

Mobilità intergenerazionale e decisioni scolastiche in Italia

Daniele Checchi
Università di Milano

Luca Flabbi
Georgetown University

questa versione: 25 gennaio 2006

Introduzione

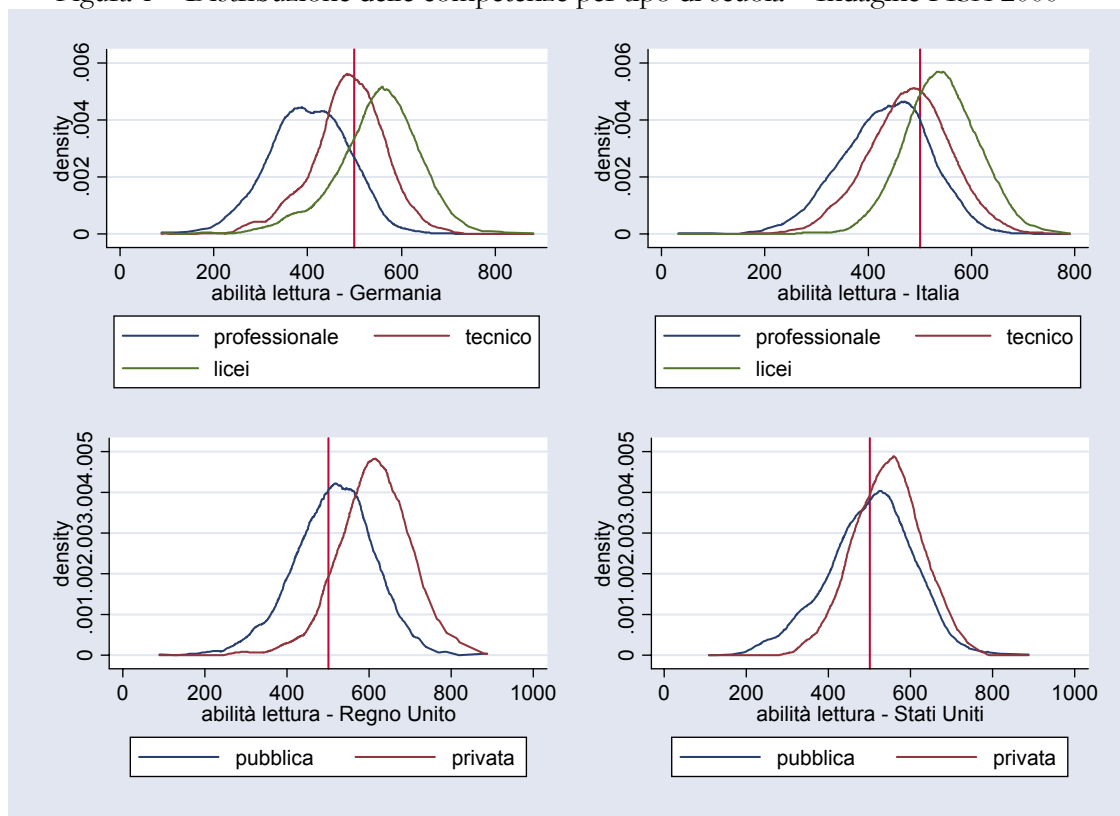
In un recente lavoro, Hanushek e Wößmann (2005) hanno analizzato in chiave comparativa gli effetti dell'esistenza di una stratificazione (*tracking*) a livello di scuola secondaria, mettendo in luce come i paesi caratterizzati da canalizzazione precoce degli studenti siano anche caratterizzati da minor performance e maggior dispersione a livello di competenze acquisite negli studenti quindicenni. Tuttavia la loro analisi considera solo gli ordinamenti istituzionali, ignorando il fatto che anche in sistemi scolastici formalmente unitari (*comprehensive*) le scelte delle famiglie tendono a riprodurre una stratificazione a posteriori che può risultare altrettanto efficace di quella prodotta per via istituzionale dalla stratificazione. Si osservi la figura 1: essa riporta la distribuzione delle competenze letterarie (*reading ability*) per quattro paesi e per tipologia di scuola, tratte dall'indagine PISA (*Programme for International Student Assessment*) condotta in ambito OCSE nel 2000 su campioni di studenti quindicenni.¹ In due di questi paesi, Italia e Germania, la scuola secondaria è organizzata secondo indirizzi distinti: vi è una filiera Licei, che abbiamo indicato come "licei" (corrispondenti ai licei classici, scientifici o linguistici nel caso italiano, e al *gymnasium* nel caso tedesco) che garantisce il pieno accesso all'istruzione universitaria; vi è una seconda filiera, che abbiamo indicato come "tecnico" (corrispondente agli istituti tecnici nel caso italiano e alle *Realschule* e *Gesamtschule* nel caso tedesco) che fornisce competenze direttamente spendibili nel mercato del lavoro pur garantendo l'accesso (pieno nel caso italiano e parziale nel caso tedesco) al livello di istruzione terziario. Vi è infine una terza filiera, indicata come "professionale" (corrispondente alle scuole di formazione professionale in Italia e alle *Hauptschule*, *Koop.Gesamtschule*, *Berufsschule* e *Berufsfachschule* nel caso tedesco), normalmente di durata inferiore alle due filiere precedenti (2-3 anni) e caratterizzate dalla impossibilità di accesso al livello successivo. La figura mostra con chiarezza che gli studenti iscritti nei *licei* hanno in media abilità linguistiche superiori² a quelle degli studenti delle scuole *tecniche*, che a loro volta sono in media superiori a quelle delle scuole *professionali*. Poiché questo livello di competenza è misurato all'inizio di ciascun curriculum (e risente quindi solo marginalmente della diversa formazione ricevuta in ciascun tipo di scuola), possiamo interpretare queste figure come il risultato di una allocazione degli studenti basata sulle capacità: i migliori (intesi come gli studenti con capacità linguistiche più elevate) vengono indirizzati nelle scuole che hanno una prospettiva di prosecuzione più elevata (di diritto, come in Germania, o di fatto,

¹ La densità di frequenza riportata in figura corrisponde alla stima con tecniche non parametriche del *kernel*.

² Si tratta cioè di dominanza in senso statistico.

come in Italia), agli studenti con capacità nella media viene lasciata la possibilità di uscire dal percorso scolastico al termine della secondaria, e agli studenti con minori abilità viene imposta (di nuovo di diritto, come in Germania, o di fatto, come in Italia) la fuoriuscita verso il mercato del lavoro. Confrontando le distribuzioni nel caso tedesco con quelle del caso italiano, notiamo però che la capacità di selezione (*sorting*) del sistema scolastico tedesco è più precisa di quello italiano. Mentre nel primo caso le tre distribuzioni sono sufficientemente distanti, nel secondo caso vi è un'ampia zona di sovrapposizione. Questo suggerisce che almeno nel caso italiano gli studenti che finiscono in una scuola professionale, caratterizzata da una bassissima probabilità di transizione all'università al suo termine, all'età di quindici anni dal punto di vista delle competenze linguistiche possedute non sono particolarmente diversi dagli studenti che si iscrivono ai licei, i quali con elevata probabilità si iscriveranno all'università.

Figura 1 – Distribuzione delle competenze per tipo di scuola – Indagine PISA 2000



Situazione molto diversa si ha nel mondo anglo-sassone, dove generalmente si incontra una sola tipologia di scuola secondaria. In questo contesto l'unica differenziazione istituzionale prevista è nella scelta di alcune materie che può in pratica generare profili scolastici abbastanza diversi. Tuttavia il riconoscimento di questi diversi profili scolastici non è un fattore così determinante nel processo di ammissione al College come la reputazione della scuola frequentata. La reputazione della scuola frequentata costituisce il principale canale di stratificazione a livello secondario ed è organizzato secondo due direttrici principali: la differenza tra scuole pubbliche e private e la differenza tra scuole pubbliche con diversa collocazione geografica. In questa situazione, gli studenti con capacità più elevate tendono ad

essere concentrati in scuole private oppure in scuole pubbliche della periferia urbana abbiente. Entrambe le tipologie di scuole sono in larga misura accessibili solo alle famiglie dotate di un reddito sufficiente a coprire le tasse di iscrizione o in grado di comprare o affittare abitazioni in zone residenziali ad alto reddito. La Figura 1 si concentra solo su uno di questi due meccanismi riportando differenziali tra scuole pubbliche e private.

Questi esempi ci suggeriscono che la stratificazione del sistema scolastico secondario avviene con modalità diverse (competenze degli studenti, sicuramente correlate con l'istruzione acquisita dai genitori, ma anche reddito o ricchezza delle famiglie, come nel caso delle scuole private) e anche con gradi di efficacia diversi. Per questo motivo ci sembra interessante analizzare, dal punto di vista empirico, quali siano le determinanti della allocazione degli studenti alle diverse tipologie di scuola, in quanto esse hanno un impatto sulle opportunità scolastiche successive (a livello terziario). Il caso italiano è da questo punto di vista un caso molto interessante: seppur consigliati dagli insegnanti durante la fase dell'orientamento scolastico, i genitori sono formalmente liberi di iscrivere i propri figli presso qualunque tipo di scuola secondaria.³ Il fatto che i livelli di competenze non siano troppo dissimili tra diversi tipi di scuole suggerisce che la decisione sia poco influenzata dalle capacità dello studente, e possa quindi riflettere altri fattori a livello familiare.

La modalità che governa la scelta scolastica ha ovviamente delle implicazioni in termini di mobilità intergenerazionale. Se infatti l'allocazione degli studenti è basata esclusivamente sulle capacità possedute, che tendono a distribuirsi in modo casuale ad ogni passaggio di generazione, le possibilità di ascesa sociale (attraverso l'acquisizione di istruzione) per individui capaci ma provenienti dalle classi inferiori sono più elevate che in un sistema dove l'allocazione degli studenti alle diverse scuole dipende principalmente dalle caratteristiche dell'ambiente familiare. Nel resto di questo capitolo noi studiamo la scelta della scuola secondaria in due diversi campioni di giovani italiani. Nel primo utilizziamo l'indagine PISA condotta nel 2003 (dove quindi non conosciamo le scelte scolastiche a livello successivo, ma possiamo analizzare le aspirazioni alla prosecuzione), mentre nel secondo caso utilizziamo l'indagine svolta da ISTAT sui diplomati dopo tre anni dal termine della scuola secondaria (dove quindi sappiamo se i giovani abbiano proseguito o meno la carriera scolastica a livello universitario). Mentre il primo campione è rappresentativo della pressoché intera popolazione studentesca (in quanto i tassi di transizione dalla scuola media inferiore alla scuola media superiore eccedono il 90%), il secondo è distorto dalla condizione che uno studente abbia completato la scuola secondaria, e non tiene quindi conto del fenomeno dei *drop-out*. I due campioni sono quindi complementari nel fornirci un quadro sufficientemente completo di come opera la stratificazione a livello di scuola media superiore in Italia.

Background familiare e scelta della scuola secondaria

Iniziamo con l'analizzare la scelta della scuola secondaria come emerge dai dati dell'indagine PISA. Per analizzare dal punto di vista statistico la scelta del tipo di scuola secondaria abbiamo bisogno di tre gruppi di variabili. Per prima cosa è necessaria una variabile dipendente che contenga informazioni riguardo il tipo di Scuola Secondaria frequentata, invece che indicare semplicemente il massimo titolo di studio conseguito (come

³ A differenza che nel caso tedesco, dove in diversi *Länder* l'orientamento degli insegnanti è vincolante e i genitori sono costretti ad attenersi al tipo di scuola secondaria indicato. Si veda Schnepf 2002.

tipicamente viene riportato nella maggioranza delle banche dati). L'indagine PISA contiene questo tipo di informazione anche se, per quanto riguarda l'Italia, non ad un livello di disaggregazione molto dettagliato. È comunque possibile individuare tre tipi di scuole: le scuole *Professionali*, le scuole *Tecniche* (gli Istituti Tecnici ed il liceo artistico) ed i *Licei* (Licei Classici, Scientifici e Linguistici).

Il secondo gruppo di variabili di cui abbiamo bisogno è relativo al background familiare. A questo proposito vorremo avere dettagliate informazioni riguardo salari e retribuzioni dei genitori e più in generale riguardo la posizione socio-economica della famiglia in cui lo studente è cresciuto. Nonostante PISA contenga molte informazioni riguardo alcuni aspetti del rapporto tra i membri della famiglia, sono del tutto assenti variabili relative a reddito e ricchezza familiare. Siamo pertanto costretti a focalizzarci semplicemente sul livello di istruzione dei genitori e a tenere conto di altri fattori socio-economici tramite un indice dello *status socio-economico*, calcolato dagli analisti di PISA in base a diverse domande relative non solo alla posizione occupazionale ma anche al possesso di beni durevoli. Abbiamo anche tenuto conto, nella specificazione estesa, di ulteriori caratteristiche della situazione familiare (quali la presenza in casa di un familiare non occupato, la presenza di famiglie regolari, e persino il numero di libri posseduti a casa). Focalizzarci sul livello di istruzione dei genitori presenta tuttavia alcuni vantaggi in quanto l'istruzione non è soggetta alle fluttuazioni che influenzano il reddito mensile o settimanale usualmente rilevato da indagini campionarie.

Il terzo gruppo di variabili di interesse, in aggiunta ai tradizionali controlli demografici (età e genere), riguarda il possesso di abilità di tipo cognitivo che rendono il proseguimento degli studi meno costoso sul piano dello sforzo, più efficiente e con maggiori probabilità di successo. Questo tipo di controlli per l'abilità degli studenti è essenziale per ridurre l'eterogeneità nella stima dell'impatto dei background familiari sulle scelte scolastiche. Da questo punto di vista, l'indagine PISA è la banca dati attualmente migliore per quanto riguarda l'Italia, in quanto fornisce i risultati di quattro tipi di test cognitivi relativi a competenze matematiche, linguistiche, scientifiche e di problem solving.⁴ Fornisce altresì informazioni riguardo la carriera scolastica in termini di bocciature o anni completati in anticipo. I risultati dei test sulle abilità matematiche sono utilizzati nella specificazione di base come misura principale dell'abilità, mentre tutti e quattro i test sono utilizzati nella specificazione estesa.⁵

Il modello econometrico che abbiamo scelto di utilizzare è una standard *Random Utility Model (RUM)*, ovvero un modello di scelta ottimale tra alternative numerabili e finite. Il modello assume che le alternative, nel nostro caso scuole professionali, scuole tecniche, e licei, siano associate con tre livelli di utilità potenzialmente diversi per ogni individuo e per ogni alternativa. L'utilità è effettivamente un'utilità attesa, modellata da una componente osservabile (deterministica) e una non osservabile (aleatoria). Denotando con $k \in (1,2,3)$ le

⁴ Questo è il principale vantaggio dell'Indagine PISA rispetto all'Indagine Bankitalia (*Indagine sui Bilanci delle Famiglie Italiane*) usualmente utilizzata per stime microeconomiche sul caso italiano. Per un esempio applicato alle scelte scolastiche a livello secondario con metodologia simile al presente lavoro, si veda Flabbi 2001.

⁵ I risultati dei test sono ovviamente molto correlati tra di loro, come appare evidente dalla matrice di correlazione seguente

	cap.matematiche	cap.linguistiche	con.scientifiche	prob.solving
cap.matematiche	1.0000			
cap.linguistiche	0.7325	1.0000		
con.scientifiche	0.8126	0.7916	1.0000	
problem solving	0.8711	0.7882	0.7897	1.0000

possibili alternative compiute e con $i = 1, 2, \dots, N$ i diversi individui, possiamo rappresentare questa utilità attesa come:

$$U_{k,i} = x_i' \beta_k + \varepsilon_{k,i} \quad (1)$$

dove $x_i' \beta_k$ denota la componente deterministica e $\varepsilon_{k,i}$ la componente aleatoria. La strategia ottimale è derivata dall'usuale massimizzazione dell'utilità e in questo caso si riduce a scegliere l'alternativa a cui è associata utilità attesa più elevata. Perciò, ad esempio, la probabilità che l'individuo i scelga l'alternativa 1 dall'insieme $\{1, 2, 3\}$ risulterà uguale a:

$$P_i(1 | 1, 2, 3) = P(U_{1,i} > U_{2,i}, U_{1,i} > U_{3,i}) \quad (2)$$

ovvero alla probabilità che l'utilità attesa associata con la prima alternativa sia al contempo maggiore di quella associata con la seconda alternativa e con la terza alternativa. Se le componenti aleatorie $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ sono indipendentemente e identicamente distribuite secondo una distribuzione Gumbel (o *extreme value*) standard, McFadden (1973) mostra che questa probabilità si riduce a:⁶

$$P_i(1 | 1, 2, 3) = \frac{\exp(x_i' \beta_1)}{\sum_{k=1}^3 \exp(x_i' \beta_k)} \quad (3)$$

ovvero il *Modello Multinomiale Logistico (MNL)*.

Questa breve rassegna della specificazione empirica è utile per mettere in luce i tre punti seguenti. Primo, l'identificazione richiede che l'utilità di una delle tre alternative sia normalizzata a zero. Questa è una immediata conseguenza della parametrizzazione (3). Secondo, per quanto riguarda l'inferenza è più interessante guardare all'impatto di un dato regressore sulla probabilità di scegliere una data alternativa, i cosiddetti *effetti marginali*, che alla stima dei coefficienti. Chiameremo effetto marginale anche il cambio nella probabilità di scelta dovuto al salto discreto di una variabile dummy da 0 a 1. Infine, il modello e la sua assunzione parametrica implicano *Indipendenza dalle Alternative Irrilevanti (IIA)*, ovvero che il rapporto tra due probabilità sia indipendente dall'insieme di alternative a cui è condizionata la scelta, dato che entrambe le alternative originali sono presenti in tutti i sottoinsiemi considerati. Questa proprietà genera un test di specificazione per il modello *MNL*, sviluppato in Hausman e McFadden 1984.⁷

I risultati riguardo gli effetti marginali dell'istruzione dei genitori sul tipo di Scuola Secondaria scelta sono presentati nella Tabella 1. Abbiamo scelto di dividere il campione in base al genere perché per quanto riguarda la mobilità intergenerazionale e le scelte scolastiche è possibile vi siano marcate differenze di comportamento tra uomini e donne.⁸ L'effetto marginale su cui concentriamo l'attenzione è l'impatto di avere almeno uno dei due genitori

⁶ In particolare, McFadden 1973 mostra che il modello MNL può essere generato da un modello di scelta ottimale come quello appena descritto solo se si assume una distribuzione Gumbel standard per le componenti aleatorie. In lavori seguenti mostra che alcune generalizzazioni sono possibili.

⁷ L'intuizione del test è la seguente: se la proprietà *IIA* è valida, allora le stime del modello includendo o escludendo le "alternative irrilevanti" non dovrebbero dare risultati molto diversi riguardo i parametri delle alternative rimanenti. La statistica è costruita come un semplice test di specificazione di Hausman tra un modello non ristretto (*unrestricted*) che include le alternative irrilevanti e un modello ristretto (*restricted*) che le esclude. McFadden 1987 mostra come sia possibili attuare il medesimo test usando un test dei Moltiplicatori di Lagrange.

⁸ Ad esempio questo è quanto trovato da Dustmann 2004 sulle scelte scolastiche in Germania.

in possesso di una laurea. Dopo aver stimato il modello MNL precedentemente presentato, abbiamo calcolato l'effetto marginale per ogni individuo nel campione come:

$$\Delta \hat{P}_i(k | \{1,2,3\}, x_i) = \hat{P}_i(k | \{1,2,3\}, d_i = 1, x_i) - \hat{P}_i(k | \{1,2,3\}, d_i = 0, x_i) \quad (4)$$

dove $(d_i = 1, x_i)$ significa che la dummy che segnala almeno uno dei due genitori in possesso di una laurea è posta uguale a uno mentre tutti gli altri regressori sono mantenuti al valore originale per l'individuo i . Una volta ottenuti gli effetti marginali, non è del tutto ovvio quale misura sintetica utilizzare per descrivere un valore medio nel campione. Una misura tradizionalmente utilizzata è il valore dell'effetto marginale alla media di tutti i regressori, ovvero:

$$\Delta \hat{P}_i \left(k | \{1,2,3\}, \frac{1}{N} \sum_{i=1}^N x_i \right) \quad (5)$$

Tuttavia questa misura è di non facile interpretazione in quanto un individuo con tutti i valori dei regressori pari alla media del campione non è necessariamente un buon candidato per rappresentare una media di comportamento, in particolare quando la maggior parte dei regressori siano variabili dummy come nel nostro caso. Abbiamo scelto perciò di presentare anche la media campionaria dell'effetto marginale e la sua deviazione standard. Si noti che la differenza tra l'effetto alla media dei regressori (5) e la media campionaria dell'effetto marginale

$$\frac{1}{N} \sum_{i=1}^N \Delta \hat{P}_i(k | \{1,2,3\}, x_i) \quad (6)$$

dipende dalla non-linearità della funzione $\Delta \hat{P}_i$. L'indicatore (5) ha il vantaggio di offrire un modo ovvio per ottenere gli errori standard e quindi per valutare la precisione delle stime mentre l'indicatore (6) è una misura più vicina alla nozione di "effetto marginale medio" nel campione.

La terza riga della Tabella 1 riporta la statistica (6), la quarta riga riporta la deviazione standard campionaria e la quinta riga riporta la statistica (5) con sottostante il rispettivo errore standard in parentesi. Per il campione maschile nella specificazione di base stimiamo un fortissimo effetto dell'istruzione dei genitori sulla scelta della scuola secondaria: ad esempio avere almeno un genitore laureato⁹ più che raddoppia la probabilità di scegliere la filiera liceale, che passerebbe da poco più del 20% a quasi il 60%. Questo impatto è dello stesso ordine di grandezza usando entrambi gli indicatori. La stima è anche sufficientemente precisa con uno standard error pari a circa un decimo della stima puntuale. Questo forte effetto positivo nella scelta del ramo accademico diventa un forte effetto negativo, anche se proporzionalmente minore, nella scelta di *Scuole Tecniche* o *Professionali*. Per quanto riguarda la *Scuola Tecnica* l'impatto negativo è riscontrabile per la gran parte del campione, anche se una minoranza (pari a circa il 14%) sperimenta un impatto positivo dell'istruzione dei genitori nello scegliere questo corso di studi.¹⁰ La stima puntuale degli effetti marginali ha un'alta varianza nel campione per quanto riguarda *Scuole Tecniche* e *Licei* mentre è maggiormente

⁹ La dummy per l'istruzione dei genitori è costruita in base al massimo titolo di studio conseguito tra i due genitori. Questo valore è poi codificato in tre variabili unitarie a seconda se il massimo titolo conseguito sia la Scuola Primaria (livello ISCED 1 e 2), la Scuola Secondaria (livello ISCED 3 e 4), oppure il Livello Terziario (livello ISCED 5 e 6) che nel testo, semplificando, chiamiamo laurea.

¹⁰ Per costruzione, le altre due alternative sperimentano un impatto negativo nella scelta *Scuola Professionale*, o positivo nella scelta *Licei*, per la totalità del campione.

concentrata intorno alla media per quanto riguarda l'effetto negativo nel scegliere una *Scuola Professionale*.

Questi risultati sono ottenuti in una specificazione di base che contiene solo controlli minimi: età in mesi come controllo demografico, risultato del test cognitivo per abilità matematiche per quanto riguarda il controllo di abilità, e le dummy per il massimo livello di istruzione conseguito dai genitori per quanto riguarda i controlli del background familiare. Si tratta di una specificazione molto sintetica che ha essenzialmente valore comparativo. La specificazione estesa risulta dal compromesso tra l'utilizzare tutta l'informazione disponibile nella banca dati e la necessità di non perdere troppe osservazioni a causa di valori mancanti (in modo da evitare una distorsione da selezione del campione). La specificazione estesa aggiunge dettagliati controlli per background familiare e abilità cognitiva¹¹. I risultati, riportati nella parte inferiore della Tabella 1, stimano un impatto inferiore dell'istruzione dei genitori sulla scelta del tipo di secondaria. L'effetto è tuttavia ancora molto significativo con un aumento della probabilità di scegliere la filiera liceale pari a circa il 50%. Le stime rimangono precise.¹²

Le stime sul campione femminile riportano un effetto leggermente inferiore sia in valore assoluto che relativo, visto che la proporzione di donne che sceglie il ramo accademico è superiore a quella degli uomini di circa venti punti percentuali (27.9% per gli uomini rispetto al 48.9% per le donne). È questo un dato interessante e relativamente recente: le donne scelgono in misura decisamente superiore il ramo accademico e l'istruzione dei genitori è meno importante in questa scelta. I risultati sono naturalmente troppo superficiali per trarre conclusioni su questo punto anche perché il differenziale praticamente scompare quando lo stime sono ottenute tramite la specificazione estesa. Tuttavia ci sembrano indicare un interessante direzione per un approfondimento.

Il risultato più importante di questa analisi è riscontrare che l'effetto dell'istruzione dei genitori sopravvive anche quando si cerchi di controllare per tutte le caratteristiche individuali dello studente, come fatto nella specificazione estesa. Due studenti identici in termini di competenze possedute, esperienza scolastica precedente, situazione familiare, numero di libri a casa ma diversi in quanto solo uno dei due ha almeno un genitore laureato, fronteggeranno probabilità di iscrizione scolastica significativamente diverse. Prendendo l'effetto medio calcolato secondo l'espressione (6) della tabella 1, vediamo che il figlio del laureato ha 25 punti percentuali di probabilità di finire in un liceo (rispetto al finire in una scuola tecnica o una professionale) in più di un figlio di un diplomato; la stessa percentuale sale al 27% per una studentessa figlia di un genitore laureato rispetto al finire in una scuola tecnica (il divario diventa addirittura 32 punti percentuali rispetto alla probabilità di finire in una scuola professionale). *Tutto ciò suggerisce che l'iscrizione alla scuola secondaria in Italia non riflette dei criteri esclusivamente meritocratici, ovvero basati sulla abilità individuale dello studente.* Se si osservano i coefficienti relativi alle abilità (vedi tabella A.2 in appendice) si nota come essi siano statisticamente significativi nell'aumentare la probabilità di scelta di un liceo (e in misura più forte nel caso dei maschi che di quello delle femmine). Nella specificazione estesa

¹¹ La specificazione completa è riportata in appendice, dove la tabella A.1 riporta le statistiche descrittive e la tabella A.2 riporta la stima del modello completo.

¹² Per quanto riguarda il test di specificazione *IIA* otteniamo per tutte le specificazioni in entrambi i campioni che il modello *MNL* non è rigettato a valori del P-value decisamente elevati. Per quanto riguarda il fit del modello, mostriamo semplicemente la distribuzione predetta alla media dei regressori. I valori sono riportati nella seconda riga e indicano un fit abbastanza preciso anche se tendiamo a sottostimare la proporzione nel ramo accademico nel campione maschile e a sovrastimarla nel campione femminile.

possiamo anche studiare l'impatto delle diverse capacità: da esse notiamo che le competenze matematiche sono le uniche ad influenzare in modo significativo la probabilità della scelta di una scuola tecnica, mentre sia queste che quelle linguistiche influenzano positivamente la probabilità di una scelta liceale. Le competenze scientifiche o della capacità di problem solving risultano non significative ma probabilmente si tratta di un problema di multicollinearità. Possiamo quindi riassumere questa sezione affermando che la stratificazione della scuola secondaria italiana opera attraverso un doppio regime: da un lato gli studenti più capaci, e dall'altro gli studenti che provengono dalle famiglie di genitori laureati, vengono indirizzati verso la filiera accademica, che punta all'iscrizione universitaria e al conseguimento di una laurea a loro volta. Se contasse solo l'abilità individuale, e se questa si formasse in modo indipendente dalla famiglia di provenienza, avremmo una situazione di perfetta mobilità intergenerazionale; se contasse invece soltanto l'istruzione dei genitori (solo i figli dei laureati accedono ai licei e quindi hanno elevata probabilità di laurearsi), avremmo una situazione di perfetta immobilità intergenerazionale nei livelli di istruzione.

Background Familiare e Livello di Istruzione Terziario

Dal momento che l'indagine PISA viene condotta su giovani quindicenni, si ignora quali siano le scelte scolastiche successive, ed in particolare se ed in che misura il tipo di secondaria scelta abbia effetto sulla probabilità di transizione all'università. Tuttavia la stessa indagine chiede agli studenti quali siano le loro intenzioni future, sia in termini occupazionali che in termini di carriera scolastica.¹³ Che gli studenti nutrano aspirazioni diverse a seconda della scuola frequentata è indubbio: l'80.55% degli studenti iscritti nei licei dichiara di volersi laureare, contro il 34.5% degli iscritti nelle scuole tecniche e il 15.9% degli iscritti nelle scuole professionali. Quando si osserva però il livello delle competenze possedute (vedi tabella 2) ci si accorge di come la scuola frequentata "modell" le aspirazioni: all'interno di coloro che intendono proseguire gli studi a livello universitario pur frequentando una scuola professionale vi sono sicuramente elementi brillanti (il punteggio massimo non si discosta eccessivamente da quello degli altri ordini di scuola), ma la maggioranza viene scoraggiata dall'ambiente circostante (il punteggio minimo di chi intende andare all'università è più elevato nelle scuole professionali che nei licei). Questo può ovviamente riflettere le caratteristiche dell'ambiente familiare, ed è per questo che è appropriato utilizzare l'analisi multivariata. La tabella 3 riporta le stime di un modello probit della probabilità di rispondere affermativamente alla intenzione di iscrizione all'università.¹⁴ Anche in questo caso usiamo una specificazione di base ed una estesa, distinta per generi. Il risultato principale della tabella è che la scuola di iscrizione ha un effetto molto più forte di quello di avere un genitore laureato (pari a 3 volte nella specificazione di base e 5 volte nella specificazione estesa, addirittura non essendo significativo l'effetto del genitore laureato nella componente femminile). Si noti che le competenze possedute tendono ad essere solo marginalmente significative, mentre mantiene un impatto forte anche la velocità con cui si sta percorrendo la

¹³ Domanda 8 del modulo del questionario *Educational career*: "Quale tipo di lavoro ti aspetti di svolgere quando avrai 30 anni?". Domanda 23 del questionario *Student questionnaire*: "Quale di questi livelli di istruzione ti aspetti di completare?".

¹⁴ In tabella 3 riportiamo solo i coefficienti di interesse, mentre in tabella A.3 sono riportate le stime complete.

carriera scolastica.¹⁵ Diventa altresì significativo il prestigio socio-economico della famiglia, oltre che la dotazione di risorse culturali, approssimate dal numero di libri in casa.

Combinando questi risultati con quelli della sezione precedente, si ha la fondata impressione che la scelta della scuola secondaria sia l'esito di un effetto congiunto di background familiare e abilità individuali, ma una volta entrati in una filiera l'effetto dominante diventa quello della filiera stessa, eventualmente rinforzato dall'ambiente familiare. In questo caso il destino individuale sarebbe fortemente predeterminato: o uno studente quattordicenne riesce ad emergere in modo così brillante che, anche a dispetto dell'istruzione dei propri genitori riesce ad iscriversi in un liceo, oppure da allora in avanti lotterà contro uno svantaggio che si accresce invece che attenuarsi.

Tuttavia stiamo sempre discutendo di aspirazioni, che nel caso di giovani quindicenni possono essere abbastanza volatili. Per questo motivo siamo passati ad analizzare un altro campione, che invece riporti le scelte effettivamente effettuate dai giovani. In assenza di un campione longitudinale che ci permetta di seguire una coorte rappresentativa di studenti, ci siamo riferiti all'indagine ISTAT sugli sbocchi lavorativi dei diplomati (nel 1998) intervistati a 3 anni dal diploma (nel 2001).¹⁶ Questo campione è ovviamente distorto, in quanto non riporta informazioni su coloro che non si iscritti alla scuola superiore o non sono riusciti a terminarla. Per avere un'idea del grado di distorsione, basti considerare che i diplomati dalle scuole secondarie superiori nell'anno scolastico 2001/2002 in rapporto alla corrispondente coorte di 19enni erano pari al 72.8% (come risultato del 68.5% per i maschi e del 77.4% per le femmine, con oscillazione tra il 62.5% dei maschi della Lombardia e l'88.7% delle femmine nel Lazio).¹⁷ Stiamo quindi lasciando fuori dall'analisi un terzo della popolazione rilevante, che alla luce dei risultati precedenti è caratterizzata da bassa scolarità dell'ambiente familiare e/ da basso rendimento scolastico, due fattori che plausibilmente influenzano la probabilità di completare la scuola secondaria. I risultati che troveremo in merito all'impatto dell'ambiente familiare sono quindi distorti verso il basso, e rappresentano quindi un limite inferiore del reale effetto nell'intera popolazione.

L'obiettivo in questa sezione è quello di studiare l'impatto del background familiare sulle scelte scolastiche immediatamente successive alla scuola secondaria. L'alternativa principale su cui focalizziamo la nostra attenzione è se gli individui scelgano di proseguire o meno gli studi a livello universitario. In particolare siamo interessati a stimare se l'impatto dell'istruzione dei genitori si riduca significativamente, o al limite scompaia completamente, quando teniamo conto del tipo di scuola secondaria frequentata. Questo risultato segnalerebbe la scelta del tipo di scuola secondaria come uno dei canali principali dell'impatto del background familiare sulle scelte scolastiche nel lungo periodo. In questa prospettiva la scelta della scuola secondaria, che viene compiuta a tredici anni rivelerebbe tutta la propria drammaticità, in quanto contribuirebbe a determinare l'intera vita scolastica successiva, e di conseguenza anche i destini lavorativi futuri.

Come abbiamo detto, l'Indagine ISTAT sugli sbocchi lavorativi dei diplomati raccoglie dati su un campione rappresentativo di diplomati a tre anni dal conseguimento del diploma. Ci concentreremo su individui intervistati nel 1998. Il campione è eccezionalmente esteso e il nostro campione di stima finale è costituito da un totale di 20929 individui. Alcune

¹⁵ Rispetto ai cosiddetti "primini" (cioè coloro che hanno iniziato la scuola a 5 anni), gli studenti in ritardo di almeno un anno hanno una probabilità del 26.4% inferiore di rispondere affermativamente alla domanda sulla intenzione di accedere all'università, tale svantaggio salendo al 34% nel caso delle studentesse (vedi tabella A.3).

¹⁶ Si veda la descrizione dei dati in ISTAT 2002.

¹⁷ Si veda la tavola 7.3 dell'Annuario Statistico Italiano 2004, curato da ISTAT.

statistiche descrittive sono riportate in Appendice, tabella A.4. Questo dataset ha il vantaggio di riportare informazioni relative al background familiare (istruzione ed occupazione dei genitori), oltre che permettere una ricostruzione sufficientemente dettagliata della carriera scolastica precedente (tipo di scuola secondaria frequentata, votazione all'uscita dalla scuola media e dalla scuola superiore, presenza di bocciature), caratteristiche che permettono di replicare quasi fedelmente le analisi prima effettuate sulla popolazione dei quindicenni.

Analogamente al caso precedente lavoreremo con un modello di scelta discreta consistente con la massimizzazione dell'utilità individuale. Le alternative considerate sono nuovamente tre: non iscriversi ad alcun corso di studi terziario (alternativa denotata con *Mai iscritto/a*), iscriversi ma ritirarsi (denotata con *Drop-out*) e infine iscriversi ed essere ancora iscritto a tempo pieno dopo tre anni dal conseguimento del diploma (denotata con *Iscritto/a*). In tabella 4 riportiamo la distribuzione del campione secondo le scelte scolastiche appena illustrate e per scuola secondaria di provenienza. Ci è sembrato interessante specificare esplicitamente l'alternativa *drop-out* perché l'abbandono scolastico a livello universitario è un problema notevole del sistema di istruzione superiore italiano.¹⁸ Ammettere questa alternativa cambia però in modo fondamentale l'albero decisionale in quanto le tre alternative non sono più sullo stesso piano temporale. In concreto, gli individui decidono se iscriversi a un corso di istruzione superiore una volta terminata la scuola secondaria e solo in un secondo tempo decidono se eventualmente ritirarsi (*drop-out*) o proseguire a tempo pieno. Valutiamo che l'iscrizione a tempo pieno a tre anni dal conseguimento del diploma sia una buona proxy per individui che otterranno un titolo di studio terziario. Tuttavia, se questo è l'albero decisionale di riferimento, il Modello Logit Multinomiale con errori indipendenti non è più appropriato. Il test *IIA* può ancora una volta essere utilizzato per verificare formalmente questa affermazione.

La parte inferiore delle Tabelle 5a (per il campione maschile) e 5b (per il campione femminile) riporta il *p-value* del test *IIA* rispetto alle tre alternative descritte. Ancora una volta utilizziamo due specificazioni, una specificazione di base con controlli minimi e una specificazione estesa che include tutta l'informazione disponibile nella banca dati (vedi più sotto). Per il campione maschile l'ipotesi *IIA* sottesa al modello *MNL* è rigettata nella specificazione estesa che contiene i controlli per il tipo di secondaria. Questo avviene quando viene eliminata una delle due alternative *Drop-out* e *Iscritto*. Non viene invece generalmente rigettata quando l'alternativa eliminata è *Mai Iscritto*. Questi risultati sono conformi alla scelta sequenziale prima descritta: le due alternative *Drop-out* e *Iscritto* sono sullo stesso piano temporale perciò quando l'alternativa precedente (*Mai Iscritto*) è eliminata non si ha un impatto significativo; quando invece una di queste due è eliminata le due alternative rimanenti non sono sullo stesso piano temporale e perciò le scelte non si ridistribuiranno indipendentemente. Per il campione femminile i risultati sono meno netti (il test è marginalmente non rigettato nella specificazione estesa) ma generalmente simili alla descrizione appena descritta per il campione maschile.¹⁹

Abbiamo perciò scelto di modellare esplicitamente questa scelta sequenziale utilizzando un *Modello Logit Sequenziale*. Si tratta della medesima struttura di un modello logit ma che tiene conto che la scelta di drop-out o continuare gli studi è condizionata al fatto di aver precedentemente scelto di iscriversi a un corso di livello di istruzione terziario. Formalmente, definendo i seguenti indicatori: $\chi_i = 1$ se l'individuo i sceglie di iscriversi a un corso di livello

¹⁸ Si veda l'analisi contenuta in Checchi 2003.

¹⁹ Da notare come nella Tabella 1 il test *IIA* sia non rigettato per livelli abbondantemente superiori a 0.5. Il test conferma quindi l'intuizione che il modello *MNL* non sia appropriato in questo contesto.

universitario dopo aver completato il livello secondario e $\lambda_{i1} = 1$ se l'individuo i è ancora iscritto a tempo pieno dopo tre anni dal conseguimento del diploma, il modello genera le seguenti probabilità:

$$\begin{aligned}
 P_i(\text{Non Iscritto}) &= P_i(\chi_i = 0) = \frac{1}{1 + \exp(x'_i\beta)} \\
 P_i(\text{Drop-out}) &= P_i(\lambda_i = 0 \mid \chi_i = 1)P_i(\chi_i = 1) = \frac{1}{1 + \exp(x'_i\alpha)} \frac{\exp(x'_i\beta)}{1 + \exp(x'_i\beta)} \\
 P_i(\text{Iscritto}) &= P_i(\lambda_i = 1 \mid \chi_i = 1)P_i(\chi_i = 1) = \frac{\exp(x'_i\alpha)}{1 + \exp(x'_i\alpha)} \frac{\exp(x'_i\beta)}{1 + \exp(x'_i\beta)}
 \end{aligned} \tag{7}$$

da cui è poi possibile stimare i coefficienti α e β col metodo della massima verosimiglianza. Nuovamente concentreremo l'attenzione sugli effetti marginali invece che sui parametri dell'utilità.

I risultati sono presentati nella parte superiore della Tabella 5a per quanto riguarda il campione maschile e nella parte superiore della Tabella 5b per quanto riguarda quello femminile. Gli indicatori riportati sono i medesimi della sezione precedente. Ad esempio, per quanto riguarda la specificazione Base, la terza riga riporta la media campionaria dell'effetto marginale, la quarta la deviazione standard, la quinta l'effetto marginale calcolato alla media dei regressori e infine la sesta riga riporta la proporzione di individui nel campione per cui l'effetto marginale è positivo. Il segno dell'effetto marginale è quello atteso per quanto riguarda *Iscritto* (effetto positivo di una più elevata istruzione dei genitori) e *Mai iscritto* (effetto negativo di una più elevata istruzione dei genitori) mentre i risultati sono misti ed imprecisi per quanto riguarda l'alternativa *Drop-out*. L'ordine di grandezza dell'impatto dei genitori è ancora estremamente elevato: nel campione maschile avere almeno un genitore laureato aumenta di circa il 50% la probabilità di essere stabilmente iscritto a un corso di studi di livello universitario. L'impatto è però marcatamente ridotto quando introduciamo nella specificazione i controlli per il tipo di scuola secondaria frequentata. Per le due alternative *Iscritto* e *Mai iscritto* la riduzione dell'effetto marginale è nell'ordine di più di un terzo (per esempio nella specificazione di base del sottocampione maschile si passa da un effetto di un genitore laureato pari al 51% in termini di probabilità ad un effetto pari al 36%). È questo un risultato importante per quanto riguarda i nostri obiettivi iniziali: conferma che la scelta del tipo di scuola secondaria è un canale fondamentale dell'impatto del background familiare sulle scelte scolastiche. Tuttavia, il significativo effetto marginale dell'istruzione dei genitori, che sopravvive anche nella specificazione estesa e anche controllando per il tipo di secondaria, segnala una diretta e persistente influenza del background familiare anche in scelte che avvengono in età più matura.

Come già trovato nella sezione precedente, l'impatto dei genitori è meno importante per quanto riguarda il campione femminile. Ci sembra che ottenere il medesimo risultato usando una banca dati differente e studiando una scelta diversi rafforzati, anche se in modo molto preliminare ed incompleto, la congettura precedentemente proposta che le donne scelgano in modo maggiormente efficiente degli uomini il proprio investimento in istruzione.

Le stime sono generalmente precise ad eccezione dell'effetto marginale per l'alternativa *Drop-out*. Nel campione femminile questa imprecisione determina anche un cambio di segno nella stima puntuale che passa da un valore negativo nella specificazione senza controlli per

la scuola secondaria ad un valore positivo quando i controlli sono inseriti nella regressione. Una possibile interpretazione per l'instabilità dei risultati su questo gruppo è che si tratti di un effetto di composizione dovuto alla maggiore eterogeneità degli individui che attuano questa scelta rispetto alle altre due.

Infine, per quanto riguarda una comparazione rispetto al modello econometrico utilizzato, notiamo che le stime con Modello Logit Multinomiale implicano valori leggermente superiori dell'impatto dei genitori sulle scelte considerate. Le conclusioni principali sono però identiche: avere almeno un genitore laureato implica un forte effetto positivo sulla probabilità di rimanere iscritto e un forte effetto negativo sulla probabilità di non iscriversi; l'effetto dei genitori è fortemente ridotto quando il tipo di scuola secondaria scelta è preso in considerazione; e infine i risultati riguardo la scelta drop-out sono imprecisi.

Spunti conclusivi

In questo capitolo abbiamo cercato di mettere in luce l'importanza del background familiare nella scelte scolastiche dei giovani italiani. Siamo partiti dall'illustrare come, in presenza di un sistema scolastico secondario ad organizzazione tripartita, possa rappresentare uno dei fattori esplicativi della elevata persistenza intergenerazionale che si osserva nelle scelte scolastiche della popolazione italiana.²⁰ Abbiamo poi illustrato come sia centrale il meccanismo di allocazione degli studenti ai diversi curricula, e come questo in Italia non avvenga soltanto sulla base delle capacità individuali, ma risenta anche del livello di istruzione dei genitori. Una volta che gli studenti siano allocati ad un indirizzo preciso, le capacità individuali diventano meno importanti, in quanto i diversi indirizzi sono molto efficaci nell'orientare le diverse aspirazioni, oltre che a fornire formazione più o meno generalista che costituisce un vantaggio nella prosecuzione degli studi. Al termine del processo formativo della scuola secondaria, gli studenti escono caratterizzati da probabilità di prosecuzione universitaria molto differenziate, ma ciò nonostante l'istruzione dei genitori continua a contare anche a questo livello. Facciamo fatica a credere che questo possa dipendere dal fatto che i genitori aiutino e sostengano lo studio dei ventiduenenni; verrebbe piuttosto da pensare che questo rifletta la stratificazione del mercato del lavoro: se i posti migliori vengono assegnati sulla base delle reti familiari, avere un genitore laureato può dare accesso a circuiti migliori, ed in quanto tale rappresenta un incentivo per i figli al conseguire titolo di accesso ai medesimi circuiti.

Se l'istruzione dei genitori conta persino nella transizione universitaria, e non si dissipa con il procedere della carriera,²¹ ne dobbiamo inferire che l'Italia è ancora molto lontana dall'offrire una uguaglianza nelle opportunità di accesso, così come anche si riscontra nel mercato del lavoro successivo (Checchi e Peragine 2005). Dal punto di vista delle politiche, ci domandiamo se una diversa organizzazione del sistema scolastico secondario nella linea di una progressiva scomparsa della differenziazione tra le scuole possa migliorare la situazione in senso egualitario, dal momento che le esperienze estere in questo senso danno indicazioni

²⁰ In Checchi, Ichino e Rustichini 1999 si è confrontato il caso italiano con quello statunitense, da cui emerge una maggior immobilità dell'Italia in confronto con altri paesi. Analisi comparative più recenti si trovano in Comi 2004, Chevalier et al. 2005 e Moccetti 2005. Flabbi 2001 conferma, limitatamente al caso italiano, la correlazione tra anni di scolarità dei genitori e scelta del tipo di scuola secondaria su dati tratti dall'Indagine sui Bilanci delle Famiglie Italiane (Banca d'Italia).

²¹ Conclusioni diverse in merito alla dissipazione dell'effetto del background nel corso della carriera in Schizzerotto 2002.

contrastanti.²² Dal nostro punto di vista, l'obiettivo deve essere la riduzione dell'importanza del background familiare nelle scelte scolastiche, e questo può essere perseguito o rafforzando il carattere meritocratico della scelta della scuola secondaria (seguendo il modello tedesco) o abolendo la differenziazione tra gli indirizzi (seguendo il modello inglese) o rendendo gli indirizzi più permeabili tra loro. La prima alternativa è discutibile per via della precoce età in cui avviene la scelta, e si espone al rischio di scelte errate in quanto basate su imperfetta informazione. La seconda scelta non fa altro che posticipare ad una età successiva il problema di selezionare gli individui sulla base delle loro capacità, richiedendo quindi l'introduzione di una qualche forma di stratificazione del sistema universitario, di diritto o di fatto. In questa seconda alternativa la scelta di un indirizzo verrebbe effettuata intorno ai 18-19 anni, quando quindi gli studenti sono in grado da un lato di aver appieno espresso le loro potenzialità, nonché dall'altro la scuola ha potuto esercitare la sua azione compensatrice degli eventuali divari all'ingresso.

Bibliografia

Cecchi, D. 2003. "The Italian educational system: family background and social stratification", ISAE Conference on Monitoring Italy.

Cecchi, D., A. Ichino e A. Rustichini. 1999. "More equal but less mobile? Intergenerational mobility and inequality in Italy and in the US". *Journal of Public Economics*, 74: 351-393.

Cecchi, D. e V. Peragine 2005. "Regional disparities and inequality of opportunity: the case of Italy" IZA Discussion Paper No. 1874/2005

Chevalier, A., K. Denny and D. McMahon. 2005. "A multicountry study of intergenerational education mobility". mimeo.

Comi, S. 2004. Intergenerational mobility in Europe: evidence from ECHP. CESifo conference on Schooling and Human Capital Formation in the Global Economy (Munich September 2004).

Dustmann, C. 2004. "Parental background, secondary school track choice, and wages", *Oxford Economic Papers*, 56: 209-230.

Flabbi, L. 2001 "Secondary Schooling Decisions in Italy", *Rivista di Politica Economica*, 91(7/8): 85-114.

²² L'esperienza di *destreaming* nel caso inglese è stata analizzata da Galindo-Rueda e Vignoles 2004, mentre quella svedese è discussa in Meghir e Palme 2004

Galindo-Rueda, F. e A.Vignoles. 2004. The Heterogeneous Effect of Selection in Secondary Schools: Understanding the changing role of ability. CESifo conference on Schooling and Human Capital Formation in the Global Economy (Munich September 2004).

Hanushek, E e L.Woessman. 2005. "Does educational tracking affect performance and inequality ? Differences-in-differences evidence across countries." NBER wp.11124, February.

Hausman, J. e D. McFadden. 1984. "A specification Test for the Multinomial Logit Model", *Econometrica*, 52: 1219-1240.

ISTAT 2002. *Diplomati e mercato del lavoro Percorsi di studio e di lavoro dei diplomati. Indagine 2001*. Roma

McFadden, D. 1987. "Regression Based Specification tests for the Multinomial Logit Model", *Journal of Econometrics*, 34: 63-82.

McFadden, D. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior", in P. Zarembka (ed.) *Frontiers in Econometrics*, New York: Academic Press.

Meghir, C. e M.Palme. 2004. Educational reform, ability and family background. IFS wp 04/10.

Mocetti, S. 2005. Intergenerational mobility in Italy. mimeo

Schizzerotto, A. 2002. *Vite ineguali – Diseguaglianze e corsi di vita nell'Italia contemporanea*. Bologna, Il Mulino

Schnepf, S.V. 2002. "A Sorting Hat that Fails? The Transition from Primary to Secondary School in Germany", Innocenti Working Paper No. 92, Florence: UNICEF Innocenti Research Center.

Tabella 1: Effetto marginale di almeno un genitore laureato sulla scelta del tipo di Scuola Secondaria – PISA 2003

	Campione Maschile			Campione Femminile		
	Professionale	Tecnica	Licei	Professionale	Tecnica	Licei
Specificazione Base:						
Distribuzione osservata	0.273	0.448	0.279	0.238	0.273	0.489
Distribuzione predetta	0.241	0.526	0.233	0.184	0.308	0.508
Effetti marginali:						
Media	-0.1307	-0.2252	0.3559	-0.1409	-0.2014	0.3422
Dev. standard	0.0576	0.1607	0.1214	0.0505	0.0464	0.0286
Alla media dei regressori	-0.1880 (0.0325)	-0.2258 (0.0428)	0.4138 (0.0416)	-0.1517 (0.0256)	-0.2070 (0.0324)	0.3587 (0.0369)
Proporzione con effetto positivo	0.000	0.144	1.000	0.0000	0.0003	1.0000
Test IIA: P-value	0.6616	0.6960	0.4451	0.9396	0.5252	0.8104
Logverosimiglianza		-4761.34			-5574.72	
Numero osservazioni		5517			5973	
Specificazione Estesa:						
Distribuzione osservata	0.263	0.453	0.284	0.236	0.272	0.492
Distribuzione predetta	0.226	0.580	0.195	0.178	0.306	0.516
Effetti marginali:						
Media	-0.0561	-0.0882	0.1443	-0.1167	-0.0549	0.1715
Dev. standard	0.0289	0.0904	0.0789	0.0519	0.0443	0.0363
Alla media dei regressori	-0.0863 (0.0284)	-0.0843 (0.0394)	0.1706 (0.0394)	-0.1243 (0.0235)	-0.0756 (0.0316)	0.1999 (0.0352)
Proporzione con effetto positivo	0.000	0.256	1.000	0.0000	0.1313	1.0000
Test IIA: P-value	0.5971	0.6386	0.8728	0.8617	0.7606	0.8203
Logverosimiglianza		-4152.97			-5189.93	
Numero osservazioni		5322			5895	

Note: in parentesi, errori standard robusti a cluster a livello di scuola. Viene riportato solo l'impatto di madre e/o padre laureati. I risultati completi e i dettagli delle due specificazioni sono in tabella A.2 in Appendice. Le stime sono da un modello logit multinomiale pesato per i pesi di stratificazioni forniti da PISA. Il test IIA (*independence of irrelevant alternatives*) è relativo all'eliminazione dell'alternativa nella colonna corrispondente.

Tabella 2: Competenze ed aspirazione di iscrizione all'università, per tipo di scuola secondaria – PISA 2003

competenze matematiche (media 500 dev.st.100)	scuole professionali	scuole tecniche	licei
intendono terminare l'università			
mediana	450.3	519.1	535.1
massimo	717.4	745.0	771.1
minimo	251.6	253.2	197.6
numero casi	472	1423	3604
non intendono completare l'università			
mediana	430.6	499.0	515.7
massimo	706.3	747.5	734.9
minimo	109.1	216.4	288.9
numero casi	2485	2703	870

Tabella 3: Effetto marginale sulla aspirazione di iscriversi all'università – PISA 2003

	campione maschile		campione femminile	
	specificazione base	specificazione estesa	specificazione base	specificazione estesa
genitore livello secondario	0.131 [0.035]**	0.073 [0.034]*	0.079 [0.020]**	0.02 [0.023]
genitore livello terziario	0.222 [0.032]**	0.116 [0.035]**	0.158 [0.021]**	0.052 [0.027]
madre titolo più alto	-0.062 [0.026]*	-0.045 [0.028]	-0.026 [0.020]	-0.002 [0.019]
capacità matematiche	0.001 [0.000]**	-0.001 [0.000]	0.001 [0.000]**	0 [0.000]
scuola tecnica	0.239 [0.038]**	0.22 [0.040]**	0.077 [0.035]*	0.077 [0.033]*
liceo	0.667 [0.024]**	0.628 [0.029]**	0.465 [0.035]**	0.435 [0.035]**
Numero osservazioni	5341	5168	5866	5759
Pseudo R ²	0.33	0.35	0.26	0.29
Log verosimiglianza	-2488.11	-2314.41	-2736.95	-2565.82

Note: la variabile dipendente riporta il desiderio degli studenti intervistati di proseguire gli studi fino all'università. In parentesi, errori standard robusti a cluster a livello di scuola. * indica significatività statistica al 5%; ** significatività al 1%. La versione estesa controlla per le diverse competenze e per la carriera pregressa, oltre che per prestigio occupazionale, condizione familiare e possesso di libri a casa. I risultati completi e i dettagli delle specificazioni sono in tabella A.3 in Appendice. Le stime sono da un modello probit pesato per i pesi di stratificazioni forniti da PISA.

Tabella 4: Scelte scolastiche postsecondarie per genere e scuola di provenienza ISTAT 2001

scuola di provenienza	campione maschile			campione femminile		
	mai iscritto	drop out	università	mai iscritto	drop out	università
licei	9.74%	6.15%	84.10%	11.13%	4.23%	84.64%
scuole tecniche	63.39%	6.92%	29.69%	57.61%	6.42%	35.97%
scuole professionali	83.43%	4.66%	11.92%	77.23%	4.90%	17.87%
Numero osservazioni	6959	615	3039	7018	667	4759

**Tabella 5a: Effetto marginale di almeno un genitore laureato
sulle scelte di istruzione universitaria – ISTAT 2001 – Campione maschile**

	Senza controlli per il tipo di Scuola Secondaria			Con controlli per il tipo di Scuola Secondaria		
	Mai Iscritto	Drop-out	Iscritto	Mai Iscritto	Drop-out	Iscritto
MODELLO LOGIT SEQUENZIALE						
Specificazione Base:						
Distribuzione osservata	0.6503	0.0580	0.2918	0.6503	0.0580	0.2918
Distribuzione predetta	0.6692	0.0655	0.2653	0.6595	0.0713	0.2692
Effetti marginali:						
Media	-0.4584	0.0119	0.4465	-0.3035	0.0202	0.2833
Dev. standard	0.0891	0.0261	0.0737	0.0900	0.0245	0.0789
Alla media dei regressori	-0.5187 (0.0135)	0.0097 (0.0116)	0.5090 (0.0190)	-0.3841 (0.0266)	0.0200 (0.0125)	0.3641 (0.0281)
Proporzione con effetto positivo	0.0000	0.6837	1.0000	0.0000	0.8027	1.0000
Logverosimiglianza		-6707.97			-6269.49	
Numero osservazioni		9956			9956	
Specificazione Estesa:						
Distribuzione osservata	0.6501	0.0579	0.2921	0.6501	0.0579	0.2921
Distribuzione predetta	0.6770	0.0661	0.2570	0.6693	0.0728	0.2579
Effetti marginali:						
Media	-0.4491	0.0069	0.4421	-0.3082	0.0137	0.2945
Dev. standard	0.0194	0.0124	0.0212	0.0238	0.0116	0.0274
Alla media dei regressori	-0.3700 (0.0857)	0.0075 (0.0255)	0.3625 (0.0806)	-0.2202 (0.0809)	0.0121 (0.0217)	0.2082 (0.0769)
Proporzione con effetto positivo	0.0000	0.6305	1.0000	0.0000	0.7255	1.0000
Logverosimiglianza		-6813.91			-6270.32	
Numero osservazioni		10453			10453	
MODELLO LOGIT MULTINOMIALE						
Specificazione Base:						
Effetti marginali alla media dei regressori	-0.5226 (0.0156)	0.0103 (0.0104)	0.5123 (0.0178)	-0.3860 (0.0228)	0.0214 (0.0131)	0.3646 (0.0239)
Test IIA: P-value	0.4175	0.4164	0.3614	0.1008	0.4633	0.0809
Specificazione Estesa:						
Effetti marginali alla media dei regressori	-0.4527 (0.0204)	0.0069 (0.0111)	0.4458 (0.0221)	-0.3092 (0.0273)	0.0156 (0.0136)	0.2937 (0.0276)
Test IIA: P-value	0.8554	0.5009	0.529	0.7656	0.0443	0.3669

Note: In parentesi, errori standard ottenuti con bootstrap. Viene riportato solo l'impatto di madre e/o padre laureati. I dettagli delle due specificazioni sono in tabella A.4a in Appendice.

Tabella 5b: Effetto marginale di almeno un genitore laureato sulle scelte di istruzione universitaria – ISTAT 2001 – Campione femminile

	Senza controlli per il tipo di Scuola Secondaria			Con controlli per il tipo di Scuola Secondaria		
	Mai Iscritto	Drop-out	Iscritto	Mai Iscritto	Drop-out	Iscritto
MODELLO LOGIT SEQUENZIALE						
Specificazione Base:						
Distribuzione osservata	0.5567	0.0525	0.3908	0.5567	0.0525	0.3908
Distribuzione predetta	0.5545	0.0591	0.3863	0.5370	0.0632	0.3997
Effetti marginali:						
Media	-0.4520	0.0007	0.4514	-0.3281	0.0176	0.3105
Dev. standard	0.1147	0.0206	0.0983	0.1138	0.0212	0.0976
Alla media dei regressori	-0.4921 (0.0108)	-0.0044 (0.0071)	0.4964 (0.0120)	-0.3868 (0.0205)	0.0112 (0.0107)	0.3755 (0.0217)
Proporzione con effetto positivo	0.0000	0.4921	1.0000	0.0000	0.7259	1.0000
Logverosimiglianza		-8351.27			-7782.38	
Numero osservazioni		11527			11527	
Specificazione Estesa:						
Distribuzione osservata	0.5568	0.0530	0.3903	0.5568	0.0530	0.3903
Distribuzione predetta	0.5587	0.0597	0.3816	0.5393	0.0643	0.3963
Effetti marginali:						
Media	-0.3703	-0.0051	0.3754	-0.2536	0.0069	0.2467
Dev. standard	0.0968	0.0203	0.0859	0.0936	0.0183	0.0841
Alla media dei regressori	-0.4293 (0.0176)	-0.0089 (0.0085)	0.4382 (0.0180)	-0.3255 (0.0254)	0.0022 (0.0118)	0.3233 (0.0264)
Proporzione con effetto positivo	0.0000	0.4112	1.0000	0.0000	0.6293	1.0000
Logverosimiglianza		-8731.71			-8094.68	
Numero osservazioni		12343			12343	
MODELLO LOGIT MULTINOMIALE						
Specificazione Base:						
Effetti marginali alla media dei regressori	-0.4955 (0.0117)	-0.0048 (0.0084)	0.5004 (0.0139)	-0.3901 (0.0182)	0.0105 (0.0115)	0.3796 (0.0203)
Test IIA: P-value	0.5964	0.0067	0.5206	0.0709	0.0639	0.0017
Specificazione Estesa:						
Effetti marginali alla media dei regressori	-0.4333 (0.0158)	-0.0075 (0.0087)	0.4408 (0.0174)	-0.3295 (0.0217)	0.0036 (0.0112)	0.3259 (0.0232)
Test IIA: P-value	0.2799	0.1404	0.3969	0.2998	0.2423	0.2021

Note: In parentesi, errori standard ottenuti con bootstrap. Viene riportato solo l'impatto di madre e/o padre laureati. I dettagli delle due specificazioni sono in tabella A.4b in Appendice.

Appendice

Tabella A.1: Statistiche descrittive del campione PISA 2003 - Italia

	campione maschile				campione femminile			
	Specificazione Base		Specificazione Estesa		Specificazione Base		Specificazione Estesa	
	Media	<i>Dev. Standard</i>	Media	<i>Dev. Standard</i>	Media	<i>Dev. Standard</i>	Media	<i>Dev. Standard</i>
età (in mesi)	15.71	0.28	15.71	0.29	15.72	0.29	15.72	0.29
straniero			0.03	0.18			0.03	0.17
parla dialetto in casa			0.24	0.43			0.19	0.40
capacità matematiche	508.15	93.10	510.58	91.90	487.02	83.35	487.65	83.15
capacità linguistiche			484.97	92.08			521.75	80.46
conoscenze scientifiche			522.61	99.16			514.30	89.84
cap.problem solving			501.61	93.69			500.73	83.88
percorso lento			0.19	0.39			0.11	0.31
percorso regolare			0.79	0.41			0.87	0.34
percorso accelerato			0.02	0.13			0.02	0.15
bocciato almeno 1 volta			0.14	0.35			0.07	0.25
titolo di studio più elevato tra i genitori:								
livello primario	0.24	0.43	0.23	0.42	0.27	0.44	0.26	0.44
livello secondario	0.39	0.49	0.39	0.49	0.43	0.50	0.43	0.50
livello terziario	0.38	0.48	0.38	0.49	0.30	0.46	0.31	0.46
madre titolo più alto	0.81	0.39	0.81	0.39	0.80	0.40	0.80	0.40
famiglia regolare			0.79	0.40			0.80	0.40
punteggio (0-100)								
associato alla occupazione più prestigiosa dei genitori			47.93	16.50			47.45	16.08
un familiare a casa			0.30	0.46			0.32	0.47
meno 25 libri in casa			0.23	0.42			0.19	0.39
tra 25 e 200 libri in casa			0.51	0.50			0.56	0.50
più di 200 libri in casa			0.24	0.43			0.24	0.43
Numerosità campione	5517		5322		5973		5859	

Nota: Il programma PISA (*Programme for International Student Assessment*) è una indagine coordinata dall'OCSE e condotta materialmente da Canada Statistics (www.pisa.org). La prima indagine è stata condotta nel 2000 in 35 paesi, mentre la seconda, da cui sono tratti questi dati, è stata condotta nel 2003 in 48 paesi. L'obiettivo dell'indagine è quello di misurare in modo comparabile tra paesi le capacità degli studenti quindicenni su diversi terreni (capacità di comprensione letteraria, capacità matematiche, conoscenza scientifica e capacità di problem solving). Nel 2003 alcune regioni (Piemonte, Lombardia, Veneto, Toscana e province autonome di Trento e Bolzano) hanno contribuito all'ampliamento del campione al fine di ottenere dei campioni regionali rappresentativi: per questa ragione abbiamo fatto uso dei pesi campionari al fine di ottenere dei dati che fossero rappresentativi a livello nazionale.

Tabella A.2: Scelta del tipo di scuola secondaria: risultati completi – PISA 2003

	Campione maschile				Campione femminile			
	Specificazione Base (5517)		Specificazione Estesa (5322)		Specificazione Base (5973)		Specificazione Estesa (5859)	
	Log-ver=- 4761.34		Log-ver=- 4152.97		Log-ver=- 5574.72		Log-ver=- 5189.93	
	Coef.	Err. St.	Coef.	Err. St.	Coef.	Err. St.	Coef.	Err. St.
<i>Tecnica</i>								
età (in mesi)	0.073	0.171	0.015	0.180	-0.033	0.248	-0.211	0.251
straniero			-0.741	0.322			0.507	0.348
parla dialetto in casa			-0.488	0.230			0.019	0.288
capacità matematiche	0.012	0.002	0.006	0.002	0.007	0.002	0.003	0.002
capacità linguistiche			0.003	0.001			0.004	0.002
conoscenze scientifiche			0.001	0.001			0.002	0.001
cap.problem solving			0.003	0.002			0.000	0.002
percorso lento			-0.366	0.649			-0.688	0.579
percorso regolare			-0.176	0.629			-0.241	0.501
bocciato almeno 1 volta			-0.058	0.199			0.431	0.236
genitore livello secondario	0.132	0.151	-0.019	0.153	0.200	0.160	0.427	0.149
genitore livello terziario	0.435	0.174	0.255	0.168	0.206	0.190	0.540	0.192
madre titolo più alto	0.160	0.133	0.063	0.138	-0.109	0.156	0.089	0.164
famiglia regolare			0.441	0.120			0.156	0.170
prestigio socio-economico			0.007	0.005			-0.002	0.006
un familiare a casa			0.192	0.165			0.504	0.173
tra 25 e 200 libri in casa			-0.048	0.135			-0.349	0.143
più di 200 libri in casa			0.037	0.182			-0.672	0.238
costante	-6.550	2.823	-5.824	3.073	-2.214	3.794	-0.551	4.040
<i>Licei</i>								
età (in mesi)	0.349	0.221	0.268	0.249	0.194	0.214	-0.005	0.220
straniero			-0.852	0.500			0.079	0.454
parla dialetto in casa			-0.882	0.278			-0.087	0.391
capacità matematiche	0.021	0.002	0.004	0.002	0.011	0.002	0.002	0.002
capacità linguistiche			0.008	0.002			0.009	0.002
conoscenze scientifiche			0.002	0.002			0.002	0.001
cap.problem solving			0.008	0.002			-0.001	0.002
percorso lento			-1.940	0.682			-1.414	0.542
percorso regolare			-0.768	0.611			-0.425	0.429
bocciato almeno 1 volta			-0.111	0.337			0.381	0.273
genitore livello secondario	1.190	0.215	0.495	0.239	0.981	0.167	0.780	0.148
genitore livello terziario	2.440	0.231	1.203	0.235	1.668	0.203	1.172	0.182
madre titolo più alto	0.587	0.151	0.377	0.165	0.282	0.143	0.394	0.157
famiglia regolare			0.569	0.195			0.095	0.159
prestigio socio-economico			0.046	0.006			0.016	0.006
un familiare a casa			0.367	0.192			0.402	0.168
tra 25 e 200 libri in casa			0.611	0.171			0.247	0.175
più di 200 libri in casa			0.927	0.232			0.493	0.258
costante	-17.226	3.682	-17.769	4.198	-8.169	3.296	-6.562	3.610

Notes: Stime dei parametri dell'utilità da Modello Logit Multinomiale pesato per i pesi campionari riportati in PISA e con errori standard robusti a clusters a livello di scuola. Si veda la tabella A.1 per la definizione completa delle variabili. Normalizzazione rispetto all'alternativa Scuola Professionale.

Tabella A.3: Aspirazione di iscriversi all'università: risultati completi – PISA 2003

	campione maschile		campione femminile	
	specificazione base	specificazione estesa	specificazione base	specificazione estesa
età (in mesi)	0.001 (0.039)	-0.036 (0.041)	0.038 (0.028)	0.016 (0.029)
straniero		0.078 (0.068)		0.074 (0.060)
capacità matematiche	0.001 (0.000)**	-0.001 (0.000)	0.001 (0.000)**	0 (0.000)
capacità linguistiche		0.001 (0.000)		0.001 (0.000)*
conoscenze scientifiche		0.001 (0.000)**		0 (0.000)
capacità problem solving		0 (0.000)		0 (0.000)
percorso lento		-0.264 (0.074)**		-0.34 (0.071)**
percorso regolare		-0.187 (0.060)**		-0.148 (0.032)**
bocciato almeno 1 volta		-0.042 (0.065)		0.016 (0.040)
genitore livello secondario	0.131 (0.035)**	0.073 (0.034)*	0.079 (0.020)**	0.02 (0.023)
genitore livello terziario	0.222 (0.032)**	0.116 (0.035)**	0.158 (0.021)**	0.052 (0.027)
madre titolo più alto	-0.062 (0.026)*	-0.045 (0.028)	-0.026 (0.020)	-0.002 (0.019)
scuola tecnica	0.239 (0.038)**	0.22 (0.040)**	0.077 (0.035)*	0.077 (0.033)*
liceo	0.667 (0.024)**	0.628 (0.029)**	0.465 (0.035)**	0.435 (0.035)**
famiglia regolare		-0.011 (0.029)		0.021 (0.023)
prestigio socio-economico		0.003 (0.001)**		0.003 (0.001)**
tra 25 e 200 libri in casa		0.051 (0.030)		0.086 (0.018)**
più di 200 libri in casa		0.166 (0.038)**		0.159 (0.020)**
Numero osservazioni	5341	5168	5866	5759
Pseudo R ²	0.33	0.35	0.26	0.29
Log verosimiglianza	-2488.11	-2314.41	-2736.95	-2565.82

Note: la variabile dipendente riporta il desiderio degli studenti intervistati di proseguire gli studi fino all'università. In parentesi, errori standard robusti a cluster a livello di scuola. * indica significatività statistica al 5%; ** significatività al 1%. Le stime sono da un modello probit pesato per i pesi di stratificazioni forniti da PISA.

Tavola A.4a: Statistiche Descrittive del campione ISTAT – diplomati 1998 intervistati nel 2001 – Italia – campione maschile

	Specificazione base N= 9956		Specificazione estesa N= 10453	
	Media	Dev. St.	Media	Dev. St.
istruzione genitori livello secondario	0.3820	0.4859	0.3800	0.4854
istruzione genitori livello terziario	0.0902	0.2865	0.0905	0.2869
madre titolo più alto	0.7475	0.4345	0.7473	0.4346
padre lavora	0.9518	0.2142		
padre lavora (operaio)			0.1655	0.3717
padre lavora (operaio specializzato)			0.2910	0.4543
padre lavora (impiegato)			0.3393	0.4735
padre lavora (impiegato direttivo)			0.1563	0.3632
ha fratelli o sorelle			0.8621	0.3448
madre a casa			0.5903	0.4918
residenza nord-est	0.1412	0.3483	0.1417	0.3487
residenza centro	0.2123	0.4090	0.2134	0.4097
residenza sud	0.2866	0.4522	0.2834	0.4507
residenza isole	0.1479	0.3550	0.1487	0.3558
non sposato/a			0.9752	0.1555
voto scuola media= buono	0.2598	0.4386		
voto scuola media= distinto	0.1241	0.3298		
voto scuola media= ottimo	0.0860	0.2803		
voto medie= sufficiente; maturità=42-47			0.1573	0.3641
voto medie= sufficiente; maturità=48-53			0.0479	0.2136
voto medie= sufficiente; maturità=54-60			0.0184	0.1343
voto medie= buono; maturità=36-41			0.0902	0.2865
voto medie= buono; maturità=42-47			0.0867	0.2814
voto medie= buono; maturità=48-53			0.0518	0.2215
voto medie= buono; maturità=54-60			0.0256	0.1581
voto medie= distinto; maturità=36-41			0.0284	0.1662
voto medie= distinto; maturità=42-47			0.0365	0.1876
voto medie= distinto; maturità=48-53			0.0284	0.1662
voto medie= distinto; maturità=54-60			0.0249	0.1557
voto medie= ottimo; maturità=36-41			0.0094	0.0964
voto medie= ottimo; maturità=42-47			0.0176	0.1315
voto medie= ottimo; maturità=48-53			0.0196	0.1387
voto medie= ottimo; maturità=54-60			0.0396	0.1950
bocciato durante scuola secondaria			0.2999	0.4582
scuola tecnica	0.4233	0.4941	0.4241	0.4942
liceo	0.1318	0.3383	0.1301	0.3364

Note: Dati tratti da *Indagine 2001: Percorsi di studio e lavoro dei diplomati*, ISTAT. Si tratta di un'indagine triennale, condotta su una singola leva di studenti a circa tre anni dal conseguimento del Diploma.

Tavola A.4b: Statistiche Descrittive del campione ISTAT – diplomati 1998 intervistati nel 2001 – Italia – campione femminile

	Specificazione Base N= 11527		Specificazione Estesa N=12343	
	Media	Dev. St.	Media	Dev. St.
istruzione genitori livello secondario	0.3504	0.4771	0.3509	0.4773
istruzione genitori livello terziario	0.0847	0.2784	0.0854	0.2795
madre titolo più alto	0.7496	0.4332	0.7494	0.4334
padre lavora	0.9451	0.2278		
padre lavora (operaio)			0.1637	0.3701
padre lavora (operaio specializzato)			0.3075	0.4615
padre lavora (impiegato)			0.3270	0.4691
padre lavora (impiegato direttivo)			0.1466	0.3537
ha fratelli o sorelle			0.8675	0.3391
madre a casa			0.5788	0.4938
residenza nord-est	0.1647	0.3710	0.1645	0.3708
residenza centro	0.1805	0.3846	0.1842	0.3876
residenza sud	0.2470	0.4313	0.2494	0.4327
residenza isole	0.1647	0.3710	0.1636	0.3699
non sposato/a			0.9472	0.2237
voto scuola media= buono	0.2907	0.4541		
voto scuola media= distinto	0.1686	0.3744		
voto scuola media= ottimo	0.1574	0.3642		
voto medie= sufficiente; maturità=42-47			0.1394	0.3464
voto medie= sufficiente; maturità=48-53			0.0524	0.2229
voto medie= sufficiente; maturità=54-60			0.0182	0.1338
voto medie= buono; maturità=36-41			0.0735	0.2609
voto medie= buono; maturità=42-47			0.0934	0.2910
voto medie= buono; maturità=48-53			0.0818	0.2741
voto medie= buono; maturità=54-60			0.0352	0.1842
voto medie= distinto; maturità=36-41			0.0258	0.1584
voto medie= distinto; maturità=42-47			0.0462	0.2099
voto medie= distinto; maturità=48-53			0.0476	0.2128
voto medie= distinto; maturità=54-60			0.0379	0.1910
voto medie= ottimo; maturità=36-41			0.0106	0.1025
voto medie= ottimo; maturità=42-47			0.0263	0.1601
voto medie= ottimo; maturità=48-53			0.0391	0.1937
voto medie= ottimo; maturità=54-60			0.0809	0.2727
bocciato durante scuola secondaria			0.1832	0.3868
scuola tecnica	0.3558	0.4788	0.3563	0.4789
liceo	0.2047	0.4035	0.2011	0.4008

Note: Dati tratti da *Indagine 2001: Percorsi di studio e lavoro dei diplomati*, ISTAT. Si tratta di un'indagine triennale, condotta su una singola leva di studenti a circa tre anni dal conseguimento del Diploma.