1: Distribuzioni di probabilità pesate ex post dell'abilità di uno studente



Fonte: tratto da Wu, 2004.

Calcolare, per ogni studente, la media dei PV (EAP; scelta standard) equivale a prendere il valore corrispondente a una delle sette linee verticali nella figura; considerare invece un solo PV significa prendere *casualmente* uno dei possibili valori ricompresi in ognuna delle distribuzioni.

A differenza di PISA, nelle RN non vengono calcolati i PV, ma viene fornita, come *raw score*, la percentuale di risposte esatte. Rispetto a PISA vi sono due differenze rilevanti. La prima è che non si tiene conto – almeno nei dati riferiti all’universo censuario che qui adopereremo⁵ – del fatto che le singole domande hanno un diverso grado di difficoltà. La seconda è che si considera direttamente un solo valore e non una combinazione di più valori plausibili: essendo le singole domande invariante da uno studente all’altro, le ragioni di tale pluralità di valori vengono quindi a mancare. All’atto pratico, la mancata ponderazione per la difficoltà delle domande dovrebbe produrre un possibile *rounding* dei risultati degli studenti, con addensamenti artificiali in prossimità dei valori i/n , con $i=1, \dots, n$, essendo n il numero di domande esistenti; l’uso di più PV, invece, dovrebbe rendere le misurazioni in PISA – pur attesa la presenza di differenze tra le domande riferite ai singoli studenti – meno affette dall’effetto di aver considerato una specifica domanda anziché un’altra.

⁵ Dal 2012 i Rapporti nazionali sulle RN, diffusi nel mese di luglio dall’Invalsi e basati sulle sole classi campione, riportano i risultati nella metrica di Rasch, ponendo la media dell’Italia pari a 200 e la standard deviation pari a 40, in perfetto parallelismo con PISA che ha una media internazionale, nell’anno base, pari a 500 e una standard deviation pari a 100. I risultati per tutte le singole scuole – a queste poi restituiti – al momento sono invece espressi in termini di percentuale di risposte corrette, tale metrica essendo di più immediata e universale comprensione.

La costruzione delle singole domande. – Nella costruzione dei singoli quesiti da sottoporre agli studenti, PISA e RN seguono invece sostanzialmente la stessa linea. Rimandando per i dettagli ai rapporti tecnici delle due rilevazioni, ci limitiamo qui a segnalare che in entrambe le rilevazioni si procede nel modo seguente:

1. costruzione delle domande da parte degli autori;
2. revisione qualitativa degli esperti;
3. pre-test;
4. analisi dei risultati dell'intera prova con la teoria classica dei test (calcolo dell'*alpha* di Cronbach e del coefficiente di correlazione biseriale per ogni distrattore) e della prova e delle singole domande mediante *Item Response Theory* (IRT), per testare la "corrispondenza" tra la difficoltà delle domande e dell'intera prova e l'abilità dei rispondenti;
5. modifica e nuovo pre-test per le domande che presentano problemi, alla luce dell'analisi dei risultati effettuata al punto precedente.

Entrambe le prove sfruttano soprattutto domande a scelta multipla con 4 alternative di risposta. Nel formularle, in entrambe si seguono gli stessi criteri: le quattro alternative (una risposta corretta e tre *distrattori*, che debbono risultare comunque plausibili; Osterlind, 1998) vanno presentate in modo da evitare che lo studente arrivi alla soluzione corretta per approssimazioni successive, assicurando così una certa *robustezza* psicometrica. I distrattori, pertanto, non devono essere troppo vicini alla risposta corretta ma comunque sufficientemente attrattivi e plausibili, e non devono essere costruiti per trarre in inganno il rispondente. Le opzioni di risposta debbono in definitiva saper discriminare tra chi padroneggia un certo tipo di abilità (o "tratto latente") e chi lo padroneggia meno.⁶

Differenze vi sono però nel peso delle domande a risposta aperta, che sono circa il 40 per cento dei quesiti in PISA e solo il 20-25 per cento nelle RN. La differenza è legata al fatto che la correzione in PISA è maggiormente effettuata in maniera controllata dal centro, il che consente per l'appunto di sfruttare questi quesiti con maggiore capacità "discriminatoria".

La conduzione e la correzione dei test. – Le due rilevazioni differiscono, infine, nelle modalità di conduzione e correzione del test. Mentre il test PISA viene vigilato e poi corretto da somministratori appositamente retribuiti, le RN sfruttano osservatori esterni solo per un sottoinsieme di classi "campione". Queste diverse modalità di conduzione della rilevazione hanno implicazioni per la precisione e l'affidabilità dei risultati, inferiori nel caso delle RN, specie ove non si tratti di classi "campione", ove la presenza di un osservatore esterno riduce alla radice i fenomeni di *cheating*, che in

⁶ In altre parole, se la risposta giusta venisse sempre preferita rispetto ai distrattori, a qualunque livello di abilità, questa indicherebbe, secondo le regole dell'IRT, che i distrattori non funzionano e che vanno riformulati o sostituiti.

quanto tali tendono a innalzare i risultati medi e ridurre la variabilità interna alla classe ove il fenomeno si realizzi.⁷ Nel prosieguo dell'analisi si farà riferimento in prevalenza a dati corretti dalla presenza di *cheating*, tramite una procedura statistica ad hoc (cfr. Quintano, Castellano e Longobardi, 2009). La procedura, pur essendo precisa in media, è però tutt'altro che priva di difetti nel singolo caso, il che introduce una componente aggiuntiva di *noise*. Essa viene inoltre posta in essere a livello dell'intera classe, col plausibile risultato di mantenere indebitamente compressa la varianza dei risultati all'interno della singola classe ove il *cheating* sia stato presente e di creare una fonte aggiuntiva di variabilità tra classi e tra scuole. Va anche detto che la procedura è stata migliorata dall'Invalsi a partire dal 2013, ma i miglioramenti non sono ancora stati applicati ai dati 2012 qui adoperati. Affronteremo più direttamente questa questione nel prossimo paragrafo.

4 Un confronto tra gli studenti “core” delle due popolazioni

Nella sezione 2 abbiamo descritto le due diverse popolazioni, evidenziandone le differenze. In questo paragrafo si porranno a confronto le due fonti, tenendo conto di tali diversità nella composizione della popolazione di riferimento. Ciò può essere fatto parametricamente, all'interno di un modello in cui la composizione della popolazione di ciascuna fonte sia tra i fattori di controllo dei risultati, oppure concentrando l'attenzione sul gruppo di studenti in comune tra le due fonti, quelli che abbiamo definito come *core* delle due indagini, pari a circa i $\frac{3}{4}$ del totale in ciascuna di esse. L'obiettivo dell'esercizio è di capire se e quanto i *pattern* del sistema italiano restituiti dalle due indagini siano coerenti tra loro: in quest'ottica, si guarderà, nell'ordine, alla variabilità complessiva (e sua scomposizione) dei risultati di PISA e RN e ai *pattern* rilevabili tramite un semplice esercizio di regressione.

4.1 La variabilità dei risultati

In PISA la varianza dei punteggi è complessivamente più bassa rispetto a RN, il coefficiente di variazione essendo pari al 18 per cento (della media) nel primo caso, al 28 per cento (in italiano) e addirittura al 37 per cento (in matematica) nel secondo. Considerare i soli studenti *core* elimina una fonte di variabilità dei risultati e, come atteso si riduce il coefficiente di variazione. Tale riduzione risulta peraltro più pronunciata in PISA che in RN, senza perciò che si riduca più di tanto la discrepanza tra le due fonti (figg. a2-a3 e tav. a6).

Quanto alla scomposizione della varianza, è innanzitutto da notare che in RN, ove è possibile individuare sia la scuola che i singoli plessi scolastici (questi essendo invece l'unica unità di base identificabile in PISA), il passaggio dalla prima ai secondi tende ovviamente ad accrescere la quota della

⁷ Per un'analisi generale del fenomeno si rimanda a Paccagnella-Sestito (2014).

varianza *between* (che ora include anche differenze tra plessi diversi della stessa scuola, che tradizionalmente nei rapporti resi noti dall'Invalsi sono accorpate invece alle differenze tra classi all'interno della stessa scuola). Concentrandosi sulla scomposizione della varianza totale *tra* plessi e *all'interno* dei singoli plessi, assicurando così piena comparabilità al *pattern* di RN e PISA, si evidenzia come la componente *between* prevalga in PISA e quella *within* in RN. Il peso di quest'ultima componente cresce in entrambi i casi quando ci si concentra sul solo gruppo *core*, in comune tra le due indagini, ma la differenza in questo aspetto tra le due indagini non si riduce, ma semmai si accentua.

Nella tav. a7, concentrando l'attenzione sui soli studenti *core* in comune tra le due fonti, si dettaglia il confronto a livello di singole aree geografiche. In entrambe le fonti, e ferme restando le differenze generali tra le stesse prima esposte, si evidenzia una variabilità complessiva dei risultati più elevata nel Mezzogiorno. In PISA la quota di varianza *within* risulta più bassa nel Mezzogiorno rispetto al Centro Nord, mentre in RN ciò è vero per l'italiano ma non per la matematica.

La differenza *tra* plessi scolastici è in parte legata alle differenze tra le diverse tipologie di scuola (Licei, Tecnici e Istituti professionali, nel nostro caso). Ciò che differisce maggiormente tra PISA e RN è proprio quest'ultima, molto più elevata (come quota del totale) in PISA che nelle RN. La componente "strettamente" *between* plessi dello stesso indirizzo è molto più simile tra le due rilevazioni, intorno al 30 per cento della varianza complessiva, e generalmente più elevata nel Mezzogiorno (tav. a7).

Vari sono i fattori che possono portare a questi differenti *pattern*. Per meglio comprenderne le ragioni, si è provato ad annullare taluni dei diversi fattori (descritti nel precedente paragrafo) che potrebbero contribuirvi. Un primo esercizio prova a vedere se vi possa influire la natura campionaria di PISA all'interno di ciascuna unità scolastica, che rende la stima del dato medio di scuola plausibilmente più ballerina, visto che si considerano solo sino a circa 40 studenti. In particolare si è provato a estrarre casualmente da ciascuna scuola presente in RN (quelle, ovviamente, con più di 40 studenti) un sample di soli 40 studenti (come in PISA). I risultati ottenuti (non riportati per brevità ma disponibili presso gli autori) non sono, tuttavia, molto diversi dalla considerazione di RN in quanto tale: la variabilità complessiva dei risultati non muta e anche la quota di varianza *between* rimane costante (aumenta in matematica ma addirittura cala in italiano).

Lungo le attese, ma poco dirimente, è anche la considerazione, nel caso di PISA, dei risultati che si otterrebbero se si usasse solo uno dei 5 *plausible values*, anziché la media di tutti e cinque: la variabilità totale aumenta seppur di poco (come ci si attenderebbe in base al fatto che la considerazione di 1 solo PV comporta un maggior *noise* nella stima del risultato di ciascun singolo studente) e si riduce la quota di variabilità tra plessi, a favore di quella all'interno di ciascun plesso (tav. a8). Va aggiunto, per completezza di informazione – anche in questo caso non si riportano i risultati per brevità – che l'ampiezza dei PV di ogni singolo studente, ossia la loro varianza, differisce a seconda del livello di

abilità dello studente: in particolare, al crescere del punteggio *medio* (EAP, ossia la media dei PV) diminuisce la varianza dei PV. In altre parole, la stima del livello di abilità dello studente diviene più precisa man mano che si procede verso punteggi più elevati.

Più interessante ai nostri fini è però il caso in cui si considerino, per le RN, le sole classi “campione”, in alternativa all’intera popolazione studentesca. Il concentrare l’attenzione alle sole classi campione consente di ridurre gli effetti della presenza di *cheating* e delle procedure di correzione dello stesso adoperate nelle RN. Come già detto, le procedure di correzione del *cheating* adottate dall’Invalsi, benché efficaci in media, possono risultare imprecise nel singolo caso: esse intervengono a livello dell’intera classe, su base statistica e quindi con possibilità di errore⁸, senza differenziare, all’interno della classe, tra i singoli alunni, per cui la variabilità interna alla classe – che il *cheating* plausibilmente e artificialmente riduce, in quanto avvicina gli studenti meno bravi a quelli più bravi – non viene mutata per effetto della correzione. Restringendo l’attenzione alle sole classi “campione” (tav. a7, ultima riga), il *pattern* della variabilità diviene assai più omogeneo tra le due fonti. Soprattutto per l’italiano, i risultati stimati delle RN presentano una variabilità complessiva inferiore e un peso della componente *tra plessi scolastici* maggiore, avvicinando il pattern restituito dai risultati di PISA. In altri termini, la presenza del *cheating*, e la inevitabilmente imperfetta correzione operata per tener conto della sua presenza⁹, sembrano essere la *magna pars* delle apparenti diversità tra PISA e RN nel pattern della variabilità complessiva dei risultati.

4.2 Un’analisi multivariata

Tra le due rilevazioni, quanto differisce il *pattern* generale che ci restituisce l’analisi multivariata delle determinanti della *performance*? Consideriamo qui i punteggi in PISA 2012 e RN, espressi in entrambi i casi in valori percentuali, confrontandoli per materia (*reading literacy*/italiano e matematica) e considerando le stesse (o analoghe) variabili esplicative. In questo semplice esercizio, si assume che i livelli di apprendimento dello studente si spieghino con il tipo di scuola frequentata, la propria condizione scolastica (se regolare, anticipatori o posticipatorio), il genere, lo status di cittadinanza, il *background* familiare proprio e quello medio della scuola frequentata, l’area geografica (macroarea o regione, alternativamente).

Nel complesso, i risultati (da interpretare sia in PISA 2012 sia in RN come differenziali in punti percentuali rispetto alle variabili di confronto, quindi su una medesima scala) sono del tutto in linea con quelli attesi (tavv. a9-a10). Ma quello che qui più ci interessa non è tanto il *pattern* generale, quanto il

⁸ La variabilità tra classi incorpora, in altri termini, una componente ascrivibile al *cheating* e una componente ulteriore legata al fatto che il *cheating* può essere in alcuni casi corretto *in eccesso* e in altri corretto *per difetto*.

⁹ Se si considerano i risultati delle RN 2012-13, ove la procedura di correzione è stata cambiata per discriminare meglio tra effettivi casi di *cheating* e presenza di “falsi positivi”, il pattern peraltro non cambia di molto.

fatto che le due rilevazioni ci restituiscono un quadro assolutamente coerente e univoco, tranne che negli effetti della condizione scolastica dello studente (anche a parità di altre condizioni, in RN gli studenti anticipatori ottengono risultati statisticamente non dissimili da quelli dei “regolari”, mentre vanno meglio in PISA) e dell’ESCS medio di plesso, che in RN incide molto di meno in italiano (tav. a9) e molto di più in matematica; per inciso, si conferma che in RN i differenziali di area sono particolarmente ampi proprio in quest’ultima materia (tav. a10). Per il resto, PISA e RN presentano risultati molto simili, sia considerando tutti gli studenti, sia circoscrivendo l’analisi ai soli studenti *core*, a sostanziale conferma di quanto già emerso dall’analisi descrittiva.

Particolarmente interessante è poi vedere quanto differiscano i coefficienti degli effetti fissi regionali. In generale, le due rilevazioni restituiscono un quadro dei divari assai coerente, che muta solo marginalmente quando si considerano i soli studenti *core*, nel senso che per questi i divari territoriali sono ancora più ampi. Una volta tenuto conto di tutti gli altri fattori, le regioni che presentano i livelli di apprendimenti più elevati sono il Veneto, la Provincia autonoma di Trento, la Lombardia, il Friuli-Venezia Giulia; tra le regioni che fanno peggio, la Calabria, la Sicilia, la Campania, la Sardegna. Tra le regioni del Centro, solo le Marche ottengono risultati superiori alla media nazionale in entrambe le rilevazioni e in entrambe le materie (soprattutto in matematica-RN); al Nord, solo la Liguria non riesce a conseguire risultati superiori alla media italiana; nel Mezzogiorno, infine, si distingue positivamente la Puglia (figg. a4-a5). Il profilo del *pattern* restituito dalle stime è ancora più simile, tra le due rilevazioni, sia in matematica che in *reading literacy*/italiano, anche se in RN le differenze tra le regioni appaiono ancora più ampie di quanto non emerga in PISA.

4.3 Le regressioni quantiliche

Gli effetti delle variabili esplicative sui livelli di apprendimento possono essere diversi a seconda del punto della distribuzione che si considera. In altre parole, tali effetti possono essere maggiori o minori per gli studenti più abili (quelli posizionati nella parte più alta della distribuzione delle competenze) o per quelli meno abili (quelli posizionati nella parte più bassa della distribuzione delle competenze). È possibile modellare la relazione funzionale tra un percentile condizionato della distribuzione (anziché la media condizionata dei punteggi) e i regressori, mediante la tecnica della regressione quantilica.

Le tavv. a11-a12 riportano i risultati di questo esercizio sia per PISA sia per le RN, distinguendo tra l’ambito di lettura/italiano (tav. a11) e quello di matematica (tav. a12). In generale, si ha evidenza del fatto che nelle RN l’impatto dei regressori sia più differenziato lungo la distribuzione dei punteggi rispetto a PISA; in altre parole, se in PISA i coefficienti stimati risultano analoghi per i diversi quantili, nelle RN essi possono essere anche molto diversi tra loro. Un caso interessante è quello che riguarda il

coefficiente relativo agli istituti professionali. In lettura/italiano, il ritardo di questo indirizzo rispetto ai licei è molto ampio, ma se in PISA esso rimane tra il 13 e il 14 per cento per i diversi percentili, nelle RN è (assai) più accentuato tra gli studenti meno abili (oltre il 20 per cento in corrispondenza del decimo e del venticinquesimo percentile) e più contenuto tra gli studenti più abili. In matematica il quadro è diverso: in entrambe le rilevazioni, infatti, il gap tra istituti professionali e licei cresce man mano che ci si sposta verso la parte alta della distribuzione dei punteggi, ma in misura assai più pronunciata nelle RN (dove supera il 30 per cento in corrispondenza del novantesimo percentile).

Altro caso interessante è quello relativo alle differenze di genere. In lettura/italiano i maschi fanno peggio delle femmine, sia in PISA sia nelle RN, ma questo divario va attenuandosi man mano che si procede lungo la distribuzione; in matematica i maschi fanno invece meglio, ma mentre in PISA il vantaggio rimane pressoché costante lungo la distribuzione, nelle RN si amplia al crescere dell'abilità degli studenti.

Infine, vi è il caso dei divari territoriali. In lettura/italiano il gap delle aree meridionali appare profondo in entrambe le rilevazioni (anche se più accentuato nelle RN) e pressoché costante lungo la distribuzione; in matematica, invece, pur essendo analogo nei valori mediani per le due rilevazioni, esso è nelle RN molto più ampio tra gli studenti meno abili, attenuandosi invece nei quantili più elevati (tavv. a11-a12).

5 Il confronto a livello di scuola e individuale

La coerenza nel pattern complessivo evidenziato da PISA e RN non implica di per sé che i risultati a livello micro, nel sottoinsieme di studenti e/o di scuole che hanno partecipato a entrambe le indagini, siano coerenti oppure che, volendo capire la situazione di una singola scuola, si possa indifferentemente fare riferimento all'una o all'altra fonte. Questo ulteriore grado di coerenza è però quel che occorre verificare ove si intenda integrare maggiormente le due indagini, ad esempio al fine di ancorare i risultati nelle RN – disponibili per tutti gli studenti e tutte le scuole – alla metrica internazionale fornita da PISA.

Nell'esaminare questa ulteriore questione, purtroppo si deve tener conto del fatto che, benché tutti gli studenti *core* di PISA fossero in teoria tenuti a partecipare alle RN, non per tutti è stato all'atto pratico possibile rintracciarne i risultati nelle RN. Dei 24.245 studenti che avrebbero dovuto essere rintracciabili nelle RN, solo 10.364 (43 per cento del totale) sono stati poi effettivamente ritrovati nei dati disponibili. La forte caduta è attribuibile per meno di un quarto al mancato abbinamento dei plessi (sui 956 plessi considerati nella nostra analisi e che fanno parte del segmento *core* di PISA, solo 830 sono stati rintracciati nelle RN); il resto della caduta è invece riconducibile al mancato *matching* di studenti all'interno dei plessi *abbinati*. Questo mancato abbinamento degli studenti, tuttavia, non è tanto legato a

vere e proprie assenze o mancate partecipazioni alle RN di studenti che hanno invece svolto la prova PISA – che comunque sono sempre possibili a livello di singolo studente¹⁰ – quanto al fatto che non in tutti i casi si è reso possibile effettuare in maniera univoca e precisa la codifica del singolo studente, che in PISA ha viaggiato in parallelo rispetto al codice adoperato nelle RN.

Evidentemente, peraltro, la maggior precisione organizzativa e amministrativa con cui la singola scuola partecipa alle diverse rilevazioni non è una caratteristica irrilevante. Le scuole, e quindi poi gli studenti, effettivamente abbinati tra le due fonti sono, infatti, mediamente quelli con risultati migliori, tanto in PISA quanto nelle RN (tav. a13). La maggiore efficienza degli istituti scolastici nel “mappare” gli studenti (e i propri plessi, unità di rilevazione in PISA 2012) è in sostanza di per sé un indicatore di “qualità” della scuola, che si riflette anche sui risultati nei test. All’interno della singola scuola abbinata nelle due fonti, i punteggi riportati dai circa 10mila studenti abbinati e dai circa 76mila che, pur partecipando alle RN, non sono presenti o comunque identificabili in PISA, non sono invece statisticamente diversi, se non marginalmente (di un punto percentuale) in italiano (tav. a14). Ciò sostanzialmente conferma la natura casuale del campione di studenti che abbiano partecipato a PISA all’interno della singola scuola.

Per il sottoinsieme di scuole abbinate, l’ordinamento nei risultati è sostanzialmente analogo nelle due fonti. La correlazione tra le due fonti nei punteggi medi di scuola è di .58 per lettura/italiano e .80 in matematica (fig. a6).¹¹ Anche se leggermente meno intensa, la correlazione è parecchio elevata anche a livello di singolo studente (rispettivamente .45 e .72 per lettura/italiano e matematica; fig. a7).

Guardando alla relazione tra i punteggi medi (di scuola e per studente) nei diversi decili della distribuzione, per lettura/italiano solo il 23 per cento delle scuole e il 21 per cento degli studenti abbinati si collocano lungo la diagonale della matrice, con risultati cioè del tutto analoghi nelle due rilevazioni (tavv. a15 e a17), mentre il livello di coerenza si innalza nel caso della matematica (il 31 per cento delle scuole e il 25 per cento degli studenti sono nello stesso decile in entrambe le fonti; tavv. a16-a18).

¹⁰ Da un lato, l’assenza potrebbe essere “strategica”, nel senso di dipendere da una sorta di boicottaggio delle RN. Si tenga conto che nel giorno delle RN alcune organizzazioni sindacali degli insegnanti avevano proclamato un giorno di sciopero (con adesioni stimate nell’ordine dell’1 per cento delle classi; cfr. il comunicato stampa Invalsi del 16 maggio 2012) e che anche nelle classi non integralmente coinvolte dalla protesta singoli studenti potrebbero aver scelto di boicottare la prova, anche perché ininfluenti sui loro scrutini. Dall’altro, l’assenza potrebbe invece essere la naturale conseguenza del tasso di assenteismo che nel sistema scolastico italiano appare piuttosto elevato: i risultati dell’indagine PISA 2012 (OCSE, 2013) mostrano che l’Italia si contende con la Turchia il non invidiabile primato per percentuale di studenti (oltre il 60 per cento, a fronte di una media OCSE di circa il 25 per cento) che hanno saltato una lezione o un intero giorno di scuola almeno una volta nelle due settimane prima dell’indagine.

¹¹ Vale la pena precisare che i valori medi di queste scuole (“abbinate”) sono ottenuti considerando anche i punteggi dei loro studenti che non facevano parte del campione PISA 2012.

6 Conclusioni

L'esame comparativo dei risultati di PISA 2012 e delle RN condotte dall'Invalsi nell'anno scolastico 2011-12 ha evidenziato un notevole grado di coerenza tra le due fonti, coerenza che invece non sussiste tra rilevazioni standardizzate sugli apprendimenti e risultati degli scrutini effettuati a livello decentrato nelle singole scuole (per meglio dire, rilevazioni standardizzate e voti attribuiti dagli insegnanti ai loro alunni sono coerenti all'interno di ogni singola classe, ma lo sono meno quando si confrontano classi e scuole diverse). Tale coerenza si accresce, in generale, quando ci si concentra sui risultati dei soggetti in comune tra le due fonti, ovverosia gli studenti 15enni che siano in II secondaria di secondo grado. Stiamo parlando dei $\frac{3}{4}$ circa di ciascuna delle due rilevazioni, l'una (PISA) riferita agli studenti 15enni in generale (compresi gli anticipatori in III, gli studenti in ritardo con gli studi e quelli in formazione professionale), l'altra (RN) riferita a chi frequenta la II secondaria di secondo grado (compresi, quindi, gli studenti con più di 15 anni ma in ritardo e quelli anticipatori, che sono in II pur avendo meno di 15 anni).

Più in particolare, entrambe le fonti restituiscono un quadro simile dei differenziali di performance esistenti tra gli studenti, sia in connessione con il diverso background familiare sia con riferimento al *pattern* territoriale. Simile è anche il quadro della variabilità complessiva, sebbene criticità da questo punto di vista si riscontrino nelle RN nelle classi cosiddette "non campione", per via della presenza di una significativa presenza di fenomeni di *cheating* e di una imperfetta correzione dello stesso, che è fonte di noise supplementare nei risultati delle RN. La coerenza tra le due fonti si estende al livello dei dati relativi ai singoli studenti e soprattutto alle singole scuole, il cui ordinamento è molto simile nelle due fonti, l'una potendo, in altri termini, essere una valida *proxy* dell'altra.

Tali evidenze paiono un buon viatico per un uso – e per un ridisegno – congiunto delle due rilevazioni. In particolare, esse suggeriscono l'opportunità di un disegno *ex ante* di meccanismi di ancoraggio dei risultati delle RN nella metrica fornita da PISA, che offre l'indubitabile vantaggio di consentire comparazioni nel tempo – laddove le rilevazioni attualmente condotte dall'Invalsi, almeno per il livello scolare in questione¹², non consentono di dire se da un anno all'altro la singola scuola o il sistema nel suo complesso abbia conseguito o meno dei miglioramenti – e nello spazio, rispetto agli altri paesi OCSE.

¹² Per la V primaria, da quest'anno l'Invalsi dovrebbe diffondere dati con ancoraggio nel tempo, da un anno all'altro, dei risultati, in modo da poterne ricavare una sorta di *chain index* a livello di sistema, anche a livello regionale o di singola scuola.

Bibliografia

- Invalsi (2013), *OCSE PISA 2012 – Rapporto nazionale*, Roma, dicembre 2013.
- OCSE (2009), *PISA Data Analysis Manual*, Parigi, 2009.
- OCSE (2013), *PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices*, Vol. IV, Parigi, 2013.
- Osterlind, S.J. (1998), *Constructing test items: Multiple-choice, constructed-response, performance, and other formats*, Kluwer Academic Publishers, Boston.
- Paccagnella, M. e Sestito, P. (2014), “School Cheating and Social Capital, Banca d’Italia”, Banca d’Italia, Temi di discussione, n. 952.
- 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*, Vol.VII, n.2, Vita e Pensiero, ISSN 1824-6672.
- Wu, M.L. (2004), “Plausible Values”, *Rasch Measurement Transactions*, 2004, 18:2 p. 976-978, <http://www.rasch.org/rmt/rmt182.pdf>
- Wu, M.L. e Adams, R. J. (2002), “Plausible values: Why they are important”, paper presentato all’11th International Objective Measurement Workshop, New Orleans, Aprile, 2002.

Appendice statistica

Tav. a1 – Composizione del campione PISA 2012 per genere, tipologia di scuola, cittadinanza e area geografica
(valori percentuali e indici)

	Popolazione di studenti 15enni (per cento sul totale)				ESCS (valori assoluti)			
	In sec. 1° grado	In I sec. 2° grado	In II sec. 2° grado	In III sec. 2° grado (anticipatari)	In sec. 1° grado	In I sec. 2° grado	In II sec. 2° grado	In III sec. 2° grado (anticipatari)
Femmine	0,7	6,8	39,3	1,4	-0,840	-0,452	-0,029	0,592
Maschi	1,4	10,0	39,2	1,2	-0,887	-0,295	0,048	0,721
Licei	-	4,0	41,2	2,0	-	0,020	0,319	0,881
Istituti tecnici	-	5,5	23,5	0,5	-	-0,356	-0,212	-0,025
Istituti professionali	-	5,1	11,0	0,1	-	-0,574	-0,533	-0,508
Scuole medie	2,1	-	-	-	-0,871	-	-	-
Centri di formazione professionale	-	2,1	2,7	-	-	-0,562	-0,588	-
Nativi	1,1	13,4	75,3	2,6	-0,922	-0,311	0,029	0,656
Immigrati	0,9	3,2	3,3	0,0	-0,846	-0,597	-0,427	-0,683
prima generazione	0,8	2,6	2,1	-	-1,466	-0,517	-0,319	-0,683
seconda generazione	0,1	0,6	1,2	0,0	-0,759	-0,616	-0,491	-
Nord Ovest	0,3	5,1	19,2	0,2	-0,563	-0,267	0,092	1,167
Nord Est	0,4	3,5	14,3	0,1	-0,920	-0,340	0,067	0,536
Centro	0,5	3,3	14,4	0,3	-0,866	-0,249	0,176	0,933
Sud (1)	0,3	2,7	18,0	1,1	-1,012	-0,573	-0,176	0,450
Sud Isole (2)	0,5	2,2	12,5	0,9	-0,951	-0,504	-0,105	0,696
Totale	2,1	16,8	78,5	2,6	-0,871	-0,359	0,009	0,650

Fonte: elaborazioni su dati OCSE-PISA. Eventuali mancate quadrature sono dovute ad arrotondamenti.

(1) Abruzzo, Molise, Campania e Puglia. – (2) Basilicata, Calabria, Sicilia e Sardegna.

**Tav. a2 – Risultati di PISA 2012 per genere, tipologia di scuola, cittadinanza e area geografica
(punteggi)**

	Punteggi rispetto alla media complessiva=100 (LETTURA)				Punteggi rispetto alla media complessiva=100 (MATEMATICA)			
	In sec. 1° grado	In I sec. 2° grado	In II sec. 2° grado	In III sec. 2° grado (anticipatari)	In sec. 1° grado	In I sec. 2° grado	In II sec. 2° grado	In III sec. 2° grado (anticipatari)
Femmine	75	93	106	111	73	87	100	105
Maschi	68	84	100	108	76	91	105	111
Licei	-	99	111	115	-	96	108	112
Istituti tecnici	-	90	99	92	-	93	102	95
Istituti professionali	-	80	87	87	-	81	87	84
Scuole medie	71	-	-	-	75	-	-	-
Centri di formazione professionale	-	81	90	-	-	84	91	-
Nativi	73	89	103	110	75	89	103	108
Immigrati	70	85	96	96	75	88	98	100
prima generazione	71	84	94	-	77	88	96	-
seconda generazione	66	86	99	96	65	89	100	100
Nord Ovest	72	92	109	115	79	92	109	117
Nord Est	68	92	109	121	77	94	110	119
Centro	75	86	103	115	78	88	103	113
Sud (1)	67	84	99	110	67	84	97	107
Sud Isole (2)	70	80	95	105	71	82	94	104
Totale	71	88	103	109	75	89	103	108

Fonte: elaborazioni su dati OCSE-PISA.

(1) Abruzzo, Molise, Campania e Puglia. – (2) Basilicata, Calabria, Sicilia e Sardegna.

Tav. a3 – Risultati di PISA 2012, per regione
(valori percentuali, punteggi e indici)

	Popolazione di studenti 15enni (per cento) (1)				Punteggi rispetto alla media complessiva=100 (LETTURA)			Punteggi rispetto alla media complessiva=100 (MATEMATICA)			ESCS (valori assoluti)		
	Al più in I sec. 2° grado (posticipatari) (2)	In II sec. 2° grado (3)	In III sec. 2° grado (anticipatari)	Totale	Al più in I sec. 2° grado (posticipatari) (2)	In II sec. 2° grado (3)	In III sec. 2° grado (anticipatari)	Al più in I sec. 2° grado (posticipatari) (2)	In II sec. 2° grado (3)	In III sec. 2° grado (anticipatari)	Al più in I sec. 2° grado (posticipatari) (2)	In II sec. 2° grado (3)	In III sec. 2° grado (anticipatari)
Valle d'Aosta	35,4	64,3	0,2	100	90	109	128	90	108	123	-0,536	-0,023	1,558
Piemonte	23,4	75,6	0,9	100	89	107	118	89	107	121	-0,435	0,037	1,257
Liguria	21,8	77,6	0,7	100	83	105	109	85	105	111	-0,437	0,150	0,918
Lombardia	21,2	78,0	0,8	100	93	110	115	94	110	115	-0,182	0,108	1,149
Veneto	20,4	79,5	0,1	100	92	110	124	95	111	117	-0,420	0,024	0,292
Prov. Aut. Trento	19,4	80,5	0,0	100	91	110	146	95	111	127	-0,293	0,074	1,770
Prov. Aut. Bolzano	29,3	70,4	0,3	100	89	106	112	93	109	120	-0,268	0,018	0,145
Friuli-Venezia Giulia	19,8	79,8	0,4	100	90	110	123	93	111	126	-0,291	0,129	0,781
Emilia-Romagna	22,1	77,1	0,8	100	85	106	121	88	107	118	-0,454	0,109	0,564
Toscana	23,8	74,9	1,3	100	83	104	118	87	107	118	-0,500	0,139	0,993
Umbria	16,7	82,3	0,9	100	83	104	120	85	105	118	-0,248	0,172	1,165
Marche	18,7	80,9	0,4	100	85	105	118	89	105	117	-0,375	0,038	1,161
Lazio	19,7	78,3	2,0	100	85	101	114	86	100	110	-0,203	0,238	0,882
Abruzzo	18,4	80,9	0,7	100	89	100	119	90	100	117	-0,270	0,097	0,711
Molise	13,3	85,3	1,4	100	77	100	115	79	99	112	-0,474	-0,082	0,998
Campania	14,3	78,2	7,5	100	84	95	109	84	94	106	-0,551	-0,180	0,463
Puglia	11,1	86,4	2,5	100	79	103	117	78	101	112	-0,874	-0,251	0,347
Basilicata	12,3	85,5	2,1	100	81	99	113	82	98	114	-0,527	-0,188	0,665
Calabria	12,4	83,3	4,3	100	72	91	101	72	90	100	-0,801	-0,178	0,776
Sicilia	16,0	75,8	8,2	100	78	95	105	80	93	104	-0,615	-0,083	0,664
Sardegna	27,9	71,0	1,0	100	81	100	117	83	99	115	-0,435	-0,021	1,130
Totale	18,9	78,5	2,6	100	86	103	109	88	103	108	-0,414	0,009	0,650

Fonte: elaborazioni su dati OCSE-PISA.

(1) Fatto pari a 100 il totale per regione (per riga). – (2) Include gli studenti della scuola secondaria di I grado e quelli del I anno di formazione professionale. – (3) Include gli studenti del II anno di formazione professionale.

Tav. a4 – Composizione della popolazione RN 2011-12 e risultati per genere, tipologia di scuola, cittadinanza e area geografica
(valori percentuali, punteggi e indici)

	Popolazione di studenti in Il sec. 2° grado (per cento sul totale)			Punteggi rispetto alla media complessiva=100 (ITALIANO)			Punteggi rispetto alla media complessiva=100 (MATEMATICA)			ESCS (valori assoluti)		
	1. Posticip. (più di 15 anni)	2. Regolari (15enni)	3. Anticip. (meno di 15 anni)	1. Posticip. (>15 anni)	2. Regolari (15enni)	3. Anticip. (<15 anni)	1. Posticip. (>15 anni)	2. Regolari (15enni)	3. Anticip. (<15 anni)	1. Posticip. (>15 anni)	2. Regolari (15enni)	3. Anticip. (<15 anni)
Femmine	7,9	39,7	2,5	93	104	101	80	99	98	-0,331	0,029	0,456
Maschi	10,8	37,0	2,1	88	100	99	87	109	109	-0,233	0,091	0,490
Licei	5,2	43,9	3,5	100	106	103	92	112	109	0,101	0,339	0,684
Istituti tecnici	6,8	23,0	0,8	94	102	94	91	100	90	-0,279	-0,209	-0,097
Istituti professionali	6,8	9,7	0,3	80	85	80	71	74	71	-0,561	-0,573	-0,494
Nativi	13,9	70,1	4,4	92	103	100	84	104	103	-0,206	0,079	0,501
Immigrati	4,8	6,5	0,3	88	100	94	85	100	97	-0,469	-0,157	-0,019
prima gener.	2,3	2,2	0,1	85	98	90	84	97	91	-0,346	-0,058	0,150
seconda gener.	2,5	4,3	0,2	90	101	96	86	101	100	-0,616	-0,361	-0,359
Nord Ovest	5,3	18,6	0,3	97	107	102	91	114	120	-0,211	0,159	0,711
Nord Est	3,7	13,6	0,2	97	109	105	92	115	118	-0,193	0,176	0,590
Centro	3,4	12,9	0,6	90	105	106	81	103	108	-0,152	0,222	0,698
Sud (1)	3,4	18,8	2,2	83	97	98	77	95	100	-0,476	-0,153	0,354
Sud Isole (2)	3,0	12,7	1,4	81	95	100	73	89	99	-0,395	-0,060	0,503
Totale	18,7	76,6	4,7	91	102	100	84	104	103	-0,275	0,059	0,471

Fonte: elaborazioni su dati Invalsi-RN. Dati corretti per il *cheating*. Eventuali mancate quadrature sono dovute ad arrotondamenti.

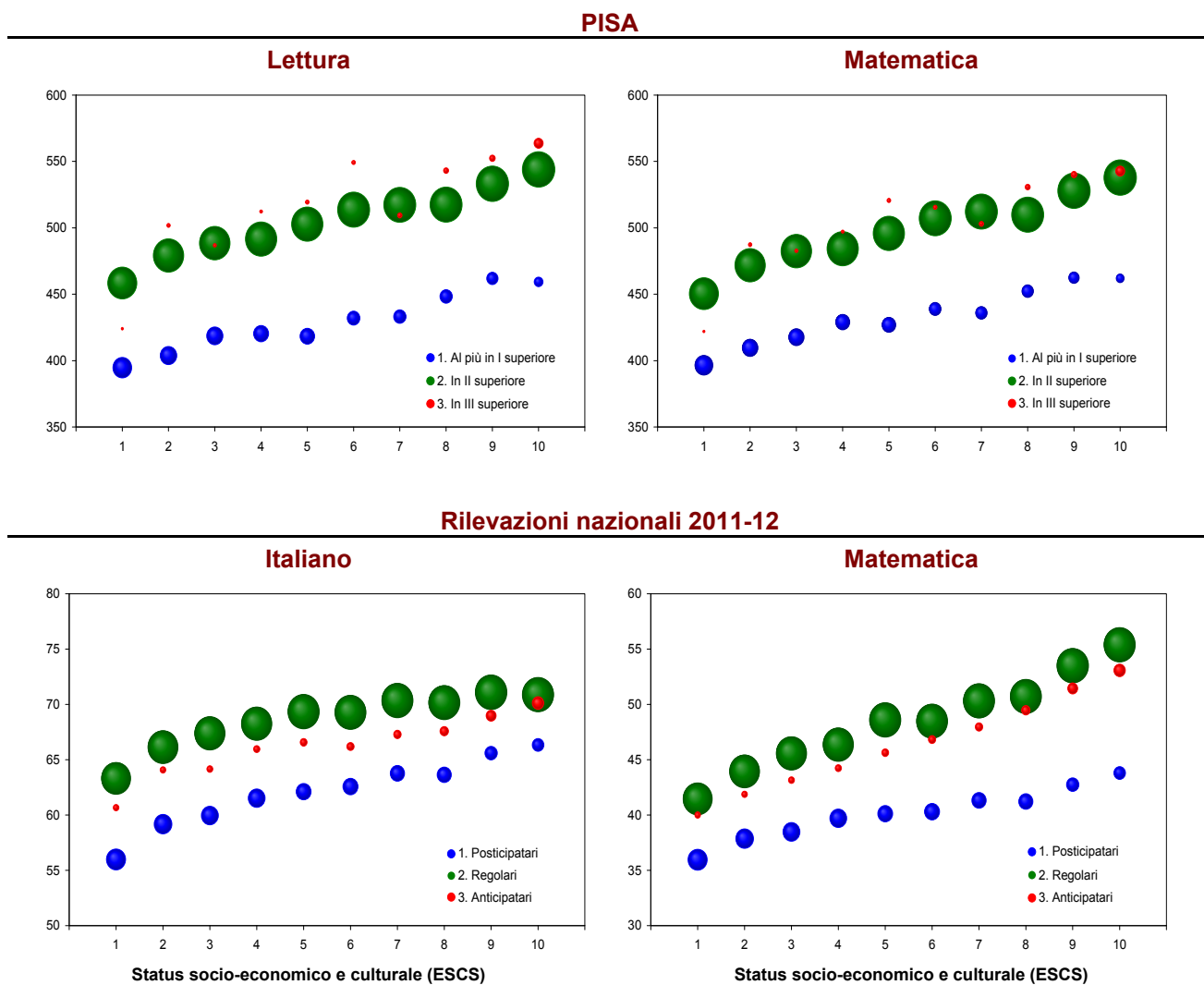
(1) Abruzzo, Molise, Campania e Puglia. – (2) Basilicata, Calabria, Sicilia e Sardegna.

Tav. a5 – Composizione della popolazione RILEVAZIONI NAZIONALI 2011-12 e risultati per regione
(valori percentuali, punteggi e indici)

	Popolazione di studenti in II sec. 2° grado (1)				Punteggi rispetto alla media complessiva=100 (ITALIANO)			Punteggi rispetto alla media complessiva=100 (MATEMATICA)			ESCS (valori assoluti)		
	1. Posticip. (più di 15 anni)	2. Regolari (15enni)	3. Anticip. (meno di 15 anni)	Totale	1. Posticip. (più di 15 anni)	2. Regolari (15enni)	3. Anticip. (meno di 15 anni)	1. Posticip. (più di 15 anni)	2. Regolari (15enni)	3. Anticip. (meno di 15 anni)	1. Posticip. (più di 15 anni)	2. Regolari (15enni)	3. Anticip. (meno di 15 anni)
Valle d'Aosta	27,3	72,0	0,7	100	102	111	116	86	106	103	-0,121	0,120	0,054
Piemonte	21,3	77,3	1,3	100	97	107	102	90	112	122	-0,232	0,140	0,854
Liguria	21,6	75,8	2,6	100	91	108	104	81	108	116	-0,132	0,295	0,603
Lombardia	21,8	77,1	1,1	100	97	107	102	93	116	121	-0,214	0,149	0,677
Veneto	20,7	78,5	0,8	100	99	109	107	94	116	118	-0,186	0,105	0,601
Prov. Aut. Trento	16,1	83,5	0,4	100	104	111	97	101	119	111	0,068	0,255	0,493
Prov. Aut. BZ (ital.)	27,3	71,6	1,1	100	91	112	109	86	110	116	-0,212	0,299	0,350
Prov. Aut. BZ (ted.)	25,4	74,4	0,2	100	-	-	-	96	106	119	0,008	0,201	0,567
Friuli-Ven. Giulia	23,4	75,7	0,9	100	98	111	106	92	115	120	-0,135	0,229	0,685
Emilia-Romagna	21,7	77,1	1,2	100	94	108	103	88	114	118	-0,247	0,238	0,575
Toscana	21,7	76,5	1,8	100	90	105	104	82	104	107	-0,188	0,233	0,676
Umbria	18,8	78,7	2,5	100	89	107	109	79	104	117	-0,208	0,204	0,715
Marche	18,8	80,2	1,0	100	90	107	108	82	108	115	-0,292	0,086	0,456
Lazio	19,7	75,2	5,1	100	90	103	106	81	100	107	-0,072	0,273	0,719
Abruzzo	17,8	79,9	2,3	100	87	104	110	77	101	112	-0,170	0,123	0,605
Molise	13,9	82,7	3,4	100	83	103	101	76	98	112	-0,337	0,053	0,674
Campania	14,1	73,1	12,8	100	80	92	96	75	90	97	-0,527	-0,202	0,315
Puglia	13,0	81,4	5,6	100	86	99	103	80	100	111	-0,514	-0,174	0,442
Basilicata	16,3	80,5	3,2	100	82	100	103	74	93	97	-0,375	-0,087	0,667
Calabria	13,1	77,7	9,3	100	78	93	100	72	86	96	-0,426	-0,082	0,584
Sicilia	17,3	73,0	9,7	100	80	94	99	75	90	101	-0,409	-0,073	0,464
Sardegna	27,2	70,1	2,7	100	83	100	101	69	88	94	-0,319	0,094	0,526
Totale	18,7	76,6	4,7	100	91	102	100	84	104	103	-0,275	0,059	0,471

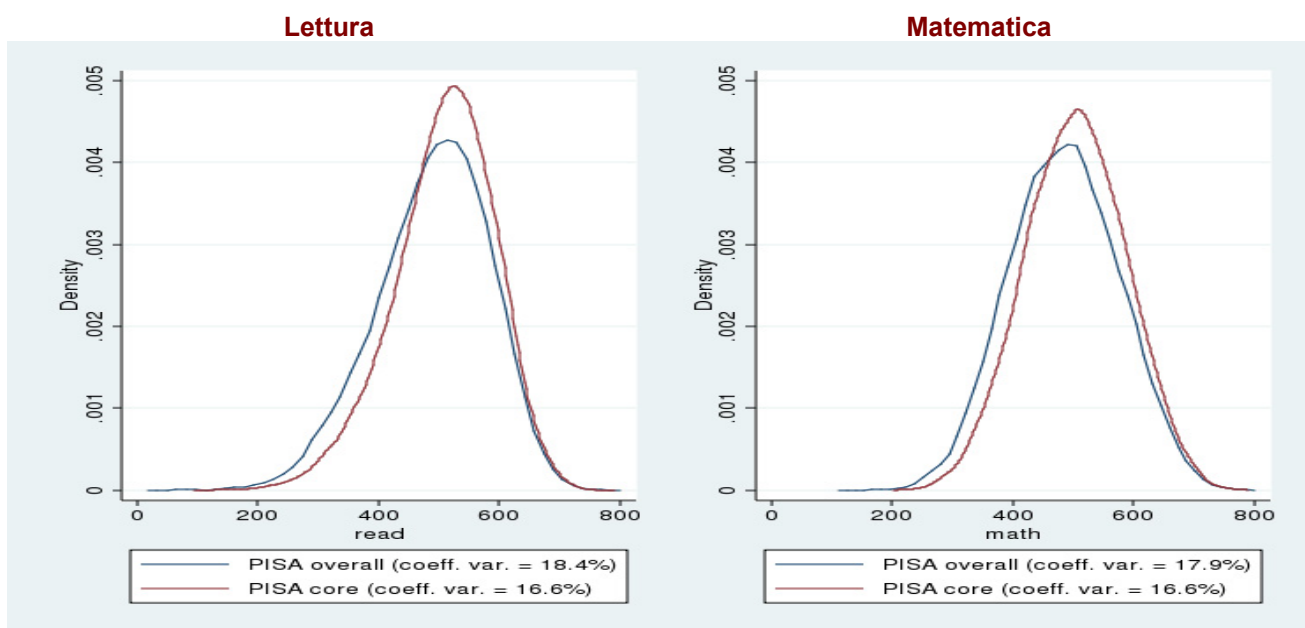
Fonte: elaborazioni su dati Invalsi-RN. Dati corretti per il *cheating*. Eventuali mancate quadrature sono dovute ad arrotondamenti. (1) Fatto pari a 100 il totale per regione (per riga).

Fig. a1 – Background socio-economico e culturale della famiglia di origine dello studente e punteggi riportati in PISA 2012 e in RN 2011-12, per condizione scolastica (punteggi)



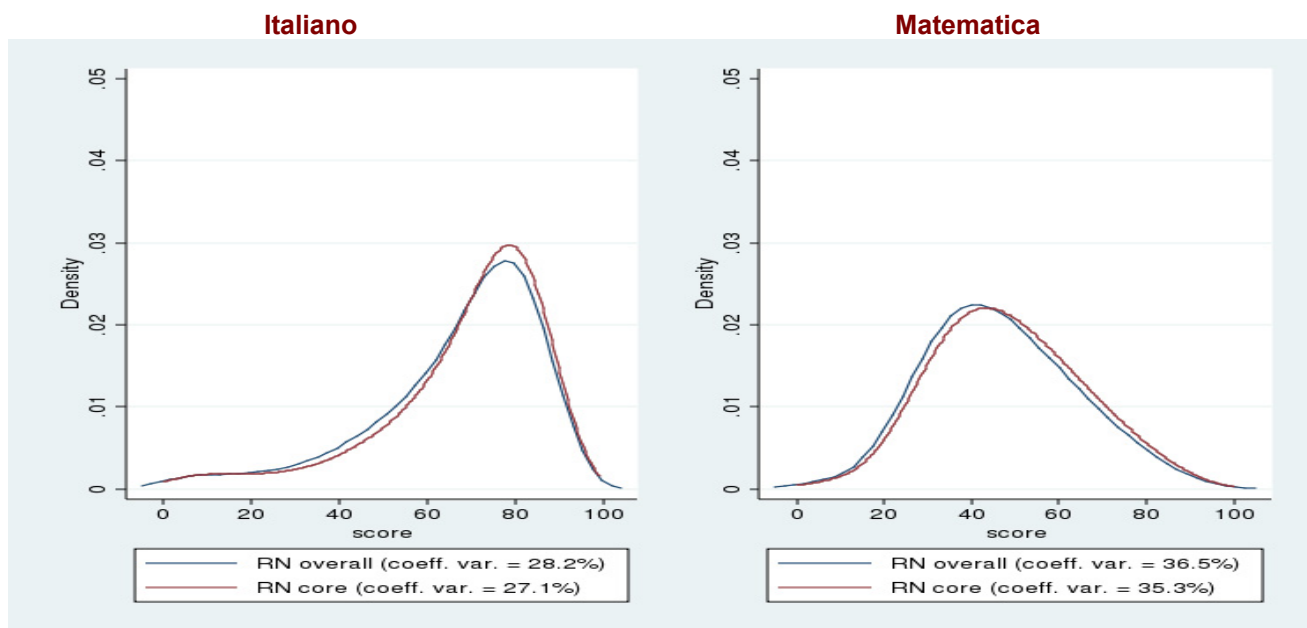
Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN. I dati Invalsi-RN sono corretti per il *cheating*. L'area delle bolle indica la numerosità degli studenti. I valori da 1 a 10 corrispondono ai decili della distribuzione del *background* familiare, per cui le bolle di ciascun decile hanno un'area complessivamente pari a un decimo della popolazione totale.

Fig. a2 – Confronto tra la distribuzione dei punteggi PISA 2012, nel complesso e per i soli 15enni in *Grade 10*



Fonte: elaborazioni su dati OCSE-PISA. Tra gli studenti *core* non sono ricompresi quelli dei Corsi di Formazione Professionale (CFP)

Fig. a3 – Confronto tra la distribuzione dei punteggi RILEVAZIONI NAZIONALI 2011-12, nel complesso e per i soli studenti “regolari”



Fonte: elaborazioni su dati Invalsi-RN. Dati corretti per il *cheating*.

Tav. a6 – Scomposizione della varianza dei punteggi PISA 2012 e RILEVAZIONI NAZIONALI 2011-12, nel complesso e per i soli 15enni in Grade 10 / “regolari”
(valori percentuali)

	Letture / Italiano				Matematica			
	Varianza TRA plessi scolastici	Varianza INTERNA ai plessi scolastici	Varianza TOTALE	Coefficiente di variazione	Varianza TRA plessi scolastici	Varianza INTERNA ai plessi scolastici	Varianza TOTALE	Coefficiente di variazione
PISA 2012								
Tutti gli studenti	57,5	42,5	100	18,4	54,6	45,4	100	17,9
Solo 15enni in Grade 10 (1)	56,3	43,7	100	16,6	54,3	45,7	100	16,6
Rilevazioni nazionali								
Tutti gli studenti	41,9	58,1	100	28,2	45,8	54,2	100	36,5
Solo studenti "regolari"	39,4	60,6	100	27,1	44,9	55,1	100	35,3
Rilev. nazionali (2)								
Tutti gli studenti	37,6	62,4	100	28,2	41,0	59,0	100	36,5
Solo studenti "regolari"	34,0	66,0	100	27,1	39,8	60,2	100	35,3

Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN. I dati Invalsi-RN sono corretti per il *cheating*. Stima HLM-Linear Mixed Model, con (per PISA) pesi-studente normalizzati (la somma dei pesi restituisce il numero complessivo di studenti) e pesi-scuola pari alla somma dei pesi-studente di ogni singola scuola (plesso). (1) Per consentire un confronto più omogeneo con i dati RN, sono esclusi gli studenti dei Corsi di Formazione Professionale (CFP). - (2) Prendendo come unità di analisi la scuola e non il plesso.

Tav. a7 – Scomposizione della varianza dei punteggi PISA 2012 e RILEVAZIONI NAZIONALI 2011-12, per i soli 15enni in Grade 10 / “regolari”
(valori percentuali)

	Letture / Italiano					Matematica				
	Varianza <i>between</i>		Varianza INTERNA ai plessi (<i>within</i>)	Varianza TOTALE	Coeff. di variazione	Varianza <i>between</i>		Varianza INTERNA ai plessi (<i>within</i>)	Varianza TOTALE	Coeff. di variazione
	Varianza TRA indirizzi	Varianza TRA plessi negli indirizzi				Varianza TRA indirizzi	Varianza TRA plessi negli indirizzi			
PISA 2012 (15enni in Grade10)(1)										
Nord Ovest	17,2	22,4	60,4	100	14,4	14,5	25,3	60,2	100	14,7
Nord Est	29,3	19,4	51,3	100	15,6	23,3	23,2	53,5	100	15,2
Centro	30,3	21,7	47,9	100	15,5	20,6	29,2	50,2	100	15,6
Sud (2)	32,4	21,9	45,7	100	17,1	24,4	25,0	50,5	100	16,8
Sud Isole (3)	30,2	24,5	45,2	100	17,3	22,6	22,9	54,5	100	16,7
Italia	27,1	29,2	43,7	100	16,6	20,7	33,7	45,7	100	16,6
Rilevazioni nazionali										
Nord Ovest	7,1	27,8	65,1	100	25,0	18,2	28,3	53,5	100	30,8
Nord Est	10,7	22,8	66,4	100	22,2	19,1	24,5	56,5	100	30,6
Centro	13,0	22,7	64,4	100	24,8	13,9	28,5	57,6	100	34,9
Sud (2)	10,1	26,3	63,6	100	30,5	9,7	29,9	60,3	100	37,7
Sud Isole (3)	13,0	26,3	60,7	100	29,7	7,3	30,3	62,4	100	37,4
Italia	9,8	29,6	60,6	100	27,1	12,1	32,8	55,1	100	35,3
Italia (solo classi “campione”)	19,7	32,5	47,7	100	20,6	16,2	37,3	46,5	100	34,2

Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN. I dati Invalsi-RN sono corretti per il *cheating*. Stima *HLM-Linear Mixed Model*, con (per PISA) pesi-studente normalizzati (la somma dei pesi restituisce il numero complessivo di studenti) e pesi-scuola pari alla somma dei pesi-studente di ogni singola scuola (plesso). La varianza tra plessi nel totale nazionale include la variabilità tra aree geografiche.

(1) Per consentire un confronto più omogeneo con i dati RN, sono esclusi gli studenti dei Corsi di Formazione Professionale (CFP). – (2) Abruzzo, Molise, Campania e Puglia. – (3) Basilicata, Calabria, Sicilia e Sardegna.

**Tav. a8 – Scomposizione della varianza dei punteggi PISA 2012 per i soli 15enni in
Grade 10
(valori percentuali)**

	Letture / Italiano				Matematica			
	Varianza TRA plessi scolastici	Varianza INTERNA ai plessi scolastici	Varianza TOTALE	Coefficiente di variazione	Varianza TRA plessi scolastici	Varianza INTERNA ai plessi scolastici	Varianza TOTALE	Coefficiente di variazione
<i>PISA 2012 (15enni in Grade10)(1)</i>								
Media dei PV (EAP)	56,3	43,7	100	16,6	54,3	45,7	100	16,6
1° PV	51,6	48,4	100	17,1	50,9	49,1	100	17,0
1 Random PV	51,6	48,4	100	17,1	51,3	48,7	100	17,0

Fonte: elaborazioni su dati OCSE-PISA. Stima HLM-Linear Mixed Model, con pesi-studente normalizzati (la somma dei pesi restituisce il numero complessivo di studenti) e pesi-scuola pari alla somma dei pesi-studente di ogni singola scuola (plesso).

(1) Per consentire un confronto più omogeneo con i dati RN, sono esclusi gli studenti dei Corsi di Formazione Professionale (CFP).

Tav. a9 – Punteggi in LETTURA/ITALIANO
(punti percentuali)

	PISA 2012 (Lettura)		Rilevazioni nazionali 2011-12 Italiano	
	Tutti gli studenti	Solo 15enni in Grade10	Tutti gli studenti	Solo "regolari"
Intercetta	113,44***	109,88***	112,06***	109,06***
Istituti tecnici	-5,44***	-5,60***	-3,42***	-3,41***
Istituti professionali	-13,45***	-13,63***	-18,27***	-18,69***
Scuole secondarie di I grado	-15,85***			
Centri di formazione professionale	-14,84***			
Studenti posticipatari (1)	-8,29***		-5,81***	
Studenti anticipatari (2)	4,38***		-0,31	
Maschi	-4,63***	-3,89***	-2,85***	-2,61***
Immigrati di prima generazione	-6,23***	-6,21***	-5,86***	-4,30***
Immigrati di seconda generazione	-5,02***	-3,06***	-2,80***	-1,71***
ESCS	0,54***	0,44***	0,68***	0,46***
ESCS di plesso scolastico	8,60***	8,22***	1,40***	0,85***
Nord Est	0,75***	1,08***	2,08***	2,19***
Centro	-6,47***	-6,76***	-2,63***	-1,67***
Sud (3)	-8,10***	-8,14***	-9,71***	-8,85***
Sud Isole (4)	-12,15***	-12,08***	-11,84***	-11,07***
N. osservazioni	30.853	24.245	400.286	307.607
R-squared	0,469	0,396	0,139	0,107

Fonte: elaborazioni su dati OCSE-PISA. Stime WLS (*Weighted Least Squares*). I dati INVALSI-RN sono corretti per il *cheating*. Le variabili di confronto sono, nell'ordine: licei; studenti 15enni in *Grade 10* in PISA; femmine; nativi; Nord Ovest.

(1) In PISA gli studenti in ritardo sono i 15enni iscritti alle scuole secondarie di primo grado o in I secondaria di secondo grado; nelle Rilevazioni nazionali sono quelli che frequentano la II secondaria di secondo grado e sono stati respinti almeno una volta. – (2) In PISA gli studenti anticipatari sono i 15enni iscritti in III secondaria di secondo grado; nelle Rilevazioni nazionali sono quelli che frequentano la II secondaria di secondo grado e hanno meno di 15 anni. Gli asterischi indicano una significatività statistica, rispettivamente, all'1 (***) , al 5 (**) e al 10 per cento (*). – (3) Abruzzo, Molise, Campania e Puglia. – (4) Basilicata, Calabria, Sicilia e Sardegna.

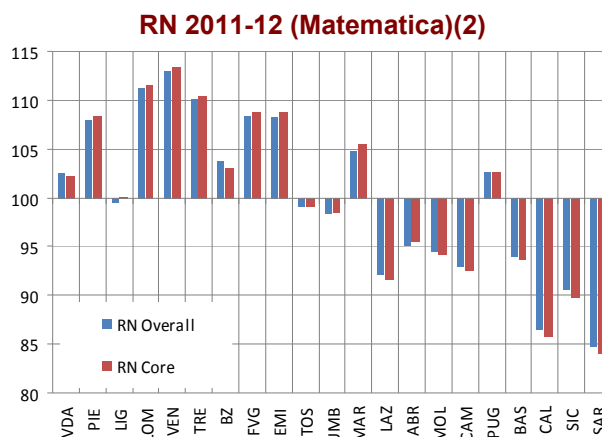
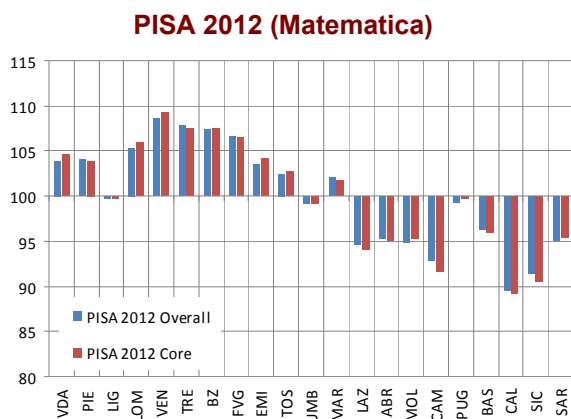
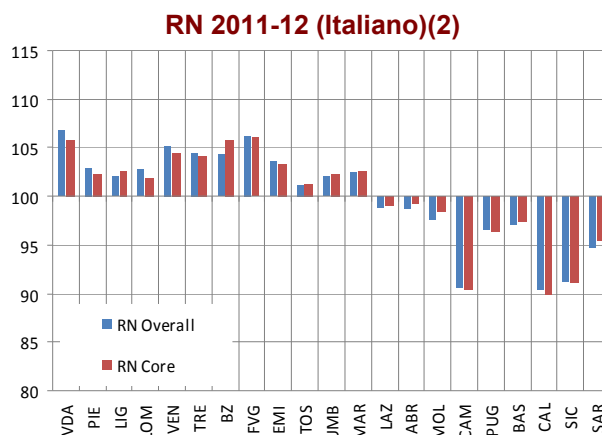
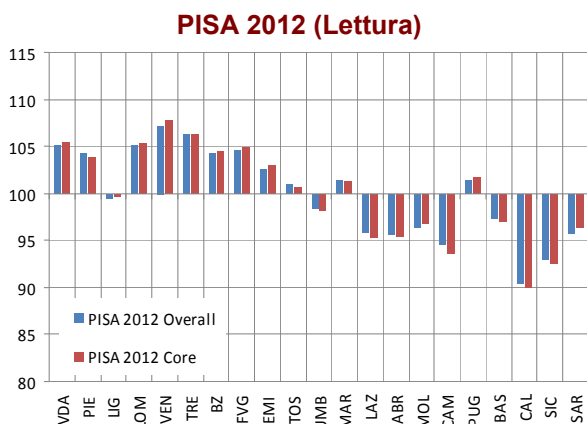
Tav. a10 – Punteggi in MATEMATICA
(punti percentuali)

	PISA 2012		Rilevazioni nazionali 2011-12	
	Tutti gli studenti	Solo 15enni in Grade10	Tutti gli studenti	Solo "regolari"
Intercetta	106,59***	103,90***	109,00***	105,42***
Istituti tecnici	-3,61***	-3,82***	-4,13***	-4,48***
Istituti professionali	-12,57***	-12,88***	-19,04***	-20,37***
Scuole secondarie di I grado	-13,87***			
Centri di formazione professionale	-14,54***			
Studenti posticipatari (1)	-8,44***		-11,21***	
Studenti anticipatari (2)	4,43***		0,15	
Maschi	6,51***	6,87***	12,15***	12,55***
Immigrati di prima generazione	-3,64***	-4,50***	-2,07***	-3,65***
Immigrati di seconda generazione	-3,69***	-2,00***	-1,54***	-2,18***
ESCS	0,67***	0,58***	1,31***	1,30***
ESCS di plesso scolastico	8,09***	7,94***	17,47***	17,91***
Nord Est	1,99***	2,09***	1,05***	1,12***
Centro	-6,16***	-6,56***	-12,42***	-12,26***
Sud (3)	-9,46***	-9,69***	-12,74***	-12,58***
Sud Isole (4)	-12,87***	-13,22***	-20,02***	-20,43***
N. osservazioni	30.853	24.245	407.135	312.637
R-squared	0,436	0,380	0,272	0,258

Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN. Stime WLS (*Weighted Least Squares*). I dati Invalsi-RN sono corretti per il *cheating*. Le variabili di confronto sono, nell'ordine: licei; studenti 15enni in *Grade 10* in PISA; femmine; nativi; Nord Ovest.

(1) In PISA gli studenti in ritardo sono i 15enni iscritti alle scuole secondarie di primo grado o in I secondaria di secondo grado; nelle Rilevazioni nazionali sono quelli che frequentano la II secondaria di secondo grado e sono stati respinti almeno una volta. – (2) In PISA gli studenti anticipatari sono i 15enni iscritti in III secondaria di secondo grado; nelle Rilevazioni nazionali sono quelli che frequentano la II secondaria di secondo grado e hanno meno di 15 anni. Gli asterischi indicano una significatività statistica, rispettivamente, all'1 (***) al 5 (***) e al 10 per cento(*). – (3) Abruzzo, Molise, Campania e Puglia. – (4) Basilicata, Calabria, Sicilia e Sardegna.

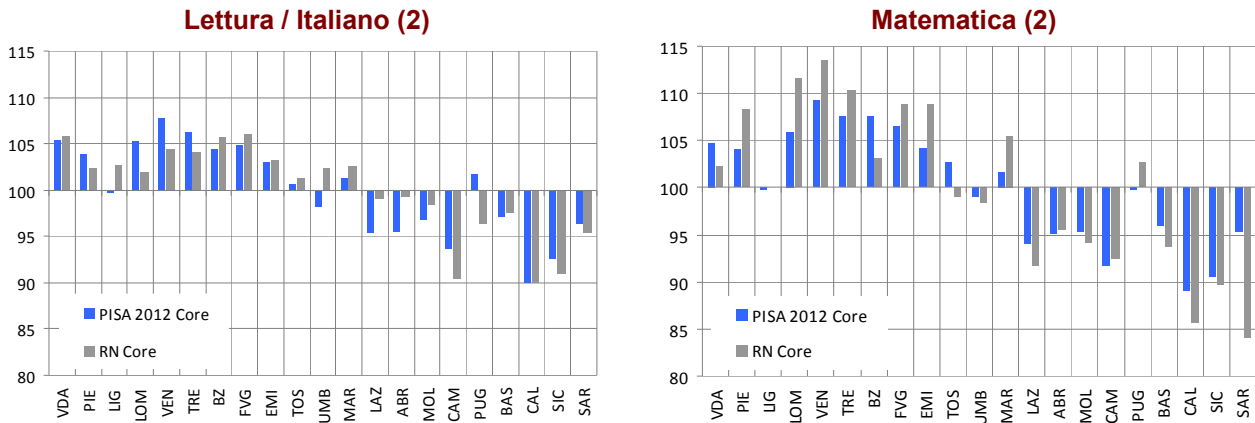
**Fig. a4 – Coefficienti stimati degli effetti fissi regionali (1)
(punti percentuali)**



Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN, questi ultimi corretti per il *cheating*.

(1) I coefficienti sono quelli risultanti da un modello WLS (secondaria di secondo grado) che tiene conto anche del tipo di scuola, della condizione dello studente (se in ritardo, regolare o anticipatorio), del genere e della cittadinanza dello studente, dell'indicatore ESCS (individuale e di scuola). Per ciascuna regione, i coefficienti vanno intesi rispetto a un individuo di riferimento medio nazionale, con le seguenti caratteristiche: frequenta il liceo, è regolare nel percorso scolastico, è femmina, è nativo, presenta un livello medio di ESCS ed è iscritto a una scuola (plesso) con un ESCS nella media. – (2) Nelle Rilevazioni nazionali, il coefficiente della provincia autonoma di Bolzano è quello degli studenti di lingua italiana per la prova di italiano, mentre è quello medio (lingua italiana e lingua tedesca) per la prova di matematica.

**Fig. a5 – Coefficienti stimati degli effetti fissi regionali, solo per gli studenti 15enni in Grade 10 o “regolari” (1)
(punti percentuali)**



Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN, questi ultimi corretti per il *cheating*.

(1) I coefficienti sono quelli risultanti da un modello WLS (*Weighted Least Squares*) che tiene conto anche del tipo di scuola, della condizione dello studente (se in ritardo, regolare o anticipatorio), del genere e della cittadinanza dello studente, dell'indicatore ESCS (individuale e di scuola). Per ciascuna regione, i coefficienti vanno intesi rispetto a un individuo di riferimento medio nazionale, con le seguenti caratteristiche: frequenta il liceo, è regolare nel percorso scolastico, è femmina, è nativo, presenta un livello medio di ESCS ed è iscritto a una scuola (plesso) con un ESCS nella media. – (2) Nelle Rilevazioni nazionali, il coefficiente della provincia autonoma di Bolzano è quello degli studenti di lingua italiana per la prova di italiano, mentre è quello medio (lingua italiana e lingua tedesca) per la prova di matematica.

Tav. a11 – Regressioni quantiliche sui punteggi in Lettura / Italiano in PISA 2012 e nelle RILEVAZIONI NAZIONALI 2011-12, per i soli 15enni in Grade 10 / “regolari” (punti percentuali)

	PISA 2012 (1)					Rilevazioni nazionali				
	Q10	Q25	Q50	Q75	Q90	Q10	Q25	Q50	Q75	Q90
Intercetta	95,11***	102,57***	109,89***	117,62***	124,08***	82,44***	104,44***	115,39***	123,19***	128,73***
Istituti tecnici	-6,62***	-5,49***	-4,99***	-5,46***	-5,74***	1,18**	-4,25***	-4,62***	-4,12***	-3,49***
Istituti professionali	-14,87***	-13,85***	-13,01***	-13,45***	-13,42***	-21,77***	-21,27***	-19,27***	-16,19***	-13,14***
Maschi	-5,34***	-4,68***	-3,47***	-3,08***	-2,29***	-6,32***	-3,69***	-2,23***	-1,54***	-1,22***
Immigrati di 1° gen.	-7,44***	-7,35***	-6,15***	-5,32***	-5,02***	-7,47***	-6,41***	-4,39***	-3,09***	-2,17***
Immigrati di 2° gen.	-2,57**	-2,91***	-2,80***	-3,10***	-3,08***	-3,19***	-2,80***	-1,81***	-1,41***	-0,93***
ESCS	0,47**	0,26*	0,39***	0,67***	0,87***	-0,14	0,66***	0,76***	0,71***	0,70***
ESCS plesso scolastico	7,91***	8,97***	8,94***	7,42***	6,60***	-11,66***	4,43***	5,90***	4,96***	3,85***
Nord Est	0,84*	0,98**	1,09***	1,50***	1,38***	5,92***	1,70***	0,71***	0,44***	0,36***
Centro	-8,56***	-7,53***	-6,72***	-5,82***	-4,98***	-2,02***	-4,52***	-3,71***	-2,72***	-2,00***
Sud (2)	-9,87***	-8,52***	-7,88***	-7,19***	-6,75***	-21,55***	-12,57***	-7,42***	-4,80***	-3,34***
Sud Isole (3)	-12,96***	-12,45***	-11,86***	-11,61***	-10,96***	-20,71***	-16,91***	-11,29***	-7,64***	-5,30***
<i>N. osservazioni</i>	24.245	24.245	24.245	24.245	24.245	307.607	307.607	307.607	307.607	307.607
<i>Pseudo R-squared</i>	0,2431	0,2417	0,2287	0,2127	0,1932	0,0535	0,0995	0,1113	0,0948	0,0761

Fonte elaborazioni su dati OCSE-PISA e Invalsi-RN. Stime WLS (*Weighted Least Squares*). I dati Invalsi-RN sono corretti per il *cheating*. Le variabili di confronto sono, nell'ordine: licei; femmine; nativi; Nord Ovest.

(1) Per consentire un confronto più omogeneo con i dati RN, sono esclusi gli studenti dei Corsi di Formazione Professionale (CFP). Gli asterischi indicano una significatività statistica, rispettivamente, all'1 (***), al 5 (**), e al 10 per cento (*). – (2) Abruzzo, Molise, Campania e Puglia. – (3) Basilicata, Calabria, Sicilia e Sardegna.

Tav. a12 – Regressioni quantiliche sui punteggi in Matematica in PISA 2012 e nelle RILEVAZIONI NAZIONALI 2011-12, per i soli 15enni in *Grade 10* / “regolari” (punti percentuali)

	PISA 2012 (1)					Rilevazioni nazionali				
	Q10	Q25	Q50	Q75	Q90	Q10	Q25	Q50	Q75	Q90
Intercetta	87,59***	94,98***	103,73***	112,69***	121,56***	67,72***	83,13*	103,38***	126,33***	147,30***
Istituti tecnici	-3,07***	-3,04***	-3,75***	-3,98***	-5,56***	1,37***	0,36***	-2,77***	-7,44***	-11,33***
Istituti professionali	-10,89***	-11,64***	-13,01***	-13,77***	-16,73***	-9,84***	-12,86***	-18,06***	-25,63***	-32,56***
Maschi	5,56***	6,26***	7,28***	7,13***	7,22***	6,34***	9,61***	12,92***	14,75***	14,74***
Immigrati di 1° gen.	-2,97***	-4,74***	-5,17***	-5,32***	-4,87***	-3,18***	-3,38***	-4,09***	-3,63***	-3,56***
Immigrati di 2° gen.	-1,45	-1,60	-2,62***	-0,98	-2,50**	-1,67***	-2,02***	-2,52***	-2,35***	-2,64***
ESCS	0,23	0,31**	0,62***	0,81***	0,86***	0,76***	1,03***	1,38***	1,51***	1,51***
ESCS plesso scolastico	7,37***	8,01***	8,13***	8,27***	6,50***	13,08***	17,01***	20,50***	21,06***	19,18***
Nord Est	1,11**	2,20***	2,49***	2,21***	2,34***	1,35***	1,08***	1,05***	0,80***	0,76***
Centro	-8,76***	-7,15***	-6,13***	-5,67***	-4,75***	-11,36***	-12,17***	-13,03***	-12,68***	-11,57***
Sud (2)	-9,74***	-9,69***	-9,67***	-9,15***	-9,59***	-13,87***	-12,06***	-11,44***	-11,60***	-11,41***
Sud Isole (3)	-13,23***	-13,01***	-13,43***	-13,15***	-14,12***	-18,56***	-18,53***	-19,92***	-21,10***	-21,46***
<i>N. osservazioni</i>	24.245	24.245	24.245	24.245	24.245	312.637	312.637	312.637	312.637	312.637
<i>R-squared</i>	0,2014	0,2136	0,2233	0,2232	0,2160	0,0937	0,1211	0,1519	0,1748	0,1799

Fonte elaborazioni su dati OCSE-PISA e Invalsi-RN. Stime WLS (*Weighted Least Squares*). I dati Invalsi-RN sono corretti per il *cheating*. Le variabili di confronto sono, nell'ordine: licei; femmine; nativi; Nord Ovest.

(1) Per consentire un confronto più omogeneo con i dati RN, sono esclusi gli studenti dei Corsi di Formazione Professionale (CFP). Gli asterischi indicano una significatività statistica, rispettivamente, all'1 (***), al 5 (***) e al 10 per cento(*). – (2) Abruzzo, Molise, Campania e Puglia. – (3) Basilicata, Calabria, Sicilia e Sardegna.

Tav. a13 – Differenze tra i punteggi riportati dagli studenti (plessi) “abbinati” e da quelli “non abbinati” in PISA e nelle RILEVAZIONI NAZIONALI
(unità, punteggi in centesimi e statistiche)

	N	Letture (PISA)	Matematica (PISA)	Italiano (RN)	Matematica (RN)
PISA 2012					
Scuole non "matchate"	126	95,6	98,0	-	-
Scuole "abbinare"	830	100,7	100,3	-	-
Statistica "t" Student		-4,0271	-1,8450	-	-
Prob. T > t		0,0001	0,0653	-	-
Studenti non "abbinati"	13.881	98,8	98,8	-	-
Studenti "abbinati"	10.364	101,6	101,6	-	-
Statistica "t" Student		-13,5781	-13,3395	-	-
Prob. T > t		0,0000	0,0000	-	-
Rilevazioni nazionali (Invalsi)					
Scuole non "abbinare"	3.285	-	-	98,9	98,7
Scuole "abbinare"	830	-	-	104,4	105,2
Statistica "t" Student		-	-	-7,2498	-6,5374
Prob. T > t		-	-	0,0000	0,0000
Studenti non "abbinati"	301.514	-	-	99,9	99,9
Studenti "abbinati"	10.364	-	-	103,4	102,6
Statistica "t" Student		-	-	-13,3931	-7,7367
Prob. T > t		-	-	0,0000	0,0000

Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN, questi ultimi corretti per il *cheating*. Sono state escluse le scuole che hanno partecipato alle Rilevazioni nazionali (Invalsi-RN) con un solo studente.

Tav. a14 – Differenze tra i punteggi riportati dagli studenti “abbinati” e “non abbinati” nelle scuole che hanno partecipato sia a PISA sia alle RILEVAZIONI NAZIONALI

(unità, punteggi in centesimi e statistiche)

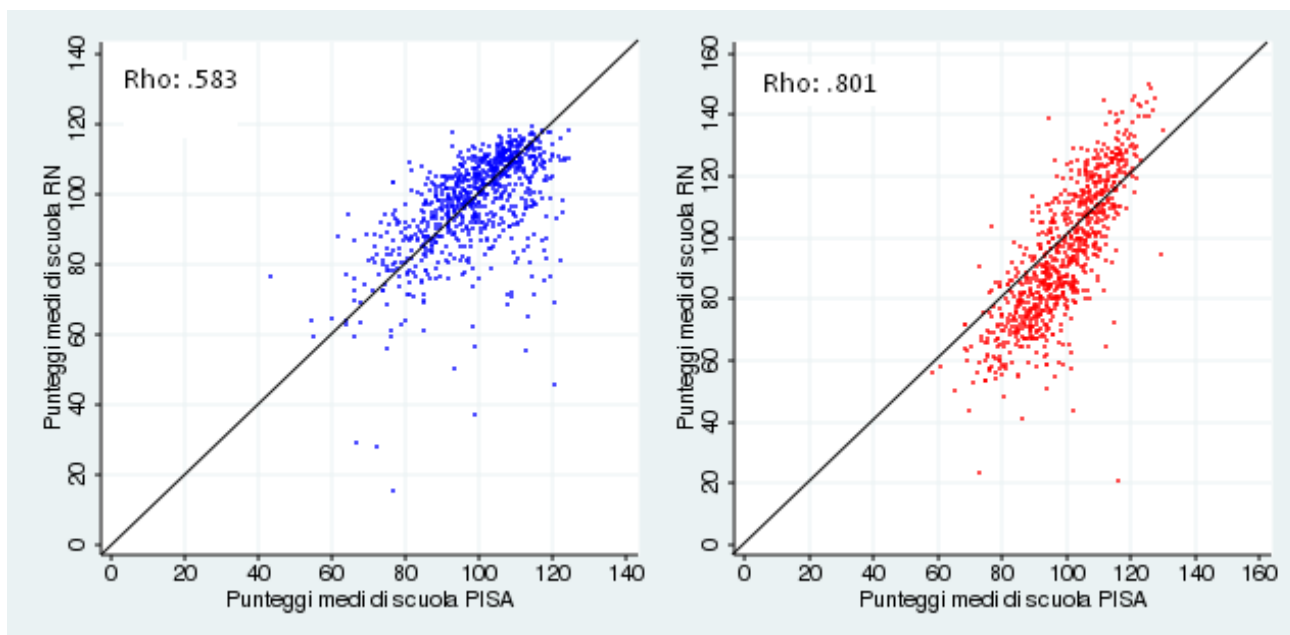
	N	Italiano (RN)	Matematica (RN)	Media italiano e matematica (RN)
<i>Rilevazioni nazionali (Invalsi)</i>				
Studenti non "abbinati"	75.931	99,9	100,1	100,0
Studenti "abbinati"	10.096	100,9	99,6	100,3
Statistica "t" Student		-4,1542	1,2195	-1,2472
Prob. $ T > t $		0,0000	0,2227	0,2123

Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN, questi ultimi corretti per il *cheating*. Sono state escluse le scuole che hanno partecipato alle Rilevazioni nazionali (Invalsi-RN) con un solo studente.

Fig. a6 – Relazione tra i punteggi medi di scuola in PISA e nelle RIL. NAZIONALI
(valori percentuali)

Lettura

Matematica

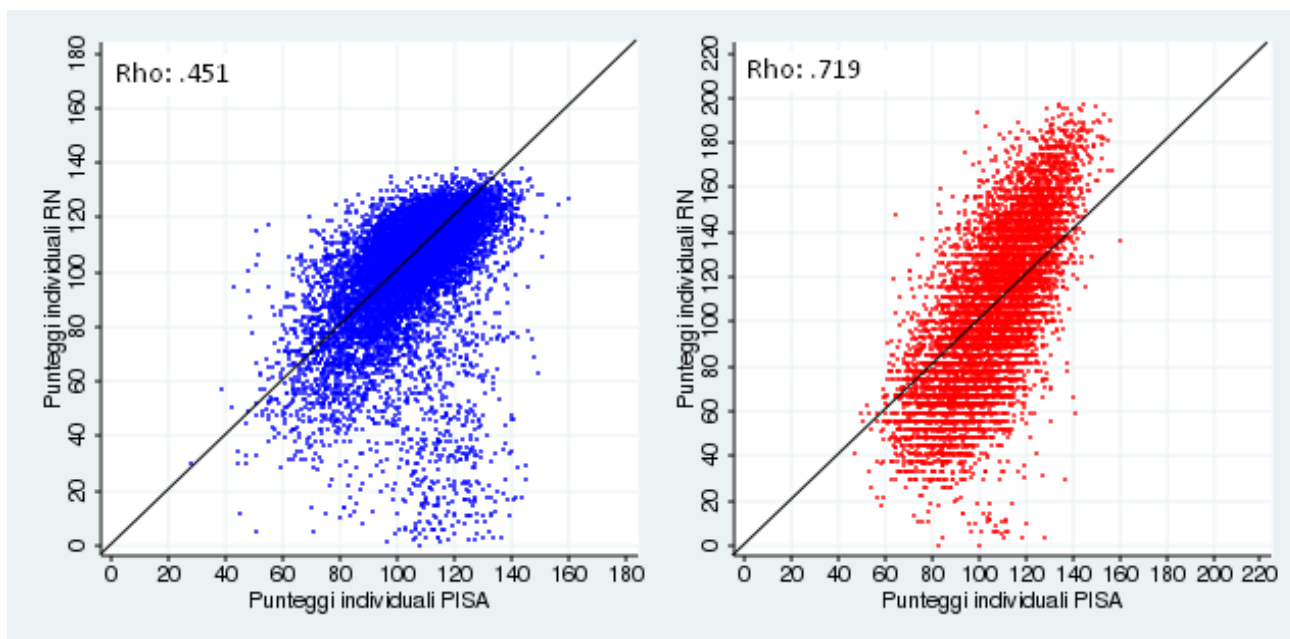


Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN, questi ultimi corretti per il *cheating*. Si tratta di 830 scuole (plessi scolastici) i cui studenti hanno partecipato a entrambe le rilevazioni, essendo 15enni in *Grade 10*. Sono state escluse le scuole che hanno partecipato alle Rilevazioni nazionali (Invalsi-RN) con un solo studente.

Fig. a7 – Relazione tra i punteggi individuali in PISA e nelle RIL. NAZIONALI
(valori percentuali)

Lettura

Matematica



Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN, questi ultimi corretti per il *cheating*. Si tratta di 10.096 studenti che hanno partecipato a entrambe le rilevazioni, essendo 15enni in *Grade 10*, e per i quali sono disponibili tutte le informazioni.

Tav. a15 – Relazione tra i punteggi medi di scuola in Lettura / Italiano riportati in PISA e nelle RILEVAZIONI NAZIONALI, per decile (unità)

		Punteggio in Lettura (PISA 2012)									
		< Q10	Q10 – Q20	Q20 – Q30	Q30 – Q40	Q40 – Q50	Q50 – Q60	Q60 – Q70	Q70 – Q80	Q80 – Q90	> Q90
Punteggio in Italiano (Invalsi-RN)	< Q10	39	17	4	3	6	2		4	2	6
	Q10 – Q20	22	21	8	8	3	6	2	2	6	5
	Q20 – Q30	14	17	17	9	7	7	4	5	2	1
	Q30 – Q40	4	8	18	15	12	5	4	4	6	7
	Q40 – Q50	1	9	14	17	14	8	7	3	3	7
	Q50 – Q60	1	6	12	13	10	18	10	4	7	2
	Q60 – Q70		3	6	9	15	13	16	8	7	6
	Q70 – Q80		2	1	5	8	9	18	13	13	14
	Q80 – Q90			1	3	4	13	13	20	16	13
	> Q90	4	2	4	2	6	4	10	22	23	23

Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN, questi ultimi corretti per il *cheating*. Si tratta di 830 scuole (plessi scolastici) i cui studenti hanno partecipato a entrambe le rilevazioni, essendo 15enni in *Grade 10*. Sono state escluse le scuole che hanno partecipato alle Rilevazioni nazionali (Invalsi-RN) con un solo studente.

Tav. a16 – Relazione tra i punteggi medi di scuola in Matematica riportati in PISA e nelle RILEVAZIONI NAZIONALI, per decile (unità)

		Punteggio in Matematica (PISA 2012)									
		< Q10	Q10 – Q20	Q20 – Q30	Q30 – Q40	Q40 – Q50	Q50 – Q60	Q60 – Q70	Q70 – Q80	Q80 – Q90	> Q90
Punteggio in Matematica (Invalsi-RN)	< Q10	46	13	6	7	4	5			1	1
	Q10 – Q20	15	27	25	10	4	1			1	
	Q20 – Q30	10	17	22	12	12	6	3	1		
	Q30 – Q40	8	9	10	21	13	18	4			
	Q40 – Q50	1	10	7	18	17	14	10	3	3	
	Q50 – Q60	1	5	5	7	15	16	18	10	4	2
	Q60 – Q70	2		3	4	10	12	17	23	10	2
	Q70 – Q80		1	4	3	5	7	14	20	22	7
	Q80 – Q90				1	2	4	10	19	27	20
	> Q90	2	3	3	1	3	2	8	9	17	52

Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN, questi ultimi corretti per il *cheating*. Si tratta di 830 scuole (plessi scolastici) i cui studenti hanno partecipato a entrambe le rilevazioni, essendo 15enni in *Grade 10*. Sono state escluse le scuole che hanno partecipato alle Rilevazioni nazionali (Invalsi-RN) con un solo studente.

Tav. a17 – Relazione tra i punteggi individuali in Lettura /Italiano riportati in PISA e nelle RILEVAZIONI NAZIONALI, per decile (unità)

		Punteggio in Lettura (PISA 2012)									
		< Q10	Q10 – Q20	Q20 – Q30	Q30 – Q40	Q40 – Q50	Q50 – Q60	Q60 – Q70	Q70 – Q80	Q80 – Q90	> Q90
Punteggio in Italiano (Invalsi-RN)	< Q10	403	148	65	52	42	62	54	61	59	101
	Q10 – Q20	304	261	175	100	71	26	33	21	20	36
	Q20 – Q30	154	217	192	178	113	69	58	27	24	15
	Q30 – Q40	74	171	193	175	147	142	90	75	46	20
	Q40 – Q50	50	96	115	153	123	139	104	88	58	34
	Q50 – Q60	21	66	106	125	162	146	154	134	84	50
	Q60 – Q70	22	44	88	93	146	141	166	159	131	65
	Q70 – Q80	10	24	57	75	118	139	177	181	190	148
	Q80 – Q90	5	11	36	65	73	105	125	163	193	190
	> Q90	4	9	20	31	54	75	89	136	241	387

Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN, questi ultimi corretti per il *cheating*. Si tratta di 10.096 studenti che hanno partecipato a entrambe le rilevazioni, essendo 15enni in Grade 10, e per i quali sono disponibili tutte le informazioni.

Tav. a18 – Relazione tra i punteggi individuali in Matematica riportati in PISA e nelle RILEVAZIONI NAZIONALI, per decile (unità)

		Punteggio in Matematica (PISA 2012)									
		< Q10	Q10 – Q20	Q20 – Q30	Q30 – Q40	Q40 – Q50	Q50 – Q60	Q60 – Q70	Q70 – Q80	Q80 – Q90	> Q90
Punteggio in Matematica (Invalsi-RN)	< Q10	436	248	148	89	47	29	31	7	8	4
	Q10 – Q20	269	236	191	145	84	70	29	12	7	4
	Q20 – Q30	142	189	186	183	146	106	52	31	8	9
	Q30 – Q40	90	139	170	177	158	117	104	59	26	3
	Q40 – Q50	51	100	141	153	188	180	141	109	60	13
	Q50 – Q60	30	55	86	114	124	139	153	136	92	27
	Q60 – Q70	13	41	57	97	114	153	176	188	156	52
	Q70 – Q80	11	25	37	44	103	116	175	212	201	123
	Q80 – Q90	3	12	24	34	55	98	131	188	256	246
	> Q90	2	3	7	10	29	38	55	104	234	564

Fonte: elaborazioni su dati OCSE-PISA e Invalsi-RN, questi ultimi corretti per il *cheating*. Si tratta di 10.096 studenti che hanno partecipato a entrambe le rilevazioni, essendo 15enni in Grade 10, e per i quali sono disponibili tutte le informazioni.