

novembre 2017

eduscopio.it
esiti universitari

documento tecnico



Fondazione
Agnelli

Trattamento e analisi dei dati

Il progetto Eduscopio è coordinato da Gianfranco De Simone che ha curato anche la redazione del presente documento tecnico.

L'analisi dei dati è a cura di Martino Bernardi e Gianfranco De Simone.

Il lavoro di predisposizione del *dataset* è a cura di Martino Bernardi con l'assistenza di Chiara Genova.

Ringraziamenti

La Fondazione Giovanni Agnelli è grata agli uffici statistici del Ministero dell'Istruzione, dell'Università e della Ricerca e del Cineca per l'assistenza offerta sui dati.

Su diversi aspetti legati alla metodologia di analisi la Fondazione Giovanni Agnelli ha tratto beneficio dal confronto con Massimiliano Bratti (Università degli Studi di Milano), Daniele Checchi (Università degli Studi di Milano), Stefano Fachin (Università di Roma "La Sapienza"), Andrea Ichino (Istituto Universitario Europeo), Kevin Lang (Boston University), Fabio Schiantarelli (Boston College), Alessandro Sembenelli (Università di Torino).

La responsabilità di eventuali errori o imprecisioni rimanenti è da attribuirsi esclusivamente alla FGA.

Sommario

Introduzione.....	1
I dati.....	3
Fonti dei dati	3
Predisposizione dataset integrato	5
Metodologia statistica: una procedura a due step.....	9
Le misure di <i>outcome</i>	9
Primo step: la standardizzazione delle misure di outcome	10
Secondo <i>step</i> : il calcolo delle medie di scuola con effetti fissi e dati <i>pooled</i>	14
Un intervallo di confidenza per l'Indice FGA	18
Analisi della stabilità nel tempo dell'Indice FGA e dei <i>rankings</i>.....	20
Quali confronti sono appropriati?	21
 Riferimenti bibliografici.....	25
 APPENDICE A: Cream skimming e Indice FGA.....	27

1. Introduzione

L'apprendimento è un processo cumulativo: quanto riusciamo ad apprendere oggi dipende in larga parte da quanto abbiamo appreso in passato. Si possono costruire conoscenze più estese e competenze più robuste solo a partire da basi solide. Non a caso, il reale valore dell'istruzione ricevuta a scuola si manifesta proprio quando ci si trova davanti alla complessità di un esame universitario o di una mansione da svolgere sul lavoro. Chi ha ricevuto un'istruzione di qualità possiede le risorse culturali e il metodo per superare l'ostacolo, viceversa chi ha basi fragili fa più fatica a trovare soluzione ai problemi e ad espandere il proprio bagaglio di conoscenze in funzione di tale obiettivo. Per questa ragione, all'istruzione secondaria è richiesto di creare le condizioni per le quali gli studenti possano intraprendere con successo il passo successivo nelle loro traiettorie di vita.

L'idea di fondo del progetto **eduscopio.it** è proprio quella di valutare uno degli esiti successivi della formazione secondaria - i risultati universitari degli studenti - per trarne un'indicazione di qualità sull'offerta formativa delle scuole da cui essi provengono. I risultati universitari (esami, voti, crediti) riflettono e danno informazioni anche sulla qualità delle "basi" formative, la bontà del metodo di studio e l'utilità dei suggerimenti orientativi acquisiti presso le scuole secondarie. In altre parole, i risultati universitari ci permettono di formulare un giudizio sulla qualità delle scuole secondarie superiori sulla base di informazioni che provengono da enti – gli atenei – che sono "terzi" rispetto alle scuole stesse, cioè imparziali, ma al tempo stesso molto interessati alla qualità delle competenze e delle conoscenze degli studenti.

Con tutte le cautele del caso, **eduscopio.it** vuole offrire informazioni e dati comparabili utili:

- agli studenti che terminano le scuole medie e alle loro famiglie di modo che, una volta scelto l'indirizzo di scuola secondaria superiore, possano individuare quali istituti nella propria area di residenza soddisfino meglio le proprie aspettative di apprendimento;
- alle scuole stesse di modo che siano maggiormente responsabilizzate rispetto a una delle loro missioni formative – la preparazione e l'orientamento agli studi universitari - e possano finalmente conoscere gli esiti di questo lavoro: quali università e quali corsi di laurea scelgono i loro diplomati e quali risultati conseguono. Si tratta di informazioni spesso non disponibili in forma chiara e

comparabile, ma che possono offrire spunti per riflettere sulla bontà dei propri sistemi di orientamento in uscita.

Ogni anno, circa il 58% dei diplomati transita verso corsi universitari. Questo dato medio nasconde una forte differenziazione tra indirizzi di istruzione secondaria di secondo grado e può oscillare tra il 91% di transizioni per i diplomati del liceo classico e il 15% di transizioni per quelli dell'istruzione professionale.

Dunque, guardare agli esiti universitari dei diplomati per inferirne delle informazioni comparabili sulla qualità dell'offerta formativa delle loro scuole di provenienza è un'idea sicuramente appropriata per i percorsi liceali e di quelli tecnici (fino al 40% di transizioni), per i quali il passaggio successivo verso studi di livello universitario è considerevole e comunque più probabile. Mentre altre metriche, basate essenzialmente sugli esiti lavorativi, sembrano essere le più appropriate per i percorsi più professionalizzanti (istituti professionali)¹.

Lo scopo di questo documento è quello di fornire in forma sintetica ma il più possibile esauriente i dettagli relativi alle fonti dei dati utilizzate, alle ipotesi e alla logica alla base delle scelte metodologiche, alle tecniche e ai risultati di stima delle performance delle scuole per come sono riportate in **eduscopio.it**. Inoltre, il documento dà conto dello stato corrente della riflessione della Fondazione Giovanni Agnelli rispetto alla metrica utilizzata e ai criteri di comparazione tra scuole più appropriati. Tale riflessione era stata avviata nel 2009 ed era proseguita nel 2012 nell'ambito di due edizioni preliminari e sperimentali dell'analisi, limitate rispettivamente alle sole scuole del Piemonte e alle scuole di quattro regioni (Piemonte, Lombardia, Emilia Romagna e Calabria)².

Al di là dell'estensione del perimetro d'analisi all'intero territorio nazionale, nel corso del tempo, anche gli aspetti più tecnici e metodologici hanno subito una sostanziale evoluzione grazie ai commenti e ai suggerimenti ricevuti da più parti. In virtù, di questi cambiamenti oltre che per il naturale evolversi dei fenomeni osservati può non esserci una stretta assonanza tra i risultati dei precedenti esperimenti e quelli di **eduscopio.it**. Allo stesso modo, le divergenze di tipo metodologico dovute alle diverse finalità delle analisi

¹ Per questi ultimi si rimanda alla sezione del portale relativa agli esiti lavorativi dei diplomati che coinvolge le scuole di sette regioni italiane (Piemonte, Lombardia, Veneto, Emilia-Romagna, Lazio, Basilicata e Sardegna) e al relativo rapporto tecnico.

² A tal proposito si vedano De Simone – Monastero – Stanchi (2009) e FGA 2012 (<http://www.fga.it/interventi-e-analisi/tutti-gli-interventi/dettaglio/article/valutare-la-qualita-delle-scuole-superiori-attraverso-i-risultati-dei-loro-diplomati-immatricolati-n-1.html>).

suggeriscono cautela nella comparazione dei risultati di questo lavoro con quelli di altre esperienze prodotte in tempi recenti su scala più limitata³.

2. I dati

2.1. Fonti dei dati

Le analisi di **eduscopio.it** si basano sulle informazioni contenute in tre *dataset* di origine amministrativa (MIUR):

- A. Anagrafe delle Scuole
- B. Scuola in Chiaro
- C. Anagrafe Nazionale dello Studente universitario

Dall'Anagrafe delle Scuole abbiamo tratto le informazioni necessarie a identificare le scuole secondarie di II grado operanti sul territorio italiano (statali e paritarie)⁴. In particolare, da questa fonte traiamo le informazioni relative a:

- Codice meccanografico identificativo dell'istituto principale;
- Codici meccanografici relativi alle articolazioni scolastiche dei singoli istituti (indirizzi di studio, plessi, etc);
- Denominazione delle scuole;
- Localizzazione delle scuole (indirizzo, comune, provincia).

Nell'edizione 2017/18, gli anni scolastici rilevanti per la nostra analisi sono: 2011/12, 2012/13, 2013/14. Per ogni singola scuola identificata attraverso i codici meccanografici di istituto e di scuola, abbiamo ricavato dalla Scuola in Chiaro informazioni di maggior dettaglio su:

- Indirizzi di studio di ogni singolo istituto;
- Numero di diplomati per indirizzi di studio;
- Voto di maturità medio dei diplomati per indirizzo di studio.

³ Si vedano, ad esempio, l'esperienza dell'indicatore IRIS dell'Università di Genova (http://www.studenti.unige.it/sites/www.studenti.unige.it/files/Iris_2014tot.pdf) o il lavoro di Bratti-Checchi-Filippin (2011) sulle scuole lombarde.

⁴ Esiste l'eccezione della Provincia Autonoma di Bolzano che raccoglie e gestisce indipendentemente dal MIUR i dati relativi alle proprie scuole. Non disponendo di tali informazioni, non possiamo prendere in considerazione gli istituti presenti sul suo territorio.

A partire dall'Anagrafe delle Scuole e dalla Scuola in Chiaro siamo dunque in grado di circoscrivere il campo dell'analisi ai soli istituti, o meglio ai soli indirizzi di studio offerti dai singoli istituti, oggetto dell'analisi di **eduscopio.it**, quelli cioè afferenti al comparto liceale e tecnico.

Infine, gli esiti al primo anno di università dei diplomati immatricolati in corsi di laurea sono stati ricostruiti a partire dalle informazioni contenute nella banca dati dell'Anagrafe Nazionale degli Studenti universitari (ANS) del MIUR. Questa banca dati raccoglie i dati amministrativi sui singoli iscritti inviati dalle segreterie di ateneo al CINECA (Bologna). I dati utilizzati per l'analisi sono stati forniti secondo criteri di 'opacità' in modo da impedire la riconoscibilità del singolo individuo, in osservanza delle norme sulla tutela della privacy.

Nei tre anni accademici di riferimento (2012/13, 2013/14, 2014/15), grazie all'ANS disponiamo dei dati relativi a tutti gli studenti iscritti in atenei italiani (pubblici e privati). Tuttavia l'ANS non copre per il momento altri enti di istruzione avanzata come l'Alta Formazione Artistica, Musicale e Coreutica o gli Istituti Tecnici Superiori. Inoltre, non esistono banche dati integrate che tengano conto anche delle immatricolazioni in corsi di primo livello in atenei esteri che, in ogni caso, riguardano un numero molto limitato di diplomati.

Dunque per gli immatricolati negli atenei italiani riusciamo ad ottenere dati relativi a:

- caratteristiche demografiche – genere, luogo di residenza, origine italiana/straniera, anno di nascita;
- studi scolastici compiuti - titolo di scuola media superiore conseguito (tipo maturità), istituto che lo ha rilasciato (codice meccanografico), anno solare di conseguimento del diploma, votazione all'esame di maturità, età al diploma (bocciature);
- esiti universitari al primo anno - anno di immatricolazione, tipo di iscrizione (tempo pieno o parziale), crediti formativi acquisiti, crediti acquisiti per stage, crediti riconosciuti validi per il corso, esami sostenuti, votazione ottenuta per ogni singolo esame;
- caratteristiche del corso di studi - ateneo, classe del corso, sede didattica, denominazione corso, eventuale presenza di numero chiuso o programmato.

2.2. Predisposizione *dataset* integrato

Nei tre anni accademici considerati (2012/13, 2013/14, 2014/15), il numero di immatricolati negli atenei italiani è stato pari a 895.945 studenti. Tra questi abbiamo individuato quelli riconducibili ai diplomati negli anni scolastici di riferimento (2011/12, 2012/13, 2013/14). Abbiamo altresì escluso quelli con maturità conseguite all'estero e quelli con un'età al diploma superiore ai 22 anni (pluri-bocciati)⁵. In definitiva, il numero finale degli immatricolati rilevanti per la nostra analisi è pari a 703.601 unità.

Le consistenze numeriche e le distribuzioni delle diverse caratteristiche individuali sono riportate nella Tabella 1.

Tabella 1: Variabili descrittive e consistenze numeriche degli immatricolati
(eduscopio.it – edizione 2017/18)

Variabili	Valore assoluto	Valore Percentuale
Genere		
<i>Maschio</i>	305955	43,48
<i>Femmina</i>	397646	56,52
Origine		
<i>Nato in italiano</i>	682105	96,94
<i>Nato all'estero</i>	21496	3,06
Età al diploma		
<i>18</i>	37518	5,33
<i>19</i>	568598	80,81
<i>20</i>	79512	11,30
<i>21</i>	15278	2,17
<i>22</i>	2695	0,38
Diploma		
<i>Liceo Classico</i>	105326	14,97
<i>Liceo Scientifico</i>	298264	42,40
<i>Liceo delle Scienze Umane</i>	61867	8,80
<i>Liceo Linguistico</i>	53683	7,63
<i>Liceo Artistico</i>	8858	1,26
<i>Tecnico Economico</i>	75831	10,78
<i>Tecnico Tecnologico</i>	67409	9,58
<i>Istruzione professionale</i>	32156	4,57
Voto medio diploma	79,54	
Regione di diploma		
<i>Piemonte</i>	44708	6,35
<i>Valle d'Aosta</i>	397	0,06

⁵ Questi ultimi sono tipicamente studenti che presentano percorsi di studio spezzettati tra scuole diverse e sarebbe improprio attribuire il loro livello di preparazione (o impreparazione) alla sola scuola nella quale hanno in fine conseguito il diploma.

<i>Lombardia</i>	102745	14,60
<i>Trentino-Alto Adige</i>	8696	1,24
<i>Veneto</i>	53684	7,63
<i>Friuli-Venezia Giulia</i>	12479	1,77
<i>Liguria</i>	15974	2,27
<i>Emilia-Romagna</i>	45384	6,45
<i>Toscana</i>	38639	5,49
<i>Umbria</i>	10176	1,45
<i>Marche</i>	20564	2,92
<i>Lazio</i>	70655	10,04
<i>Abruzzo</i>	18376	2,61
<i>Molise</i>	4717	0,67
<i>Campania</i>	86170	12,25
<i>Puglia</i>	54669	7,77
<i>Basilicata</i>	8954	1,27
<i>Calabria</i>	28860	4,10
<i>Sicilia</i>	60802	8,64
<i>Sardegna</i>	16952	2,41
Anno Accademico di immatricolazione		
<i>2009/10</i>	222160	31,57
<i>2010/11</i>	238517	33,90
<i>2011/12</i>	242924	34,53
Anni di ritardo tra diploma e immatricolazione		
<i>0</i>	661098	93,96
<i>1</i>	37870	5,38
<i>2</i>	4633	0,66
Settore disciplinare di immatricolazione		
<i>Economico-Statistico</i>	105528	15,00
<i>Giurisprudenza-Scienze Politiche</i>	90212	12,82
<i>Medicina e Chirurgia</i>	23850	3,39
<i>Professioni Sanitarie</i>	39675	5,64
<i>Scienze matematiche, fisiche e naturali</i>	125800	17,88
<i>Scienze Motorie</i>	14280	2,03
<i>Sociale</i>	49945	7,10
<i>Tecniche</i>	127560	18,13
<i>Umanistiche</i>	126750	18,01
Distanza media tra comune di diploma e comune di immatricolazione (km)	91,66	
Numero di crediti nel primo anno accademico		
<i>Nessun CFU</i>	106880	15,19
<i>Almeno 1 CFU</i>	596720	84,81
Percentuale di CFU nel primo anno accademico	55,54	
Voto medio esami nel primo anno accademico	24,81	

A partire dai tati individuali possiamo calcolare le informazioni aggregate sui diplomati delle scuole secondarie di II grado transitati all'università (tasso d'immatricolazione, % di immatricolati con libretto vuoto al primo anno, voto medio di maturità degli immatricolati, scelte d'immatricolazione). Tali informazioni, insieme a quelle attinte dalla Scuola in Chiaro, ci permettono di definire gli indicatori descrittivi riportati nelle *schede scuola* di **eduscopio.it**.

Come spiegato nella sezione introduttiva, comparare le scuola sulla base degli esiti universitari dei loro diplomati ha senso solo se parliamo di indirizzi liceali (classico, scientifico, scienze umane, linguistico) e tecnici (economico e tecnologico, così come riclassificati dopo la riforma della scuola secondaria di II grado del 2010). Infatti, è da queste scuole che un numero significativo di studenti transita verso i corsi universitari.

Diverso è il caso degli artistici per i quali l'università assorbe solo il 30% dei diplomati. Un'altra quota presumibilmente significativa è assorbita dalle Accademie delle Belle Arti per le quali tuttavia non disponiamo di dati, non essendo questi raccolti nell'ANS. Nelle prime edizioni abbiamo preferito non elaborare né fornire i dati relativi ai licei artistici proprio a causa della loro parzialità. Dall'edizione dello scorso anno, anche in virtù delle richieste pervenute da docenti e famiglie, proviamo a dare conto della capacità manifestata dai licei artistici nella preparazione degli studenti ai corsi universitari. Benché parziale, si tratta di un'informazione comunque utile per le scelte delle famiglie e l'autovalutazione delle scuole, anche se necessita di maggior cautela nell'interpretazione.

In generale, per dare una maggiore solidità statistica ai risultati, abbiamo deciso di considerare solo le scuole per le quali osserviamo che un congruo numero di diplomati (almeno 1 su 3) prosegue gli studi a livello universitario. Quando i tassi di prosecuzione sono così significativi, le scuole non possono non avere tra i propri obiettivi formativi quello di preparare per l'istruzione terziaria. Sempre per ragioni di solidità statistica dei risultati, per evitare che il risultato complessivo dipenda dalle performance di pochi studenti particolarmente brillanti o carenti, abbiamo introdotto un'altra condizione: consideriamo solo le scuole che per il singolo indirizzo di studio mandino all'università un numero non inferiore a 21 diplomati (la dimensione media di una classe quinta) nell'arco del triennio considerato. Si tratta di due soglie ricavate da un'analisi visiva delle discontinuità nelle distribuzioni, rispettivamente, dei tassi di transizione e delle numerosità medie di studenti

giunti al diploma per ogni scuola. L'adozione di una soglia relativa (almeno un terzo di diplomati per indirizzo di studio devono proseguire) e di una soglia assoluta (almeno 21 studenti per indirizzo di studio in un triennio) attenua considerevolmente il rischio di distorsione nelle misurazioni.

Le consistenze numeriche delle scuole (indirizzi di studio) ammesse al confronto e gli indicatori descrittivi delle loro caratteristiche essenziali sono riportate nella Tabella 2.

Tabella 2: Variabili descrittive a livello scuola
(eduscopio.it – edizione 2017/18)

Variabili	Valore assoluto	Valore Percentuale
Indirizzi di diploma		
<i>Liceo Classico</i>	669	15,27
<i>Liceo Scientifico</i>	1355	30,93
<i>Liceo delle Scienze Umane</i>	506	11,55
<i>Liceo Linguistico</i>	597	13,63
<i>Liceo Artistico</i>	77	1,76
<i>Tecnico Economico</i>	652	14,88
<i>Tecnico Tecnologico</i>	525	11,98
Scuola statale/paritaria		
<i>Statale</i>	3764	85,92
<i>Paritaria</i>	617	14,08
Regione scuola		
<i>Piemonte</i>	314	7,17
<i>Lombardia</i>	687	15,68
<i>Trentino-Alto Adige</i>	55	1,26
<i>Veneto</i>	329	7,51
<i>Friuli-Venezia Giulia</i>	80	1,83
<i>Liguria</i>	95	2,17
<i>Emilia-Romagna</i>	269	6,14
<i>Toscana</i>	280	6,39
<i>Umbria</i>	65	1,48
<i>Marche</i>	134	3,06
<i>Lazio</i>	443	10,11
<i>Abruzzo</i>	110	2,51
<i>Molise</i>	36	0,82
<i>Campania</i>	455	10,39
<i>Puglia</i>	291	6,64
<i>Basilicata</i>	68	1,55
<i>Calabria</i>	174	3,97
<i>Sicilia</i>	356	8,13
<i>Sardegna</i>	140	3,20
Numero medio diplomati annuo	66,15	
Percentuale diplomati femmine		58,25

Percentuale diplomati stranieri		3,13
Media voto diplomati	77,08	
Media voto immatricolati	79,21	
Tasso di immatricolazione		70,65
Percentuale immatricolati con zero CFU nel primo anno acc.		10,49
Percentuale di CFU nel primo anno accademico		54,31
Voto medio esami nel primo anno accademico	24,66	

3. Metodologia statistica: una procedura a due step

3.1. Le misure di *outcome*

Per ogni studente i appartenente alla scuola j e immatricolato nell'anno t abbiamo preso in considerazione due indicatori di performance al primo anno⁶ di università:

- **Media dei voti** conseguiti agli n esami universitari sostenuti con successo, ponderata per i crediti formativi (CFU) di ciascun esame (così come stabilito dal singolo corso di laurea c nell'ateneo u), secondo la formula seguente⁷

$$media_t^{ij} = \frac{\sum_{k=1}^{n_i} voto_{tk}^{ij} * CFU_k^{ij}}{\sum_{k=1}^{n_i} CFU_k^{ij}}$$

- **Crediti formativi universitari ottenuti**, in percentuale sul totale previsto per il primo anno⁸ nel corso di laurea c dello specifico ateneo u , secondo la formula seguente

$$rendimento_t^{ij} = \frac{\sum_{k=1}^{n_i} CFU_k^{ij}}{CFU^{ij}}$$

Ci siamo limitati a considerare il primo anno del corso di laurea universitario per due ordini di ragioni.

⁶ Date le diverse articolazioni dei corsi di laurea, consideriamo come primo anno di università una sequenza di quattro sessioni d'esame: inverno (1a sessione), primavera/estate (2a sessione), fine estate/ inizio autunno (3a sessione), inverno successivo (4a sessione). Dunque il limite temporale concesso agli studenti per completare gli esami del primo anno è il 30 aprile dell'anno accademico successivo a quello di immatricolazione.

⁷ Per ridurre la complessità della notazione abbiamo soppresso l'indicazione del corso di laurea e dell'ateneo di appartenenza.

⁸ Si tiene ovviamente conto del fatto che lo studente possa essere iscritto a tempo pieno o parziale (studenti-lavoratori) e, dunque, che il numero di crediti totale previsto per il primo anno possa variare da individuo a individuo. Il rendimento è calcolato sulla base di quanto scelto dallo studente al momento dell'immatricolazione in termini di "impegno dichiarato" (CFU^{ij}).

In primo luogo, la performance al primo anno è fortemente predittiva degli esiti universitari nel medio-lungo periodo. Dai dati dell'ANS risulta che oltre il 60% di coloro che abbandonano non ha sostenuto alcun esame. Se si considerano anche coloro che hanno sostenuto al massimo due esami nel primo biennio, la percentuale sale all'80%. In un lavoro recente, Aina et al. (2016) dimostrano che la performance al primo anno determina gran parte della probabilità di conseguire la laurea, della probabilità di conseguire la laurea con il massimo dei voti, della probabilità di laurearsi in regola e senza andare fuoricorso.

In secondo luogo, ipotizziamo che l'effetto della scuola di origine (conoscenze e competenze acquisite, metodo di studio, orientamento in uscita) si espliciti quasi interamente nel primo anno di istruzione terziaria; dopodiché è probabile che predominino le conoscenze fondamentali apprese "sul campo", nell'ambito del corso di studi. Pertanto sarebbe ingiusto attribuire alla scuola secondaria ciò che in realtà dipende in larga misura dall'università stessa e dall'esperienza individuale. Si tratta ovviamente di un'assunzione arbitraria, mancando studi specifici su questo tema; ma abbiamo avviato uno studio per testare la robustezza di questa ipotesi e per stimare il profilo temporale della persistenza dell'effetto scuola durante gli anni universitari. Presto saremo in grado di appurare se questa ipotesi possa essere mantenuta o debba essere superata, modificando di conseguenza anche le metriche di **eduscopio.it**.

3.2. Primo step: la standardizzazione delle misure di *outcome*

Le performance universitarie dei singoli studenti possono essere considerate come il risultato di un insieme di fattori che operano congiuntamente. A parità di abilità, preparazione e impegno, di certo contano le scelte universitarie. Gli atenei e i corsi di studio differiscono tra loro in termini di severità, difficoltà intrinseca, qualità della didattica. Oltre a questi vi sono molti altri aspetti legati all'offerta formativa delle università che possono avere un impatto sulla capacità degli studenti di superare gli esami e conseguire buone valutazioni. Si tratta di fattori fuori dal controllo delle scuole secondarie di II grado che non possono essere chiamate a risponderne. Pertanto, qualsiasi tipo di considerazione sui meriti dei singoli studenti e delle scuole che li hanno preparati alla prosecuzione universitaria deve prescindere dai possibili effetti distorsivi delle scelte universitarie degli individui.

Per standardizzare gli *outcomes* universitari calcoliamo dei fattori di correzione dei punteggi legati ai corsi di laurea (c), di ciascun ateneo, (u). Tali fattori sono stimati a partire dai dati relativi a tutti gli immatricolati iscritti negli atenei italiani, attraverso un modello di regressione definito come segue:

$$y_{cut}^i = \alpha_0 + \beta_1' X_{cut}^i + \beta_2' Z_{cut}^i + \sum \varphi_{cu} + \sum \tau_t + \varepsilon_{cut}^i$$

dove

- $y_{cut}^i \equiv \{media_t^{ij}, rendimento_t^{ij}\}$ è di volta in volta l'*outcome* d'interesse;
- α_0 è una costante;
- X_t^i è un vettore di caratteristiche demografiche dello studente; in particolare consideriamo il genere, l'origine italiana o straniera, la distanza tra il comune della scuola frequentata e quello sede del corso di laurea (per controllare per l'effetto del pendolarismo e dell'essere "fuorisede" sui rendimenti universitari)⁹;
- Z_t^i è un vettore che contiene le variabili relative alla carriera scolastica dello studente; in particolare, voto di maturità conseguito, tipo di scuola (statale non statale), tipo di maturità (indirizzo di studi) e regione nel quale è stato conseguito il diploma (per controllare per eventuali differenze nei *grading standards* tra scuole di diverso tipo e/o su base territoriale)¹⁰, età al diploma (bocciature), età all'immatricolazione (per tener conto delle immatricolazioni ritardate rispetto all'anno di diploma);
- $\sum \varphi_{cu}$ è il set di dummies "corso di laurea-ateneo" che restituirà il valore dei fattori di correzione;
- $\sum \tau_t$ è un set di dummies che identifica l'anno accademico di immatricolazione;
- ε_t^i è un termine residuale distribuito normalmente.

⁹ Quest'ultima variabile è inclusa nel modello anche al quadrato per cogliere eventuali effetti non lineari della distanza.

¹⁰ La questione della scarsa affidabilità dei voti di maturità quali indicatori delle competenze e delle conoscenze degli studenti ha avuto in anni recenti ampio risalto nel dibattito pubblico (si veda la polemica sui "bonus maturità") e in quello scientifico. Le differenze negli standard di valutazione su base territoriale e tra indirizzi formativi sono note e i tentativi di irrigidire i criteri di calcolo delle voto finale di maturità non hanno prodotto risultati tangibili. La variabilità dei *grading standards* delle scuole è indagata e discussa anche in Bratti-Checchi-Filippin (2011).

Le caratteristiche individuali degli studenti sono incluse nella regressione per attenuare la distorsione nelle stime dei fattori di correzione (φ_{cu}). Infatti con l'ipotesi di *selection on observables* possiamo tenere conto per quanto possibile della "selezione favorevole" (autoselezione) di studenti con particolari profili di abilità in specifici corsi di laurea.

I risultati delle stime sotto diverse specificazioni sono riportati in Tabella 3. Più precisamente per ognuno dei due *outcomes* (rendimento, media) riportiamo le stime di una specificazione con le sole caratteristiche individuali (demografiche e di carriera) nelle colonne (1) e (4), rispettivamente; una specificazione che include le dummies "corso di laurea-ateneo" ($\sum \varphi_{cu}$) nelle colonne (2) e (5)¹¹. Dal confronto tra specificazioni con e senza dummies "corso di laurea-ateneo" si evince che le scelte di immatricolazione giocano un ruolo importante nei rendimenti scolastici osservati: il potere esplicativo del modello aumenta del 42% nel caso dei rendimenti e del 58% nel caso delle medie. Dunque i corsi di studio universitari sono in effetti molto differenziati tra loro per qualità e *grading standards* e le correzioni che ci troveremo ad operare sugli esiti universitari per renderli comparabili saranno importanti.

Per dare maggiore flessibilità al modello ed attenuare ulteriormente le distorsioni nelle stime dei fattori di correzione, abbiamo testato un'ultima specificazione che oltre alle dummies "corso di laurea-ateneo" sostituisce i due gruppi di dummies sul tipo di maturità e sulla regione di diploma fin qui considerate separatamente, con il set di dummies che scaturisce dalla loro intersezione¹². I risultati riportati nelle colonne (3) e (6) della Tabella 3 rivelano che tali ulteriori differenziazioni non modificano il quadro rispetto a quanto visto in colonna (2) e (4).

¹¹ Per esigenze di spazio omettiamo i coefficienti stimati per le variabili relative al tipo di maturità e alle regioni di diploma e indichiamo soltanto la loro presenza nella specificazione. Analogamente, omettiamo le 496 dummies "corso di laurea-ateneo" indicando tuttavia se sono presenti nella specificazione.

¹² Se prima ipotizzavamo che le differenze territoriali nelle politiche di voto alla maturità prescindessero dagli indirizzi di studio (e viceversa), adesso cerchiamo di capire se invece le differenze negli standard valutativi delle scuole possano invece variare da indirizzo a indirizzo in modo differenziato, a seconda dell'area territoriale.

Tabella 3: Stima dei fattori di correzione
(eduscopio.it – edizione 2017/18)

Variabili	Rendimento (% CFU)			Media voti		
	(1)	(2)	(3)	(4)	(5)	(6)
Genere (rif. Maschio):						
<i>Femmina</i>	0.0296*** (0.000826)	0.00289*** (0.000814)	0.00311*** (0.000814)	0.400*** (0.00719)	-0.0135** (0.00688)	-0.0115* (0.00687)
Origine: (rif. Nato in Italia)						
<i>Nato all'estero</i>	-0.0376*** (0.00220)	-0.0283*** (0.00209)	-0.0253*** (0.00209)	-0.413*** (0.0200)	-0.215*** (0.0184)	-0.190*** (0.0184)
Età al diploma	-0.0165*** (0.000483)	-0.0172*** (0.000461)	-0.0169*** (0.000461)	0.0229*** (0.00480)	-0.0346*** (0.00443)	-0.0337*** (0.00443)
Voto di maturità	0.00908*** (3.34e-05)	0.0103*** (3.29e-05)	0.0103*** (3.29e-05)	0.0925*** (0.000292)	0.105*** (0.000281)	0.105*** (0.000281)
Ente gestore (rif. Scuola statale): <i>Scuola non statale</i>	-0.0945*** (0.00149)	-0.112*** (0.00142)	-0.113*** (0.00143)	-0.573*** (0.0138)	-0.669*** (0.0127)	-0.681*** (0.0128)
Anni di ritardo tra diploma e immatricolazione (rif. 0 anni):						
<i>1 anno di ritardo</i>	0.00103 (0.00170)	-0.00946*** (0.00162)	-0.00916*** (0.00162)	0.437*** (0.0154)	0.259*** (0.0142)	0.260*** (0.0142)
<i>2 anni di ritardo</i>	0.0206*** (0.00465)	0.00699 (0.00440)	0.00729* (0.00440)	0.788*** (0.0429)	0.558*** (0.0394)	0.558*** (0.0394)
Distanza tra comune di diploma e sede universitaria	0.000166*** (6.76e-06)	9.34e-05*** (7.72e-06)	9.22e-05*** (7.74e-06)	0.000489*** (5.81e-05)	-0.000335*** (6.46e-05)	-0.000382*** (6.48e-05)
Distanza tra comune di diploma e sede universitaria²	-8.91e-08*** (8.10e-09)	-9.77e-08*** (8.92e-09)	-9.49e-08*** (8.95e-09)	-1.73e-06*** (6.97e-08)	1.61e-07** (7.46e-08)	2.20e-07*** (7.48e-08)
Anno acc. di immatr. (rif. 2012/13):						
<i>2013/14</i>	-0.000135 (0.000933)	-0.00205** (0.000884)	-0.00196** (0.000883)	-0.0205** (0.00813)	-0.0449*** (0.00746)	-0.0441*** (0.00746)
<i>2014/15</i>	0.0152*** (0.000935)	0.0127*** (0.000888)	0.0128*** (0.000887)	0.0473*** (0.00812)	0.0210*** (0.00747)	0.0222*** (0.00747)
Costante	0.167*** (0.0101)	0.0794*** (0.00989)	0.0933*** (0.0106)	17.29*** (0.0987)	17.68*** (0.0931)	17.80*** (0.0979)
Tipo maturità (dummies)	si	si	no	si	si	no
Regione diploma (dummies)	si	si	no	si	si	no
Tipo maturità I Regione diploma (dummies)	no	no	si	no	no	si
Corso I Ateneo (dummies)	no	si	si	no	si	si
N. Osservazioni	705,394	705,394	705,394	586,479	586,479	586,479
R ²	0.201	0.286	0.288	0.223	0.349	0.350

Note: Standard errors robusti in parentesi. Significatività: *** p<0.01, ** p<0.05, * p<0.1

Dunque, per i due *outcomes* scegliamo le specificazioni riportate in colonna (2) e colonna (4) e da queste estraiamo i fattori di correzione che verranno sottratti ai punteggi individuali di ogni studente secondo la seguente formula:

$$\tilde{y}_t^i = y_{cut}^i - \varphi_{cu}.$$

A questo punto chi frequenta un corso particolarmente penalizzante in termini valutativi (segno stimato negativo) vedrà i propri punteggi aumentati di un valore equivalente al fattore di correzione. Viceversa, chi è iscritto ad un corso di studi particolarmente generoso vedrà i propri punteggi diminuiti di un valore equivalente al fattore di correzione.

Tale standardizzazione rende le performance universitarie dei diplomati comparabili (vedi Tabella 4). A parità di valore medio di rendimenti e medie dei voti, gli *outcomes* corretti ($\widetilde{media}_t^{ij}, \widetilde{rendimento}_t^{ij}$) possono assumere valori fuori dall'intervallo di variazione originario, proprio in virtù dell'applicazione del fattore di correzione. In alcuni corsi di laurea a un 18 corrisponde un voto che sarebbe insufficiente in altri corsi di laurea dello stesso tipo offerti da altri atenei; analogamente, alcuni 30 valgono fino a 3,5 punti in più se messi a confronto con i 30 di corsi analoghi presso altri atenei.

Tabella 4: Statistiche descrittive delle misure di outcome prima e dopo la standardizzazione (eduscopio.it – edizione 2017/18)

Variabili	N. Oss.	Media	Std. Dev.	Min	Max
$rendimento_t^{ij}$	705394	0,5548	0,3504	0	1
$media_t^{ij}$	586479	24,81	28,23	18	31
$\widetilde{rendimento}_t^{ij}$	705394	0,5548	0,3430	-0,4133	1,3772
$\widetilde{media}_t^{ij}$	586479	25	27,43	14,29	34,28

3.3. Secondo step: il calcolo delle medie di scuola con effetti fissi e dati *pooled*

La performance di ogni scuola è ottenuta attraverso una stima su dati *pooled* nel triennio considerato con una specificazione di questo tipo:

$$\tilde{y}_t^{ij} = \mu_0 + \sum \theta^j + \sum \pi_t + \rho_t^i$$

dove

- $\tilde{y}_t^{ij}$ è di volta in volta l'*outcome* d'interesse per lo studente i diplomatosi nella scuola j ;
- μ_0 è una costante;
- $\sum \theta^j$ è il set di dummies, una per indirizzo in ciascuna scuola, che restituirà la performance di ogni singola scuola rispetto all'*outcome* considerato;
- $\sum \pi_t$ è un set di dummies che identifica l'anno scolastico nel quale lo studente si è diplomato;
- ρ_t^i è un termine residuale distribuito normalmente.

Una volta ottenuti i risultati delle regressioni, i valori delle medie e dei rendimenti a livello scuola ($\theta^j \equiv \{media^j, rendimento^j\}$) sono stati riscalandi tra 0 e 100 e utilizzati per il calcolo di un indice che tiene conto congiuntamente e in pari misura (50/50) dei due *outcomes*¹³:

$$indice\ FGA^j = media^j * 0.5 + rendimento^j * 0.5$$

ottenendo così le misure finali per il confronto tra scuole le cui caratteristiche descrittive sono riportate in Tabella 5.

**Tabella 5: Statistiche descrittive degli *outcomes* a livello scuola
(eduscopio.it – edizione 2017/18)**

Variabili	N. Oss.	Media	Std. Dev.	Min	Max
<i>rendimento^j</i>	4381	58,87	14,53	0	100
<i>media^j</i>	4381	54,56	15,32	0	100
<i>Indice FGA^j</i>	4381	56,72	14,09	0	97,68

Nella nostra analisi su dati *pooled*, la capacità di una scuola di preparare bene agli studi universitari è rivelata dalla media delle performance standardizzate dei suoi studenti diplomatisi in tre anni successivi. Tenere conto simultaneamente delle tre coorti di diplomati fa sì che le scuole vengano valutate per la capacità educativa espressa da un numero maggiore di consigli di classe (insegnanti). La misura sarà dunque più realistica

¹³ Si noti che poiché il rendimento vale 0 per gli studenti con libretto d'esami vuoto al primo anno, l'indicatore tiene in qualche modo conto del fenomeno del *dropout* universitario e dunque restituisce indirettamente anche un giudizio sulla resilienza media degli studenti di ogni singola scuola.

poiché meno dipendente dalle particolarità di una singola “annata” di diplomati e/o di un piccolo gruppo di docenti. Per dare un’idea del grado di *smoothing* ottenuto con una media su tre coorti rispetto a quanto rivelato dalle singole annualità, proponiamo in Tabella 6 un’analisi della correlazione tra le posizioni ottenute da ciascuna scuola nel *ranking* stimato su ogni singolo anno e sul triennio. In pratica, andiamo a valutare qual è il grado di stabilità delle performance osservate di coorte in coorte.

Dalla Tabella 6 appare evidente che sebbene la correlazione sia positiva, elevata e significativa, non vi è una piena convergenza tra quanto emerge nei singoli anni: le differenze tra coorti di studenti e probabilmente tra consigli di classe (team di docenti) sono tali da comportare degli aggiustamenti non trascurabili nel *ranking*. Una media sul triennio appiana queste oscillazioni e attenua il rischio di una misurazione troppo dipendente dalle specificità di un’annualità.

Tabella 6: Correlazione tra le posizioni occupate dalle singole scuole in ognuno degli anni considerati (eduscopio.it – edizione 2017/18)

Rendimento				
	Ranking 2011/12	Ranking 2012/13	Ranking 2013/14	Ranking Triennio
Ranking 2011/12	1.000			
Ranking 2012/13	0,646	1.000		
Ranking 2013/14	0,63	0,6332	1.000	
Ranking Triennio	0,8754	0,8696	0,8538	1.000
Media				
	Ranking 2011/12	Ranking 2012/13	Ranking 2013/14	Ranking Triennio
Ranking 2011/12	1.000			
Ranking 2012/13	0,5871	1.000		
Ranking 2013/14	0,5893	0,5649	1.000	
Ranking Triennio	0,8475	0,8454	0,8316	1.000

Nota: sig. $p < 0.000$

In Tabella 7 proponiamo dei confronti tra le performance delle scuole sulla base degli indirizzi di studio e della localizzazione territoriale. È facile osservare come queste riflettano i divari già messi in luce dalle indagini sugli apprendimenti nazionali (INVALSI) e internazionali (OCSE-PISA). È bene rimarcare che anche in questo caso le diverse performance tra i comparti educativi non riflettono unicamente differenze nella qualità dell’offerta formativa, ma anche altri fattori, tra i quali la tipica stratificazione dei diversi

indirizzi di studio per abilità scolastica ed estrazione socio-culturale degli studenti (autoselezione). Peraltro, quando si guarda agli esiti universitari, le scuole con curricula più prettamente accademici (licei) sono per costruzione avvantaggiate.

Tabella 7: Correlazione tra outcomes e caratteristiche delle scuole (indirizzo di studio, territorio)
(eduscopio.it – edizione 2017/18)

Variabili	Rendimenti			Medie		
	(1)	(2)	(3)	(4)	(5)	(6)
Indirizzi di diploma (rif. Liceo Classico)						
Liceo Scientifico	0.00443 (0.00619)		-0.00100 (0.00582)	-0.0635*** (0.00625)		-0.0734*** (0.00562)
Liceo delle Scienze Umane	-0.134*** (0.00771)		-0.142*** (0.00725)	-0.209*** (0.00780)		-0.221*** (0.00701)
Liceo Linguistico	-0.0557*** (0.00737)		-0.0647*** (0.00693)	-0.102*** (0.00745)		-0.118*** (0.00670)
Liceo Artistico	-0.129*** (0.0158)		-0.153*** (0.0149)	-0.0217 (0.0159)		-0.0614*** (0.0144)
Tecnico Economico	-0.147*** (0.00721)		-0.158*** (0.00680)	-0.225*** (0.00728)		-0.245*** (0.00658)
Tecnico Tecnologico	-0.112*** (0.00763)		-0.131*** (0.00724)	-0.150*** (0.00772)		-0.179*** (0.00700)
Regione scuola (rif. Piemonte)						
Lombardia		0.00291 (0.00952)	0.00360 (0.00836)		0.0199** (0.00982)	0.0201** (0.00808)
Prov. Aut. Trento		0.0539*** (0.0204)	0.0664*** (0.0179)		0.0576*** (0.0211)	0.0731*** (0.0173)
Veneto		0.0284** (0.0110)	0.0303*** (0.00968)		0.0194* (0.0114)	0.0214** (0.00936)
Friuli-Venezia Giulia		-0.00208 (0.0175)	-7.77e-05 (0.0154)		0.0465** (0.0181)	0.0507*** (0.0149)
Liguria		-0.0442*** (0.0164)	-0.0541*** (0.0144)		-0.00354 (0.0169)	-0.0168 (0.0139)
Emilia-Romagna		0.00972 (0.0116)	0.0153 (0.0102)		0.0329*** (0.0120)	0.0404*** (0.00986)
Toscana		-0.0447*** (0.0115)	-0.0421*** (0.0101)		0.0117 (0.0119)	0.0144 (0.00975)
Umbria		-0.0285 (0.0190)	-0.0340** (0.0167)		0.0111 (0.0196)	0.00592 (0.0162)
Marche		-0.0122 (0.0144)	0.00105 (0.0127)		0.0467*** (0.0149)	0.0637*** (0.0123)
Lazio		-0.0644*** (0.0103)	-0.0827*** (0.00907)		-0.0419*** (0.0106)	-0.0628*** (0.00877)
Abruzzo		-0.111*** (0.0155)	-0.109*** (0.0136)		-0.0440*** (0.0160)	-0.0383*** (0.0131)
Molise		-0.0128	-0.0147		-0.0331	-0.0342

		(0.0246)	(0.0216)		(0.0254)	(0.0209)
<i>Campania</i>		-0.0880***	-0.104***		-0.0594***	-0.0756***
		(0.0103)	(0.00903)		(0.0106)	(0.00873)
<i>Puglia</i>		-0.0291**	-0.0397***		-0.0556***	-0.0653***
		(0.0114)	(0.0100)		(0.0117)	(0.00967)
<i>Basilicata</i>		-0.0225	-0.0197		-0.0815***	-0.0741***
		(0.0187)	(0.0164)		(0.0193)	(0.0159)
<i>Calabria</i>		-0.0855***	-0.0930***		-0.161***	-0.169***
		(0.0132)	(0.0116)		(0.0136)	(0.0112)
<i>Sicilia</i>		-0.0221**	-0.0420***		-0.0810***	-0.103***
		(0.0108)	(0.00954)		(0.0112)	(0.00922)
<i>Sardegna</i>		-0.103***	-0.109***		-0.125***	-0.129***
		(0.0142)	(0.0125)		(0.0147)	(0.0121)
Costante	0.648***	0.619***	0.692***	0.655***	0.569***	0.697***
	(0.00506)	(0.00789)	(0.00830)	(0.00512)	(0.00814)	(0.00802)
N. Scuole	4,381	4,381	4,381	4,381	4,381	4,381
R ²	0.189	0.078	0.290	0.254	0.117	0.404

Note: Standard errors robusti in parentesi. Significatività: *** p<0.01, ** p<0.05, * p<0.1

4. Un intervallo di confidenza per l'Indice FGA

L'Indice FGA fotografa la performance di una scuola in un particolare triennio. Essendo l'indagine di **eduscopio.it** condotta su base censuaria questa fotografia non risente dell'incertezza derivante da possibili errori di campionamento. Tuttavia, gli effetti scuola calcolati (*point estimates*) sono pur sempre i valori attesi di una distribuzione di possibili risultati. Dunque, il valore reale della performance può oscillare tra un valore minimo e un valore massimo che dipendono da quanto diversi tra loro sono i risultati dei diplomati di quella scuola (varianza) e dalle dimensioni della scuola stessa (numero dei diplomati). Ad esempio, per una scuola di piccole dimensioni, le performance possono dipendere in modo cruciale dagli exploit o dagli esiti disastrosi di pochi studenti e quindi l'intervallo entro cui può oscillare la performance sarà ampio. Viceversa per le scuole di grandi dimensioni, la stima è più precisa perché la media è calcolata a partire dalle performance individuali di un numero elevato di studenti e, dunque, è più contenuto il rischio che i risultati di uno solo di loro condizionino in modo significativo il risultato della scuola.

Tipicamente, per dare un'idea di quanto sia ampio il campo di variazione possibile per ogni performance si calcola un **intervallo di confidenza** intorno al valore stimato. Si tratta della cosiddetta "forchetta" entro cui in 95 casi su 100 si trova il reale valore della performance della scuola.

A partire dall'edizione 2015 di **eduscopio.it** abbiamo deciso di aggiungere questa informazione nelle schede di approfondimento relative alle singole scuole. In particolare, abbiamo derivato l'intervallo di confidenza relativo ai valori assunti dall'Indice FGA di ciascuna scuola. La forchetta è dunque un'informazione di dettaglio in più per la comparazione tra scuole e si aggiunge alle altre cautele metodologiche adottate nel calcolo dell'Indice FGA (soglie dimensionali delle scuole, immatricolazione di almeno un terzo dei diplomati) per evitare comparazioni improprie o poco significative.

Per costruire la forchetta intorno alla performance di ciascuna scuola j secondo la consueta formula

$$indice\ FGA^j \pm 1.96 \frac{\sigma_{Indice\ FGA}^j}{\sqrt{n^j}}$$

abbiamo bisogno di conoscere il numero di immatricolati provenienti dalla scuola j (n) e la varianza dell'Indice FGA nelle performance individuali degli studenti di quella scuola. La prima informazione è nota. Per quanto riguarda, invece, la deviazione standard dell'indice FGA ($\sigma_{Indice\ FGA}^j$) bisogna ricordare che quest'ultimo è costruito a partire dalla media di due indicatori (rendimenti, media dei voti) riscalatati su un intervallo (0-100) secondo la formula:

$$\theta_{[0-100]}^{i,j} = \frac{\theta^{i,j} - \theta_{min}^i}{\theta_{max}^i - \theta_{min}^i}$$

dove $\theta^{i,j}$ è l'indicatore i (media voti, rendimento) riferito alla scuola j e $\theta_{max/min}^i$ sono i valori massimi o minimi che quell'indicatore assume nella distribuzione dei risultati delle scuole.

Pertanto, è necessario riscalare opportunamente anche le varianze dei due singoli indicatori i e poi calcolare una media delle varianze stesse per ottenere la varianza dell'Indice FGA e da questa la deviazione standard.

Con opportune derivazioni¹⁴ si può dimostrare che la varianza di un indicatore i riscalato tra (0-100) secondo la definizione data sopra può essere essa stessa riscalata secondo la formula:

$$var\left(\theta_{[0-100]}^{i,j}\right) = \frac{1}{(\theta_{max}^i - \theta_{min}^i)^2} var(\theta^{i,j}).$$

¹⁴ Per semplicità si è ipotizzato che i valori minimi e massimi dell'indicatore fossero dati e non estratti essi stessi da una distribuzione di valori casuali. La FGA è disponibile a fornire il dettaglio delle derivazioni della formula semplificata e di quella generalizzata a chiunque ne faccia richiesta.

In definitiva la varianza dell'Indice FGA è data dall'espressione:

$$var_{Indice\ FGA}^j = \frac{1}{4} var(\theta_{[0-100]}^{rendimento,j}) + \frac{1}{4} var(\theta_{[0-100]}^{media\ voti,j}) + \frac{1}{2} covar(\theta_{[0-100]}^{rendimento,j}, \theta_{[0-100]}^{media\ voti,j})$$

e da questa ricaviamo il valore della deviazione standard dell'Indice FGA

5. Analisi della stabilità nel tempo dell'Indice FGA e dei *rankings*

Non avendo inserito novità sui criteri di valutazione delle scuole e spostando di volta in volta il punto di osservazione di un solo anno in avanti, non si osservano variazioni di rilievo nelle performance delle scuole nell'edizione 2017/18 rispetto alle edizioni precedenti di **eduscopio.it**. Le scuole con ottimi risultati vedono confermati i propri risultati e allo stesso modo quelle con pessimi risultati restano confinate nelle parti basse delle classifiche. La correlazione tra gli indicatori di questa edizione rispetto a quelli della precedente (2016/17) è del 95,6% per i rendimenti, del 94,5% per le medie e del 93,7% per l'Indice FGA calcolato sulla base dei due indicatori (cfr. Tabella 9). Come prevedibile, la correlazione cala se confrontiamo l'edizione corrente con quella del 2015/16, ma resta comunque alta. Infatti, in quel caso, due coorti di studenti sulle tre prese in considerazione non coincidevano con quelle analizzate nell'edizione corrente.

In definitiva, la qualità a scuola è qualcosa che si costruisce nel tempo e non si acquisisce o si perde in modo significativo nell'arco di un solo anno.

Tabella 8: Correlazione tra gli indicatori (rendimento, media, Indice FGA) nelle tre edizioni di eduscopio.it

Rendimento			
	Ranking 2017/18	Ranking 2016/17	Ranking 2015/16
Ranking 2017/18	1		
Ranking 2016/17	0,956	1	
Ranking 2015/16	0,911	0,955	1
Media			
	Ranking 2017/18	Ranking 2016/17	Ranking 2015/16
Ranking 2017/18	1		
Ranking 2016/17	0,945	1	
Ranking 2015/16	0,888	0,947	1

Indice FGA			
	Ranking 2017/18	Ranking 2016/17	Ranking 2015/16
Ranking 2017/18	1		
Ranking 2016/17	0,937	1	
Ranking 2015/16	0,870	0,929	1

Anche i *rankings* stilati in base all'Indice FGA presentano una certa stabilità. Se ci focalizziamo sui centri urbani più grandi, dove vi sono numerose scuole anche di qualità comparabile e piccole oscillazioni nell'Indice FGA potrebbero causare rilevanti variazioni nelle posizioni, notiamo che la maggior parte degli spostamenti sono limitati. Ad esempio, nelle aree metropolitane di Milano e Napoli, circa l'83% delle scuole non guadagna o perde più di 3 posizioni nel proprio ranking di riferimento. Dove i numeri si fanno ancora più ampi, come nel caso di Roma, la percentuale di stabilità scende al 61% ma è comunque ancora elevata. Più si riduce la dimensione dei centri urbani e il numero di scuole, minori sono le variazioni di posizione e maggiore è la stabilità delle posizioni.

In definitiva, le distanze relative tra scuole restano non si modificano radicalmente da un anno all'altro. Tuttavia una stessa scuola, pur vedendo invariata la propria posizione, può notare una variazione di alcuni punti nel proprio indice. Oltre che dalla variazione nel tempo delle performance osservate, questo può dipendere dal fatto che, per come è costruito, l'Indice FGA non può non risentire dell'ingresso e dell'uscita (a causa dell'imposizione della soglia dimensionale) di alcune scuole dal novero di quelle considerate. Per queste ragioni, l'Indice FGA va sempre letto in un'ottica comparativa e di confronto tra le distanze relative delle scuole, mentre bisogna aver cautela nell'intraprendere una sua lettura in senso longitudinale e assoluto.

6. Quali confronti sono appropriati?

Una volta individuata una misura comparabile di efficacia nella preparazione agli studi universitari, è opportuno chiedersi quali tipi di confronti tra scuole siano più proficui per gli studenti che vogliano compiere scelte più consapevoli e per i docenti che vogliano avviare riflessioni più circostanziate sugli esiti del proprio operato.

Dal punto di vista degli studenti e delle loro famiglie, la comparazione tra scuole è tanto più rilevante quanto più è circoscritta in termini territoriali. A differenza delle scelte universitarie che risentono anche di una certa disponibilità alla mobilità da parte dei diplomati, la scelta della scuola secondaria di II grado non può non tenere conto dell'esigenza di prossimità al luogo di residenza dello studente. Per questa ragione, è del tutto irrilevante per uno studente sapere dov'è localizzata la scuola dell'indirizzo prescelto che offre in assoluto le migliori basi per gli studi universitari, se quella scuola non è per lui raggiungibile. È molto più utile, invece, avere a disposizione una comparazione delle scuole che operano in prossimità della propria abitazione o a una distanza relativamente ridotta e percorribile su base quotidiana.

Ma anche confrontare scuole che offrono indirizzi di studio differenti non fornirebbe informazioni del tutto rilevanti per una scelta consapevole: ogni indirizzo di studio presenta una proposta formativa specifica che valorizza in modo differenziato le abilità e le conoscenze pregresse degli studenti. Per avere un percorso di studio proficuo e senza pericoli di insuccesso lo studente dovrebbe optare per indirizzi formativi che valorizzino i suoi punti di forza e i suoi interessi¹⁵. Per questo serve un buon processo di orientamento alle scuole medie. Qualsiasi comparazione della qualità formativa delle scuole, come quella proposta da **eduscopio.it**, fornisce un ulteriore elemento di riflessione per la scelta, ma non può sostituirsi al processo di orientamento, che invece è presupposto.

Dunque, **eduscopio.it** offre la possibilità di mettere a confronto le scuole che propongono il medesimo indirizzo di studi (licei classici con licei classici, istituti tecnici ad indirizzo economico con istituti tecnici ad indirizzo economico, e così via) in un raggio variabile tra i 10 km e i 30 km di distanza dal luogo di residenza¹⁶, a scelta dell'utente.

Definire una cella di confronto secondo i principi di omogeneità dell'offerta formativa (stesso indirizzo di studi) e prossimità territoriale fornisce indicazioni utili anche a scuole e docenti, poiché consente comparazioni eque.

¹⁵ In Bernardi-Bratti-De Simone (2014) si dà conto dell'efficacia di un servizio di orientamento offerto dal Comune di Torino e basato su test standardizzati, colloqui con orientatori e informazioni dettagliate sulle potenzialità scolastiche degli studenti e sui percorsi di studio che meglio le valorizzano.

¹⁶ La diversa gradazione della distanza si rende opportuna per rendere più flessibile la ricerca, anche in considerazione della variazione nella densità abitativa e nell'offerta scolastica dei diversi territori. In un'area metropolitana esistono decine di alternative in un raggio di soli 10km. Viceversa, anche in un raggio di 30 km potrebbero non esserci alternative confrontabili nel caso di comuni localizzati in aree particolarmente remote.

Le condizioni in termini di ricchezza economica e di capitale umano dei contesti educativi nei quali le scuole operano possono influire sulla capacità di promuovere apprendimenti di qualità. Esistono vincoli e opportunità strettamente legati al territorio di riferimento e la qualità delle scuole si manifesta anche nella loro capacità di attenuare gli effetti dei primi e di cogliere le seconde. Quello che le scuole non possono fare – almeno, non da sole e in tempi brevi - è modificare il contesto di riferimento. Dunque, una valutazione equa e utile alla riflessione sul proprio operato è quella che mette a confronto scuole operanti nel medesimo ambito territoriale.

Ma ciò non è sufficiente. Come già discusso in precedenza, nel sistema educativo italiano ci può essere molta differenza in partenza tra gli studenti che scelgono diversi indirizzi di studio, sia in termini di risultati scolastici pregressi che di condizioni sociali, culturali ed economiche della famiglia¹⁷. Questi sono fattori che influenzano gli esiti di apprendimento indipendentemente dall'operato delle scuole; per questo, affinché i confronti siano realmente informativi, bisogna comparare grandezze omogenee, cioè scuole con offerte formative analoghe e, dunque, con condizioni di partenza analoghe.

Per avere un ordine di grandezza dell'influenza che territorio di riferimento e indirizzo di studi hanno sulla capacità formativa delle scuole, si può fare riferimento nuovamente ai risultati riportati in Tabella 7. Le stime rivelano che, se considerati congiuntamente, territorio e indirizzo di studi possono spiegare fino al 28% della variabilità nei risultati delle scuole in termini di rendimenti universitari e fino al 39% della variabilità dei risultati delle scuole in termini di media dei voti¹⁸. Dunque, consentire confronti solo a parità di indirizzo e territorio di riferimento rende le comparazioni più eque e attenua il rischio che le differenze di efficacia non siano riconducibili a differenze qualitative tra scuole.

Si potrebbe obiettare che pur a parità di indirizzo di studi e area territoriale di riferimento, le scuole potrebbero attrarre in partenza tipologie diverse di studenti con evidente vantaggio per quelle che attraggono studenti con abilità maggiori o retroterra socio-culturali più favorevoli. In definitiva, esisterebbe un grado di eterogeneità nelle

¹⁷ Si vedano tra gli altri Checchi (2010), Checchi e Flabbi (2007) e Mocetti (2011).

¹⁸ Se anziché considerare unità territoriali quali le regioni considerassimo aree più circoscritte, la capacità esplicativa del modello crescerebbe ulteriormente poiché sarebbe in grado di catturare anche le specificità territoriali intra-regionali. Pertanto, nel consentire confronti entro raggi limitati (10-20-30 km), eduscopio.it favorisce ulteriormente l'equità dei confronti, limitandoli entro aree con condizioni di contesto più omogenee.

condizioni di partenza anche tra due licei scientifici della stessa città o due tecnici ad indirizzo industriale dello stesso comprensorio, e non tener conto di queste differenze non consentirebbe di individuare la reale efficacia delle scuole e potrebbe invalidare il confronto. Questa possibilità potrebbe avere una certa rilevanza soprattutto per le scuole localizzate nei grandi centri urbani dove esistono numerose alternative tra le quali gli studenti possono scegliere, mentre, dove l'offerta è minore, l'autoselezione è probabile che avvenga in massima parte lungo le direttrici degli indirizzi di studio.

Purtroppo, non avendo a disposizione informazioni sulle abilità scolastiche pregresse degli studenti al momento dell'iscrizione alle scuole superiori né sul loro retroterra socio-culturale, non possiamo verificare la fondatezza di questa obiezione nel nostro caso. Tuttavia, per farci una prima idea dell'incidenza di questi canali di autoselezione tra scuole dello stesso tipo abbiamo utilizzato i dati della rilevazione nazionale dell'INVALSI sugli apprendimenti in italiano e matematica al grado 10 (seconda classe di scuola secondaria di II grado) per l'anno scolastico 2010/11¹⁹. Scomponendo la varianza degli apprendimenti in due livelli: studente e scuole (modello vuoto), abbiamo appurato che, sia in italiano che in matematica, oltre il 40% della variabilità dei risultati di apprendimento è dovuta a differenze tra scuole (*between-school differences*); la rimanente parte è invece dovuta a fattori individuali a livello studente (*between-students within-schools differences*). Quando nella parte fissa del modello vengono considerati l'indirizzo di studi e la localizzazione geografica (provincia) della scuola, osserviamo che la varianza nei risultati tra scuole si riduce considerevolmente, passando dal 42,6% al 27,2% in italiano e dal 40,5% al 32% in matematica. Dunque, si riconferma quanto da noi trovato relativamente all'influenza di territorio e indirizzo di studio sulle performance delle scuole. Sfruttando l'informazione disponibile sullo status degli studenti in termini di estrazione socio-culturale (Economic and socio-cultural status, Escs²⁰) osserviamo che, una volta tenuto conto anche di questo fattore, la varianza tra scuole non si riduce ulteriormente in modo significativo (-0,6% e -0,2% per italiano e matematica, rispettivamente)²¹. Dunque l'indirizzo di studi appare come il canale più rilevante per l'autoselezione degli studenti nelle

¹⁹ Tra le rilevazioni disponibili, si tratta dell'unica che investa almeno una delle tre annualità da noi prese in considerazione per eduscopio.it.

²⁰ Per la definizione, la costruzione e le caratteristiche statistiche dell'indice Escs si rimanda a Campodifiori et al. (2010).

²¹ L'indisponibilità di dati longitudinali non ci permette di controllare anche per il livello di abilità in ingresso degli studenti oltre che per l'Escs. De Simone (2013) rileva che in uscita della scuola secondaria di I grado, vi è una significativa correlazione tra i livelli di apprendimento degli studenti e la loro estrazione socio-culturale.

scuole²² e questa prima evidenza indiziaria suggerisce che le differenze residue nella composizione del corpo studentesco delle scuole potrebbero non essere tali da influenzare significativamente le performance delle scuole.

Riferimenti bibliografici

Aina, C., Cappellari, L., e M. Francesconi, 2016. *Is Well Begun Half Done? Evidence from Educational Production Functions of University Students*, mimeo.

Benadusi, L., Fornari, R., e O. Giancola, 2010. Così vicine, così lontane. La questione dell'equità scolastica nelle regioni italiane, *Scuola Democratica*, 1, 57-79.

Bernardi, M., Bratti, M. e G. De Simone, 2014. *"I wish I knew ..." - Misperceived Ability, School Track Counseling Services and Performances in Upper Secondary Education*, IZA Discussion Papers 7940, Institute for the Study of Labor (IZA).

Bratti, M., Checchi, D. e A. Filippin, 2011. *Valore di segnalazione del voto di diploma e grading standard nelle scuole secondarie superiori*, Rapporto per la Fondazione Cariplo.

Checchi, D., 2010. Percorsi scolastici e origini sociali nella scuola italiana. *Politica Economica*, 3, 359–388.

Checchi, D., e L. Flabbi, 2007. *Intergenerational mobility and schooling decisions in Germany and Italy: The impact of secondary school tracks*. IZA Discussion Paper 2876, Institute for the Study of Labor (IZA).

Campodifiori, E., Figura, E., Papini, M. e R. Ricci, 2010. Un indicatore di status socio-economico-culturale degli allievi della quinta primaria in Italia, INVALSI Working Paper n. 2.

De Simone, G., 2013. "Render unto primary the things which are primary's: Inherited and fresh learning divides in Italian lower secondary education", *Economics of Education Review*, Elsevier, vol. 35(C), 12-23.

²² Risultati analoghi, ma ottenuti su dati OCSE-PISA (2009), si trovano in Benadusi-Fornari-Giancola (2010).

De Simone, G., Monastero, B. e A. Stanchi, 2009. *Un esercizio di valutazione esterna: come le Università piemontesi giudicano (indirettamente) le scuole della Regione*, FGA Working Paper n. 18, Fondazione Giovanni Agnelli (Torino).

Mocetti, S., 2011. Educational choices and the selection process: Before and after compulsory schooling. *Education Economics*, 20(2), 189–209.

APPENDICE A: *Cream skimming* e Indice FGA: una prima analisi esplorativa

Valutare le scuole in base alla performance universitaria dei loro diplomati vuol dire prendere in considerazione solo un sottoinsieme degli studenti che le anno frequentate. Infatti, nel corso del quinquennio di durata dell'istruzione secondaria di secondo grado, alcuni studenti potrebbero aver abbandonato gli studi o potrebbero aver cambiato scuola a seguito di vicende personali (trasferimenti di residenza) o di percorsi di studio accidentati (bocciature multiple). Di fatto, si può ipotizzare che se una scuola adottasse una politica di valutazione dei propri studenti molto rigorosa e selettiva, potrebbe operare nel corso degli anni una vera e propria scrematura – nella letteratura sulla valutazione delle scuole il fenomeno è noto con l'espressione di *cream skimming* - portando al diploma solo gli studenti più brillanti e lasciando che quelli con maggiori difficoltà si trasferiscano in altre scuole o abbandonino gli studi. Se così fosse, alcune scuole particolarmente selettive potrebbero godere di un indebito vantaggio nella valutazione degli esiti universitari a discapito di altre più "inclusive". In tal caso, **eduscopio.it** sovrastimerebbe la qualità delle prime e sottostimerebbe quella delle seconde.

La questione ha una sua legittimità e per questa ragione abbiamo cominciato ad esplorare la sua fondatezza empirica.

Con l'aiuto dell'Ufficio Statistico del MIUR abbiamo ripercorso all'indietro per cinque anni l'Anagrafe Nazionale degli Studenti (ANS) per individuare i diplomati (e le loro classi di appartenenza) al momento dell'iscrizione alle scuole secondarie di secondo grado. In definitiva, ricostruendo per ogni diplomato la classe di appartenenza al primo anno, possiamo capire quanti tra gli studenti iscritti al primo anno di scuola secondaria siano effettivamente arrivati fino in fondo, ottenendo un diploma. Questa operazione di lettura longitudinale dei dati dell'ANS si è rivelata molto complicata ha dato risultati incompleti per due ragioni:

- data la relativa giovane età dell'ANS, l'unica coorte tra quelle considerate in **eduscopio.it** per la quale esistano i dati per l'intero corso di studi secondario di secondo grado è quella dei diplomati del 2012/13, i quali (salvo ripetenze) erano iscritti alla prima classe della scuola di appartenenza nel 2008/09, primo anno di implementazione dell'ANS;
- al primo anno di implementazione dell'ANS non si è riusciti a raggiungere una copertura completa delle scuole e, poiché i dati sono raccolti e caricati

dalle segreterie scolastiche, in diversi casi la qualità delle informazioni era approssimativa.

Questi due limiti hanno fatto sì che potessimo costruire un parziale *indice di inclusività*²³ solo per il 67% delle scuole classificate nell'edizione 2016/17 di **eduscopio.it**. Pertanto, le analisi che seguono non possono tener conto di una scuola su tre tra quelle classificate e vanno interpretate con estrema cautela. Dunque, con i dati a disposizione non è possibile indagare approfonditamente il legame tra *cream skimming* e performance in **eduscopio.it**, ma possiamo avviare un'analisi esplorativa che può cominciare ad offrire qualche risposta in merito e che si consoliderà nelle prossime edizioni²⁴.

Il nostro *indice di inclusività* è costruito come il rapporto tra gli studenti diplomati nell'anno di riferimento 2012/13 e il numero degli studenti iscritti al primo anno dello stesso corso nella stessa scuola cinque anni prima, nel 2008/09. Ovviamente, questo indice ha un limite inferiore teorico pari a zero (è il caso estremo che tutti gli studenti iscritti nel 2008/09 siano stati bocciati nel corso dei successivi cinque anni) e un limite superiore che può essere anche maggiore di uno. Infatti, al diploma nel 2012/13 possiamo arrivarci non solo gli studenti iscritti in quella scuola sin dal primo anno, ma anche quelli che sono arrivati in quella scuola nel corso del quinquennio, ma che avevano cominciato gli studi altrove²⁵.

Pertanto, quando l'*indice di inclusività* ha un valore basso siamo in presenza di una scuola molto selettiva, quando l'indice ha un valore alto siamo in presenza di una scuola molto inclusiva e che addirittura attrae/accoglie studenti in uscita da altre scuole.

È vero che le scuole poco inclusive (molto selettive) finiscono per avere un vantaggio indebito nelle valutazioni di **eduscopio.it**? Per rispondere a questa domanda abbiamo cercato di capire qual è la correlazione tra *indice di inclusività* e *indice FGA* a parità di indirizzo del corso di studi, di tipo di gestione (statale, paritaria non statale), di

²³ Per *inclusività* intendiamo la capacità di una scuola di condurre il maggior numero di studenti iscritti al primo anno al conseguimento del diploma, limitando il fallimento educativo (bocciature) e l'abbandono scolastico.

²⁴ Un'analisi multivariata sulle caratteristiche correlate alla probabilità di poter avere un *indice di inclusività* o meno rivela che le scuole per le quali non riusciamo ad avere il dato sono prevalentemente scuole non statali, di dimensioni relativamente piccole e localizzate in particolari regioni del Sud e insulari (Campania, Sicilia, Sardegna). Il dato non sorprende visto che la prima edizione dell'ANS ha avuto le defezioni più frequenti proprio tra le scuole con questo profilo.

²⁵ Nei nostri dati l'*indice di inclusività* oscilla tra i valori osservati (0,14 - 1,32), con una media pari a 0,66 e una standard deviation di 0,16.

area geografica (provincia). Se l'ipotesi fosse fondata dovremmo trovare un legame negativo tra *indice di inclusività* e *indice FGA*: al crescere dell'inclusività e dunque della probabilità che al diploma arrivino anche studenti meno brillanti le performance universitarie dovrebbero peggiorare.

Le stime che riportiamo nella Tabella 8, tuttavia, raccontano un'altra storia. Nella colonna (1) osserviamo i normali divari tra le scuole dei diversi indirizzi, con classici e scientifici che, anche in virtù della selezione favorevole degli studenti in uscita dalle scuole medie, manifestano performance migliori rispetto agli altri indirizzi e le scuole paritarie tendenzialmente meno performanti di quelle statali. Quando introduciamo l'*indice di inclusività* nel modello di stima (colonna 2) osserviamo che esso non presenta alcuna correlazione significativa con l'*indice FGA*.

Tuttavia, l'effetto nullo potrebbe celare effetti positivi e negativi specifici per particolari tipologie di scuole che nel modello in colonna (2) tendono a compensarsi. Per questa ragione in colonna (3) presentiamo un modello dove interagiamo l'indicatore sulla gestione della scuola (statale/paritaria) con l'*indice di inclusività*. Le stime rivelano che, in effetti, nelle scuole statali potrebbe esserci addirittura una relazione positiva tra inclusività e performance in **eduscopio.it**, mentre nelle scuole paritarie questa potrebbe invece essere negativa come ci si aspettava. Tuttavia, i due effetti sono trascurabili. Ad esempio, affinché tra due scuole paritarie possa esistere una differenza dell'1% nell'indice FGA attribuibile alle diverse politiche di selettività, è necessario che gli indici di inclusività differiscano tra loro di ben 10 punti percentuali.

Abbiamo indagato anche la possibilità che il legame tra inclusività e performance in **eduscopio.it** possa differire per indirizzi di studio (colonna 4), ma non rileviamo *patterns* particolari.

In definitiva, la questione dell'influenza sulle performance osservate in **eduscopio.it** della diversa selettività praticata *in itinere* dalle scuole secondarie, per quanto teoricamente ammissibile, non ha al momento alcun sostanziale riscontro empirico. Tuttavia, data l'incompletezza dei dati, ci riserviamo di svolgere ulteriori approfondimenti nei prossimi anni volti ad avvalorare o smentire questa conclusione preliminare.

Tabella 8: Correlazione tra Indice FGA, indice di inclusività e altre caratteristiche delle scuole
(eduscopio.it – edizione 2016/17)

Variabili	Rendimenti			
	(1)	(2)	(3)	(4)
Indirizzi di diploma (rif. Liceo Classico)				
<i>Liceo Scientifico</i>	-0.618 (0.425)	-0.562 (0.430)	-0.510 (0.429)	-5.579** (2.599)
<i>Liceo delle Scienze Umane</i>	-12.28*** (0.590)	-12.18*** (0.602)	-12.07*** (0.601)	-16.36*** (3.425)
<i>Liceo Linguistico</i>	-5.735*** (0.523)	-5.669*** (0.528)	-5.606*** (0.527)	-7.994*** (3.088)
<i>Tecnico Economico</i>	-14.39*** (0.479)	-14.14*** (0.556)	-13.79*** (0.563)	-19.54*** (2.644)
<i>Tecnico Tecnologico</i>	-12.12*** (0.527)	-11.82*** (0.628)	-11.40*** (0.636)	-15.53*** (2.735)
Paritaria (rif. Statale)	-9.543*** (0.538)	-9.662*** (0.555)	-1.632 (2.215)	-9.630*** (0.556)
Indice di inclusività		0.0105 (0.0120)	0.0295** (0.0130)	-0.0408 (0.0294)
Indice di inclusività * Paritaria (rif. Statale)			-0.106*** (0.0283)	
Indice di inclusività * Indirizzi di diploma (rif. Liceo Classico)				
<i>Liceo Scientifico</i>				0.0661* (0.0339)
<i>Liceo delle Scienze Umane</i>				0.0545 (0.0469)
<i>Liceo Linguistico</i>				0.0282 (0.0412)
<i>Tecnico Economico</i>				0.0784** (0.0382)
<i>Tecnico Tecnologico</i>				0.0462 (0.0424)
Costante	67.51*** (0.355)	66.72*** (0.975)	65.34*** (1.040)	70.67*** (2.292)
N. Scuole	2,935	2,935	2,935	2,935
R ²	0.497	0.497	0.499	0.498

Note: Standard errors robusti in parentesi. Dummies provinciali incluse nel modello e non riportate. L'indice di inclusività è stato riscalato moltiplicandolo per 100. Significatività: *** p<0.01, ** p<0.05, * p<0.1