

Fabrizio Alboni, Giorgio Tassinari

Parole, soltanto parole, parole tra noi

Quaderni di Dipartimento

Serie Ricerche 2016, n. 7

ISSN 1973-9346



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Dipartimento di Scienze Statistiche “Paolo Fortunati”

Fabrizio Alboni, Giorgio Tassinari

Parole, soltanto parole, parole tra noi

Fabrizio Alboni e Giorgio Tassinari

Dipartimento di Scienze statistiche “Paolo Fortunati”

Università di Bologna

Abstract

In recent years there has been a fair amount of debate in Italy about regional differences in student achievements at the end of the high school cycle (Esame di Stato). In particular, concerns have been expressed about the performance of students in Southern Italy, higher than the national average, while standardized tests of learning outcomes point to a very different situation. The paper investigates this issue using simple statistical methods to disentangle the geographical area effect from the effects of other characteristics. The main finding is that a regional effect is present in final grades, but it is more important in the area of Central Italy rather than in Southern Italy.

Premessa

Coloro che tra noi ricordano le trasmissioni televisive degli anni Sessanta forse avranno riconosciuto il titolo di una famosa canzone interpretata da Mina e da Alberto Lupo. La storia è semplice: una bella signora (Mina, appunto) lamenta che il suo fidanzato (Alberto Lupo) non superi mai la fase delle “parole”, non riuscendo ad arrivare ad uno sviluppo libidico completo. Con le debite proporzioni, lo stesso sentimento di irritazione interpretato da Mina ci ha colto leggendo i numerosi articoli apparsi nella prima metà del mese di agosto sui grandi quotidiani nazionali sull'apparente “eccesso” di studenti del Sud che superano l'esame di Stato con il massimo dei voti (100 e lode). A ciò hanno fatto seguito dichiarazioni di presidenti di regione ed assessori all'istruzione delle regioni del nord. La cifra è comune: antimeridionalismo e destituzione della credibilità della scuola statale.

Ma è poi vero? In altri termini, è vero che l'eccesso di 100 e lode sia imputabile al lassismo delle scuole del Mezzogiorno. Per rispondere in maniera credibile, è necessario affrontare il problema con un poco di respiro e con un approccio tecnico-razionale, senza tesi precostituite da promuovere.

1. Analisi statistica descrittiva

A supporto della tesi del lassismo meridionale, diversi commentatori portano i risultati dei test PISA (OCSE) e INVALSI (MIUR), che evidenzerebbero risultati di gran lunga peggiori degli studenti del Sud rispetto a quelli del Nord. In premessa, vale la pena soffermarsi su questa comparazione, a nostro giudizio completamente spuria e viziata da un profondo e assai grave errore di metodo.

Si tratta di un confronto a dir poco ingenuo, poiché le coorti di studenti sono diverse. Inoltre il voto dell'esame di Stato riguarda studenti di circa 18 anni, i test PISA e Invalsi studenti di 15-16. Ancora, PISA e INVALSI si concentrano su solo due discipline (lingua italiana e matematica), mentre l'esame di Stato considera

tutte le discipline, e tiene conto della carriera dello studente negli ultimi tre anni. PISA e Invalsi sono invece prove *one-shot*. Insomma, è come paragonare il Giro d'Italia con le Tre Valli Varesine, gara nobilissima e di tradizione, ma non sufficiente a costruire sui suoi risultati un ranking nazionale.

Ci siamo già soffermati sul significato, sotto il profilo politico-filosofico, degli esercizi di valutazione nella scuola, e rimandiamo a quei contributi¹ per una disamina più approfondita. Va tuttavia sottolineata l'estrema fragilità tecnico-statistica dei risultati INVALSI, che si basano su una metodologia (il modello di Rasch (Falocci et al., 2010) assai suggestiva, ma che richiede assunzioni assai stringenti per la sua applicazione. Per questi motivi utilizzeremo come riferimento soltanto i risultati dei test PISA-OCSE. Come già detto, un approccio intellettualmente onesto al problema, e non banalmente ideologico, ovvero informato dal duplice pregiudizio anti-meridionale e anti-statale, richiede un'analisi il più possibile razionale dei dati a nostra disposizione. Inoltre è a nostro giudizio importante cercare di quantificare, pur con tutti i limiti della metodologia statistica, il peso del carattere "scuola secondaria del Sud" nel determinare il risultato dell'esame di Stato. E' quanto cercheremo di fare nel proseguo.

¹ Moretto e Tassinari (2009); Tassinari (2012).

Tabella 1. Diplomati all'esame di Stato per regione e secondo la classe di voto (A.S. 2014-2015) (in percentuale)

Regione	Voto							Totale
	60	61-70	71-80	81-90	91-99	100	100 e lode	
Piemonte	7,8	30,4	30,6	19,3	7,6	3,7	0,7	100,0
Valle d'Aosta	5,5	31,0	34,6	20,0	7,1	1,7	0,1	100,0
Lombardia	9,2	32,7	30,4	17,8	6,5	2,9	0,4	100,0
Trentino A.A.(*)	5,9	26,5	31,9	22,2	8,7	4,1	0,7	100,0
Veneto	7,0	29,0	31,1	20,0	8,2	4,1	0,6	100,0
Friuli V.G.	6,7	29,9	32,6	19,3	7,5	3,7	0,4	100,0
Liguria	7,8	30,4	30,7	18,0	8,2	4,2	0,7	100,0
Emilia Romagna	6,6	28,1	30,5	20,2	9,0	4,5	1,0	100,0
Toscana	7,2	28,6	30,6	19,6	8,8	4,4	0,7	100,0
Umbria	7,6	28,0	29,2	18,8	9,5	5,4	1,6	100,0
Marche	5,5	26,3	29,9	20,5	9,8	6,3	1,7	100,0
Lazio	10,3	30,0	28,9	17,6	7,7	4,7	0,8	100,0
Abruzzo	9,0	30,2	28,8	17,5	8,1	5,5	0,9	100,0
Molise	9,1	31,8	27,9	17,5	7,6	5,3	0,8	100,0
Campania	10,4	29,1	27,5	17,6	8,9	5,7	0,8	100,0
Puglia	7,8	26,7	27,6	19,3	9,5	6,8	2,3	100,0
Basilicata	9,5	31,4	27,9	17,7	8,0	4,5	1,0	100,0
Calabria	7,7	25,6	26,6	20,0	11,0	7,6	1,6	100,0
Sicilia	10,3	28,5	27,5	18,7	8,6	5,5	0,9	100,0
Sardegna	8,0	30,3	29,9	18,1	8,1	4,8	0,8	100,0
Italia	8,6	29,3	29,2	18,7	8,4	4,9	0,9	100,0

(*) I risultati relativi alla provincia di Bolzano non sono disponibili.

Fonte: MIUR, Focus esiti esami di Stato II grado, A.S. 2014-2015

Nella Tabella 1 è riportata la distribuzione per regione degli studenti che hanno superato l'esame di Stato nell'A.S. 2014-2016 secondo la classe di voto ottenuto. Dall'esame dell'ultima colonna è evidente la grande variabilità della percentuale di studenti che hanno riportato il massimo dei voti. Tuttavia, considerando il valore medio nazionale come punto di riferimento (ovvero una percentuale di 0,9), ci accorgiamo che la percentuale è più alta non solo in alcune regioni meridionali (come la Calabria e la Puglia), ma anche nelle Marche ed in Umbria. Se, anziché la media aritmetica assumiamo come valore di riferimento la mediana della distribuzione per regione della percentuale di 100 e lode, otteniamo sempre lo stesso valore di 0,9. E' comunque significativo che tutte le regioni del Nord abbiano valori sotto la mediana, e le regioni del Sud abbiano valori intorno alla mediana o superiori (in alcuni casi di molto superiori, come in Puglia o in Calabria).

Tuttavia, dobbiamo tener presente che vi sono altre cause di variazione sistematica del voto conseguito all'esame di Stato, cause che lo stesso MIUR, nella pubblicazione citata, mette in evidenza (Tabella 2):

- il tipo di percorso di studio influenza il risultato in modo assai incisivo (gli studenti dei licei hanno una percentuale di 100 e lode circa doppia di quelli che frequentano gli istituti tecnici);
- la nazionalità dello studente influenza in modo assai forte il risultato (gli studenti con cittadinanza italiana figli di cittadini italiani hanno i risultati migliori).

Oltre a questi elementi occorre tener presente l'importanza del fattore sesso degli studenti. Le femmine infatti rappresentano circa il 50% di tutti i diplomati, ma la loro presenza sale al 61,1% per quanto riguarda i licei e scende al 34,8% negli istituti tecnici. Che le ragazze siano più brave dei ragazzi lo dimostra in modo chiarissimo il fatto che, a parità di indirizzo scolastico, le ragazze riportano sistematicamente risultati più brillanti.

Questi fattori di eterogeneità si intersecano ed interagiscono tra loro. Ad esempio è ben noto che la percentuale di studenti stranieri che frequentano i licei è assai bassa. In generale, la quota di studenti non italiani che arrivano a sostenere l'esame di Stato è assai esigua e di conseguenza la loro performance assai più modesta di quella che riguarda gli italiani non influenza in modo palese la distribuzione media dei voti all'esame di Stato.

Tabella 2. Distribuzione percentuale del voto all'esame di Stato secondo il tipo di scuola media superiore, la nazionalità dello studente, il sesso (a.s. 2014-2015)

Carattere	Voto							Totale	Mediana
	60	61-70	71-80	81-90	91-99	100	100 e lode		
Tipo di scuola									
Licei	5,4	23,3	29,5	22,3	10,9	7,1	1,5	100,0	77,2
<i>di cui Classici</i>									
<i>di cui Scienze umane</i>									
Istituti Tecnici	11,9	34,5	28,5	15,4	6,2	3,1	0,4	100,0	71,2
<i>di cui Economici</i>									
<i>di cui Tecnologici</i>									
Istituti professionali	8,6	29,3	29,2	18,7	8,4	4,9	0,9	100,0	74,1
<i>di cui Industriali</i>									
<i>di cui Servizi</i>									
Cittadinanza									
Italiana	8,4	28,9	29,3	18,9	8,5	5,0	0,9	100,0	
Non italiana e nati all'estero	12,8	37,2	28,3	14,4	4,7	2,2	0,3	100,0	
Non italiana e nati in Italia	11,7	34,7	29,3	15,6	5,5	2,8	0,4	100,0	
Italia	8,6	29,3	29,2	18,7	8,4	4,9	0,9	100,0	74,1

(*) I risultati relativi alla provincia di Bolzano non sono disponibili.

Fonte: MIUR, Focus esiti esami di Stato II grado, A.S. 2014-2015

Ci sembra palese, quindi, che il risultato “regionale” degli esami di Stato sia influenzato dalla composizione della studentesca. Il corretto confronto tra i risultati delle diverse regioni dovrebbe quindi basarsi sull’eliminazione di tali differenze. Dovrebbe essere chiaro a questo punto che confrontare i risultati della Lombardia con quelli della Sicilia, *sic et simpliciter*, non ha nessun senso. E’ la classica somma di mele e di pere. Avrebbe invece senso confrontare i risultati agli esami di Stato degli studenti lombardi che frequentano il liceo classico con quelli degli studenti siciliani che frequentano lo stesso tipo di scuola. Sarebbe inoltre necessario, “dentro” ciascun indirizzo di scuola media superiore, standardizzare la composizione degli studenti tenendo conto del sesso degli studenti e della loro cittadinanza.

Un ulteriore elemento di segmentazione del sistema scolastico italiano è quello tra scuole statali e scuole private (paritarie e non). Purtroppo il MIUR non elabora i risultati degli esami di Stato secondo questo carattere, sebbene i risultati dei test PISA mettano in evidenza differenze drammatiche tra i due sistemi².

La situazione si chiarisce notevolmente se consideriamo congiuntamente i due caratteri di stratificazione, tipo di istituto e regione. Nella tabella seguente (tabella 3) viene riportata la percentuale di 100 e lode secondo i due caratteri menzionati.

² Ad esempio, nella abilità reading, per la provincia di Trento il punteggio medio è di 520 per gli studenti delle scuole pubbliche e di 455 per gli studenti delle scuole paritarie (*PISA 2009 Results Database* disponibile on line all’Url://www.oecd.org/pisa/pisaproducts/pisa2009resultswhatstudentsknowandcandostudentperformanceinreadingmathematicsandsciencevolumei_https.htm). Per le altre regioni i dati sulle scuole paritarie non vengono rilevati.

Tabella 3 – percentuale di diplomati con 100 e lode per regione e tipo di istituto –a.s. 2013-2014

	Istituto Professionale	Ist. Tecnico	Liceo	Totale
01.Piemonte	0.085	0.329	1.162	0.508
03.Lombardia	0.060	0.135	0.473	0.219
04.Trentino Alto Adige	0.000	0.245	0.786	0.388
05.Veneto	0.049	0.322	0.703	0.345
06.Friuli	0.000	0.298	0.700	0.322
07.Liguria	0.052	0.111	1.096	0.339
08.Emilia Romagna	0.093	0.765	1.632	0.775
09.Toscana	0.050	0.317	1.368	0.537
10.Marche	0.240	0.842	2.208	1.005
11.Umbria	0.446	0.785	2.521	1.170
12.Lazio	0.107	0.332	1.542	0.698
13.Abruzzo	0.000	0.212	2.241	0.688
14.Molise	0.000	0.595	1.693	0.804
15.Campania	0.006	0.221	1.560	0.633
16.Puglia	0.116	0.849	3.638	1.437
17.Basilicata	0.000	0.218	2.045	0.721
18.Calabria	0.028	0.234	2.469	0.864
19.Sicilia	0.000	0.250	1.990	0.700
20.Sardegna	0.050	0.194	1.164	0.474
Totale	0.066	0.355	1.616	0.655

Fonte: MIUR, Focus esiti esami di Stato II grado, A.S. 2014-2015

Si conferma che, anche a livello territorialmente disaggregato, vi è un effetto sistematico dovuto al tipo di istituto e che a parità di tipo di istituto, emerge un effetto sistematico che tende a far sì che nelle regioni meridionali la percentuale di studenti che conseguono 100 e lode è più elevata, a parità di tipo di percorso di istruzione.

E' interessante contrapporre alla distribuzione doppia della percentuale di studenti che hanno riportato il voto di 100 e lode la distribuzione doppia del voto mediano (Tabella 4).

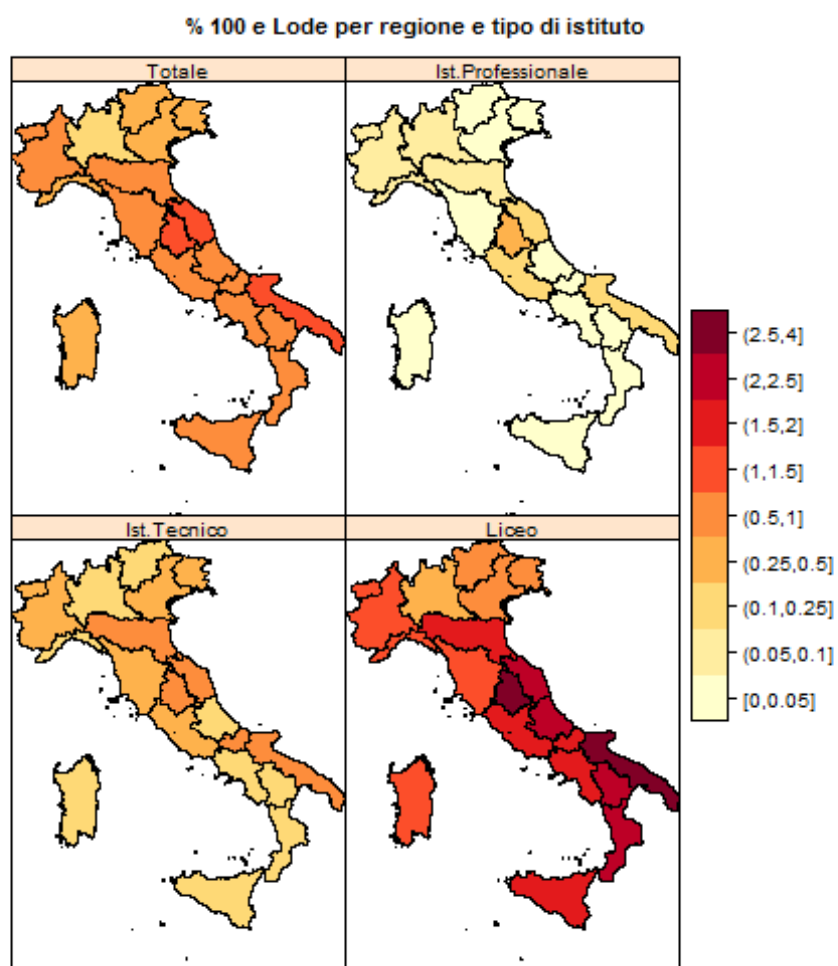
Tabella 4 – Voto mediano per regione e tipo di istituto

	Ist. Professionale	Ist. Tecnico	Liceo	Totale
01.Piemonte	69.9	72.0	77.0	72.9
03.Lombardia	69.6	71.1	75.7	72.1
04.Trentino Alto Adige	70.9	73.5	77.5	74.4
05.Veneto	70.3	73.4	77.1	73.5
06.Friuli	70.4	73.6	77.3	73.7
07.Liguria	70.2	72.4	76.7	72.8
08.Emilia Romagna	71.9	73.4	78.1	74.2
09.Toscana	70.2	72.8	77.6	73.4
10.Marche	72.2	74.8	80.5	75.5
11.Umbria	70.1	71.0	78.9	72.9
12.Lazio	70.3	71.1	76.8	72.9
13.Abruzzo	68.5	71.7	77.6	72.5
14.Molise	70.5	70.2	78.5	72.8
15.Campania	70.0	73.2	79.7	74.6
16.Puglia	70.6	72.8	80.9	74.5
17.Basilicata	68.6	72.1	78.8	73.2
18.Calabria	70.7	73.0	82.4	75.2
19.Sicilia	69.0	72.1	79.7	73.6
20.Sardegna	69.5	71.7	77.0	72.9
Totale	70.2	72.4	78.3	73.6

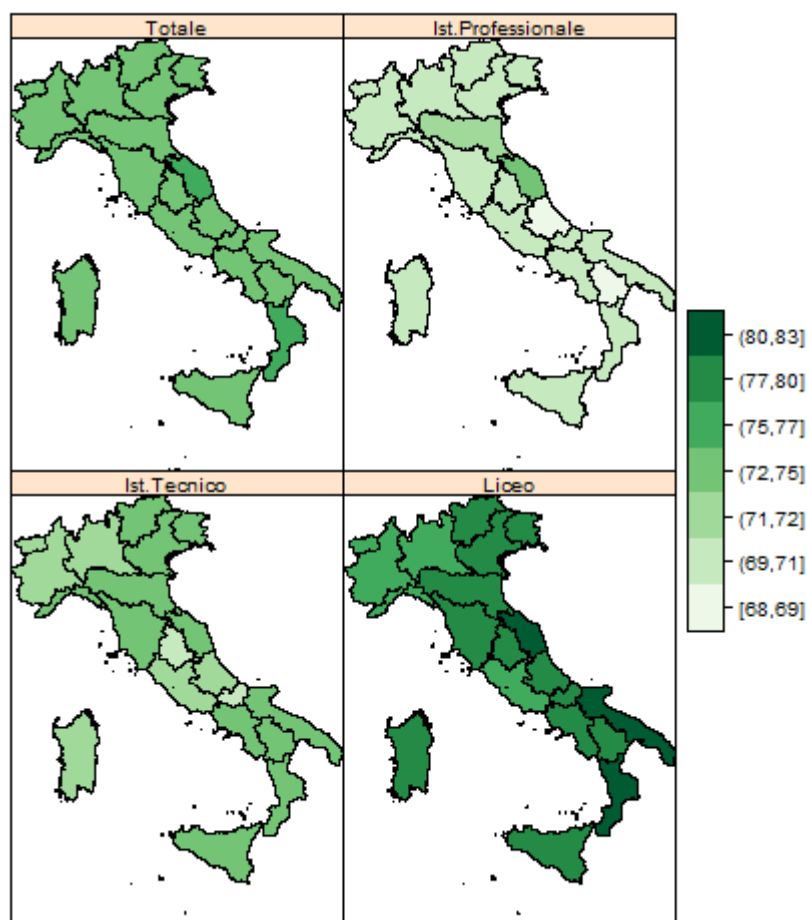
Fonte: MIUR, Focus esiti esami di Stato II grado, A.S. 2014-2015

In questo secondo caso le differenze nel voto mediano, a parità di tipo di istituto, sono assai meno accentuate. L'effetto meridione, se esiste, sembra riguardare soltanto la percentuale di 100 e lode. Una rappresentazione grafica delle due distribuzioni è riportata di seguito (Figura 1).

Figura 1



Voto mediano per regione e tipo di istituto



Un esercizio ermeneutico

Come ben noto, il modello di regressione lineare può essere impiegato per calcolare delle medie condizionate. In questa sede lo utilizzeremo a tale scopo, per cercare di “quantificare” l’impatto che ha sulla percentuale di diplomati all’esame di Stato con 100 e lode la circostanza di aver affrontato e superato lo stesso in un istituto

superiore localizzato nel Mezzogiorno. Si tratta di una regressione “ecologica”, in quanto le variabili esplicative sono riferite al contesto socio-economico di ciascuna regione. Poiché inoltre abbiamo a disposizione dati censuari e non campionari, preferiamo non ricorrere alle tecniche di inferenza statistica standard³.

Le variabili utilizzate come predittori sono le seguenti:

- a) Il reddito pro-capite regionale (in migliaia di euro), che dovrebbe catturare il grado di sviluppo economico. Sul segno del coefficiente atteso del coefficiente c'è ambiguità. Un segno positivo implicherebbe un effetto dello sviluppo delle forze produttive sul livello di civilizzazione della popolazione e dell'impegno scolastico. Al contrario un segno negativo starebbe ad indicare che i giovani percepirebbero il grado di istruzione come un indicatore debole dei redditi futuri. Inoltre, vista l'alta correlazione tra reddito e patrimonio, possiamo presumere che l'istruzione e la cultura non siano “vissute” come strumenti di emancipazione, che sarebbe veicolata al contrario dalla ricchezza della famiglia;
- b) la percentuale di popolazione in età 30-34 in possesso almeno del diploma di laurea triennale (che dovrebbe catturare sia il livello culturale delle famiglie che la qualità knowledge-intensive della struttura produttiva); anche in questo caso ci si attende un segno positivo del coefficiente;
- c) la percentuale di studenti che non completano il ciclo di istruzione secondaria (ci si attende un segno positivo del coefficiente, in quanto si presume che l'eliminazione degli

³ Ci sembra che nel nostro contesto non sia sensato immaginare la presenza, anche solo virtuale” dell'universo dei campioni. Inoltre, trattandosi di dati *cross-section*, ci pare assolutamente fuori luogo fare riferimento alla nozione di processo stocastico ergodico che giustifica l'uso delle tecniche inferenziali nel contesto dei dati time-series (Piccolo e Vitale, 1981).

studenti più “deboli” implichi un incremento della performance dei “sopravvissuti”);

- d) una variabile dummy che cattura l’effetto differenziale tra regioni del nord e del sud.

Tabella 5 – Variabili utilizzate nel modello di regressione lineare

Regione	% 100 e lode (2014-2015)	Reddito pro-capite (2014)	% Laurea (2014)	% uscita precoce	Dummy geografica
	Y	X1	X2	X3	X4
Piemonte	0,7	19,6	24,2	12,7	0
Valle d’Aosta	0,1	20,1	20,7	16,2	0
Lombardia	0,4	21,2	25,9	12,9	0
Trentino-AA	0,7	21,3	25,0	10,9	0
Veneto	0,6	18,6	23,5	8,4	0
Friuli-VG	0,4	19,3	27,1	11,1	0
Liguria	0,7	20,8	31,3	13,6	0
Emilia-R.	1,0	21,0	25,1	13,2	0
Toscana	0,7	19,3	24,8	13,8	0
Marche	1,7	17,9	24,9	10,9	0
Umbria	1,6	17,7	30,3	9,1	0
Lazio	0,8	18,4	31,6	12,5	0
Abruzzo	0,9	15,8	25,3	9,6	1
Molise	0,8	14,6	26,7	12,2	1
Campania	0,8	12,6	18,1	19,6	1
Puglia	2,3	13,2	21,2	16,9	1
Basilicata	1,0	12,9	19,8	12,2	1
Calabria	1,6	12,3	23,3	16,8	1
Sicilia	0,9	12,9	17,7	24,0	1
Sardegna	0,8	16,6	17,4	23,4	1
Italia	0,9	17,5	23,9	15,0	

FONTE: MIUR e ISTAT

Tabella 6 – Matrice di correlazione tra variabili

	Y	X1	X2	X3	W1
Y	1.00	-0.51	-0.03	0.04	-0.11
X1	-0.51	1.00	0.53	-0.46	0.11
X2	-0.03	0.53	1.00	-0.70	-0.05
X3	0.04	-0.46	-0.70	1.00	0.17
W1	-0.11	0.11	-0.05	0.17	1.00

n= 20

La tabella di correlazione tra le variabili mette in evidenza una correlazione negativa tra il reddito pro-capite e la percentuale di studenti che riportano il voto più elevato. Le altre correlazioni sono praticamente nulle. Si riscontra anche una correlazione negativa molto forte (-0,70) tra la variabile che rispecchia la diffusione del diploma di laurea e la percentuale di abbandoni, nonché una correlazione negativa tra reddito pro-capite e percentuale di abbandoni.

I risultati dell'analisi di regressione sono riportati nella tabella 7.

Tabella 7 – Risultati dell'analisi di regressione lineare

	Stima	Errore stand.	t value	Pr(> t)
(Intercept)	2.00345	2.04922	0.978	0.344
X1	-0.08836	0.08078	-1.094	0.291
X2	0.03134	0.03884	0.807	0.432
X3	-0.03047	0.03814	-0.799	0.437
X4Sud	0.34783	0.67714	0.514	0.615

Residual standard error: 768 on 15 degrees of freedom

Multiple R-squared: 0.3832, Adjusted R-squared: 0.2187

F-statistic: 2.33 on 4 and 15 DF, p-value: 0.1034

Esaminando i coefficienti emerge l'effetto negativo del reddito pro-capite (che è un effetto netto, depurato dall'influenza della zona geografica) ed anche l'effetto positivo della variabile geografica. Conforme alle attese è il segno del coefficiente della variabile "percentuale di laureati". Poiché si tratta di dati amministrativi

riferiti a tutto l'universo dei diplomati, non ci sembra giustificato il ricorso alle tecniche inferenziali. Occorre osservare comunque che la quota di variabilità spiegata dal modello è modesta (R^2 corretto pari a 0,21, valore ragionevole per una regressione ecologica).

La lettura dei risultati è ostacolata dalla diversa unità di misura delle variabili. Per eliminare tale effetto procediamo al calcolo dei cosiddetti coefficienti beta.

Tuttavia, come abbiamo mostrato sopra, l'effetto del tipo di istituto è molto forte nel determinare il valore della variabile dipendente.

Analisi micro

Un secondo approccio, a nostro giudizio più fecondo, è quello di fare ricorso direttamente ai dati micro riferiti ai singoli istituti scolastici. Anche in questo caso abbiamo utilizzato i dati MIUR riferiti all'a.s. 2013-2014.

Lo strumento statistico che abbiamo utilizzato è quello della regressione logistica multinomiale (Multinomial Logit, Brasini et. (2010)), in cui la variabile dipendente è data dalla percentuale di istituti che hanno avuto una percentuale di diplomati con 100 e Lode compresa alternativamente nelle seguenti classi:

nessun 100 e Lode

tra 0 e 1%

tra 1 e 2%

tra 2 e 3%

oltre il 3%

Abbiamo pertanto stimato un modello per ciascuna classe della variabile dipendente. Le variabili indipendenti sono date dal voto medio, dalla percentuale di abbandoni, dal tipo di istituto e dall'area geografica (Nord-Ovest, Nord-Est, Centro, Sud, Isole).

Nella tabella 8 è riportata la distribuzione degli istituti secondo la variabile dipendente, mentre i risultati sono riportati nella tabella 9.

Tabella 8. Istituti secondo il tipo, la ripartizione geografica e la percentuale di diplomati con 100 e lode.

Ist. Professionale	[0]	(0,1]	(1,2]	(2,3]	(3,12]	Sum
Nord-Ovest	250	2	4	1	1	258
Nord-Est	203	4	5	0	0	212
Centro	228	6	5	5	2	246
Sud	379	4	8	0	0	391
Isole	175	0	0	1	0	176
In complesso	1235	16	22	7	3	1283

Ist. Tecnico	[0]	(0,1]	(1,2]	(2,3]	(3,12]	Sum
Nord-Ovest	376	22	16	7	6	427
Nord-Est	238	42	33	17	6	336
Centro	320	29	29	16	18	412
Sud	475	38	34	24	25	596
Isole	275	12	11	7	6	311
In complesso	1684	143	123	71	61	2082

Liceo	[0]	(0,1]	(1,2]	(2,3]	(3,12]	Sum
Nord-Ovest	164	47	39	21	17	288
Nord-Est	87	31	43	22	13	196
Centro	90	42	64	36	57	289
Sud	154	36	66	47	148	451
Isole	88	15	34	28	47	212
In complesso	583	171	246	154	282	1436

Segue Tabella 8

In complesso	[0]	(0,1]	(1,2]	(2,3]	(3,12]	Sum
Nord-Ovest	790	71	59	29	24	973
Nord-Est	528	77	81	39	19	744
Centro	638	77	98	57	77	947
Sud	1008	78	108	71	173	1438
Isole	538	27	45	36	53	699
Sum	3502	330	391	232	346	4801

Fonte: MIUR

Va rimarcato, brevemente, che gli istituti che licenziano almeno un diplomato con 100 e lode sono assai pochi, solo 1300 su 4800. La “strage” è massima negli istituti professionali (53 istituti su 1283). Ma anche nei licei, vi sono ben 580 istituti che non hanno alcun diplomato con 100 e Lode. Questa grande eterogeneità deve pertanto farci diffidare delle analisi eccessivamente aggregate. Gli studenti sono ovviamente i “figli” di una comunità ed occorre quindi tenere debitamente conto delle variabili di contesto⁴.

⁴ La letteratura infatti suggerisce di utilizzare come unità elementare lo studente, e di articolare i modelli su più livelli, ad esempio classe-istituto-comune-provincia etc. Vedi su questo, tra gli altri, Bryk e Raudenbush (2002).

Tabella 9 – Risultati della stima del modello Multinomial Logit

	Coefficients	Std. Errors	Z	p-value	Odd ratio
(tra lo 0 e l'1%)					
(Intercept)	-10.677	1.212	-8.807	0.000	0.000
Vmediano	0.091	0.017	5.411	0.000	1.095
prcABBANDONOt	-0.350	0.130	-2.699	0.007	0.705
TIPO.ISTITUZIONE.ridIst.Tecnico	1.631	0.270	6.047	0.000	5.110
TIPO.ISTITUZIONE.ridLiceo	2.402	0.291	8.262	0.000	11.047
AREA5Nord-Est	0.517	0.183	2.820	0.005	1.676
AREA5Centro	0.428	0.181	2.367	0.018	1.534
AREA5Sud	-0.099	0.181	-0.548	0.584	0.906
AREA5Isole	-0.595	0.241	-2.469	0.014	0.552
(tra l'1% e il 2%)					
(Intercept)	-12.521	1.132	-11.057	0.000	0.000
Vmediano	0.116	0.016	7.429	0.000	1.123
prcABBANDONOt	-0.357	0.124	-2.883	0.004	0.700
TIPO.ISTITUZIONE.ridIst.Tecnico	1.111	0.238	4.662	0.000	3.038
TIPO.ISTITUZIONE.ridLiceo	2.280	0.253	9.003	0.000	9.777
AREA5Nord-Est	0.764	0.192	3.983	0.000	2.148
AREA5Centro	0.863	0.184	4.694	0.000	2.371
AREA5Sud	0.368	0.182	2.022	0.043	1.445
AREA5Isole	0.061	0.218	0.279	0.780	1.063
(tra il 2 e il 3%)					
(Intercept)	-17.620	1.423	-12.378	0.000	0.000
Vmediano	0.169	0.019	8.875	0.000	1.185
prcABBANDONOt	-0.317	0.155	-2.049	0.040	0.728
TIPO.ISTITUZIONE.ridIst.Tecnico	1.607	0.402	3.999	0.000	4.987
TIPO.ISTITUZIONE.ridLiceo	2.593	0.413	6.280	0.000	13.376
AREA5Nord-Est	0.690	0.262	2.628	0.009	1.993
AREA5Centro	0.994	0.244	4.066	0.000	2.701

Segue Tabella 9

AREA5Sud	0.507	0.241	2.106	0.035	1.660
AREA5Isole	0.419	0.271	1.546	0.122	1.520
<hr/>					
(sopra il 3%)					
(Intercept)	-24.914	1.411	-17.652	0.000	0.000
Vmediano	0.252	0.017	14.407	0.000	1.286
prcABBANDONOt	-0.271	0.138	-1.964	0.049	0.763
TIPO.ISTITUZIONE.ridIst.Tecnico	2.164	0.598	3.619	0.000	8.705
TIPO.ISTITUZIONE.ridLiceo	3.446	0.597	5.769	0.000	31.380
AREA5Nord-Est	0.127	0.326	0.390	0.696	1.136
AREA5Centro	1.436	0.256	5.609	0.000	4.205
AREA5Sud	1.274	0.245	5.211	0.000	3.577
AREA5Isole	0.718	0.278	2.585	0.010	2.050

R2 corretto di Nagelkerke 0,37; AIC = 9171,90

I coefficienti sono tutti altamente significativi, e molto buone sono anche le statistiche test complessive. Il modello è costruito in modo differenziale rispetto ad una tipologia base, costituita dagli istituti professionali del Nord-Ovest. L'interpretazione deve essere formulata con riferimento alla colonna OR (Odd Ratio) che esprime il cambiamento della probabilità rispetto alla tipologia base. Quindi, con riferimento all'ultima classe (percentuale di 100 e Lode superiore al 3%) il valore di 31,38 significa che un liceo ha una probabilità 31 volte superiore rispetto ad un istituto professionale di avere più del 3% di 100 e Lode. Va rimarcato che l'effetto del voto mediano è positivo, e quello della percentuale di abbandoni negativo (gli istituti che riescono ad evitare gli abbandoni, od in cui gli studenti hanno una minore propensione ad abbandonare hanno percentuali più alte di 100 e lode).

Ci preme sottolineare che gli Odds Ratios concernenti il Sud e le Isole sono sempre più bassi degli Odds ratios per il Centro ed in molti casi anche di quelli del Nord-Est. Pertanto, depurando i dati

dalle condizioni al contorno, l'effetto territorio è più forte negli istituti scolastici del Centro rispetto a quelli del Sud-Isole

Un tentativo di conclusione

Non va taciuto che l'analisi delle performance delle scuole è un terreno assai articolato, e andrebbe affrontato con un respiro assai più ampio di quanto si sia fatto in questa sede. Ci pare tuttavia che anche da un'analisi "short-cut" come quella che abbiamo condotto si possano trarre alcuni elementi sufficientemente solidi:

- e) La scuola italiana è una scuola severa. Infatti su 4500 istituti superiori, nel 2013-2014 ben 3500 non hanno avuto nemmeno un diplomato con 100 e Lode;
- f) vi è una tendenza uniforme degli istituti superiori del Nord-Ovest ad avere percentuali di studenti con 100 e Lode inferiori al resto d'Italia; la sovrarappresentazione di 100 e Lode caratterizza maggiormente le scuole del Centro che non quelle del Sud-Isole. Vi sono anche differenze significative tra Nord-Ovest e Nord-Est;
- g) la polemica anti scuole meridionali è frutto di un'analisi superficiale dei dati e rivela un intento fazioso ed ideologico, intendendo ideologia come "falsa coscienza" nel senso dell'"Ideologia tedesca" di Marx-Engels;

L'effetto territoriale comunque si afferma come rilevante anche ad un'analisi condotta con strumenti tecnico-razionali sofisticati. Ci pare doveroso quindi riflettere ed introdurre modificazioni all'attuale meccanismo dell'esame di maturità. Per i motivi già esplicitati sopra, non abbiamo nessuna fiducia nell'introduzione di prove standardizzate, la cui vacuità è ormai assodata nella letteratura internazionale (Harris 2011). Forse tornare alle "vecchie" commissioni esterne potrebbe rappresentare un rimedio che assicuri maggiormente l'equità orizzontale nei risultati.

Va da sé che questa seconda alternativa è più costosa della prima. E qui si introduce una dimensione, quella del “costo della giustizia sociale” (Kolm, 1969), assai interessante, ma sulla quale la riflessione, nel nostro paese, è ancora agli inizi.

Riferimenti bibliografici

S. Brasini, M. Freo, F. Tassinari, G. Tassinari, (2010), *Marketing e pubblicità. Metodi di analisi statistica*, Bologna, Il Mulino.

S. Bryk, B. Raudenbush, (2002), *Hierarchical Linear Models: Applications and Data Analysis Methods*, Thousand Oaks, Sage.

N. Falocci, M. Gualdi, M. Matteucci, S. Mignani, (2010), *La Validazione Statistica di test standardizzati di profitto: principali aspetti di metodo e due casi di studio sulla valutazione degli apprendimenti nella scuola primaria*, Invalsi, Working Paper n. 9.

D. N. Harris (2011), *Value Added Measures and the Future of Education Accountability*, Science, 33, 6044 (August), pp. 826-827.

S.C.Kolm, (1969), *The Optimal Production of Social Justice*, in H. Guitton, J. Margolis, (eds.), "Public Economics", McMillan, London.

B. Moretto, G. Tassinari, *La libertà è partecipazione*, Scuola e Costituzione, 28 febbraio 2009.

D. Piccolo, C. Vitale, (1981), *Metodi statistici per l'analisi economica*, Bologna, Il Mulino.

Tassinari G. (2012), *Due o tre cose sul progetto Vales*, Inchiesta on line, 11 febbraio 2012.