



UNIVERSITÀ DEGLI STUDI DI ANCONA

DIPARTIMENTO DI ECONOMIA

PERFORMANCE ACCADEMICA E SCELTA
DELLA FACOLTÀ UNIVERSITARIA: ASPETTI
TEORICI E EVIDENZA EMPIRICA

Massimiliano Bratti e Stefano Staffolani

QUADERNI DI RICERCA n.

18 aprile 2001

Comitato scientifico:

Sintesi

In questo articolo vengono analizzati i fattori che influenzano l'impegno di studio, e conseguentemente la *performance* degli studenti iscritti all'università e quelli che spingono gli studenti a indirizzarsi verso uno specifico corso di laurea (facoltà). In particolare, un modello teorico definisce le grandezze che incidono sull'impegno profuso durante gli studi universitari e il livello di utilità massima attesa che ogni studente presume di ottenere da ognuna delle lauree. L'applicazione empirica presentata mette in luce come il background familiare, i caratteri del percorso formativo pre-universitario, l'abilità degli individui e le aspettative di voto incidano sia sulla scelta dell'indirizzo di studio che sulla *performance* universitaria. Emerge che, anche se i fattori *forward looking* sono importanti, l'esistenza di fattori *inerziali* può portare ad una scarsa sensibilità delle iscrizioni ai diversi corsi di laurea rispetto alle condizioni del mercato del lavoro dei laureati.

JEL Class.: I21

Keywords: facoltà, istruzione, sample selection, università

Indirizzo: Massimiliano Bratti: Dipartimento di Economia, Università degli Studi di Ancona, 60121 Ancona e Department of Economics, University of Warwick, CV4 7AL, Coventry (UK). Email: massib@deanovell.unian.it.
Stefano Staffolani: Dipartimento di Economia, Università degli Studi di Ancona, Piazza Martelli, 8, 60121 Ancona. Email: staffolani@deanovell.unian.it

Performance accademica e scelta della facoltà universitaria: aspetti teorici e evidenza empirica

Massimiliano Bratti e Stefano Staffolani

1 Introduzione

Nella letteratura economica sono numerosi gli studi empirici che hanno analizzato la decisione di proseguire gli studi a livello universitario¹, meno frequenti sono invece le analisi del tipo di corso di laurea ovvero della facoltà scelta.

Tuttavia lo studio dei fattori che influenzano la scelta della facoltà è interessante per varie ragioni. Un primo motivo potrebbe essere quello di testare se le aspettative di ‘successo’ degli studenti, in termini di voto atteso, influenzano la scelta del tipo di corso di laurea. Le aspettative di successo nei diversi corsi dovrebbero influenzare la scelta della facoltà data la correlazione positiva tra voto di laurea, che può essere considerato come un indicatore di *performance* accademica, ed il reddito percepito dai laureati². Un altro importante motivo è quello di individuare se esistono dei pattern ricorrenti nella scelta di certi tipi di corsi per tipo di *background* socio-familiare. Ciò risulta utile dato che in genere un certo tipo di studi è legato all’esercizio di una data professione e pertanto ad una data *performance* sul mercato del lavoro. A questo proposito è stato sostenuto altrove che gli individui provenienti da famiglie più agiate avrebbero una maggiore propensione al rischio e tenderebbero a scegliere i corsi più difficili e pertanto più rischiosi (Montmarquette *et al.* 1996) ma in genere anche più remunerativi. Ciò limiterebbe in parte la mobilità sociale. La letteratura sulla domanda di istruzione terziaria ha evidenziato che in generale il *background* familiare ha scarsa importanza per la decisione di frequentare l’università (vedi ad esempio Heckman 1999), contrariamente al tipo di istruzione secondaria ed alla *performance* scolastica

¹Per l’Italia si vedano tra gli altri Gambetta (1987), Cobalti e Schizzerotto (1993), Checchi (2000).

²Si veda ad esempio Staffolani e Sterlacchini (2001) con riferimento allo stesso dataset utilizzato nel presente lavoro.

passata (che sono tuttavia correlate col retroterra socio-familiare dello studente). Analogamente, in questa sede è interessante indagare se il *background* familiare non è importante nemmeno per la scelta del tipo di studi intrapreso. In entrambi i casi la valutazione del peso relativo di meccanismi ‘razionali’, come nel caso delle aspettative di successo, e di meccanismi ‘inerziali’, come nel caso dell’origine socio-familiare (in grado di rendere più lento l’adeguamento qualitativo dell’offerta di laureati alla domanda di figure professionali espressa dall’economia, per le quali esistono vantaggi relativi in termini di occupabilità e/o di redditi) è di estrema importanza.

Le principali conclusioni cui perviene il nostro contributo indicano che, tanto da un punto di vista teorico che empirico, esistono forti elementi ‘inerziali’ che spingono i diplomati a scegliere un dato corso di laurea. Oltre questi elementi, è rilevante nelle scelte anche l’aspettativa di successo nelle diverse facoltà, misurata attraverso il voto atteso.

Il paper si articola come segue. Nel paragrafo successivo introdurremo un semplice modello di scelta della disciplina universitaria (a cui ci riferiremo nel corso del paper anche usando i termini di gruppo di lauree o facoltà), che verrà utilizzato per specificare il modello econometrico. Nel paragrafo 3 spiegheremo i problemi che la stima del modello econometrico pone. Il paragrafo 4 introduce brevemente l’indagine IPLAM³, utilizzata per le analisi empiriche. Il paragrafo 5 illustra le variabili incluse nel modello econometrico. I risultati della stima delle equazioni di *performance* sono riportati nel paragrafo 6 e quelli della stima della forma che abbiamo definito ‘semi-strutturale’ del modello di scelta del tipo di laurea nel paragrafo 7. Il paragrafo 8 conclude.

2 Il framework teorico di riferimento

Nella letteratura economica si possono trovare analisi teoriche ed empiriche della scelta di proseguire gli studi dopo il completamento dell’istruzione secondaria, meno frequenti sono invece le analisi della scelta del tipo di studi universitari, ovvero della facoltà, soprattutto in Italia⁴.

Scopo di questo paragrafo è presentare un modello di scelta della facoltà condizionato all’ipotesi che l’individuo abbia già deciso di proseguire gli studi a livello universitario dopo il conseguimento del diploma di maturità.

³L’indagine IPLAM è stata svolta nell’ambito di una ricerca finanziata dal CNR (progetto di ricerca coordinato n. 97.01357.CT10 del 1997, rifinanziato nel 1998). Ringraziamo il Dipartimento di Economia dell’Università di Ancona che ci ha messo a disposizione la banca dati e tutti coloro, in particolare Pietro Alessandrini, Alessandro Sterlacchini, Riccardo Lucchetti e Barbara Ermini che hanno contribuito alla realizzazione della banca dati.

⁴Per una rara eccezione si veda Checchi (2000).

2.1 Il modello

Ipotizziamo che l'individuo i tragga utilità da una serie di attività indicate con z :

$$V_i = \sum_{t=0}^T \beta^t E[U(z_{i,1,t}, z_{i,2,t} \dots z_{i,n,t})] \quad (1)$$

dove V_i è l'utilità vitale dell'individuo i , U la funzione di utilità uniperiodale, $\beta = 1/(1+r)$, con $r > 0$, il fattore di sconto e $E[\cdot]$ indica l'aspettativa matematica. L'orizzonte temporale considerato è quello compreso tra il periodo in cui la scelta dell'iscrizione alla facoltà deve essere effettuata ($t = 0$) e il periodo in cui si termina la vita attiva (T).

Tra le diverse attività dell'individuo i consideriamo il tempo dedicato allo studio nella disciplina universitaria j al tempo t ($x_{i,j,t}$), il tempo libero ($\ell_{j,t}$) e l'attività di consumo ($C_{i,t}$). Assumiamo per semplicità che l'individuo nel ciclo di vita non possa intraprendere gli studi universitari in più discipline, né simultaneamente né sequenzialmente. Ipotizziamo inoltre che⁵ $U'_{x_{i,j,t}}$ possa essere sia positiva che negativa, cioè che l'individuo possa trarre utilità o disutilità dall'attività di studio e che $U'_{C_{i,t}} > 0$, $U'_{\ell_{i,t}} > 0$, cioè che tanto il consumo che il tempo libero diano utilità.

Assumiamo che quando gli individui si iscrivono ipotizzino di terminare il corso nella durata media attesa del corso j , indicata con τ_j , ovvero non si ritengano mediamente più o meno capaci degli altri studenti in termini di tempo richiesto per il completamento degli studi⁶. Mentre l'individuo è iscritto all'università (per $0 \leq t \leq \tau_j$) egli può impiegare il tempo disponibile (normalizzato ad uno) per lo studio, il tempo libero ed il lavoro ($h_{i,t}$), cioè risulta $1 = x_{i,j,t} + \ell_{i,t} + h_{i,t}$, mentre nella seconda fase del ciclo vitale (per $\tau_j + 1 \leq t \leq T$), quando egli ha completato gli studi, il tempo è impiegato solo lavorando o usufruendo di tempo libero, cioè $1 = \ell_{i,t} + h_{i,t}$.

L'individuo (d'ora in poi per semplificare la notazione tralasciamo il pedice i) che ha scelto la facoltà j presenta la seguente funzione dell'utilità attesa nella disciplina j :

$$V_j = \sum_{t=0}^{\tau_j} \beta^t E[U(x_{j,t}, \ell_{j,t}, C_{j,t})] + \sum_{t=\tau_j+1}^T \beta^t E[U(\tilde{\ell}_{j,t}, \tilde{C}_{j,t})] \quad (2)$$

dove la scissione in due fasi della vita dell'individuo dipende dal fatto che

⁵Il termine y'_x indica la derivata parziale della variabile y rispetto alla variabile x

⁶Questa ipotesi è forse poco realistica per la realtà italiana, dove la distribuzione delle durate effettive di studio è molto dispersa. Assumiamo inoltre che tutti terminino gli studi, anche se con *performances* differenti. Questo anche perché nella parte empirica ci riferiremo ai soli laureati.

l'evoluzione del salario nel tempo presenta una discontinuità al momento della laurea. Le variabili con la tilde rappresentano variabili di scelta relative alla fase lavorativa. L'individuo dovrà scegliere, in ognuna delle facoltà, il tempo di studio ($x_{j,t}$), il tempo libero durante il periodo di studio ($\ell_{j,t}$) e il tempo libero durante il periodo di lavoro ($\tilde{\ell}_{j,t}$) sotto il vincolo di bilancio relativo al consumo.

Supponiamo che i comportamenti degli agenti siano costanti all'interno delle due fasi, cioè, una volta scelta una allocazione ottimale del tempo, questa rimanga costante durante la fase della vita in cui si studia e durante la fase in cui si lavora. Questo implica $x_{j,t} = x_j$, $\ell_{j,t} = \ell_j$, $C_{j,t} = C_j$ per ogni $0 \leq t \leq \tau_j$ e $\tilde{\ell}_{j,t} = \tilde{\ell}_j$, $\tilde{C}_{j,t} = \tilde{C}_j$ per ogni $\tau_j + 1 \leq t \leq T$. Inoltre, supponiamo che $R_{j,t} = R_j$, $\phi_{j,t} = \phi_j$ per ogni t ; in ogni periodo di studio si può contare sulla stessa disponibilità finanziaria⁷, in termini di redditi non da lavoro; il costo dello studio universitario è costante per ogni corso di laurea (e dipende tra l'altro dalla localizzazione degli atenei che offrono un certo tipo di corso).

Assumiamo l'esistenza di mercati dei capitali perfettamente imperfetti (PICM, *perfectly imperfect capital markets*⁸). In questo caso in ogni periodo il bilancio deve essere in pareggio. Il vincolo di bilancio per i periodi di studio sarà allora:

$$C_j = \omega_k(1 - x_j - \ell_j) + R_j - \phi_j \equiv C(x_j, \ell_j) \quad (3)$$

dove ω_k è il salario atteso del diplomato con maturità k e con voto P_k (e pertanto il primo addendo a destra dell'uguale rappresenta il reddito da lavoro). L'individuo può finanziare il proprio consumo durante gli studi lavorando o attraverso i trasferimenti familiari, dopo aver pagato le spese per gli studi. In linea di principio dovremmo continuare ad indicare l'operatore valore atteso, dato che al momento della scelta della facoltà molte delle variabili di interesse non sono conosciute. Nel seguito tralasceremo di indicare questo operatore, ma si tenga conto che nella realtà gran parte delle variabili cui facciamo riferimento (salari futuri, *performance*) sono dei valori attesi.

Una volta ultimati gli studi universitari, per i periodi di lavoro il vincolo di bilancio è:

$$\tilde{C}_j = w(j, P_j)(1 - \tilde{\ell}_j) \quad (4)$$

dove si suppone che il salario del laureato nella disciplina j , $w(j, P_j)$, sia

⁷Questi trasferimenti potrebbero essere differenziati tra corsi di laurea sia a causa di preferenze dei genitori propensi a pagare per alcuni corsi ma non per altri, sia per eventuali borse di studio.

⁸Si suppone che durante il periodo di studio non sia possibile prendere a prestito dal sistema bancario, pertanto gli studi devono essere finanziati dalla famiglia, dagli Enti pubblici tramite borse di studio o attraverso il lavoro, ipotesi che non sembra essere lontana dalla realtà italiana.

funzione positiva del voto ottenuto alla laurea P_j . In particolare supporremo la seguente funzione di *performance* universitaria:

$$P_j = P(x_j, p_k, k) \quad (5)$$

che ci dice che il voto ottenuto nella facoltà j è funzione della quantità di studio, del tipo di scuola superiore frequentata, della *performance* nella scuola superiore. L'accumulazione di capitale umano dovuto allo studio incrementa quindi il salario in misura dipendente dall'ottenimento del titolo ma anche dipendente dalla *performance* negli studi.

Dalla equazioni 4 e 5 possiamo pertanto scrivere sinteticamente:

$$\tilde{C}_j = w(x_j)(1 - \tilde{\ell}_j) \equiv \tilde{C}(x_j, \tilde{\ell}_j) \quad (6)$$

dove $w(x_j) \equiv w(j, P_j(x_j, P_k, k))$ indica le aspettative di salario di un individuo che sceglie la facoltà j e si impegna nello studio per x_j ore.

Le derivate del consumo nel periodo di studio rispetto ai due argomenti (equazione 3) sono ambedue negative, mentre le stesse derivate nel periodo di lavoro (equazione 6) sono rispettivamente positiva (tanto più si è studiato, tanto più la *performance* è buona e tanto più il salario atteso è elevato) e negativa.

La funzione di utilità intertemporale attesa, dell'equazione 2 può essere scritta nel modo seguente:

$$V_j = \sum_{t=0}^{\tau_j} \beta^t U(x_j, \ell_j, C(x_j, \ell_j)) + \sum_{t=\tau_j+1}^T \beta^t U(\tilde{\ell}_j, \tilde{C}(x_j, \tilde{\ell}_j)) \quad (7)$$

Sviluppando le sommatorie dell'equazione precedente è possibile riscriverla:

$$V_j = \frac{(1 - \beta^{\tau_j+1})U(x_j, \ell_j, C(x_j, \ell_j)) + (\beta^{\tau_j+1} - \beta^{T+1})U(\tilde{\ell}_j, \tilde{C}(x_j, \tilde{\ell}_j))}{1 - \beta} \quad (8)$$

Le variabile di scelta dell'individuo sono il tempo di studio e il tempo libero nelle due fasi, quella studentesca e quella lavorativa.

La massimizzazione dell'utilità intertemporale attesa del nostro agente (equazione 8) rispetto al tempo di studio (x_j), al tempo libero nel primo sottoperiodo (ℓ_j) e al tempo libero nel periodo in cui si lavora ($\tilde{\ell}_j$) porta, nel caso di una soluzione interna, alle seguenti condizioni di primo ordine:

$$-\frac{U'_{x_j} + U'_{C_j} C'_{x_j}}{U'_{\tilde{C}_j} \tilde{C}'_{x_j}} = \frac{\beta^{\tau_j+1} - \beta^{T+1}}{1 - \beta^{\tau_j+1}} \equiv B_j \quad (9)$$

$$U'_{\ell_j} = -U'_{C_j} C'_{\ell_j} \quad (10)$$

$$U'_{\tilde{\ell}_j} = -U'_{\tilde{C}_j} \tilde{C}'_{\ell_j} \quad (11)$$

2.2 Una analisi con forme funzionali esplicite

Supponiamo una funzione di utilità lineare nel consumo e separabile nei tre argomenti, precisamente:

$$U(x, \ell) = (m\ell - n\ell^2) + (ax - bx^2) + \kappa C \quad (12)$$

quindi κ è l'utilità marginale del consumo⁹. Sappiamo dalle equazioni 3 e 6 che $C'_{x_j} = C'_{\ell_j} = -\omega_k$ e che $C'_{\tilde{\ell}_j} = -w(x_j)$.

Date queste ipotesi, otteniamo immediatamente dalla 10 che:

$$U'_{\ell_j} = \kappa\omega_k \quad \Rightarrow \quad \ell^* = \ell(\kappa, \omega_k) \quad \forall j$$

e che pertanto la scelta ottimale del tempo libero durante gli studi universitari è indipendente dal tempo dedicato allo studio. La quantità ottimale di tempo libero durante la fase degli studi è quindi indipendente dal tipo di facoltà scelta. Dipende invece dalle preferenze e dal percorso formativo preesistente (vedi la definizione di ω_k , equazione 3). L'equazione 11 può essere scritta:

$$U'_{\tilde{\ell}_j} = \kappa w(x_j) \quad \Rightarrow \quad \tilde{\ell}_j^* = \tilde{\ell}(\kappa, x_j)$$

La scelta del tempo libero durante il periodo di lavoro dipende dal tipo di facoltà scelto e dall'impegno profuso nello studio. Supponendo, come abituale, che $U'_{\tilde{\ell}}$ sia concava, tanto più si è studiato tanto più la *performance* universitaria sarà alta e tanto più sarà basso il tempo libero e elevato l'orario di lavoro.

Supponiamo che la *performance* nella facoltà j sia linearmente dipendente dalla *performance* ottenuta alla maturità e dall'impegno nello studio:

$$P_j = \gamma_{0j}P_k + \gamma_{1j}x_j$$

⁹La forma quadratica ci permette di analizzare tanto individui che ottengono utilità dallo studio (se $a > 0$, in un intorno di $x = 0$ lo studio dà utilità) che individui che vedono sempre il tempo di studio come un 'male' ($a \leq 0$); inoltre, il tempo libero crea sempre utilità se $m > 2n$, cioè se la derivata dell'utilità rispetto al tempo libero è sempre positiva anche per il valore massimo $\ell = 1$.

e supponiamo infine che il salario che si percepisce dopo aver completato gli studi nella disciplina j sia crescente linearmente rispetto alla *performance*, in particolare che valga:

$$w_j = \alpha_j P_j$$

Pertanto, possiamo scrivere il salario come:

$$w_j = \bar{w}_{jk} + z_j x_j$$

dove $\bar{w}_{jk} = \alpha_j \gamma_{0j} P_k$ è il salario che si percepisce da laureati se l'impegno nello studio è nullo, e $z_j = \alpha_j \gamma_{1j}$ è la derivata del salario rispetto alle ore di studio. Non sembra troppo restrittivo assumere che $\bar{w}_{jk} \geq \omega_k$, ovvero anche se l'individuo ha un impegno nullo ($x_j = 0$) come laureato percepisce un salario non inferiore a quello che avrebbe percepito da diplomato.

Dall'equazione 10 e data la relazione tra salario e tempo di studio, si ottiene facilmente $m - 2n\ell_j = \kappa\omega_k$, quindi:

$$\ell_j = \frac{m - \kappa\omega_k}{2n}$$

Da cui emerge che deve valere: $\kappa\omega_k \leq m \leq \kappa\omega_k + 2n$.

Riprendiamo le equazioni 9 e 11; date la funzione di utilità degli individui dell'equazione 12 otteniamo:

$$a - 2bx_j = \kappa \left[\omega_k - B_j z_j (1 - \tilde{\ell}_j) \right]$$

e

$$m - 2n\tilde{\ell}_j = -\kappa(\bar{w}_{jk} + z_j x_j)$$

cioè un sistema a due equazioni e due incognite (x_j e $\tilde{\ell}_j$). La soluzione ci fornisce i valori ottimali di queste due variabili.

$$x_j^* = \frac{2na + \kappa[z_j B_j (2n - m + \kappa\bar{w}_{jk}) - 2n\omega_k]}{4bn - \kappa^2 z_j^2 B_j}$$

$$\tilde{\ell}_j^* = \frac{\kappa(\kappa\omega_k - a)z_j + 2b(m - \kappa\bar{w}_{jk}) - \kappa^2 z_j^2 B_j}{4bn - \kappa^2 z_j^2 B_j}$$

L'equazione che definisce x_j^* ci fornisce informazione sulle variabili che incidono sull'impegno profuso nello studio nelle diverse facoltà, e quindi sulla

performance ottenuta dagli studenti. Supponiamo che il denominatore sia positivo¹⁰.

In particolare è facile ottenere che l'impegno nello studio e la *performance* dipendono dai parametri del modello con questi segni:

$$\frac{\partial x_j^*}{\partial m} < 0 \quad \frac{\partial x_j^*}{\partial a} > 0 \quad \frac{\partial x_j^*}{\partial b} < 0 \quad \frac{\partial x_j^*}{\partial \omega_k} < 0 \quad \frac{\partial x_j^*}{\partial \bar{w}_{jk}} > 0$$

Inoltre, tenendo conto che risulta $2n + \kappa \bar{w}_{jk} > m$, si ottiene:

$$\frac{\partial x_j^*}{\partial z_j} > 0 \quad \frac{\partial x_j^*}{\partial B_j} > 0$$

e, data la definizione di B_j :

$$\frac{\partial x_j^*}{\partial \tau_j} < 0 \quad \frac{\partial x_j^*}{\partial T} > 0$$

Non è invece possibile valutare in modo univoco la variazione del tempo di studio al variare dell'utilità marginale del consumo (κ) e del parametro n .

Ovviamente, dato che la *performance* universitaria è crescente rispetto al tempo di studio, i segni attesi della relazione tra *performance* e parametri del modello sono uguali a quelli analizzati per il tempo di studio.

Proviamo a riepilogare le relazioni che ci sembrano più importanti:

$$P_{i,j,k} = P(\underset{+}{P_k}, \underset{+}{\gamma_{0j}}, \underset{+}{a_i}, \underset{-}{m_i}, \underset{-}{b_i}, \underset{-}{\omega_k}, \underset{+}{\bar{w}_{jk}}, \underset{+}{\alpha_{ij}}, \underset{+}{\gamma_{1j}}, \underset{-}{\tau_j}, \underset{+}{T_i})$$

La *performance* attesa sarà tanto più elevata quanto più l'individuo ha ottenuto una buona valutazione alla maturità (P_k) e tanto più questa valutazione è utile per il corso di laurea j (γ_{0j}). La *performance* dipenderà ovviamente dalle preferenze degli individui. Tanto più 'piace' poco il tempo libero (basso m) e tanto più piace studiare (alto a , basso b), tanto migliore sarà la *performance*.

Se il lavoro ottenibile con il diploma retribuisce bene (alto ω_k) si otterranno *performance* peggiori durante gli studi universitari.

¹⁰Se fosse negativo, avremo il risultato paradossale che $\frac{dx}{da} < 0$, cioè che se aumenta la preferenza per lo studio il tempo di studio si riduce. Lo stesso si verifica per il tempo libero: se il denominatore fosse negativo, il tempo libero si ridurrebbe all'aumentare del parametro m . Inoltre, si tenga conto che $z = \frac{\epsilon}{1-\epsilon} \frac{w}{x}$, dove ϵ è l'elasticità del salario al tempo di studio. Se, come ci sembra realistico, questa elasticità è bassa e per valori di x che non tendano a zero, anche z tende ad essere basso; questo conferma che il segno atteso del denominatore è positivo.

Lauree che premiano molto i voti alti in termini di incremento di salario atteso dovuto ad una buona *performance* (alto α_j) spingono all'ottenimento di valutazioni migliori.

Una lunga durata attesa del corso di studi (alto τ_j) spinge gli individui a studiare di meno in ogni periodo ottenendo *performance* peggiori. L'effetto di una lunga vita lavorativa (che dipende tra l'altro dall'età dell'individuo) è esattamente nella direzione opposta.

E' infine interessante sottolineare che l'impegno nello studio e il voto atteso non dipendono nè dai trasferimenti della famiglia, R_j , nè dal costo dell'istruzione¹¹, ϕ_j . Ciò non toglie che la famiglia possa influire sulla *performance* universitaria attraverso altre vie; ad esempio, l'utilità derivante dallo studio (i parametri a e b) è presumibilmente influenzata dal titolo di studio dei genitori. Inoltre è possibile che il rendimento della *performance* differisca per classe sociale. Potrebbe accadere che tramite meccanismi di *social networking* i laureati provenienti da una classe sociale elevata siano in grado di trovare una buona occupazione e percepire redditi elevati a prescindere dal voto di laurea. Al contrario gli individui provenienti da background meno privilegiati potrebbero avere un maggiore rendimento della *performance*, in termini di probabilità di trovare una buona occupazione e percepire redditi elevati. Se questo fosse il caso il modello predice che gli individui meno agiati studiano di più ed hanno *performance* migliori.

Dato che la scelta ottima di tempo libero durante gli studi non dipende da variabili influenzate dal corso di laurea (come la *performance* universitaria o il rendimento economico della stessa), questa scelta incide solo sul *trade-off* tra tempo dedicato allo studio e tempo dedicato al lavoro; corsi di laurea più remunerativi in termini di salario atteso, più difficili per una durata attesa più lunga e nei quali le prospettive salariali dipendono fortemente dalla *performance* universitaria devono portare ad un minore impegno lavorativo durante gli studi e quindi a minori consumi nel periodo di studio.

Noti i valori ottimali di x_j^* per ognuno dei corsi di laurea e noto il tempo libero di cui si dispone durante gli studi ℓ_j^* , si può calcolare il livello di consumo nella fase degli studi e nella fase lavorativa, cioè l'utilità uniperiodale $U(x^*, \ell^*, C^*)$ e, dall'equazione 8 l'utilità intertemporale attesa V_j^* per ognuno dei possibili corsi di studio. Ogni individuo sceglierà quindi il corso che permette di ottenere una utilità più elevata. Ovviamente in questa scelta tornano ad essere importanti anche quei parametri come il costo per tipo di istruzione universitaria ϕ_j che non incidevano sulla determinazione del livello ottimale dell'impegno e sulla *performance*.

Pertanto, secondo questa logica dovremmo osservare per ogni studente

¹¹Su questa conclusione incide ovviamente l'ipotesi di individui neutrali al rischio.

che la facoltà scelta è quella che permette di ottenere l'utilità attesa più elevata dall'equazione 8 una volta che sono stati sostituiti i valori ottimali delle variabili oggetto di scelta dell'individuo.

I risultati ottenuti possono essere così riepilogati:

1. il consumo nel periodo di studio dipende dalla facoltà scelta. In particolare esistono delle interazioni complesse tra caratteristiche socio-familiari, come i trasferimenti di cui si può disporre nel corso degli studi, offerta di lavoro dello studente e risultati universitari. Durante gli studi gli studenti si trovano a fronteggiare un *trade-off* tra offerta di lavoro e *performance* universitaria: se essi vogliono raggiungere dei soddisfacenti livelli di consumo devono farlo a discapito dei loro risultati accademici che determineranno tuttavia i loro redditi (e consumo) futuri;
2. la scelta della facoltà dipende altresì dall'utilità (o la disutilità) derivante dal tempo trascorso studiando, cioè dalla domanda di consumo di istruzione. A questo proposito le stesse caratteristiche socio-familiari di cui sopra potrebbero entrare nuovamente in gioco differenziando le preferenze individuali per certe discipline;
3. la scelta della facoltà dipende inoltre dal rendimento economico della *performance* universitaria nelle diverse discipline¹². Gli individui tendono *ceteris paribus* ad iscriversi alle discipline per cui i redditi attesi sono maggiori e questo dipende tra l'altro dalla loro probabilità attesa di successo nei diversi corsi, cioè dalla *performance* attesa. È inoltre interessante osservare che il rendimento economico dell'istruzione dipende dagli anni che rimangono ancora da vivere al termine degli studi ($T - \tau_j$) e pertanto dall'età degli studenti. Una possibile implicazione è che studenti 'maturi' (per cui T è piccolo) potrebbero avere una preferenza per corsi che richiedono meno tempo per essere completati (cioè con minore durata attesa) al fine di fruire del rendimento economico dell'istruzione per un periodo più lungo;

3 Analisi empirica: il problema della selezione

Un problema che si deve affrontare qualora si voglia stimare un modello di scelta del corso di laurea è quello della *self-selection*. Il modello economico il-

¹²Cioè da come il voto nella disciplina della laurea influenza la probabilità di trovare un'occupazione ed il salario ricevuto.

lustrato nel paragrafo precedente suggerisce che coloro che vengono osservati nei diversi corsi sono effettivamente quelli per cui la scelta risulta ottimale, nel senso che massimizza l'utilità attesa. Tale questione è ben spiegata da Roy (1951) relativamente al problema della scelta dell'occupazione. In ciascuna occupazione o nel nostro caso in ciascun corso di laurea osserviamo soltanto coloro che hanno dei vantaggi comparati in termini di utilità ricevuta nello svolgimento di quella determinata professione, o nello studio di quella data disciplina. In fase di stima del modello se non si tiene conto del problema della selezione si ottengono delle stime inconsistenti dei parametri di interesse. La teoria econometrica necessaria alla risoluzione del problema è stata elaborata nel caso bivariato da Lee (1978) e Heckman (1979) ed estesa al caso multivariato da Lee (1983), a cui rimandiamo per le derivazioni analitiche¹³. Alcune applicazioni empiriche si possono trovare in Willis e Rosen (1979), Trost e Lee (1984) e Berger (1988).

Iniziamo con lo specificare in dettaglio le equazioni e la logica del modello. La *earnings equation* per la disciplina j , con $j = 1...J$ è:

$$y_{ij} = X_i\alpha_{1j} + P_{ij}\alpha_{2j} + v_{ij} \quad (13)$$

La 13 va interpretata come un'equazione di reddito e non di salario (cioè la $w(x_j)h_{i,j}$ nel modello teorico), e tiene conto anche della durata del tempo di lavoro associata con la disciplina j . X_i è un vettore di caratteristiche personali dell'individuo i , P_{ij} è la *performance* ottenuta dall'individuo i nel corso di studi j , cioè una misura del successo accademico, e v_{ij} una componente d'errore non osservabile dall'econometrico e incorrelata con i regressori (ad esempio la 'fortuna'). Evidenza in favore di un rendimento positivo della *performance* accademica, tale da giustificare una formulazione di questo tipo, si può trovare in Dolton e Makepeace (1990), Smith and Naylor (2000) e Stafolani e Sterlacchini(2001)¹⁴ con specifico riferimento ai dati che utilizzeremo nell'analisi empirica. A sua volta la funzione di *performance* accademica può essere definita come:

$$P_{ij} = Z_i\gamma_j + u_{ij} \quad (14)$$

dove Z_i è un insieme di caratteristiche personali dell'individuo i e u_{ij} una componente stocastica d'errore non osservabile dall'econometrico ed incorrelata con i regressori che condiziona la *performance*, come ad esempio eventuali problemi di salute dell'individuo.

¹³Un approccio alternativo al problema della selezione in presenza di modelli di scelta policotomica è stato elaborato da Hay (1980).

¹⁴Il voto di laurea ha un impatto positivo sia sui redditi percepiti che sulla probabilità di occupazione.

Assumendo una formulazione lineare, la forma ‘strutturale’ della *value function* per l’individuo i qualora scelga di studiare la disciplina j è:

$$V_{ij} = X_i\beta_{1j} + y_{ij}\beta_{2j} + \epsilon_{ij} \quad (15)$$

Abbiamo definito questa forma ‘strutturale’ nel senso che scinde il contributo del reddito atteso degli individui, il quale dipende a sua volta dalla *performance*, da quello degli altri fattori che contribuiscono a determinare il valore massimo dell’utilità. La forma non è strutturale nel senso stretto del termine, ovvero non stimiamo i parametri della funzione di utilità vitale data dall’equazione 12.

Sostituendo y_{ij} in V_{ij} possiamo ottenere una forma che chiameremo ‘semi-strutturale’ della *value function*, mentre sostituendo a P_{ij} l’espressione relativa otteniamo la forma ridotta. La stima della forma ‘semi-strutturale’ richiede che i set di regressori (caratteristiche) X_i e Z_i non coincidano (mentre quella della forma ridotta non richiede tale restrizione). Ciò verrà ottenuto attraverso delle restrizioni di esclusione come si trova frequentemente nella letteratura sull’argomento¹⁵.

Innanzitutto, cominciamo col precisare che i dati disponibili, in cui i redditi sono espressi in classi e vi è scarsa variabilità tra individui, non ci consentono di stimare la forma che abbiamo definito ‘strutturale’ del modello. Per questa ragione ne stimeremo la forma ‘semi-strutturale’. Essa appare comunque di interesse. Infatti, sebbene la letteratura abbia mostrato che vi siano dei rendimenti ex-post dalla *performance* universitaria, che possono variare a seconda della disciplina studiata, non è altrettanto chiaro se gli studenti nella scelta del corso di laurea tengano conto della loro probabilità attesa di successo e dei redditi attesi. Si potrebbero verificare diverse situazioni. Gli individui potrebbero non tenere conto affatto delle prospettive occupazionali e/o reddituali che un corso di laurea offre e scegliere il corso di laurea privilegiando aspetti ‘non pecuniari’¹⁶ (come l’arricchimento culturale). Il test di questa ipotesi richiederebbe tuttavia la stima del modello strutturale (cioè quello che include il reddito atteso). Potrebbe accadere invece che gli individui tengano conto del rendimento economico prodotto da certi tipi di laurea, ma in maniera ‘naive’ non considerando cioè la probabilità attesa di successo, o che gli individui nella scelta della disciplina di studi tengano conto solo della loro probabilità di successo o infine sia della probabilità di successo che del reddito atteso dalla *performance* accademica. Un’implicazione interessante è che mentre nel primo e nel terzo e quarto caso l’adeguamento dell’offerta alla domanda di certe figure professionali per cui vi è scarsità in un’economia

¹⁵Vedi ad esempio Willis e Rosen (1979) o Berger (1988).

¹⁶Per l’importanza dei fattori ‘non pecuniari’ vedi Fiorito e Dauffenbach (1982).

(e pertanto alti rendimenti in termini di occupabilità e redditi) sarebbe più difficile da colmare, nel secondo caso gli individui risponderebbero molto più rapidamente alle esigenze espresse dal sistema produttivo ed ai segnali del mercato¹⁷.

4 L'indagine IPLAM

L'indagine sull'Inserimento Professionale dei Laureati degli Atenei Marchigiani (indagine IPLAM) si basa su un questionario somministrato nei mesi di novembre-dicembre 1997 e gennaio 1998 a tutti i laureati degli atenei marchigiani nell'anno solare 1992 e residenti nelle Marche al momento della laurea. Per informazioni dettagliate sulla natura dell'indagine e sui dati contenuti rimandiamo al capitolo introduttivo del volume curato da Staffolani e Sterlacchini (2001).

Il tasso di risposta ai questionari è stato del 57.2% in linea con le indagini sull'inserimento professionale svolte dall'Istat (1995) ed analoghe indagini svolte in altre regioni (vedi Santoro e Pisati 1996). Nel nostro caso specifico occorre precisare che il campione utilizzato nel presente studio è frutto di una doppia selezione che avviene in maniera sequenziale. Infatti al sottocampione dei rispondenti al 1997, 683 individui su una popolazione di 1575 laureati, sono stati inviati ulteriori questionari (in cui si chiedeva tra l'altro il voto di diploma). L'insieme dei rispondenti alla seconda ondata è stato di 543 laureati. Gli individui per cui tutte le informazioni utilizzate nel modello sono disponibili¹⁸ sono 528, che rappresenta allora la dimensione finale del campione utilizzato.

Tutte le analisi empiriche che seguiranno sono state fatte utilizzando dei pesi che esprimono l'inverso della probabilità che un individuo sia stato incluso nel campione¹⁹. Tale probabilità è stata stimata utilizzando un probit con selezione. Il probit principale si riferisce alla probabilità di rispondere ai quesiti aggiuntivi spediti nel 2000 mentre il probit che determina la selezione riguarda la probabilità di avere risposto alla prima ondata dell'indagine. Il maggiore vantaggio di stimare un probit con selezione è la possibilità di utilizzare dei regressori diversi nei due stadi della selezione e pertanto sfruttare le informazioni ricavate dalla prima ondata dell'indagine. Per esigenze di

¹⁷Almeno in termini di iscrizione a determinati corsi anche se ciò non implica necessariamente che il numero di laureati vari.

¹⁸Dal campione di 543 individui sono stati esclusi gli individui per cui il voto di diploma o le informazioni relative al tipo di laurea e facoltà non erano disponibili.

¹⁹Sull'utilizzo di pesi nelle stime che utilizzano indagini vedi ad esempio Magge, Rabb e Burbidge (1998).

brevità non riportiamo le stime del modello di selezione del campione, basti qui accennare che nel primo stadio sono stati inclusi, differenziati per genere, i seguenti regressori: l'ateneo, il corso di laurea, la provincia alla laurea ed il voto²⁰. Mentre nel secondo stadio sono stati inclusi il salario al 1997 per classi, la provincia al 1997, le ore settimanali lavorate al 1997 e la prole al 1997²¹.

Completata questa fase preliminare l'analisi empirica si articola come segue:

1. stima della forma ridotta dell'equazione di selezione attraverso un multinomial logit (MNL);
2. utilizzo degli *inverse Mill's ratios*, propriamente modificati per tenere conto della natura policotomica della scelta²², nel secondo stadio nella stima delle equazioni di *performance* per disciplina;
3. stima del voto atteso per disciplina e utilizzo della variabile per la stima della forma 'semi-strutturale' del modello di scelta della facoltà.

5 Strategia econometrica e variabili utilizzate

Ipotizzando che gli errori ϵ_{ij} in V_{ij} (vedi paragrafo 3) abbiano una distribuzione di tipo Gumbel e siano indipendenti tra le diverse alternative j il problema teorico di massimizzazione ha una controparte empirica nel modello multinomial logit²³. Nell'appendice 2 viene descritta più in dettaglio la procedura di stima.

La prima fase dell'analisi empirica consiste come già detto nella stima della forma ridotta del multinomial logit (MNL).

Omettiamo qui di riportare i risultati delle stime, che risultano del tutto strumentali alle stime successive e di scarso interesse dato che il nostro intento è quello di stimare la forma che abbiamo definito 'semi-strutturale' del MNL. La forma ridotta del modello prevede correttamente la scelta della facoltà di circa il 51% degli studenti (vedi tabella 1).

²⁰Una formulazione simile è utilizzata in Lucchetti e Venieri (2001).

²¹Si tratta di variabili che potrebbero influenzare la reperibilità degli individui o la loro attitudine a rispondere ai questionari, le stime sono disponibili su richiesta dall'autore.

²²Vedi l'appendice 2.

²³Questo risultato si deve a McFadden (1974). Per un'introduzione all'uso del modello logit in economia si veda Cramer (1991).

Tabella 1: Fit del MNL per la scelta della facoltà, forma ridotta

Casi Osservati	Casi previsti						totale
	ingegn.	med.-scient.	economico	letter.	giurid.	pol.-soc.	
ingegneria	38	8	5	1	1	0	53
med.-scient.	10	44	7	9	16	6	92
economico	14	23	83	3	15	2	140
letterario	1	19	12	46	7	5	90
giuridico	1	27	20	12	37	2	99
pol.-soc.	2	7	8	8	8	21	54
Totale	66	128	135	79	84	36	528

Come abbiamo già detto la stima della forma ‘semi-strutturale’ richiede che alcuni regressori presenti nelle equazioni di *performance* siano esclusi dalla scelta della facoltà.

In particolare, dalla stima delle equazioni di *performance* sono state escluse le variabili relative alla residenza, che influenzano i costi e pertanto la scelta della facoltà e con minore probabilità il voto di laurea. Inoltre mentre nelle equazioni di *performance* è stato inserito il voto di diploma differenziato per tipo di diploma, nella scelta della facoltà abbiamo incluso solo delle dummy per il tipo di diploma. Abbiamo fatto questa scelta poiché il voto di diploma probabilmente influenza la scelta di proseguire nell’istruzione, e, una volta fatta, è il tipo di diploma che potrebbe influenzare in maniera diretta la scelta del tipo di corso di laurea indipendentemente dalla *performance* nella scuola secondaria, mentre il voto per tipo di diploma esercita plausibilmente il suo effetto solo indirettamente per il tramite delle attese di successo (voto atteso) nei diversi corsi. Abbiamo preferito queste restrizioni di esclusione, dettate dal buon senso, a quelle normalmente utilizzate nella letteratura (vedi ad esempio Becker 1975, Willis e Rosen 1979, Berger 1988) per cui si ipotizza che le caratteristiche legate all’abilità influenzano il rendimento dell’istruzione e pertanto il salario atteso, mentre quelle relative al *background* socio-familiare influenzano le opportunità di istruzione. La ragione è che vi sono studi in cui si evidenzia che tramite meccanismi di *social-networking* l’origine socio-familiare influenza anche il rendimento dell’istruzione²⁴ (vedi ad esempio Montgomery 1991). Inoltre alla luce del modello del paragrafo 2 è altresì evidente che le variabili socio-familiari influenzano la *performance* accademica.

In breve, le equazioni di *performance* includono i seguenti regressori:

²⁴La presenza di eterogeneità nei rendimenti dell’istruzione per *background* socio-familiare rappresenta una assunzione standard nella recente letteratura sui rendimenti dell’istruzione (si veda Card 2000).

- genere: in quanto vi possono essere delle differenze tra maschi e femmine nell'utilità o disutilità derivante dallo studio di certe discipline, per cui l'input di impegno ($x_{i,j}$) o il tipo di abilità rilevanti per certi corsi potrebbe differire per genere;
- età: in virtù del fatto che gli studenti 'maturi' è più probabile che svolgono anche un'attività lavorativa ed hanno pertanto meno tempo da dedicare allo studio. Inoltre il modello teorico prevede che l'impegno dipende positivamente da T , la durata residua della vita attiva, per cui la *performance* dovrebbe ridursi con l'età;
- voto del diploma differenziato per tipo di diploma: evidentemente il livello ed il tipo di preparazione acquisita durante gli studi secondari superiori ($P_{i,k}$) potrebbero differenziare i risultati in termini di voto nel corso universitario²⁵;
- capitale culturale della famiglia. Si tratta di un indicatore del livello culturale della famiglia, al momento dell'iscrizione dello studente al corso di laurea, misurato nel seguente modo: capitale culturale 'eccellente' se entrambi i genitori hanno la laurea, 'alto' se almeno un genitore ha la laurea, 'medio' se almeno un genitore ha il diploma di maturità, 'basso' se entrambi i genitori hanno al massimo la licenza media. Tale misura è la stessa utilizzata in Santoro e Pisati (1996) e Staffolani e Sterlacchini (2001). Come abbiamo visto il capitale culturale potrebbe influenzare le preferenze degli individui, soprattutto l'utilità (o disutilità) dello studio. In generale ci aspettiamo che il 'gusto' per lo studio, e la *performance*, sia maggiore per individui con capitale culturale più elevato;
- capitale sociale della famiglia: è un indicatore della classe sociale della famiglia, al momento dell'iscrizione dello studente al corso di laurea, basato sul lavoro del genitore con professione socialmente più prestigiosa. Si tratta di una *proxy* del livello di risorse (trasferimenti) di cui lo studente può disporre dalla famiglia durante il periodo degli studi. In analogia con Staffolani e Sterlacchini (2001) abbiamo utilizzato le seguenti categorie: 1) imprenditori, liberi professionisti e dirigenti; 2) impiegati, insegnanti e quadri; 3) lavoratori autonomi (artigiani e commercianti) e coadiuvanti; 4) operai, lavoratori a domicilio 5) disoccupati. Diverse sono le ragioni per l'inclusione del capitale sociale che potrebbero influenzare le preferenze degli individui o il rendimento

²⁵Si veda Altonji (1993) per un modello sequenziale di scelte scolastiche.

della *performance*, differenziando pertanto l'impegno e la *performance* accademica tra individui con diverso background sociale. Il segno atteso è indeterminato se adottiamo le ipotesi verosimili, che individui con maggiore capitale sociale abbiano una maggiore preferenza per lo studio ma un minore rendimento della *performance* (per i già citati meccanismi di *social networking*).

La forma 'semi-strutturale' del multinomial logit include:

- genere: vi potrebbero essere delle differenze di genere nella scelta del tipo di occupazione e pertanto del tipo di istruzione universitaria²⁶;
- residenza: si tratta di un indicatore dei costi legati alla scelta di certi corsi (ϕ_j nel modello);
- età all'iscrizione: soggetti relativamente più anziani potrebbero privilegiare discipline con minore durata attesa, dato che quest'ultima influenza il periodo in cui essi possono fruire del rendimento dell'istruzione ($T - \tau_j$);
- capitale sociale e capitale culturale della famiglia: in quanto vi potrebbero essere degli effetti diretti del *background* socio-familiare sulla scelta del tipo di istruzione universitaria, oltre a quelli che passano attraverso le aspettative di *performance*. Individui provenienti da *background* meno agiati potrebbero privilegiare certi tipi di corso rispetto ad altri, ad esempio corsi con minore durata attesa e che pertanto richiedono un minore impegno finanziario (sia in termini di costi diretti che di *foregone earnings*);
- tipo di diploma: gli individui potrebbero percepire che certi tipi di istruzione secondaria hanno come sbocco 'naturale' determinati tipi di istruzione universitaria. L'effetto che non passa attraverso il voto atteso, che tiene conto dell'utilità della preparazione specifica ricevuta nella scuola superiore per il corso universitario, potrebbe essere interpretato come il ruolo di forze 'inerziali' o 'conformistiche' che spingono certi individui verso determinati percorsi scolastici;
- differenza tra il voto atteso dall'individuo nel gruppo di facoltà utilizzato come riferimento (quello rispetto a cui vengono normalizzati i coefficienti del multinomial logit al fine di ottenere l'identificazione degli stessi) ed il voto atteso negli altri gruppi di facoltà considerati.

²⁶Sulle differenze di genere nelle scelte occupazionali e conseguentemente nel tipo di studi universitari si veda ad esempio Blackmore e Low (1984).

La descrizione dettagliata delle variabili utilizzate è riportata nell'appendice 1.

6 Risultati delle equazioni di *performance* universitaria

In questa sezione descriviamo brevemente i risultati della stima delle equazioni di *performance* universitaria per disciplina.

I diversi tipi di laurea sono stati aggregati seguendo i criteri standard nella letteratura, vedi ad esempio ISTAT (1995) o MURST (1999), ed al tempo stesso cercando di garantire una numerosità sufficiente ai diversi gruppi:

- Gruppo ingegneria: ingegneria civile edile, ingegneria meccanica, ingegneria elettronica, ingegneria civile D.S.P.T.;
- Gruppo medico-scientifico: medicina e chirurgia, odontoiatria, fisica, matematica, farmacia, chimica, scienze naturali, scienze biologiche, scienze geologiche;
- Gruppo economico: economia e commercio;
- Gruppo letterario: lingue e letterature straniere, lettere, filosofia, pedagogia;
- Gruppo giuridico: giurisprudenza;
- Gruppo politico-sociale: scienze politiche, sociologia.

Per quanto concerne l'offerta dei diversi corsi il gruppo ingegneria è disponibile ad Ancona (AN), quello medico-scientifico ad Ancona (AN), Camerino (MC) ed Urbino (PS), quello economico ad Ancona (AN) ed Urbino (PS), quello letterario a Macerata (MC) ed Urbino (PS), quelli giuridico e politico-sociale a Camerino (MC), Macerata (MC) ed Urbino (PS).

I coefficienti stimati delle equazioni di *performance* sono riportati nella tabella 2.

1. **Gruppo ingegneria.** I maschi hanno uno svantaggio di circa quattro punti nel voto di laurea rispetto alle femmine. Non si nota un pattern tipico del voto al variare del livello di capitale sociale mentre il voto aumenta all'aumentare del capitale culturale della famiglia. Ciò può essere razionalizzato in termini di una maggiore preferenza allo studio per individui con capitale culturale più elevato. La differenza di voto

Tabella 2: Equazioni di *performance* universitaria

Variabili	Gruppo di laurea					
	ingegneria	medico-scientifico	economico	letterario	giuridico	politico-sociale
maschio	-3.97**	-0.36	0.35	2.51**	1.59**	-0.89**
etanini	2.06**	-0.07	-3.94**	0.33**	-0.12	0.01
misseta	41.30**	-2.26	-87.34**	6.35**	-8.21**	6.12**
sociale1	0.77**	-0.92	-1.24	-0.60**	0.35	-3.71**
sociale3	1.81**	0.36	0.30	-1.29**	3.07**	0.85
sociale4	0.49	5.64**	1.38	-0.60	4.25**	-2.31**
sociale5	1.82	6.90**	-7.94**	0.61	-1.47	2.90**
cultura1	6.17**	-3.67**	0.62	0.06	2.71**	0.17
cultura2	3.14**	5.25**	0.39	-2.03**	1.67**	-4.86**
cultura3	2.55**	4.02**	-1.02	-0.89**	3.78**	0.28
votod1	0.34**	0.31**	0.56**	0.10**	0.41**	0.02
votod2	0.33**	0.36**	0.46**	0.11**	0.45**	0.03
votod3	-	0.34**	0.48**	0.12**	0.33**	0.08**
votod4	-	0.34**	0.61**	0.11**	0.43**	0.03
votod5	0.31**	0.34**	0.44**	0.12**	0.33**	0.12**
votod6	-	0.57**	0.36**	0.08**	0.48**	0.10**
votod7	-	0.39**	0.42**	0.08**	0.47	0.04
j1	-2.79**	-	-	-	-	-
j2	-	1.53**	-	-	-	-
j3	-	-	6.40**	-	-	-
j4	-	-	-	-1.70**	-	-
j5	-	-	-	-	3.47**	-
j6	-	-	-	-	-	2.10**
N	53	92	140	90	99	54
R ²	51.50	31.66	41.89	30.18	44.30	40.82

Note. **: significativo al 5%. Gli standard error per le stime sono stati corretti per tenere conto che gli *inverse Mill's ratios* modificati sono stimati. La forma corretta della matrice varianza-covarianza è ricavata nel caso di selezione determinata da un probit da Heckman (1979) e Greene (1981) ed estesa da Greene (1998) al caso di selezione determinata da un multinomial logit.

tra il livello più basso e quello più alto di capitale culturale dello studente è di più di sei punti. I tipi di diploma di scuola superiore che forniscono la preparazione più adeguata per questo gruppo di corsi di laurea sono nell'ordine il liceo scientifico, quello classico e gli istituti tecnici industriale e per geometri. Nei primi due casi si tratta anche di diplomi che forniscono una preparazione di tipo generico, ovvero finalizzata alla continuazione degli studi e non al lavoro, che ha uno scarso valore sul mercato. Per cui il costo opportunità di studiare per individui con questi tipi di diploma è inferiore e ci aspettiamo che l'impegno allo studio sia più elevato.

2. **Gruppo medico-scientifico.** In questo caso si registra un effetto monotono negativo del capitale sociale sulla *performance*. Tale effetto

potrebbe essere spiegato dalla diversa motivazione allo studio di individui con diversa origine sociale. Infatti gli studenti provenienti dalle classi meno agiate e che hanno deciso di continuare gli studi potrebbero essere anche quelli più motivati. Inoltre tale effetto netto negativo potrebbe essere anche determinato dalla prevalenza dell'effetto negativo sull'impegno di studio del minore rendimento della *performance* per individui socialmente più agiati su quello positivo determinato da una maggiore preferenza per lo studio. Anche il capitale culturale ha un segno diverso da quello atteso dal modello teorico. Ma anche in questo caso l'effetto negativo può essere spiegato dalle diverse motivazioni ed opportunità nella continuazione degli studi, per cui gli individui provenienti da classi sociali più prestigiose potrebbero ricevere pressioni esterne verso il proseguimento degli studi indipendentemente dalle loro reali motivazioni, al contrario degli individui meno agiati, ciò che spiegherebbe il segno dell'effetto osservato. Curiosamente il gruppo di diplomi linguistico ed artistico è quello che garantisce la migliore *performance*, ma dato l'esiguo numero di questi individui (2 su un totale di 92 laureati) anche in questo caso l'effetto potrebbe essere spiegato in termini delle particolari caratteristiche degli individui che hanno scelto un tipo di laurea non in linea con la loro formazione scolastica precedente e rappresentare una mera *proxy* dell'abilità²⁷. Per questa ragione anche nel seguito commenteremo l'effetto soltanto di quei diplomi per cui il numero di casi è sufficiente da escludere che esso possa essere attribuito non tanto al tipo di diploma ma alle caratteristiche particolari di certi individui.

3. **Gruppo economico.** In questo caso soltanto gli studenti con un'origine sociale particolarmente disagiata (con entrambi i genitori che non lavorano) hanno uno svantaggio in termini di voto (di circa 8 punti rispetto agli altri) mentre il capitale culturale non ha alcun effetto sul voto di laurea. Tra i tipi di diploma quelli più utili sono nell'ordine quello di istituto tecnico commerciale e quello di liceo scientifico. La *performance* si riduce al crescere dell'età, come previsto dal modello teorico (T minore).

²⁷In questo caso tali categorie di diploma vanno considerate come delle *dummy* incluse per tenere conto di possibili *outliers*. Non si dimentichi poi che il nostro campione include solo i laureati, per cui gli studenti in possesso dei diversi tipi di diploma che osserviamo nel campione finale sono solo quelli che hanno completato gli studi. Se i tassi di abbandono per corsi di laurea fossero diversi a seconda del tipo di diploma posseduto l'effetto rilevato potrebbe essere spiegato dal fatto che soltanto individui straordinariamente capaci e motivati in possesso dei diplomi linguistico ed artistico conseguono una laurea nel gruppo medico-scientifico.

4. **Gruppo letterario.** I maschi hanno un vantaggio di 2 punti e mezzo nel voto di laurea. Anche in questo caso non si rileva un effetto monotono crescente del capitale sociale o di quello culturale, anzi spesso gli studenti con origine socio-culturale più bassa sono quelli che ricevono voti migliori. Anche in questo caso tuttavia il segno dell'effetto del capitale sociale e di quello culturale potrebbero risentire del problema della selezione, dato che la decisione di proseguire nell'istruzione universitaria e il livello di *drop-out* sono con elevata probabilità condizionati dall'origine socio-culturale degli individui. Non vi sono differenze sostanziali nel tipo di preparazione fornito dai diversi tipi di scuola secondaria.
5. **Gruppo giuridico.** I maschi hanno un vantaggio di circa un punto e mezzo nel voto di laurea. In questo caso l'effetto del capitale sociale è addirittura monotono decrescente, mentre quello del capitale culturale non è definito. Il diploma più utile è quello di liceo classico.
6. **Gruppo politico-sociale.** I maschi hanno uno svantaggio in termini di voto di meno di un punto. Anche in questo caso l'effetto del capitale sociale e di quello culturale non sono ben definiti. Il diploma più utile è quello magistrale.

È evidente dall'analisi che sebbene abbiamo corretto le equazioni di *performance* accademica per l'endogenità della scelta del tipo di gruppo di laurea, l'analisi risente comunque del fatto che la selezione degli individui nell'istruzione terziaria in termini sia di iscrizione all'università che in termini di *drop-out* potrebbe essere, e con elevata probabilità è, non casuale rispetto ai livelli di quelli che abbiamo definito capitale culturale e capitale sociale dello studente, per cui i coefficienti stimati per queste ultime variabili, anche se spesso non contraddicono le previsioni dell'analisi teorica non sono comunque soggetti ad un'interpretazione univoca²⁸. Inoltre occorre osservare che soprattutto per il capitale sociale, inteso come *proxy* del livello di risorse a disposizione degli studenti per gli studi, non abbiamo informazioni su eventuali borse di studio di cui gli studenti meno agiati ma più capaci potrebbero fruire.

²⁸A prescindere dalla non disponibilità dei dati individuali sul proseguimento nell'istruzione universitaria si tratterebbe comunque di un modello sequenziale di selezione, dove la prima decisione è quella bivariata di continuare negli studi e la seconda quella multivariata del gruppo di laurea, difficile da implementare dal punto di vista econometrico (infatti per quanto a nostra conoscenza la totalità di modelli implementati in letteratura prevede due decisioni sequenziali di tipo binario, vedi Maddala 1986).

Ai fini della nostra analisi ciò rappresenta tuttavia solo una minore preoccupazione dato che le equazioni di *performance* sono state stimate solo per ottenere i voti previsti nei diversi gruppi di laurea, necessari per la stima del modello ‘semi-strutturale’ e la validità delle previsioni, condizionate alla decisione di proseguire nell’istruzione universitaria e portare a termine gli studi, rimane sebbene i coefficienti siano difficilmente interpretabili alla luce del modello teorico illustrato a causa del suddetto problema di selezione.

Da notare che il ‘regressore di selezione’ (i j) risulta significativo in tutti i gruppi di lauree considerati, ciò a riprova del fatto che esiste effettivamente un problema di endogeneità nella scelta del tipo di laurea.

7 Stima del modello ‘semi-strutturale’

La stima del modello ‘strutturale’ richiederebbe una quantità di informazioni di gran lunga superiore a quella di cui disponiamo, tra cui di particolare interesse sarebbero i redditi familiari e quelli provenienti da borse di studio, per distinguere l’effetto dei vincoli economici dalle pressioni di carattere sociale, ed i redditi puntuali dei laureati per la stima delle funzioni di reddito. In mancanza di questi dati nell’analisi empirica non testeremo l’ipotesi che gli individui nella scelta della facoltà considerano i redditi attesi (che risultano funzione della *performance* attesa), ma quella che essi tengono conto nel determinare tale scelta della loro probabilità di successo (*performance* attesa). La stima della forma ‘semi-strutturale’ appare comunque di interesse. Dato che tale forma è ottenuta sostituendo la *earnings function* nelle funzioni di utilità vitale relative alle diverse discipline i coefficienti delle variabili presenti in entrambi i modelli, quello di *performance* e quello di scelta della facoltà, risultano difficili da interpretare, dato che riassumono gli effetti dei regressori sui due modelli. Così ad esempio i coefficienti relativi al tipo di occupazione dei genitori stimati nella forma ‘semi-strutturale’ potrebbero riflettere effetti di pressione sociale o differenze nelle preferenze, oppure differenze nel rendimento dell’istruzione. Il nostro modello non consente di distinguere gli effetti cosiddetti ‘pecuniari’ da quelli ‘non pecuniari’. Una implicazione è che soltanto i regressori esclusi da uno dei due modelli, quelli che determinano l’identificazione, possono essere interpretati alla luce del modello teorico. Essi sono nel modello ‘semi-strutturale’ la residenza, il tipo di diploma, ed il voto atteso di laurea. Per questa ragione ci concentreremo nel commento su di essi.

Un’altra precisazione necessaria è che nel modello ‘semi-strutturale’ è stato inserito come regressore la differenza tra il voto atteso di laurea in una determinata facoltà e quello atteso nella disciplina rispetto alla quale i coeffi-

cienti del multinomial logit sono normalizzati ²⁹ (il gruppo ‘politico-sociale’). Ciò implica che qualora il rendimento della *performance* fosse uguale in tutte le discipline le differenze di voto rifletterebbero anche differenze di reddito e pertanto il relativo coefficiente può essere considerato come l’effetto delle differenze nel reddito atteso nelle diverse discipline. In caso di eterogeneità nei rendimenti della *performance* tra discipline³⁰ questa corrispondenza non è più assicurata ed il coefficiente va interpretato semplicemente come l’effetto delle aspettative di successo nei diversi corsi di laurea, che rimane comunque di interesse. La tabella 3 mostra i coefficienti stimati dal MNL ‘semi-strutturale’ e la tabella 4 l’effetto dei regressori che abbiamo considerato di interesse data la presunta univocità, alla luce del modello, dell’interpretazione dell’effetto stimato. Da notare che il voto atteso ha il segno previsto dalla teoria: una riduzione del *gap* di voto di un gruppo rispetto a quello politico-sociale aumenta la probabilità di iscriversi al primo.

Le probabilità di iscriversi alle diverse facoltà se tutti gli individui risiedessero in provincia di Ancona sarebbero: 13.6% gruppo ingegneria, 19.2% gruppo medico-scientifico, 20.1% gruppo economico, 13% gruppo letterario, 21.1% gruppo giuridico e 13.1% gruppo politico-sociale. Se tutti gli individui risiedessero in provincia di Macerata la probabilità di iscriversi al gruppo ingegneria si ridurrebbe a 5.6% ed al gruppo economico a 15.1%. Se tutti gli individui risiedessero in provincia di Pesaro la probabilità di iscriversi al gruppo ingegneria diverrebbe 7.1%. Tale evidenza sembra avvalorare l’ipotesi che la distanza della residenza degli studenti (prima dell’iscrizione) dagli atenei che offrono certi tipi di corso, che incide ovviamente sui costi dell’istruzione, influenza la scelta della facoltà.

Per quanto concerne il tipo di diploma di scuola secondaria superiore, se tutti gli studenti avessero il diploma di liceo scientifico (a parità di voto atteso nelle diverse facoltà) la probabilità di iscrizione al gruppo ingegneria sarebbe del 7.6%, al gruppo medico-scientifico del 33.4%, a quello economico 21.3%, a quello letterario 9.6%, a quello giuridico 21% ed a quello politico-sociale 7.2%. Se tutti gli individui avessero il liceo classico la probabilità di iscriversi al gruppo ingegneria aumenterebbe di poco, 7.7%, quella di iscriversi al gruppo medico-scientifico e a quello economico scenderebbe notevolmente, rispettivamente a 20.8% e 10.2%, mentre quelle di iscriversi ai gruppi let-

²⁹ Ciò è stato fatto in quanto considerando solo il voto atteso nelle discipline si avrebbe l’indesiderabile effetto che variazioni nel voto atteso nella facoltà di riferimento non determinerebbero alcuna variazione della probabilità di iscriversi a questa ed alle altre facoltà considerate.

³⁰ A nostra conoscenza, complice la carenza di dati, non esistono in Italia degli studi sulle differenze nei rendimenti economici della *performance* universitaria (i coefficienti α_{2j} nell’equazione 13 tra corsi di laurea).

Tabella 3: MNL per la scelta della facoltà, forma ‘semi-strutturale’

Variabili	Gruppo di laurea				
	ingegneria	medico-scientifico	economico	letterario	giuridico
maschio	4.17**	-1.11**	0.32	-2.39**	-0.79*
resid1	-31.81*	0.58	-38.04**	1.50	-0.88
resid3	1.88**	1.14	2.29**	1.18	1.59**
resid6	-1.33*	0.30	-0.16	0.39	0.39
resid9	-1.24*	-0.14	-0.09	0.54	-0.76
resido	-0.64	-1.41	-0.08	0.79	-0.54
etaini	-2.08**	-0.06	-0.53*	-0.17**	-0.17**
misseta	-36.52**	-0.07	-9.02	-2.09	-1.67
sociale1	-3.71**	-1.13	-0.33	-1.04	-0.17
sociale3	-0.98	0.03	-0.45	-0.03	-0.01
sociale4	-2.04**	-2.24**	-0.90	-0.66	-0.67
sociale5	-1.29	-0.94	-0.51	-0.29	1.16
cultura1	-0.92	2.80	1.29	0.95	2.43
cultura2	-5.81**	-1.27	0.44	1.42	1.41
cultura3	-1.70**	-0.84	0.14	-0.25	-0.08
dipl2	-0.61	-0.98	-1.35*	0.23	0.01
dipl3	-25.33**	-3.66**	-4.01**	-0.48	-1.50**
dipl4	-26.38**	-4.23**	-0.24**	-0.75	-1.09*
dipl5	4.51**	1.38	0.63	0.69	0.38
dipl6	-26.27**	-3.58**	-2.02	0.57	-1.36
dipl7	6.69**	-1.63**	-1.05	-0.64	-2.74**
evoto1	-0.72**	-	-	-	-
evoto2	-	-0.21**	-	-	-
evoto3	-	-	-0.11**	-	-
evoto4	-	-	-	-0.32	-
evoto5	-	-	-	-	-0.02
N	528				
Pseudo R^2	30.42				

Note: i coefficienti con due asterischi sono significativi al 5%, quelli con un asterisco al 10%.

terario (17.5%), giuridico (33%) e politico-sociale (10.8%) aumenterebbero. Mentre se tutti gli studenti provenissero dalle scuole magistrali la probabilità di iscriversi ai gruppi ingegneria, medico-scientifico ed economia risulterebbe fortemente abbattuta a favore dei gruppi letterario (28%), giuridico (29.7%) e politico-sociale (32.8%). Qualora tutti gli studenti provenissero da istituti tecnici commerciali la probabilità di iscriversi al gruppo economico più che raddoppierebbe mentre quella cumulata di iscriversi ai gruppi ingegneria e medico-scientifico sarebbe inferiore al 2%. Nel caso in cui tutti gli studenti fossero in possesso del diploma di istituto tecnico industriale o per geometri sarebbe soprattutto la probabilità del gruppo ingegneria a crescere (che risulta quasi quintuplicata), mentre se avessero diplomi di tipo linguistico o artistico è la probabilità di iscriversi al gruppo letterario che subirebbe un forte incremento (passando a 41.5%). Infine se tutti avessero diplomi di istituto professionale, altro tecnico o altri è la probabilità di iscriversi al gruppo

ingegneria che aumenterebbe fortemente passando a 56.4%. Ciò evidenzia che indipendentemente dalle aspettative di voto la provenienza scolastica ha un forte impatto sulle decisioni di iscriversi a certi corsi, e che elementi che possiamo definire di tipo ‘inerziale’, in quanto dipendono dalle scelte fatte in passato, risultano importanti nella scelta della facoltà.

Di seguito viene fatto un esperimento simile relativamente ai voti attesi nelle diverse discipline. Per il complesso degli studenti la probabilità media di iscriversi ai diversi gruppi è del 10.6% al gruppo ingegneria, 19.8% a quello medico-scientifico, 21.9% a quello economico, 15.6% a quello letterario, 20.7% a quello giuridico e 11.5% a quello politico-sociale. Se tutti gli studenti avessero la media dei voti attesi del gruppo ingegneria le probabilità suddette diverrebbero 15.4%, 19%, 20.1%, 14.5%, 20.9% e 10.2%. Degli effetti simili si rilevano anche attribuendo alla variabile voto atteso le media degli altri gruppi: in tutti i casi la probabilità di iscriversi al gruppo considerato aumenta.

Tabella 4: Probabilità previste dal MNL, forma ‘semi-strutturale’

Caratteristiche	Gruppo laurea					
	ingegn.	med.-scient.	economico	letter.	giurid.	pol.-soc.
resid3	0.15	0.14	0.32	0.12	0.23	0.04
resid6	0.06	0.23	0.18	0.15	0.27	0.11
resid9	0.07	0.20	0.25	0.22	0.12	0.14
ddipl1	0.08	0.33	0.21	0.10	0.21	0.07
ddipl3	0.08	0.21	0.10	0.18	0.33	0.11
ddipl4	0.00	0.06	0.04	0.28	0.30	0.33
ddipl5	0.00	0.02	0.44	0.14	0.22	0.18
ddipl6	0.32	0.39	0.11	0.06	0.09	0.03
ddipl7	0.56	0.12	0.10	0.09	0.02	0.10
evoto	0.11	0.20	0.22	0.16	0.21	0.11
evoto (ing.)	0.15	0.19	0.20	0.15	0.21	0.10
evoto (med.-sc.)	0.16	0.20	0.13	0.16	0.24	0.12
evoto (econ.)	0.04	0.25	0.26	0.15	0.21	0.10
evoto (lett.)	0.02	0.25	0.17	0.17	0.25	0.14
evoto (giur.)	0.12	0.23	0.15	0.16	0.23	0.11
evoto (pol.-soc.)	0.04	0.25	0.07	0.19	0.29	0.16
etaini 19	0.15	0.18	0.22	0.15	0.19	0.10
etaini 25	0.00	0.35	0.04	0.13	0.21	0.26

Note: la tabella mostra la variazione nella probabilità prevista di iscriversi ai diversi gruppi di lauree cambiando alcune caratteristiche del campione. Le probabilità potrebbero non sommare ad uno per effetto degli arrotondamenti.

Abbiamo infine considerato l’effetto di un aumento dell’età di tutti gli

studenti dalla media all'inizio degli studi, 19 anni, a 25 anni: in questo caso è la probabilità di iscriversi al gruppo ingegneria ed economia a ridursi notevolmente a vantaggio di quella di iscriversi soprattutto ai gruppi politico-sociale e medico-scientifico. Questi risultati sono coerenti con il modello teorico: all'aumentare dell'età gli studenti tendono a parità di risultati attesi ad iscriversi ai corsi con durata attesa inferiore, dato che potranno fruire del rendimento dell'istruzione per un periodo più lungo. Infatti la durata media dei corsi di studio è 8.38 anni per il gruppo ingegneria, 6.54 per quello medico-scientifico, 7.25 per quello economico, 6.94 per quello letterario, 6.48 per quello giuridico e 5.78 per quello politico-sociale.

8 Conclusioni

In questo paper è stato svolto uno studio della *performance* accademica e della scelta della facoltà universitaria a cui iscriversi. L'analisi evidenzia che il background socio-familiare è in grado di influenzare la *performance* accademica degli studenti e che sia variabili di tipo *forward looking* che fattori 'inerziali', legati a decisioni prese in passato, influenzano la scelta della facoltà. Tra gli elementi 'inerziali' non-pecuniari vanno citati il tipo di diploma di scuola superiore conseguito, che influenza la scelta del corso indipendentemente dalle aspettative di successo, tra quelli *forward looking* invece il voto atteso nei diversi gruppi di laurea. Questo suggerisce che gli individui quando devono scegliere il percorso universitario tengono conto delle loro possibilità di successo nei diversi corsi, che dipendono tra l'altro dalla *performance* nella scuola superiore, ed inoltre sono mossi da fattori 'inerziali', come il tipo di scuola superiore frequentato, che li spingono verso determinati percorsi accademici. Ciò implica che anche se nell'economia si presentasse un eccesso di domanda sull'offerta di certe figure di laureati, un esempio tra tutti quelli legati al settore IT di cui tanto si parla, e conseguentemente un forte aumento dei rendimenti economici per i laureati nel gruppo ingegneria, l'adeguamento spontaneo della domanda all'offerta tenderebbe ad essere piuttosto lento. Se le autorità di *policy* volessero cercare di influenzare l'offerta di laureati in certe discipline dovrebbero agire allora nei cicli iniziali dell'istruzione, influenzando la scelta e la *performance* nella scuola secondaria superiore. Per colmare l'eventuale gap sarebbero allora necessari degli interventi lungimiranti della autorità di politica economica che dovrebbero avere come target i cicli iniziali del percorso formativo. In questo contesto, indipendentemente dall'elevamento dell'età dell'obbligo scolastico, la realizzazione di un sistema di scuola secondaria superiore omogeneo negli anni iniziali e che ritarda

pertanto l'età della scelta della specializzazione, la quale presumibilmente condizionerà anche il tipo di studi universitari intrapresi, appare adeguata.

Riferimenti bibliografici

- [1] Altonji, J.G. (1993), 'The Demand for and Return to Education When Education Outcomes Are Uncertain', *Journal of Labor Economics*, 11: 48-83.
- [2] Becker, G. (1975), *Human Capital*, Second Edition, Chicago, The University of Chicago Press.
- [3] Berger, M.C. (1988), 'Predicted Future Earnings and Choice of College Major', *Industrial and Labor Relations Review*, 41: 417-429.
- [4] Blakmore, A.E. and Low, S.A. (1984), 'Sex Differences in Occupational Selection: The Case of College Majors', *Review of Economics and Statistics*, 66: 157-163.
- [5] Card, D. (2000), 'Estimating the Return to Schooling: Progress on Some Persistent Econometrics Problems', *NBER Working Paper*, n. 7769.
- [6] Checchi, D. (2000), 'University Education in Italy', *International Journal of Manpower*, 21: 177-205.
- [7] Cobalti, A. e Schizzerotto, A. (1993), 'Inequality of Educational Opportunity in Italy' in Shavit Y. e Blossfeld H.-P. (ed.), *Persistent inequality. Changing Educational Attainment in Thirteen Countries*, Boulder, Westview Press.
- [8] Cramer, J.S. (1991), *The Logit Model for Economists*, Londra, Edward Arnold.
- [9] Dolton, P. e Makepeace, G.H. (1990), 'The Earnings of Economic Graduates', *Economic Journal*, 107: 710-726.
- [10] Fiorito, J. e Dauffenbach, R.C. (1982), 'Market and Nonmarket Influences on Curriculum Choice by College Students', *Industrial and Labor Relations Review*, 36: 88-101.
- [11] Gambetta, D. (1987), *Were they pushed or did they jump? Individual decision mechanism in education*, Cambridge, Cambridge University Press.
- [12] Greene, W.H. (1981), 'Sample Selection Bias as a Specification Error: Comment', *Econometrica*, 49: 795-798.

- [13] Greene, W.H. (1998), *LIMDEP Version 7.0 User's Manual*, Econometric Software Inc.
- [14] Hay, J. (1980), 'Selectivity Bias in a Simultaneous Logit-OLS Model: Physician Specialty Choice and Specialty Income', Manoscritto, University of Connecticut Health Center.
- [15] Heckman, J.J. (1979), 'Sample Selection Bias as a Specification Error', *Econometrica*, 47: 153-161
- [16] Heckman, J.J. (1999), 'Policies to Foster Human Capital', *NBER Working Paper* n. 7288.
- [17] Istat (1995), *Inserimento professionale dei laureati. Indagine 1995*, Roma, Istat.
- [18] Lee, L.F., (1978), 'Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables', *International Economic Review*, 19: 415-433.
- [19] Lee, L.F. (1983), 'Generalized Econometric Models with Selectivity', *Econometrica*, 51: 507-512
- [20] Lucchetti, R. e Venieri, C. (2001), 'La metodologia dell'indagine' in (a cura di) Staffolani S. e Sterlacchini A. (2001).
- [21] Maddala, G.S. (1986), 'Disequilibrium, self selection, and switching models', in Z. Griliches e M.D. Intriligator (ed.), *Handbook of Econometrics*, vol. III, 1671-1688.
- [22] Magee, L., Robb, A. L. and Burbidge, J.B. (1998), 'On the Use of Sampling Weights When Estimating Regressions Models with Survey Data', *Journal of Econometrics*, 84: 251-271.
- [23] McFadden, D. (1974), 'Analysis of Qualitative Choice Behaviour' in Paul Zarembka (ed.), *Econometrics*, New York, Academic Press.
- [24] Montmarquette, C., Cannings, K. e Mahseredjian, S. (1997), 'How Do Young People Choose College Majors?', *Cirano*, Scientific Series n.97s-38.
- [25] MURST (1999), *Valutazione dell'efficacia dell'istruzione universitaria rispetto al mercato del lavoro*, Ministero dell'Università e della Ricerca Scientifica e Tecnologica, Osservatorio per la valutazione del sistema universitario.

- [26] Roy, A.D. (1951), ‘Some Thoughts on the Distribution of Earnings’, *Oxford Economic Papers*, 3: 135-146.
- [27] Santoro, M. e M. Pisati (1996), *Dopo la laurea: Status, sfide e strategie*, Bologna, il Mulino.
- [28] Smith, J. e Naylor, R.A. e Mcknight, A. (2000), ‘Occupational Earnings of Graduates. Evidence for the 1993 University Population from the USR’, mimeo, University of Warwick.
- [29] Staffolani, S. e Sterlacchini, A. (2001), *Istruzione universitaria, occupazione e reddito. Un’analisi delle condizioni professionali e dei redditi dei laureati marchigiani*, Milano, Franco Angeli.
- [30] Trost, R.P. e Lee, L.F. (1984), ‘Technical Training and Earnings: A Polychotomous Choice Model with Selectivity’, *Journal of Economics and Statistics*, 66: 151-156.
- [31] Willis, R.J. e Rosen, S. (1979), ‘Education and Self-Selection’, *Journal of Political Economy*, 87: S7-S36.

APPENDICI

1 Elenco delle variabili utilizzate

maschio	dummy per studente maschio
femmina	dummy per studente femmina (categoria di riferimento)
resid1	residenza prima dell'iscrizione omessa
resid2	residenza prima dell'iscrizione in provincia di Ancona (AN, gruppo di riferimento)
resid3	residenza prima dell'iscrizione in provincia di Ascoli Piceno (AP)
resid6	residenza prima dell'iscrizione in provincia di Macerata (MC)
resid9	residenza prima dell'iscrizione in provincia di Pesaro (PS)
resido	residenza prima dell'iscrizione fuori regione
etaini	età iniziale
misseta	dummy per età iniziale omessa
cultura1	capitale culturale 'eccellente' (vedi paragrafo 5)
cultura2	capitale culturale 'elevato'
cultura3	capitale culturale 'medio'
cultura4	capitale culturale 'basso' (categoria di riferimento)
sociale1	il genitore con occupazione socialmente più prestigiosa è imprenditore, libero professionista e dirigente
sociale2	il genitore con occupazione socialmente più prestigiosa è impiegato, insegnante o quadro (categoria di riferimento)
sociale3	il genitore con occupazione socialmente più prestigiosa è un lavoratore autonomo o coadiuvante
sociale4	il genitore con occupazione socialmente più prestigiosa è operaio o lavorante a domicilio
sociale5	entrambi i genitori sono disoccupati
dipl1	diploma di liceo scientifico (gruppo di riferimento)
dipl2	diploma di liceo classico
dipl3	diploma di istituto magistrale
dipl4	diploma di istituto tecnico commerciale
dipl5	diploma di istituto tecnico industriale o istituto tecnico per geometri
dipl6	diploma di liceo linguistico o artistico
dipl7	diploma di istituto professionale, altro istituto tecnico o altro
votom1	voto in dipl1
votom2	voto in dipl2
votom3	voto in dipl3
votom4	voto in dipl4
votom5	voto in dipl5
votom6	voto in dipl6
votom7	voto in dipl7
evoto1	differenza tra i voti attesi nei gruppi politico-sociale e ingegneria
evoto2	differenza tra i voti attesi nei gruppi politico-sociale e medico-scientifico
evoto3	differenza tra i voti attesi nei gruppi politico-sociale e economico
evoto4	differenza tra i voti attesi nei gruppi politico-sociale e letterario
evoto5	differenza tra i voti attesi nei gruppi politico-sociale e giuridico
j1	<i>inverse Mill's ratio</i> trasformato per la probabilità di iscriversi al gruppo ingegneria
j2	<i>inverse Mill's ratio</i> trasformato per la probabilità di iscriversi al gruppo medico-scientifico
j3	<i>inverse Mill's ratio</i> trasformato per la probabilità di iscriversi al gruppo economico
j4	<i>inverse Mill's ratio</i> trasformato per la probabilità di iscriversi al gruppo letterario
j5	<i>inverse Mill's ratio</i> trasformato per la probabilità di iscriversi al gruppo giuridico
j6	<i>inverse Mill's ratio</i> trasformato per la probabilità di iscriversi al gruppo politico-sociale.

2 Un modello policotomico di selezione

Per completezza riportiamo qui i principali risultati ottenuti da Lee (1983). Nel modello di selezione della facoltà l'individuo deve scegliere tra J ($j = 1, \dots, J$) facoltà. Abbiamo già visto nel paragrafo 2 che l'utilità indiretta connessa alla scelta della facoltà j può essere espressa come:

$$V_{ij} = X_i\beta_{1j} + y_{ij}\beta_{2j} + \epsilon_{ij} \quad (a)$$

Esprimendo la stessa funzione in forma ridotta, sostituendo cioè al reddito ed alla *performance* le relative espressioni indicate nel paragrafo 2 abbiamo:

$$V_{ij} = \Gamma_i\theta_j + \omega_{ij} \quad (b)$$

dove Γ_i è un insieme di regressori che include tutte le covariate incluse in Z_i o in X_i .

L'individuo i sceglierà l'alternativa s se e solo se tale alternativa offre la maggiore utilità indiretta tra tutte quelle possibili, ovvero $V_{is} > V_{ij}$ per $j = 1, \dots, J$ e $j \neq s$. La probabilità che l'alternativa s sia scelta è data da:

$$P_{is} = \text{prob}(V_{is} > \max V_{ij} | j = 1, \dots, J; j \neq s) \quad (c)$$

Definiamo ora $\eta_{is} = \max (V_{ij} - \omega_{ij})$, per $j = 1, \dots, J$ e $j \neq s$. Ne consegue che l'alternativa s è scelta se e solo se $\Gamma_i\theta_s > \eta_{is}$. Se ipotizziamo che le componenti stocastiche ω_{ij} siano identicamente indipendentemente distribuite in modo Gumbel abbiamo che la distribuzione $F(\eta_{is})$ di η_{is} è:

$$F(\eta_{is}) = \frac{\exp(\eta_{is})}{\exp(\eta_{is}) + \sum_{j=1, j \neq s}^J \exp(\Gamma_i\theta_j)} \quad (d)$$

e la probabilità che venga scelta l'alternativa s è:

$$P_s = \frac{\exp(\Gamma_i\theta_s)}{\sum_{j=1}^J \exp(\Gamma_i\theta_j)}. \quad (e)$$

Si tratta evidentemente del modello multinomial logit (MNL) di McFadden (1974). Indichiamo con Φ la funzione di distribuzione normale standard e con ϕ la funzione di densità normale standard. La trasformazione $G = \Phi^{-1}(F)$ è strettamente crescente e la variabile casuale trasformata $\eta_{is}^* = G(\eta_{is})$ è normale standard. Dato che G è una trasformazione strettamente crescente ne deriva che l'alternativa s è scelta se e solo se $G(\Gamma_i\theta_s) > \eta_{is}^*$. Se ipotizziamo che la componente stocastica u_{ij} nell'equazione di *performance* sia normale possiamo altresì assumere la normalità congiunta di u_{ij} e

η_{is}^* . In questo caso Lee (1983) ha mostrato che condizionatamente alla scelta dell'alternativa s risulta:

$$P_{is} = Z_i\gamma_s - \sigma_s\rho_s\phi(G(\Gamma_i\theta_s))/F(\Gamma_i\theta_s) + \xi_{is} \quad (f)$$

dove σ_s è la deviazione standard dell'errore u_{is} e ρ_s la correlazione tra u_{is} ed η_{is}^* . La relazione (f) può essere stimata in maniera consistente in due stadi. Nel primo stadio viene stimato il modello MNL in forma ridotta, dopo di che si stima il regressore $\hat{J}_{is} = \phi(G(\Gamma_i\hat{\theta}_s))/F(\Gamma_i\hat{\theta}_s)$, che chiameremo regressore di selezione o *inverse Mill's ratio* trasformato, che viene utilizzato nel secondo stadio per la stima delle equazioni di *performance* universitaria.

Stimate le equazioni di *performance* immuni dall'effetto del *selection bias* è possibile ottenere per ogni individuo i voti attesi nei diversi gruppi di laurea. Questi voti attesi vengono poi utilizzati in un terzo stadio per la stima della forma 'semi-strutturale' del modello MNL.