



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

WORKING PAPER N. 02/2010

**Un indicatore di status socio-economico-culturale degli
allievi della quinta primaria in Italia**

Emiliano Campodifiori, Elisabetta Figura, Monica Papini, Roberto Ricci

INVALSI

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

Abstract

Il presente lavoro illustra il metodo e le tecniche utilizzati per il calcolo di un indicatore dello status socio-economico-culturale (ESCS) degli allievi della V primaria che hanno partecipato alle rilevazioni del Servizio Nazionale di Valutazione (SNV) dell'anno scolastico 2008-2009. Il calcolo dell'ESCS si basa su indicatori discreti come il livello d'istruzione dei genitori (HISEI) e la loro condizione occupazionale (PARED), ma anche su un indicatore continuo in grado di esprimere una misura di prossimità delle condizioni materiali in cui vive l'allievo al di fuori della scuola (HOMEPOS). Quest'ultimo indicatore è calcolato partendo dai dati del campione stratificato su base regionale degli allievi della V primaria e utilizzando tecniche afferenti all'ambito metodologico della *Rasch Analysis*. Il lavoro illustra inoltre alcune soluzioni operative per superare, almeno in parte, il problema dei dati mancanti, sempre presenti nelle rilevazioni su larga scala, specie per le cosiddette variabili di contesto. Infine, si propongono alcune prime analisi circa la relazione tra l'ESCS e i livelli di apprendimento raggiunti in Italiano e Matematica.

Parole chiave: Indicatore di status socio-economico-culturale, *Rasch Analysis*, analisi delle componenti principali, imputazione dati mancanti.

Introduzione

Le caratteristiche socio-culturali ed economiche degli studenti e delle famiglie dalle quali essi provengono giocano un ruolo molto importante sui livelli di apprendimento conseguiti, sino a partire dai primi anni di scuola. E' assodato, infatti, che gli allievi che vivono in condizioni di maggiore vantaggio economico, ma anche sociale e culturale, hanno migliori possibilità di conseguire risultati più soddisfacenti durante il loro percorso formativo. E' noto ormai in letteratura che la condizione socio-culturale ed economica, il cosiddetto *background*, ha un notevole valore predittivo sui risultati conseguiti dagli allievi (INVALSI, 2008; OECD 2007). Acclarato questo aspetto, è fondamentale disporre di uno strumento, ovvero di un indicatore, che permetta di misurare in un qualche modo il *background* degli allievi oggetto d'interesse. La definizione di un indicatore dello status socio-economico-culturale e il relativo calcolo pone dei problemi di natura generale e tecnica. Essi sono affrontati nel presente lavoro prendendo a riferimento alcune soluzioni note in letteratura, in particolare nel contesto delle ricerche internazionali volte alla misurazione dei livelli di apprendimento raggiunti dagli studenti del ciclo d'istruzione primario e secondario. L'indicatore del *background* di seguito illustrato si riferisce al campione stratificato su base regionale degli allievi della scuola primaria dell'anno scolastico 2008/09 frequentanti le scuole campione¹ che hanno preso parte al Servizio Nazionale di Valutazione (SNV). Gli allievi hanno sostenuto due prove, una d'Italiano e una di Matematica, e hanno compilato un questionario studente finalizzato alla raccolta di un insieme di dati di contesto attinenti ad alcune caratteristiche socio-economico-culturali e ad aspetti di tipo motivazionale e metacognitivo. Inoltre, per ogni allievo le segreterie delle scuole hanno fornito, tra le altre, alcune informazioni sulle caratteristiche delle famiglie di provenienza, in particolare sullo stato occupazionale dei genitori e sul loro livello d'istruzione. Dopo una descrizione delle variabili sulle quali si basa il calcolo dell'indicatore di status socio-economico-culturale (ESCS²), nel presente lavoro si prende in esame il problema dei dati mancanti che, tipicamente, si pone in indagini ed analisi di questo tipo.

¹ Per la realizzazione del Servizio Nazionale di Valutazione 2008-2009 l'INVALSI ha individuato *ex ante* un campione con rappresentatività regionale di 1052 scuole primarie in cui le prove sono state somministrate alla presenza di un osservatore esterno che ha garantito il rispetto di un protocollo di somministrazione controllato e predefinito.

² Denominazione internazionale dell'indice di *background* dello studente: *Economic, Social and Cultural Status* (ESCS).

1. La costruzione dell'indicatore di status socio-economico-culturale (ESCS)

Diverse ricerche nazionali e internazionali mostrano che le condizioni personali, il cosiddetto *background* familiare, con le quali gli allievi, anche della scuola primaria, si affacciano al loro percorso formativo esercitano un ruolo importante sui livelli di apprendimento conseguiti.

Scopo del presente lavoro è la costruzione di un indice di status socio-economico-culturale, noto nella letteratura internazionale come ESCS o SES, in grado di dare ragione dell'entità dell'effetto sugli apprendimenti degli allievi della quinta classe della scuola primaria italiana del *background* familiare. Negli ultimi anni due importanti ricerche internazionali, OCSE-PISA e IEA-TIMSS, hanno mostrato in modo inequivocabile l'importanza dell'ESCS sia come variabile esplicativa dei livelli di risultati riscontrati sia per le implicazioni in termini di possibili politiche scolastiche. Sotto il profilo più propriamente tecnico-metodologico, la metodologia proposta da OCSE-PISA pare essere più convincente e articolata, sia per il numero di variabili prese in considerazione sia per le modalità di costruzione dei singoli indicatori che compongono l'ESCS stesso. Per questa ragione il presente lavoro riprende in buona parte il metodo e le nomenclature proposti da OCSE-PISA, pur con alcune modifiche dettate dalla specificità e peculiarità dei dati presi in considerazione (OECD, 2005). L'indicatore ESCS sui dati SNV 2008-2009 della quinta classe della scuola primaria è stato definito per dare ragione dell'effetto sugli apprendimenti di alcuni aspetti comunemente ritenuti rilevanti: il *background* familiare, la condizione occupazionale dei genitori e il loro di livello d'istruzione, la presenza di alcuni beni materiali che possono essere considerati una variabile di prossimità (*proxy*) della condizione della famiglia di origine degli allievi. Più precisamente, come sarà illustrato nel prosieguo del presente lavoro, l'ESCS si basa sui seguenti indicatori: lo status occupazionale dei genitori (HISEI), il livello d'istruzione dei genitori espresso in anni d'istruzione formale seguita calcolati secondo standard internazionali (PARED) e il possesso di alcuni beni materiali intesi come variabili di prossimità di un contesto economico-culturale favorevole all'apprendimento (HOMEPOS).

2. Lo status occupazionale dei genitori

La condizione occupazionale dei genitori degli allievi esercita una certa influenza sull'ambiente economico e culturale nel quale gli allievi vivono al di fuori della scuola e ciò è particolarmente rilevante per gli alunni della scuola primaria. Infatti, è noto che a questa età il peso del contesto familiare è molto importante, forse ancora di più che per altri momenti della vita degli studenti in cui il concetto di ambiente extrascolastico assume un'accezione più ampia, anche se profondamente influenzata dalla condizione socio-economica delle famiglia di provenienza. Le informazioni utilizzate per la costruzione dell'indicatore dello status occupazionale dei genitori dello studente, di

seguito indicato come HISEI³, sono state ricavate dalla “Scheda raccolta dati studenti e famiglie” attivata dall’INVALSI e compilata a cura delle segreterie delle istituzioni scolastiche iscritte alla rilevazione degli apprendimenti dell’anno 2009 (SNV 2008 -2009). Questo strumento di rilevazione ha reso possibile l’acquisizione di informazioni in possesso della scuola (voti scolastici, orari, etc.) e importanti dati relativi al *background* familiare degli studenti che non sarebbe stato possibile chiedere direttamente agli allievi (titolo di studio dei genitori, occupazione dei genitori, etc.). Nonostante l’onerosità della rilevazione e la sua non obbligatorietà, la risposta delle scuole è stata piuttosto elevata arrivando circa all’80% delle istituzioni iscritte all’SNV 2008-2009. I genitori hanno potuto definire la propria professione scegliendo tra 12 distinte categorie di impiego riportate nella scheda; i dati raccolti sono stati poi aggregati in 6 gruppi ordinati in base al crescente livello di status occupazionale⁴. Tale ricodifica ha, pertanto, definito due variabili categoriche ordinali, BFMJ⁵ (livello occupazionale del padre) e BMMJ⁶ (livello occupazionale della madre), il cui massimo definisce l’indicatore di status occupazionale della famiglia dell’allievo (HISEI).

3. Il livello d’istruzione dei genitori

Un’altra variabile relativa al *background* familiare dello studente, spesso utilizzata nelle analisi degli apprendimenti, è il livello d’istruzione dei genitori. Il titolo di studio dei genitori è infatti ritenuto in letteratura ancor più influente dello stato occupazionale sui livelli di apprendimento conseguiti dagli allievi. La definizione di una variabile o di un indicatore in grado di misurare adeguatamente il livello d’istruzione dei genitori pone alcuni problemi di tipo metodologico, specie se si opera in un contesto internazionale in cui le comparazioni sono più difficili. Nel caso d’interesse, il calcolo dell’indicatore è stata effettuato prendendo la classificazione internazionale ISCED⁷ (OECD, 1999). Il punto di partenza per la costruzione dell’indicatore d’interesse, denominato PARED, è stata la rilevazione del titolo di studio posseduto da entrambi i genitori e la sua classificazione entro sei categorie ordinate⁸. Analogamente alla costruzione dell’indice HISEI, le informazioni per l’indicatore PARED sono state acquisite con la collaborazione della segreteria della scuola interessata. Il titolo di studio secondo la codifica ISCED è stato rilevato per entrambi i genitori e tra i due titoli è stato scelto quello più elevato e poi convertito in una stima degli anni di scolarizzazione⁹ così da ottenere l’indicatore PARED propriamente detto.

³ Highest Parental Occupational Status.

⁴ Cfr. Appendice 1.

⁵ Father’s Occupational Status.

⁶ Mother’s Occupational Status.

⁷ International Standard Classification of Education.

⁸ Cfr. Appendice 1.

⁹ Cfr. Appendice 2.

4. Il benessere economico-culturale

L'utilizzo di dati relativi al possesso di determinati beni materiali è ritenuto da più ricercatori lo strumento migliore per misurare il benessere economico familiare dello studente, specie negli studi nell'ambito dell'istruzione (OECD, 2007). L'attenzione è stata focalizzata sulla rilevazione della presenza o meno nella casa dell'allievo di un certo tipo di beni; ciò, infatti, sembra catturare meglio del reddito il livello di benessere economico della famiglia in quanto il possesso di determinati beni è il riflesso di una più stabile solidità patrimoniale della famiglia. La fonte dei dati è un insieme di item tratti dal questionario somministrato agli studenti in occasione della rilevazione degli apprendimenti del 2009. Allo scopo di verificare la supposta unidimensionalità sottostante gli item presi in considerazione, sono state effettuate una serie di analisi fattoriali confermatrice su diversi sottoinsiemi dei 10 quesiti disponibili¹⁰. Alla luce dei risultati ottenuti è stato poi scelto il gruppo di domande che mostra le migliori statistiche di adattamento ai dati. In altre parole, sono state individuate le variabili ritenute più affidabili per misurare il cosiddetto *benessere economico-culturale familiare*. Gli item utilizzati per la costruzione dell'indicatore oggetto d'interesse, alcuni dei quali sono stati ricodificati in forma dicotomica (cfr. appendice 3), sono stati sintetizzati mediante una procedura di *scaling* riconducibile alla metodologia di *Rasch*, così come avviene nelle principali indagini internazionali sugli apprendimenti (Bond & Fox, 2007). Nel caso in esame il costrutto latente oggetto d'interesse è il benessere economico-culturale della famiglia e si vuole quindi definire un indicatore (HOMEPOS) espresso da un punteggio per ogni allievo in grado di dare ragione del livello di benessere di cui egli può godere. In virtù di un principio generale di parsimonia, si è preferito utilizzare variabili dicotomiche o ricodificate in questa forma poiché le analisi fattoriali condotte non hanno fornito indicazioni di significative variazioni in termini di potere esplicativo rispetto ai modelli in cui le poche variabili politomiche non sono state rese dicotomiche. Lo *scaling* delle domande di un questionario solitamente è realizzato mediante il modello di Rasch, in questo caso nella sua versione più semplice, ovvero per variabili dicotomiche:

$$P(X_i = 1 | \theta) = \frac{\exp(\theta - b_i)}{1 + \exp(\theta - b_i)} \quad (1)$$

dove $P(X_i = 1 | \theta)$ rappresenta la probabilità che un dato rispondente scelga l'opzione 1 dell'item i -esimo, dato un certo livello del tratto latente θ , mentre b_i rappresenta il posizionamento dell'item i -

¹⁰ Cfr. Appendice 3.

esimo sulla dimensione latente θ . In termini più espliciti, questa metodologia di *scaling* permette di collocare sulla stessa scala di misura sia i rispondenti sia le domande. Nel caso in esame, sulla stessa metrica è possibile esprimere sia il grado individuale di benessere economico-culturale di ciascun allievo sia il livello dell'indicatore in cui si posiziona un determinato item. In questo modo è possibile comprendere il potere segnaletico di ciascun item rispetto al costrutto che si vuole misurare.

Tavola 1: Stime WLE del parametro b_i ($i = 1, \dots, 9$) (posizionamento dell'item i -mo sul tratto latente θ) per ciascuno dei 9 quesiti costituenti l'indice HOMEPOS.

Nella tua abitazione possiedi:	Stime di b_i
a) Un posto tranquillo per studiare	-2,196
b) Una scrivania per fare i compiti	-2,157
c) Enciclopedie (composte da libri o su CD o DVD)	-1,125
d) Un collegamento ad Internet per fare ricerche	-1,102
e) Un allarme antifurto	0,850
f) Una camera solo per te	-0,137
g) Più di un bagno	-0,570
h) La tua famiglia possiede più di un'automobile	-0,998
i) Più di cento libri	1,061

Nella seconda colonna della Tavola 1 sono riportate le stime basate sulla funzione di verosimiglianza e calcolate mediante il metodo di Warm (*Weighted Likelihood Estimates*, WLE) che, come noto, permette di correggere in modo soddisfacente la distorsione propria del metodo di massima verosimiglianza (Warm, 1989). Tali stime sono centrate in linea teorica in modo che il loro valore medio sia 0 e la deviazione standard sia unitaria. Da ciò segue che item con un b_i basso sono indicatori di livelli bassi del *background* familiare e, viceversa, quelli con un b_i più alto. Per ogni item è stato inoltre valutata la bontà di adattamento del modello teorico con i dati¹¹ mediante l'indice *weighted mean-square statistic (infit)* che si mantiene all'interno dei margini ritenuti comunemente accettabili (Wu *et al.*, 1997), ovvero non assumono valori significativamente superiori ad 1.

¹¹ Esula dagli scopi del presente lavoro la distinzione teorico-metodologica tra la *Rasch Analysis* e l'*Item Response Theory* circa le assunzioni teoriche per la valutazione della corrispondenza tra dati e modelli. In questa sede ci si è limitati al controllo dell'indice *weighted mean-square statistic* all'interno delle comuni soglie di accettabilità.

5. I dati mancanti

In tutte le rilevazioni su larga scala i dati mancanti rappresentano un problema delicato che richiede particolare attenzione per contenere l'effetto distorsivo che essi producono sulle stime. La presenza di dati mancanti è solitamente più accentuata quando le informazioni raccolte afferiscono ad alcune sfere, come, ad esempio, quelle relative al possesso di beni materiali o al titolo di studio. Se da un lato è vero che il fenomeno è tanto più contenuto, quanto maggiore è l'adesione alle finalità della ricerca, d'altro canto il problema riveste sempre e comunque una certa importanza. Anche la raccolta delle informazioni di contesto affidate alle segreterie delle scuole per il SNV 2008-2009 non fa eccezione rispetto ai suddetti problemi, anche se i tassi di risposta ottenuti, se paragonati con ricerche analoghe, si sono mantenuti su livelli soddisfacenti, sia per le scuole campione che per quelle volontarie. Se si circoscrive l'attenzione sulle scuole campione in cui tutto il protocollo di realizzazione della ricerca è stato costantemente monitorato e che sono state utilizzate per il calcolo dell'ESCS, le informazioni acquisite mediante la "Scheda raccolta dati studenti e famiglie" riguardano potenzialmente un collettivo di 43516 studenti. La Tavola 2 riporta come si sono distribuite le mancate risposte rispetto alle variabili oggetto d'interesse.

Tavola 2: Distribuzione delle mancate risposte rispetto alle variabili di interesse.

Schede per tipologia di dati mancanti	Numero di schede
Schede con dati mancanti sulla professione e sul titolo di studio di entrambi i genitori	11195
Schede con dati mancanti solo sulla professione di entrambi i genitori	294
Schede con dati mancanti solo sul titolo di studio di entrambi i genitori	1127

Diverse sono le modalità con le quali si cerca di superare il problema dei dati mancanti, cercando di contenere gli effetti distorsivi che essi producono sulle stime proposte. In particolare, nel presente lavoro si è fatto riferimento alle prescrizioni dell'ISTAT e dell'OCSE (ISTAT, 2010; OECD 2007)

L'imputazione dei dati mancanti ha seguito un approccio diverso a seconda della mancanza contemporanea per il singolo studente del titolo di studio e della professione dei genitori o di una sola delle due variabili. Nel primo caso i dati sono stati completati, per ciascun alunno, imputando come indicatore dello stato occupazionale della famiglia (HISEI) e come indicatore del livello d'istruzione familiare (PARED) i rispettivi valori modali calcolati a livello d'istituto o, qualora non fossero disponibili, a livello regionale. Nel secondo caso si è proceduto ad una imputazione deterministica per regressione: sono state calcolate le regressioni lineari delle variabili mancanti (HISEI o PARED) sulle variabili presenti (HISEI o PARED, HOMEPOS) utilizzando il collettivo

di 30900 osservazioni complete¹². La valutazione dell'impatto delle imputazioni dei dati mancanti sulla misura costruita, l'indicatore di status socio-economico-culturale, è stata effettuata *ex post* mettendo a confronto la correlazione lineare tra le stime dei livelli di apprendimento e l'ESCS ottenuto con o senza i dati mancanti. Poiché la forza associativa con le misure di prestazione in Matematica e Italiano si riduce solo in minima parte in seguito all'imputazione dei dati mancanti, si è optato per perseguire questa soluzione. In questo modo è stato possibile associare ad ogni studente un valore dell'indicatore ESCS. Tuttavia, se questa modalità di procedere trova ampia conferma in letteratura, è bene rimarcare come essa mantenga caratteristiche di affidabilità e robustezza quando, come nel caso in esame, lo scopo è quello di effettuare analisi a livello nazionale o regionale. Sarebbe ovviamente diverso il caso in cui l'obiettivo fosse un'analisi a livello di scuole. In tale situazione è certamente preferibile mantenere come mancanti i casi in cui, ad esempio, una scuola intera non abbia fornito alcuna risposta alla "Scheda raccolta dati studenti e famiglie".

6. L'indice di status socio-economico-culturale (ESCS)

Diversi studi hanno evidenziato un legame, anche forte, fra il rendimento scolastico degli studenti e il loro *background* familiare (Brunello & Checchi, 2006; Carneiro & Heckman 2003; Marks, Cresswell, & Ainley, 2006). Come richiamato in precedenza, anche le ricerche IEA-TIMSS e OCSE-PISA confermano a diversi livelli questo legame sia a livello internazionale sia nazionale (INVALSI, 2008; Martini & Ricci, 2007; 2010). E', quindi, molto importante disporre di un indicatore che sia in grado di fornire una misura robusta e adeguata della condizione socio-economico-culturale di ciascun allievo e, di conseguenza, del *background* familiare medio di ciascuna scuola. L'indicatore proposto nel presente lavoro è stato definito e, conseguentemente calcolato, seguendo le indicazioni metodologiche che hanno trovato maggiore riscontro in ambito internazionale. L'ESCS è stato costruito in funzione dei tre indicatori illustrati in precedenza: lo status occupazionale dei genitori (HISEI), il loro livello d'istruzione (PARED) e il possesso o meno di un ventaglio di particolari beni strumentali e culturali (HOMEPOS).

12 Metodo suggerito per la prima volta da S. F. Buck nel 1960; l'imputazione è definita "deterministica" poiché il residuo è posto nullo per ipotesi; in dettaglio, il valore mancante dell'*i*-mo studente è imputato con il valore predittivo della regressione:

$$\hat{Y}_i = b_0 + \sum_{j=1}^k b_j X_{ij} + e_i$$

con b_0, b_j ($j = 1, \dots, k$) coefficienti di regressione della variabile mancante Y sulle variabili presenti X_j ed e_i residuo.

L'indicatore ESCS è ricavato mediante un'analisi in componenti principali dei tre indicatori HISEI, PARED e HOMEPOS. In particolare, constatato che la prima componente estratta è in grado di spiegare oltre il 56% della varianza complessiva e che solo ad essa è associato un auto valore maggiore di 1, in linea con quanto proposto da OCSE-PISA, sono stati assunti come valori dell'ESCS i punteggi associati alla prima componente principale. In base alla procedura sinteticamente illustrata, si è ottenuto un indicatore a media nulla e deviazione standard unitaria. In questo modo, ad esempio, un allievo con un valore individuale dell'ESCS strettamente positivo è uno studente con un *background* socio-economico-culturale più favorevole della media italiana. L'entità dello scostamento dalla media nazionale può essere valutata in funzione della deviazione standard. Sempre per proseguire nell'esempio, se l'allievo d'interesse presenta un ESCS pari a 1,5, esso è superiore di un'unità e mezzo di deviazione standard rispetto alla media dell'intero Paese. La valutazione della capacità predittiva dell'indicatore ESCS rispetto alle misure degli apprendimenti è stata effettuata in termini di correlazione lineare. In particolare, l'ESCS calcolato con i dati mancanti imputati ha una correlazione pari a 0,25 per l'Italiano e a 0,21 per la Matematica, quindi valori non lontani da quelli che si possono riscontrare sui dati OCSE-PISA. Lo scopo del presente lavoro, come più volte richiamato, è quello di costruire un indice di *background* e di conseguenza vedere come esso si associ alle misure dei livelli di apprendimento. Una prima semplice idea dell'associazione tra ESCS e livelli di apprendimento si può desumere dai dati della Tavola 3.

Tavola 3: Punteggi medi nei quartili dell'indicatore ESCS per le 5 macroaree italiane.

Macroarea ¹³	Quartili dell'indicatore ESCS	Punteggi medi di Italiano	Punteggi medi di Matematica
Nord-ovest	I	473	477
	II	499	502
	III	513	511
	IV	528	529
Nord-est	I	472	486
	II	497	510
	III	511	519
	IV	530	536
Centro	I	465	466
	II	493	495
	III	506	502
	IV	526	521
Sud	I	481	486
	II	496	497
	III	520	520
	IV	540	524
Sud-Isole	I	443	441
	II	480	471
	III	499	487
	IV	529	507

La Tavola 3 riporta i valori dei punteggi¹⁴ d'Italiano e Matematica nei quattro quartili dell'indicatore ESCS. E' evidente l'andamento decisamente crescente dei punteggi medi di entrambe le discipline in relazione a migliori condizioni socio-economico-culturali degli allievi. In altri termini, come evidenziato già in molte ricerche nazionali e internazionali, il miglioramento delle condizioni culturali e materiali di cui un allievo può godere si riflette in un sensibile miglioramento dei livelli medi di apprendimento, con un divario di oltre mezza unità di deviazione standard tra i risultati medi di coloro che godono di un background familiare meno favorevole e coloro che invece possono avvantaggiarsi delle condizioni migliori. La Figura 1 e la Figura 2 permettono di visualizzare i dati della Tavola in chiave comparativa rispetto all'intero Paese. E'

¹³ *Nord-ovest*: Valle d'Aosta, Piemonte, Liguria, Lombardia. *Nord-est*: Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna. *Centro*: Toscana, Umbria, Marche, Lazio. *Sud*: Abruzzo, Molise, Campania, Puglia. *Sud-Isole*: Basilicata, Calabria, Sicilia, Sardegna.

¹⁴ I punteggi della Tavola 3 sono ottenuti mediante il modello di Rasch e standardizzati a livello nazionale con media 500 e deviazione standard 100.

interessante notare come il divario che si riscontra solitamente tra Nord e Sud del Paese sia prevalentemente identificabile nelle regione della macroarea Sud-Isole, piuttosto che nel Sud. Inoltre, si nota come tale tendenza sia più evidente per la Matematica che per l'Italiano. Infine è importante notare che le differenze tra i punteggi medi nei diversi livelli (quartili) dell'ESCS sono più rimarcate in quelli bassi piuttosto che in quelli alti.

Figura 1: Punteggi medi in Italiano nei quartili dell'indicatore ESCS per le 5 macroaree italiane

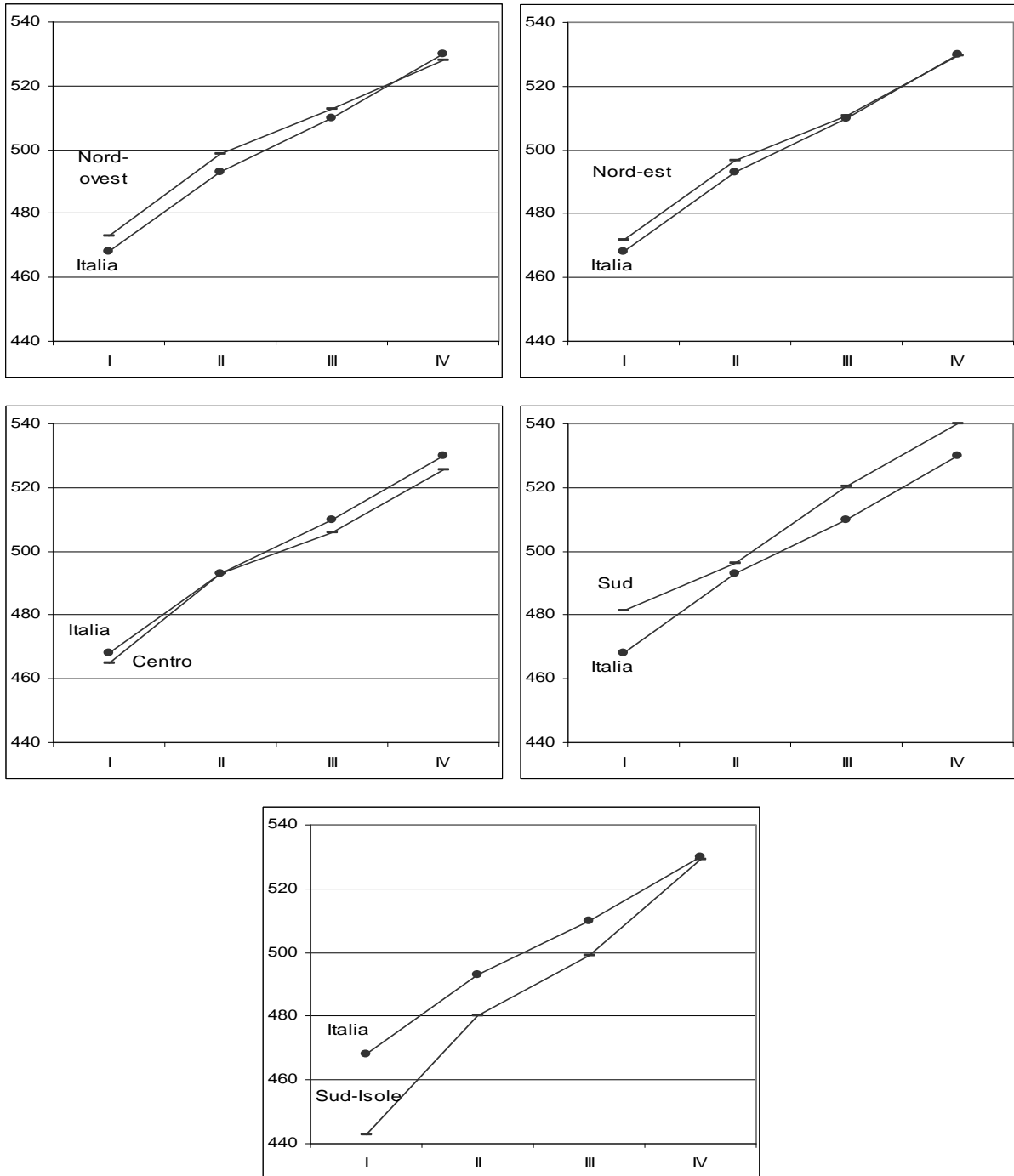
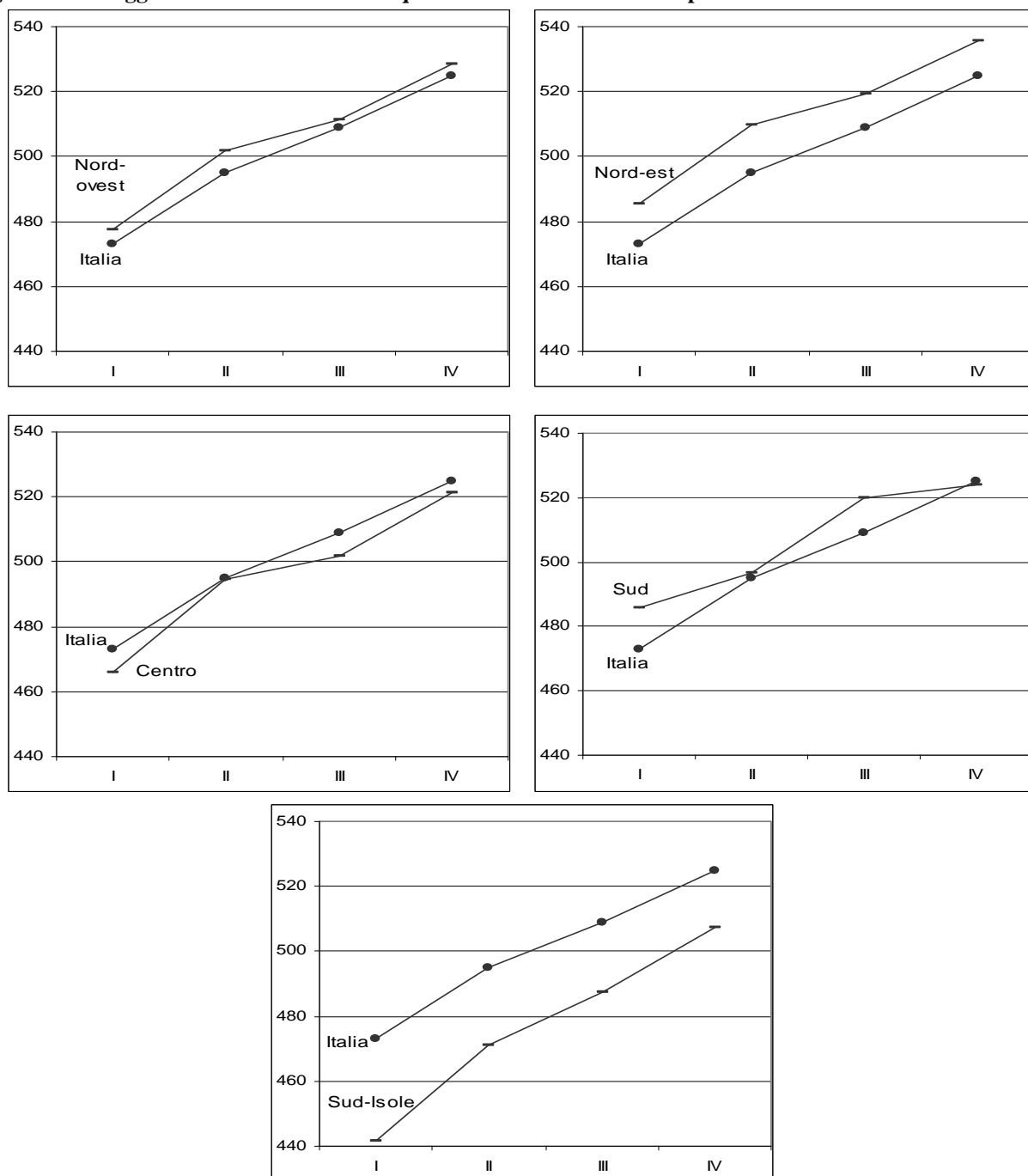


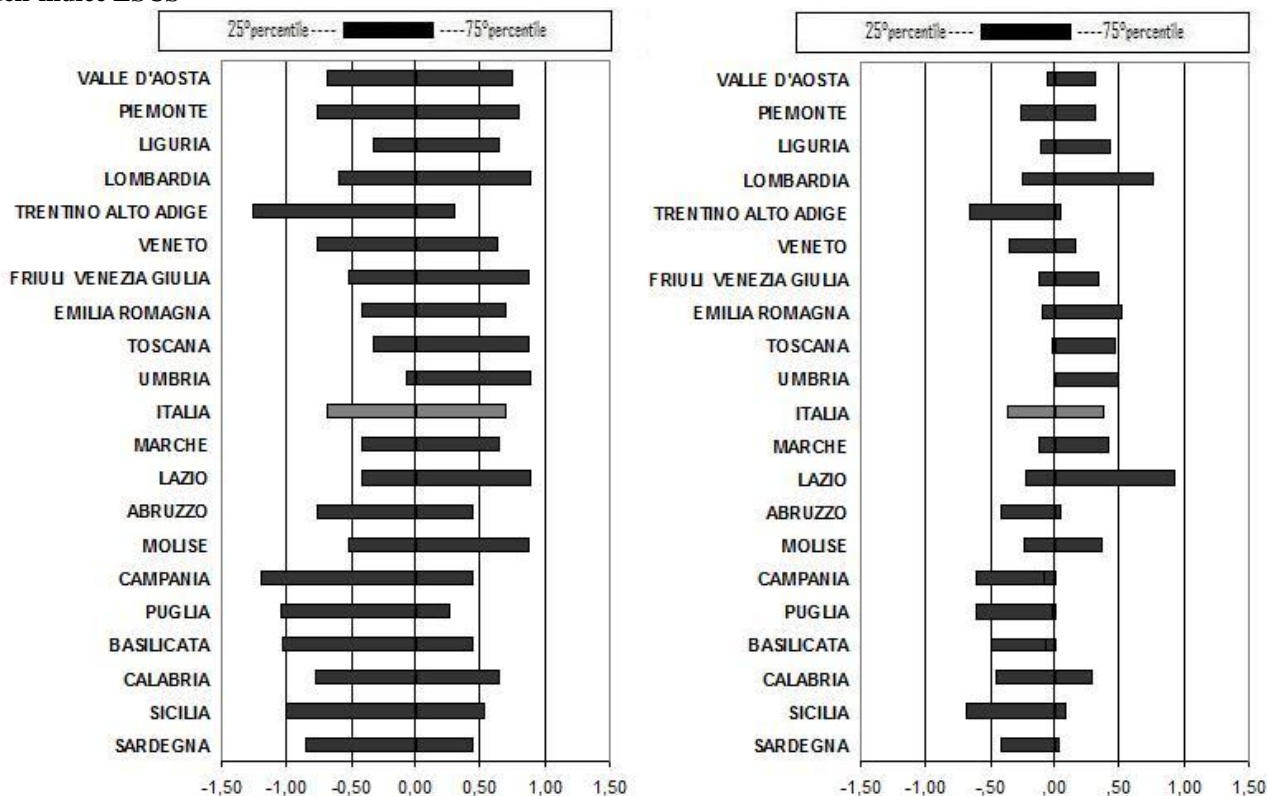
Figura 2: Punteggi medi in Matematica nei quartili dell'indicatore ESCS per le 5 macroaree italiane



Oltre ad una prima analisi dell'associazione tra valori dell'ESCS e rendimenti medi d'Italiano e Matematica, è importante analizzare come la popolazione degli alunni si distribuisce rispetto alle proprie caratteristiche socio-economico-culturali, per come esse vengono rappresentate e misurate mediante l'ESCS stesso. La Figura 3 mette a confronto il campo di variazione esistente tra il primo ed il terzo quartile dell'ESCS nel Paese e nelle regioni sia livello di singolo allievo sia a livello medio di scuola. Il restringimento delle distribuzioni alla distanza interquartile è motivato dall'esigenza di eliminare i valori estremi della distribuzione, ovvero condizioni di particolare

vantaggio o svantaggio sotto il profilo sociale, economico e culturale. In questo modo la lunghezza delle barre verticali permette di comparare in maniera più robusta la variabilità dell'ESCS nelle singole regioni e nel Paese.

Figura 3: Variabilità degli studenti (grafico di sinistra) e delle scuole (grafico di destra) nella distribuzione dell'indice ESCS



Prima di qualsiasi valutazione, è opportuno rimarcare come l'indicatore ESCS sia più saturo sotto il profilo socio-culturale che economico e ciò spiega come regioni che godono di condizioni economiche particolarmente vantaggiose, mostrano valori individuali e aggregati di scuola piuttosto bassi. A questo proposito è emblematico il caso del Trentino-Alto Adige che, notoriamente, gode di livelli di reddito molto alti, ma la cui popolazione adulta dispone in misura minore di titoli di studio elevati. Come era da attendersi, la Figura 3 mostra che le regioni dell'Italia centro-settentrionale mostrano sia a livello individuale che di scuola distanze interquartiliche dell'ESCS più favorevoli rispetto alle regioni del Mezzogiorno. Tale tendenza si riscontra in maniera ancora più evidente a livello medio di scuola rispetto al quale alcune regioni del Sud e le Isole mostrano distanze interquartiliche quasi interamente collocate sui valori negativi dell'ESCS. Anche se lo scopo del presente lavoro è principalmente limitato alla descrizione tecnico-metodologica delle procedure in base alle quali è stato costruito l'ESCS, è opportuno anticipare alcune riflessioni che saranno oggetto di ulteriori approfondimenti.

Tavola 4: Punteggi medi rispetto all'ESCS¹⁵.

Macroarea ¹⁶	(1) ESCS medio ¹⁷	(2) Percentuale di studenti sotto il 15° percentile nazionale di ESCS	(3) Punteggio medio Italiano	(4) Punteggio medio d'Italiano per ESCS uguale alla media nazionale	(5) Punteggio medio Matematica	(6) Punteggio medio di Matematica per ESCS uguale alla media nazionale
Nord-ovest	0,13	13,0	505	502	507	504
Nord-est	0,03	12,2	502	502	513	512
Centro	0,28	8,6	503	496	501	495
Sud	-0,27	21,1	503	509	502	506
Sud-Isole	-0,17	18,9	482	487	472	476

I dati della Tavola 4 suggeriscono alcune piste di riflessione e ulteriori approfondimenti, in particolare circa i legami associativi tra l'ESCS e i risultati conseguiti dagli allievi di V primaria raggruppati per macroarea, mentre la stessa analisi a livello regionale è riportata nell'appendice 4.

In particolare, si nota la profonda differenza all'interno del Paese della distribuzione degli allievi in condizioni di maggior difficoltà sotto il profilo socio-economico-culturale. Presa la soglia dell'ESCS che individua a livello nazionale il 15° percentile inferiore, si nota come nel Sud gli studenti al di sotto di questo livello sono oltre il 21% e nella area Sud-Isole quasi il 19%. Ancor più interessante è vedere quale sarebbe il punteggio in Italiano e Matematica a parità di condizioni socio-economico-culturali. Infatti le colonne (4) e (6) permettono di confrontare quale sarebbero stati i risultati in Italiano e Matematica, rispettivamente, se tutte le macroaree avessero avuto lo stesso livello di ESCS, ovvero un indicatore di status socio-economico-culturale pari alla media nazionale. In primo luogo si nota che i risultati delle macro-aree settentrionali non si modificano di molto se si considerano livelli di ESCS pari alla media nazionale, mentre quelli delle aree centro-meridionali si modificano in maniera più sensibile. E' interessante inoltre notare che se si elimina l'effetto delle differenti condizioni socio-economico-culturali, i risultati del Sud diventano leggermente superiori a quelli del resto del Paese, anche se in termini non statisticamente significativi, mentre quelli dell'area Sud-Isole permangono sia in Italiano che in Matematica significativamente più bassi di quelli italiani complessivamente considerati.

¹⁵ Cfr. appendice 4 per le disaggregazioni regionali.

¹⁶ *Nord-ovest*: Valle d'Aosta, Piemonte, Liguria, Lombardia. *Nord-est*: Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna. *Centro*: Toscana, Umbria, Marche, Lazio. *Sud*: Abruzzo, Molise, Campania, Puglia. *Sud-Isole*: Basilicata, Calabria, Sicilia, Sardegna.

¹⁷ L'ESCS medio per l'ITALIA è pari a 0 e la deviazione standard è pari a 1.

7. Considerazioni conclusive

La definizione e il calcolo di un indicatore di status socio-economico-culturale riveste un'importanza notevole e sotto il profilo strettamente metodologico e sotto quello più propriamente interpretativo. Il ricorso a tecniche che permettano di utilizzare sia indicatori discreti che continui consente di utilizzare un più ampio ventaglio di variabili che si è riscontrato essere buoni indicatori di prossimità del costrutto oggetto d'indagine. Come sovente accade, la presenza significativa di dati mancanti richiede il ricorso a tecniche d'imputazione pur nella consapevolezza che ciò introduce, anche se in misura contenuta, un effetto distorsivo delle stime proposte. L'indicatore dello status socio-economico-culturale permette inoltre di approfondire le analisi dei risultati conseguiti dagli allievi della V primaria nelle prove d'Italiano e di Matematica. Per quanto tali approfondimenti non siano l'oggetto specifico del presente lavoro, alcune piste d'indagine e di analisi appaiono in tutta evidenza. Così come già verificato in altre ricerche nazionali ed internazionali, il *background* individuale degli allievi si associa in maniera evidente con i livelli di apprendimento conseguiti. Inoltre, emerge con chiarezza che mentre in alcune aree del Paese il divario nelle condizioni economico-sociali e culturali sembrano poter spiegare buona parte delle differenze riscontrate, per altre regioni le richiamate differenze, anche rilevanti, non paiono fornire una spiegazione soddisfacente della diversità dei livelli di apprendimento raggiunti. Proprio queste differenze, riscontrabili sia a livello individuale che di scuola, suggeriscono nuove piste di approfondimento e ricerca che costituiranno oggetto di analisi specifica in altri lavori promossi dall'INVALSI.

Riferimenti bibliografici

- Bond, T. G., & Fox, R. (2007). *Applying The Rasch Model, Fundamental Measurement in the human Sciences*. Mahwah, N.J.: Lawrence Erlbaum Associates Publishers.
- Brunello, G., & Checchi, D. (2006). *Does school tracking affect equality of opportunity? New international evidence*. Bonn: Forschungsinstitut zur Zukunft der Arbeit.
<http://d-nb.info/994962398/34> .
- Carneiro, P. & Heckman, J. (2003). *Human capital policy*. Cambridge, MA: National Bureau of Economic Research.
- INVALSI (2008). *Le competenze in scienze, lettura e Matematica degli studenti quindicenni. Rapporto nazionale PISA 2006*. Roma: Armando Editore.
- Fortini, M. (2000). *Linee guida metodologiche per rilevazioni statistiche*. Roma: ISTAT.
<http://www.istat.it/strumenti/metodi/lineeguida.pdf>
- Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining Socioeconomic Inequalities in Student Achievement: The role of home and school factors. *Educational Research and Evaluation*, 12(2), 105-128.
- Martini A., & Ricci R. (2007). I risultati PISA 2003 degli studenti italiani in Matematica: un'analisi multilivello per tipologia di scuola secondaria. *Induzioni*, 34, 73-93.
- Martini A., & Ricci R. (2010). Un esperimento di misurazione del valore aggiunto delle scuole sulla base dei dati PISA 2006 del Veneto. *Rivista di Economia e Statistica del territorio*, (lavoro accettato ed in corso di pubblicazione).
- OECD (1999). *Classifying Educational Programmes: Manual for ISCED-97 Implementation in OECD Countries*. Paris: OECD.
- OECD (2005). *PISA 2003: Technical Report*. Paris: OECD.
- OECD (2007). *PISA 2006: Science Competencies for Tomorrow's World. Vol. 1 e 2*. Paris: OECD.
- Warm, T.A. (1989). Weighted Likelihood Estimation of Ability in Item Response Theory. *Psychometrika*. 54(3), 427-450.
- Wu, M.L., Adams, R.J., & Wilson, M.R. (1997). *ConQuest: Multi-Aspect Test Software*, [computer program]. Camberwell: Australian Council for Educational Research.

Appendice 1

La scheda genitori (Fig. A1.1) è stata distribuita dalle segreterie scolastiche alle famiglie degli studenti coinvolti nella rilevazione degli apprendimenti 2008-2009.

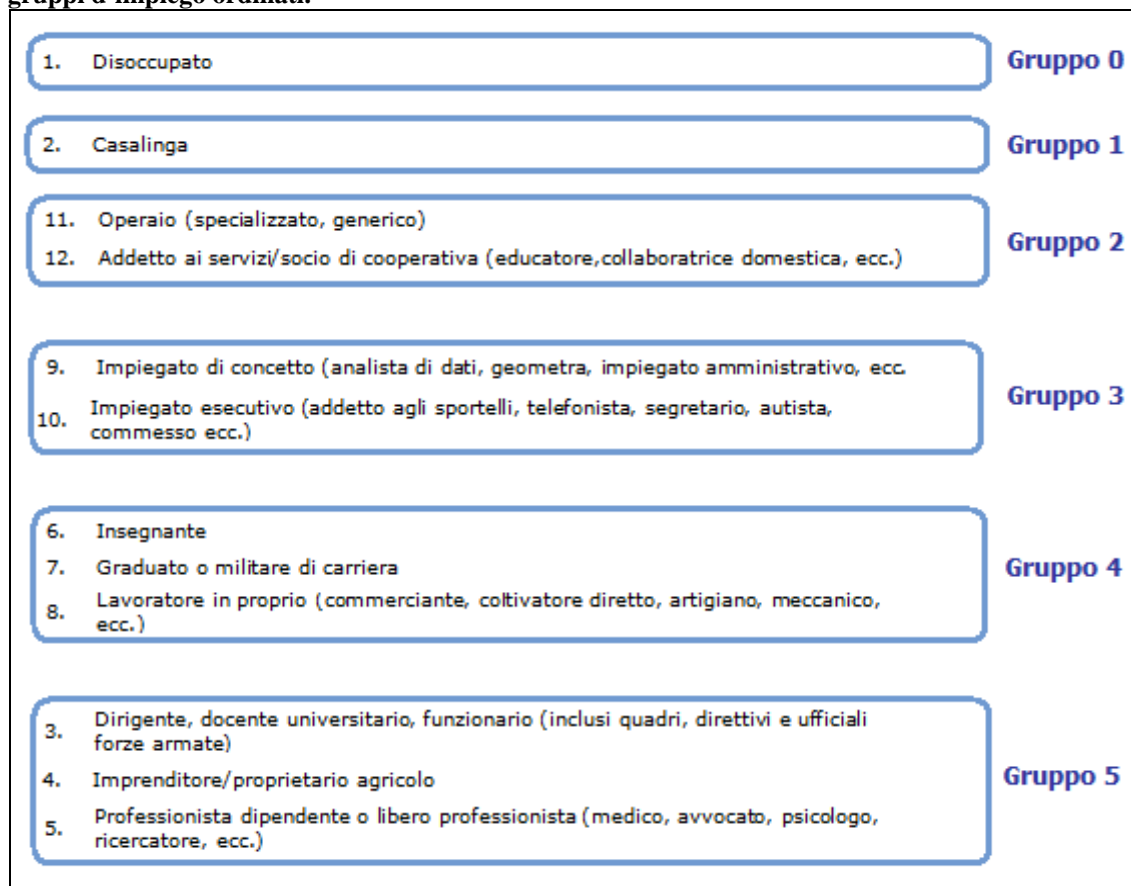
Figura A1.1 : Scheda per la raccolta dati dei genitori (estratto).

Scheda genitori (o tutori)			
Codice studente: _____			
1. Luogo di nascita			
	Studente	Madre	Padre
1. Italia	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Unione Europea (Austria, Belgio, Bulgaria, Cipro, Danimarca, Estonia, Finlandia, Francia, Germania, Grecia, Irlanda, Lettonia, Lituania, Lussemburgo, Malta, Paesi Bassi, Polonia, Portogallo, Regno Unito, Repubblica ceca, Romania, Slovacchia, Slovenia, Spagna, Svezia, Ungheria)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Paese europeo non UE	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Altro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Se lo studente non è nato in Italia, indicare l'età di arrivo in Italia			
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1 anno	2 anni	3 anni	4 anni
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5 anni	6 anni	7 anni	8 anni
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9 anni	10 anni o più	<input type="checkbox"/>	<input type="checkbox"/>
3. Titolo di studio			
		Madre	Padre
1. Licenza elementare		<input type="checkbox"/>	<input type="checkbox"/>
2. Licenza media		<input type="checkbox"/>	<input type="checkbox"/>
3. Qualifica professionale triennale		<input type="checkbox"/>	<input type="checkbox"/>
4. Diploma di maturità		<input type="checkbox"/>	<input type="checkbox"/>
5. Altro titolo di studio superiore al diploma (I.S.E.F., Accademia di Belle Arti, Conservatorio)		<input type="checkbox"/>	<input type="checkbox"/>
6. Laurea		<input type="checkbox"/>	<input type="checkbox"/>
4. Professione			
		Madre	Padre
1. Disoccupato		<input type="checkbox"/>	<input type="checkbox"/>
2. Casalinga		<input type="checkbox"/>	<input type="checkbox"/>
3. Dirigente, docente universitario, funzionario (inclusi quadri, direttivi e ufficiali forze armate)		<input type="checkbox"/>	<input type="checkbox"/>
4. Imprenditore/proprietario agricolo		<input type="checkbox"/>	<input type="checkbox"/>
5. Professionista dipendente o libero professionista (medico, avvocato, psicologo, ricercatore, ecc.)		<input type="checkbox"/>	<input type="checkbox"/>
6. Insegnante		<input type="checkbox"/>	<input type="checkbox"/>
7. Graduato o militare di carriera		<input type="checkbox"/>	<input type="checkbox"/>
8. Lavoratore in proprio (commerciante, coltivatore diretto, artigiano, meccanico, ecc.)		<input type="checkbox"/>	<input type="checkbox"/>
9. Impiegato di concetto (analista di dati, geometra, impiegato amministrativo, ecc.)		<input type="checkbox"/>	<input type="checkbox"/>
10. Impiegato esecutivo (addetto agli sportelli, telefonista, segretario, autista, commesso ecc.)		<input type="checkbox"/>	<input type="checkbox"/>
11. Operaio (specializzato, generico)		<input type="checkbox"/>	<input type="checkbox"/>
12. Addetto ai servizi/socio di cooperativa (educatore, collaboratrice domestica, ecc.)		<input type="checkbox"/>	<input type="checkbox"/>

Le informazioni riportate nella scheda, assieme a quelle inerenti alcuni aspetti didattici (ore di lezione, voti ecc.), hanno costituito l'insieme dei dati di contesto forniti dalle scuole all'INVALSI

tramite il modulo *on-line* “Scheda raccolta dati studenti e famiglie”. La sezione riguardante la scelta della professione dei genitori prevede 12 categorie generali d’impiego che, in sede di analisi dei dati raccolti, sono state aggregate in 6 gruppi tra loro ordinati rispetto al contenuto professionale e retributivo (Fig. A1.2)

Figura A1. 2: Corrispondenza tra le 12 categorie professionali della “scheda genitori” e i 6 gruppi d’impiego ordinati.



La sezione relativa alla selezione del titolo di studio posseduto prevedeva, come si vede chiaramente in figura 1, sei categorie ordinate in base al livello del percorso di studi raggiunto: licenza elementare, licenza media, qualifica professionale triennale, altro titolo di studio superiore al diploma, laurea.

Appendice 2

La Tavola in figura A2.1 riporta la corrispondenza per diversi paesi tra titolo di studio e anni di istruzione formale secondo l'ultima versione della *International Standard Classification of Education* (ISCED 97). Tale classificazione, utilizzata da OCSE, UNESCO e EUROSTAT, suddivide i programmi didattici in sette livelli ed è alla base per costruire indicatori statistici per il confronto internazionale nel campo dell'istruzione. Nel caso in esame è stata utilizzata per convertire il titolo di studio più elevato tra i due genitori in una stima degli anni d'istruzione formale.

Figura A2. 1 : Tavola per la conversione del titolo di studio in anni di scuola frequentati¹⁸

	ISCED 1	ISCED 2	ISCED 3B or 3C	ISCED 3A or 4	ISCED 5B	ISCED 5A or 6
Australia	6.0	10.0	11.0	12.0	14.0	15.0
Austria	4.0	9.0	12.0	12.5	15.0	17.0
Belgium	6.0	9.0	12.0	12.0	14.5	17.0
Canada	6.0	9.0	12.0	12.0	15.0	17.0
Czech Republic	5.0	9.0	11.0	13.0	16.0	16.0
Denmark	6.0	9.0	12.0	12.0	15.0	17.0
England, Wales & North. Ireland	6.0	9.0	12.0	13.0	15.0	16.0
Finland	6.0	9.0	12.0	12.0	14.5	16.5
France	5.0	9.0	12.0	12.0	14.0	15.0
Germany	4.0	10.0	13.0	13.0	15.0	18.0
Greece	6.0	9.0	11.5	12.0	15.0	17.0
Hungary	4.0	8.0	10.5	12.0	13.5	16.5
Iceland	7.0	10.0	13.0	14.0	16.0	18.0
Ireland	6.0	9.0	12.0	12.0	14.0	16.0
Italy	5.0	8.0	12.0	13.0	16.0	17.0
Japan	6.0	9.0	12.0	12.0	14.0	16.0
Korea	6.0	9.0	12.0	12.0	14.0	16.0
Luxembourg	6.0	9.0	12.0	13.0	16.0	17.0
Mexico	6.0	9.0	12.0	12.0	14.0	16.0
Netherlands	6.0	10.0		12.0		16.0
New Zealand	5.5	10.0	11.0	12.0	14.0	15.0
Norway	6.0	9.0	12.0	12.0	14.0	16.0
Poland		8.0	11.0	12.0	15.0	16.0
Portugal	6.0	9.0	12.0	12.0	15.0	17.0
Scotland	7.0	11.0	13.0	13.0	16.0	16.0
Slovak Republic	4.5	8.5	12.0	12.0	13.5	17.5
Spain	5.0	8.0	10.0	12.0	13.0	16.5
Sweden	6.0	9.0	11.5	12.0	14.0	15.5
Switzerland	6.0	9.0	12.5	12.5	14.5	17.5
Turkey	5.0	8.0	11.0	11.0	13.0	15.0
United States	6.0	9.0		12.0	14.0	16.0

¹⁸ Fonte: PISA 2006 TECHNICAL REPORT, OECD 2009

Appendice 3

Le informazioni utilizzate per costruire l'indice HOMEPOS sono state tratte da un insieme di item presenti nel questionario somministrato agli studenti nella rilevazione degli apprendimenti del 2008-2009.

Il gruppo di item "Q6" (fig. A3.1) si pone come obiettivo quello di indagare le condizioni familiari di supporto allo studio (presenza di un PC per studiare, collegamento ad internet, etc.):

Figura A3. 1: ITEM 6¹⁹

6. A casa tua ci sono:		<i>Metti una crocetta su un solo quadratino per ogni riga.</i>	
		<i>Si</i>	<i>No</i>
A.	Un posto tranquillo per studiare	<input type="checkbox"/>	<input type="checkbox"/>
B.	Un computer che puoi usare per lo studio	<input type="checkbox"/>	<input type="checkbox"/>
C.	Una scrivania per fare i compiti	<input type="checkbox"/>	<input type="checkbox"/>
D.	Enciclopedie (composte da libri o su CD o DVD)	<input type="checkbox"/>	<input type="checkbox"/>
E.	Un collegamento a Internet per fare ricerche	<input type="checkbox"/>	<input type="checkbox"/>
F.	Un allarme antifurto	<input type="checkbox"/>	<input type="checkbox"/>
G.	Una camera solo per te	<input type="checkbox"/>	<input type="checkbox"/>

Il gruppo di item "Q7" (fig. A3.2) presenta quesiti correlati al benessere materiale della famiglia:

Figura A3. 2 : ITEM 7¹⁶


7. Quante di ciascuna delle seguenti cose ci sono a casa tua?		<i>Metti una crocetta su un solo quadratino per ogni riga.</i>			
		<i>Nessuno/a</i>	<i>Uno/a</i>	<i>Due</i>	<i>Tre o più di tre</i>
A.	Bagni	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B.	Automobili	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

¹⁹ **Fonte:** http://www.invalsi.it/snv0809/documenti/Questionario_studente_classe_V_primaria.pdf


L'item "Q8" (fig. A3.3) richiede agli studenti una stima approssimativa del numero di libri presenti a casa dello studente:

Figura A3.3: ITEM 8²⁰


8. Quanti libri ci sono all'incirca a casa tua (esclusi i libri di scuola)?
Metti una crocetta su un solo quadratino.

 → *Questo disegno rappresenta lo spazio occupato da 5 libri*


A. Nessuno o pochissimi (0-10 libri)

Spazio occupato da 10 libri → 


B. Abbastanza da riempire una mensola (11-25 libri)

Spazio occupato da 25 libri → 


C. Abbastanza da riempire uno scaffale (26-100 libri)

Spazio occupato da 100 libri → 

D. Abbastanza da riempire due scaffali (101-200 libri)

Spazio occupato da 200 libri → 

E. Abbastanza da riempire tre o più scaffali (più di 200 libri)

Spazio occupato da più di 200 libri → 

²⁰ http://www.invalsi.it/snv0809/documenti/Questionario_studente_classe_V_primaria.pdf

Figura A3. 4: Ricodifica dell'item 8 ("Quanti libri ci sono all'incirca in casa tua -esclusi i libri di scuola-?") in variabile dicotomica.

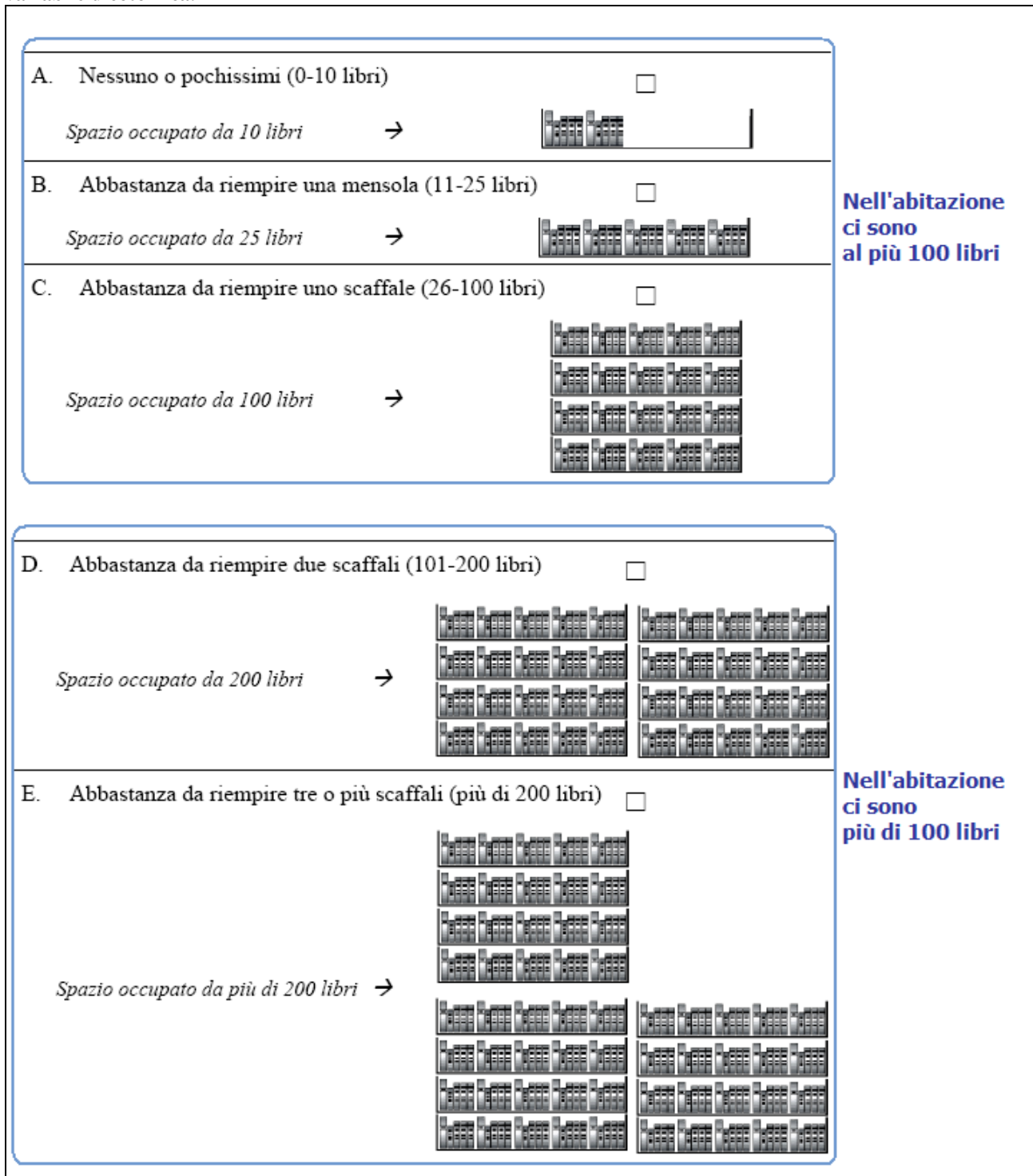


Figura A3. 5: Ricodifica dell'item 7 ("Quante di ciascuna delle seguenti cose ci sono a casa tua?") in variabile dicotomica.

A. Bagni	<p><i>Nessuno/a</i> <i>Uno/a</i></p> <p><input type="checkbox"/> <input type="checkbox"/></p>	<p><i>Due</i> <i>Tre o più di tre</i></p> <p><input type="checkbox"/> <input type="checkbox"/></p>
	<p>L'abitazione è dotata di al più un bagno</p>	<p>L'abitazione è dotata di 2 o più bagni</p>
B. Automobili	<p><i>Nessuno/a</i> <i>Uno/a</i></p> <p><input type="checkbox"/> <input type="checkbox"/></p>	<p><i>Due</i> <i>Tre o più di tre</i></p> <p><input type="checkbox"/> <input type="checkbox"/></p>
	<p>La famiglia possiede al più una automobile</p>	<p>La famiglia possiede 2 o più automobili</p>

Appendice 4

Tavola 4. 1: Punteggi medi rispetto all'ESCS.

Regione	ESCS medio	Percentuale di studenti sotto il 15° percentile nazionale di ESCS	Punteggio medio Italiano	Punteggio medio Italiano per ESCS uguale alla media nazionale	Punteggio medio Matematica	Punteggio medio Matematica per ESCS uguale alla media nazionale
Valle D'Aosta	0,12	10,7	503	500	497	493
Piemonte	0,03	14,9	508	507	512	511
Liguria	0,16	8,1	502	497	498	495
Lombardia	0,16	12,8	505	501	506	503
Trentino Alto Adige	-0,35	26,2	489	495	504	509
Veneto	-0,05	12,4	504	505	518	519
Friuli Venezia Giulia	0,14	10,6	505	501	519	515
Emilia Romagna	0,22	8,2	503	496	507	501
Toscana	0,25	7,8	513	507	517	512
Umbria	0,34	5,1	510	501	515	507
Marche	0,17	9,0	521	517	521	517
Lazio	0,32	9,5	492	484	484	476
Abruzzo	-0,16	12,7	499	504	488	492
Molise	0,14	12,2	515	512	500	496
Campania	-0,31	23,9	503	511	504	510
Puglia	-0,26	19,4	503	508	502	505
Basilicata	-0,23	19,0	504	510	498	503
Calabria	-0,08	16,9	506	508	502	504
Sicilia	-0,20	20,5	472	478	460	465
Sardegna	-0,14	15,7	479	482	467	470