

Analisi dei tempi di inserimento lavorativo dei laureati dell'Università di Bari

Francesco Delvecchio, Francesco d'Ovidio¹

Dipartimento di Scienze Statistiche - Università degli studi di Bari

Riassunto. I tempi di inserimento lavorativo dei laureati sono influenzati sia da fattori legati alle particolari richieste di questo (esperienze conseguite, flessibilità, ecc.), sia da caratteristiche proprie dei laureati stessi, come la capacità di apprendere, la disponibilità alla mobilità lavorativa (almeno nel Mezzogiorno, ove la richiesta di laureati ed in genere di forza-lavoro è tuttora meno consistente che in altre aree). Nella presente nota viene compiuta, con metodi statistici multivariati, una verifica dell'effettiva influenza dei fattori legati al lato dell'offerta di lavoro (le caratteristiche dei laureati dell'Università di Bari) sulla maggiore o minore rapidità di inserimento lavorativo. Volendo tener conto anche delle diverse opportunità offerte da settori disciplinari ben diversi fra loro, si è fatto ricorso anche a modelli multilivello, considerando come unità di secondo livello le Facoltà di provenienza dei laureati.

Parole chiave: Tempi di inserimento, Laureati, Università di Bari, Regressione categoriale, Analisi multilivello.

1. Introduzione

In un tempo non troppo lontano, chi si laureava con buoni voti (in genere sinonimo di buona preparazione) aveva di fronte a sé un certo ventaglio di scelte: insegnamento, libera professione, una buona posizione aziendale, ...spesso, anzi, erano le aziende che si attivavano presso le università per attrarre i laureati più promettenti. Tutte possibilità che, con l'andare del tempo, non sono svanite, ma sono generalmente offerte a contingenti di laureati molto meno cospicui di un tempo. Si possono ipotizzare motivazioni disparate per tale fenomeno: la tendenza alla globalizzazione dei mer-

¹ La presente nota è stata realizzata nell'ambito del progetto "Transizioni Università-Lavoro e valorizzazione delle competenze professionali dei laureati: modelli e metodi di analisi multidimensionale delle determinanti", cofinanziato dal MIUR; coordinatore nazionale è Luigi Fabbris, coordinatore del gruppo di Bari è Francesco Delvecchio. Della nota, opera congiunta dei due autori, va attribuita a F. Delvecchio la redazione finale dei paragrafi 1, 2 e 6, a F. d'Ovidio quella dei paragrafi 3, 4 e 5.

cati, la managerialità delle gestioni, ma anche la volontà (e talora la necessità) delle imprese di ridurre i costi del lavoro con la pratica dell'*under-employment*, che, nonostante talune incongruenze fra le competenze fornite dal sistema formativo universitario e quelle realmente necessarie in ambito aziendale, comunque assicura personale potenzialmente qualificato.

D'altra parte, le posizioni lavorative tradizionalmente appannaggio dei laureati costituiscono una quota minoritaria nel complesso degli addetti delle imprese e degli enti in Italia. È quindi logico che una parte dei laureati venga inserita in posizioni lavorative inferiori a quelle tipiche per un laureato, di tipo genericamente impiegatizio. Questo "accomodamento verso il basso" del mercato del lavoro, benché comune a tutte le realtà economiche italiane (ed anche estere), è più frequente nel Mezzogiorno, la cui dinamica economica è in genere meno favorevole che altrove. In detto territorio è anche maggiore la quota di laureati che non trovano un'occupazione per un tempo abbastanza rilevante dopo il conseguimento del titolo, o che si devono accontentare di forme di flessibilità lavorativa molto penalizzanti: contratti a tempo determinato inframmezzati da periodi di disoccupazione per aggirare la normativa vigente (la cosiddetta Legge Biagi), ambigue situazioni di lavoro autonomo (come quelle degli agenti di commercio mono-mandatari o dei "promotori finanziari" legati a istituti di credito o di *placement* finanziario), eccetera.

Per i laureati del Mezzogiorno più che per altri, pertanto, si potrebbe ipotizzare che un fattore importante per l'acquisizione di una occupazione in tempi brevi sia appunto la disponibilità alla flessibilità lavorativa, unitamente alla disponibilità alla mobilità territoriale, in quanto è noto che nelle regioni settentrionali vi è una minore disoccupazione. È in questo filone che si pone il presente contributo, che vuole verificare con metodologie statistiche multivariate quali siano, fra le caratteristiche dei laureati dell'Università di Bari presenti sul mercato del lavoro, quelle che risultano significativamente legate ai minori tempi di attesa. La base di partenza è un'indagine campionaria incentrata sui laureati dell'Ateneo barese, e per motivi di disponibilità di dati, si tralascia qui l'analisi della domanda da parte di imprese e istituzioni, così come di altri fattori estranei rispetto alle competenze ed alle scelte dei laureati.

2. I tempi di inserimento lavorativo dei laureati dell'Università di Bari

Partendo da una base di dati amministrativi relativi ai laureati dell'Ateneo barese negli anni dal 1995 al 2000 (elaborata dal locale Centro Servizi Informativi), nel periodo compreso fra dicembre 2003 ed aprile 2004 è stata effettuata una indagine campionaria telefonica volta a verificare, dopo almeno tre anni dal conseguimento del ti-

tolo, la situazione lavorativa dei medesimi. Per i particolari sulle modalità di indagine si rimanda ad altri contributi già pubblicati (Delvecchio e d'Ovidio, 2004; Crocetta, d'Ovidio e Toma, 2005). Le Facoltà dell'Università di Bari attive nel periodo erano: Agraria, Economia, Farmacia, Giurisprudenza, Lettere e Filosofia, Lingue e letterature straniere, Medicina e Chirurgia, Medicina veterinaria, Scienze della formazione, Scienze matematiche fisiche e naturali, Scienze politiche. Dalla popolazione di laureati provenienti da dette Facoltà è stato estratto, con procedimento casuale, un campione stratificato per corso di laurea e per genere, la cui numerosità, a valle dei controlli di coerenza sulle interviste effettuate, è risultata pari a 2.785 unità, inferiore al previsto² ma comunque sufficientemente rappresentativa dell'universo.

Purtroppo, la distribuzione per singolo corso di laurea, ben articolata a fini descrittivi, determina alcuni subcampioni con numerosità singolarmente troppo esigue per consentire inferenze affidabili. Per questo motivo, nella presente nota ci si riferirà ad aggregazioni in Facoltà ed anche (in caso di numerosità ancora insufficienti) in gruppi di Facoltà, pur nella consapevolezza che all'interno di non poche Facoltà sussistono cospicue differenze di competenze e di sbocchi professionali fra i vari corsi di laurea attivati.

Nella Tab. 1 è riportata la composizione finale del campione per Facoltà e genere degli intervistati, sia in termini assoluti che in distribuzione percentuale secondo la Facoltà; inoltre, viene riportata la percentuale relativa ai laureati che risultano essere o essere stati occupati a tre anni dalla laurea (anche a tempo determinato e pur se al momento dell'intervista risultavano disoccupati), al netto dell'eventuale servizio di leva espletato dopo la laurea (calcolo effettuato sottraendo la durata media di inattività dovuta al periodo di leva, ossia 10 mesi, al tempo che gli intervistati hanno impiegato a trovare lavoro³).

Dai dati, sembra che la tanto citata disparità di opportunità lavorative fra maschi e femmine sia molto meno ampia di quanto affermato in letteratura, nonostante che la maggior parte degli intervistati non si sia allontanata dal Mezzogiorno⁴. Si potrebbe già confermare, quindi, che la formazione universitaria fornisce, alle donne

² Il piano di campionamento, invero, prevedeva 3020 interviste telefoniche; nonostante la reiterazione fino a cinque tentativi di contatto, l'utilizzo anche dei numeri di telefonia mobile forniti dai familiari e la predisposizione di un elenco di riserva di numerosità doppia rispetto a quella dell'elenco principale, l'ammontare finale delle interviste effettuate è poco superiore a 2800; di queste, una trentina sono state scartate per manifesta inconsistenza ed incompletezza delle risposte fornite.

³ Si tenga conto che si è ritenuto opportuno inserire i laureati in Medicina e Chirurgia iscritti alle Scuole di Specializzazione di detta Facoltà, fra gli occupati a tempo determinato e non fra le persone "in formazione", poiché il loro rapporto con la struttura formativa è regolamentato, anche i fini contributivi, appunto in modo pari ai contratti di lavoro a tempo determinato.

⁴ Territorio ove, più che altrove, le donne trovano o conservano un lavoro meno facilmente della controparte maschile, per via del conclamato pregiudizio da parte dei datori di lavoro legato alla ovviamente maggiore discontinuità lavorativa dovuta a motivi familiari che la natura o gli usi associano alla figura femminile in modo esclusivo (gravidezze) o comunque prevalente (malattie infantili, ecc.).

Tabella 1. Distribuzione del campione di laureati presso l'Ateneo barese dal 1995 al 2000, per Facoltà e genere; percentuali di occupati nei tre anni dalla laurea.

Facoltà	Genere			% per Facoltà	% occupati nei 3 anni		
	F	M	MF		F	M	MF
Agraria	11	42	53	1,9	90,9	95,2	94,3
Economia	239	195	434	15,6	87,9	92,8	90,1
Farmacia	53	25	78	2,8	84,9	92,0	87,2
Giurisprudenza	356	244	600	21,5	63,2	73,0	67,2
Lettere e filosofia	298	55	353	12,7	79,9	78,2	79,6
Lingue e letterature straniere	207	20	227	8,2	88,4	85,0	88,1
Medicina e chirurgia	108	110	218	7,8	72,2	74,5	73,4
Medicina veterinaria	12	13	25	0,9	66,7	76,9	72,0
Scienze della formazione	198	17	215	7,7	86,9	100,0	87,9
Scienze matematiche, fisiche e naturali	221	150	371	13,3	86,0	94,0	89,2
Scienze politiche	113	98	211	7,6	83,2	82,7	82,9
Università di Bari	1.816	969	2.785	100,0	80,0	83,4	81,2

che si presentano sul mercato del lavoro, almeno gli strumenti indispensabili per ridurre l'impatto dei vincoli sociali a cui esse sono spesso sottoposte.

Tuttavia, esistono ruoli che privilegiano il genere maschile ed altri a cui le donne hanno più facilmente accesso, e tale evenienza è riflessa dalle consistenti differenze fra le percentuali di occupazione entro tre anni dalla laurea di laureati e laureate delle varie Facoltà. Limitando la discussione ai dati di una certa consistenza numerica, appare evidente la maggiore occupabilità maschile fra chi possiede una laurea in Giurisprudenza, in Scienze MM.FF.NN. (in special modo se laureati in Scienze dell'Informazione), in Medicina Veterinaria ed in Farmacia.

Per venire all'argomento di cui alla presente nota, ossia ai tempi di inserimento nel mercato del lavoro, bisogna innanzitutto distinguere fra le diverse tipologie di approccio al mercato stesso da parte dei laureati: alcuni di essi, infatti, già lavoravano durante la frequenza degli studi universitari ed hanno proseguito nella medesima occupazione (213 intervistati, ossia il 7,6%), mentre altri hanno proseguito la propria formazione universitaria (al momento dell'intervista, 128 laureati, pari al 4,6% del campione, senza differenze significative fra i sessi); infine, 239 laureati, pari all'8,6% degli intervistati, per vari motivi che esulano dall'argomento qui studiato (ridotta propensione alla mobilità, scarsa richiesta delle loro competenze, motivi familiari, ecc.) sono rimasti fuori dal mercato del lavoro, ed al momento dell'indagine risultano ancora "in cerca di prima occupazione", per quanto si possa esprimere dubbi sull'effettiva insistenza di detta ricerca. Tutti costoro vanno esclusi dall'analisi.

In definitiva, 2.205 intervistati (il 79,2% del campione) precedentemente non occupati nella medesima posizione sono entrati nel mercato del lavoro entro tre anni dalla laurea, sia con occupazioni a tempo indeterminato che a tempo determinato; pe-

Tabella 2. Distribuzione percentuale degli intervistati con esperienze lavorative post-laurea, secondo il tempo intercorso fra laurea e prima occupazione (al netto dell'eventuale servizio militare o civile adempiuto dopo la laurea), per Facoltà e genere.

	Tempo netto fra laurea e primo lavoro				Totale
	fino a 12 mesi	13-24 mesi	25-36 mesi	Oltre 36 mesi	
<i>Facoltà</i>					
Agraria	88,4	11,6	-	-	100,0
Economia	77,5	16,1	4,0	2,4	100,0
Farmacia	88,7	8,1	1,6	1,6	100,0
Giurisprudenza	31,6	22,6	37,7	8,1	100,0
Lettere e filosofia	46,2	28,6	13,8	11,4	100,0
Lingue e letterature straniere	70,4	19,0	4,8	5,8	100,0
Medicina e chirurgia	56,2	35,1	1,5	7,2	100,0
Medicina veterinaria	78,9	15,8	-	5,3	100,0
Scienze della formazione	71,1	20,1	5,4	3,4	100,0
Scienze matem., fisiche e nat.	80,2	12,3	3,9	3,6	100,0
Scienze politiche	74,8	16,1	6,5	2,6	100,0
<i>Genere</i>					
Femmine	61,2	20,8	11,4	6,6	100,0
Maschi	66,6	19,1	10,8	3,5	100,0
Università di Bari	63,1	20,2	11,2	5,5	100,0

raltro, alcuni di essi, in grande prevalenza di sesso femminile, risultano già in stato di disoccupazione (245 persone, ossia l'8,8% del campione).

La base di dati dell'analisi condotta nel seguito è costituita da questi 2.205 laureati che, dopo il conseguimento del titolo, hanno avuto esperienze lavorative di tipo diverso rispetto a quelle eventualmente pregresse. La distribuzione percentuale del subcampione così selezionato, secondo il tempo occorso fra laurea e primo lavoro (al netto dell'eventuale servizio militare o civile), è riportata nella Tabella 2.

Nella tabella si nota immediatamente che la distribuzione è, per fortuna degli interessati, fortemente sbilanciata verso i tempi più brevi, in quanto il 63% del subcampione fa registrare non oltre un anno di attesa, il 20,2% fra uno e due anni, il 12,3% fra due e tre anni, mentre solo il 5,5% ha intrapreso una occupazione oltre il limite dei 3 anni. È evidente un certo divario fra i sessi, in quanto la quota di laureate che trovano lavoro entro un anno dal conseguimento del titolo è un po' meno cospicua della relativa quota maschile, mentre incide maggiormente la quota relativa al *placement* oltre i tre anni (ma ciò può essere anche dovuto al correttivo utilizzato per ridurre lo svantaggio temporale di chi ha svolto il servizio di leva dopo la laurea).

Molto più consistenti sono le differenze che si rilevano per le diverse Facoltà, a cui peraltro afferiscono corsi di laurea di diversa estrazione, ma generalmente affetti da problemi comuni: Agraria e Farmacia risultano le Facoltà le cui competenze so-

no più immediatamente collocabili sul mercato del lavoro, Giurisprudenza la meno favorita, a causa presumibilmente del grande peso della componente libero-professionale (per la quale è d'obbligo il praticantato in uno studio) sul complesso degli sbocchi preferenziali fra i laureati di questa Facoltà.

Prima di proseguire l'analisi statistica, è opportuno specificare che, a causa della esigua numerosità dei gruppi di intervistati laureati presso le Facoltà di Agraria e di Medicina veterinaria, esplicitata in Tabella 1 (numerosità che si è ulteriormente ridotta nel sottogruppo su cui è condotta l'analisi stessa), si è reso necessario accorpare i due gruppi di laureati in uno solo. Fortunatamente, le caratteristiche delle due Facoltà, in termini di *placement* dei propri laureati, sono abbastanza simili, e per questa ragione (unitamente al ridotto peso numerico delle Facoltà stesse) l'operazione effettuata dovrebbe distorcere l'analisi in misura presumibilmente minore della distorsione che sarebbe stata introdotta dall'utilizzo di subcampioni troppo esigui.

In definitiva, per quanto riguarda il gruppo-Facoltà Agraria+Veterinaria, la quota di laureati che trova occupazione entro un anno dal conseguimento del titolo (al netto di un eventuale servizio militare post-laurea) assomma all'85,5% del sub-campione, mentre il 12,9% di essi trova lavoro fra uno e due anni e solo l'1,6% risulta ancora in cerca di occupazione allo scadere dei tre anni dal momento della laurea.

3. Base di partenza per l'analisi statistica dei tempi di inserimento

Con analisi esplorative di diversa natura (osservazione grafica, analisi di tabelle di contingenza, analisi loglineare) sono state individuate, fra le variabili disponibili nel *dataset* (dati di indagine + dati amministrativi d'Ateneo), le seguenti, influenti in modo significativo sulla variabile risposta "tempo di inserimento":

<i>Variabili esplicative su scala nominale:</i>
Facoltà, Genere, Diploma di scuola superiore conseguito, Eventuale iscrizione universitaria, Informazione pregressa sulle prospettive lavorative post-laurea, Esperienze lavorative precedenti alla laurea, Abilitazione professionale o all'insegnamento, Competenze di cui si è rilevata carenza, Laurea come requisito per il lavoro, Prosecuzione attività lavorativa dei genitori, Ramo di attività economica, Disponibilità al trasferimento per lavoro.
<i>Variabili esplicative su scala ordinale:</i>
Livello di frequenza delle lezioni, Numero di lingue conosciute almeno discretamente, Grado di conoscenza della lingua inglese, Grado di conoscenza della lingua francese, Livello delle conoscenze informatiche, Livello di adeguatezza del lavoro alle aspirazioni.
<i>Variabili esplicative su scala di intervalli:</i>
Valutazione della qualità degli insegnamenti specialistici, Valutazione della qualità delle attività pratiche, Valutazione della qualità delle attività professionalizzanti, Valutazione complessiva della preparazione universitaria ai fini lavorativi, Voto medio negli esami di profitto, Voto di laurea.

Assecondando lo scopo della presente nota, la Facoltà di laurea verrà considerata distintiva del livello di gruppo, anche perché l'influenza di detta variabile sui tempi di inserimento appare, dall'analisi descrittiva, eccezionalmente maggiore di qualsiasi altra (pur non assorbendo completamente la variabilità totale), come si evince dalle osservazioni presentate nelle pagine seguenti.

A causa dei già citati limiti editoriali, comunque, non si ritiene opportuno né indispensabile alla presente trattazione mostrare qui la distribuzione dei tempi per ogni categoria delle numerose variabili nominali ed ordinali o per le variabili numeriche (eventualmente divise in classi), pur se detta analisi descrittiva risulta di per sé interessante e degna di un approfondimento in altra sede, come è evidente leggendo le tabelle presentate nel paragrafo precedente. Va solo riferito per chiarezza di interpretazione che, nell'insieme delle variabili su scala nominale, le uniche dicotome sono *Genere* (M/F) e *Prosecuzione attività lavorativa dei genitori* (Sì/No), mentre le altre hanno una maggiore articolazione di modalità.

Si è ritenuto inopportuno complicare in modo irrealistico il modello di partenza inserendo nel successivo modello multivariato variabili non statisticamente connesse alla variabile risposta a livello univariato. Tale posizione appare preferibile (dati alcuni problemi computazionali connessi all'incremento delle variabili considerate) anche perché generalmente la struttura gerarchica delle relazioni viene rispettata ed accade di rado che una caratteristica singolarmente non legata ad una variabile risposta possa esserlo in concomitanza con altre.

Un problema esiziale che ci si trova ad affrontare in uno studio sui tempi di inserimento è il tipo di analisi applicabile.

Si consideri, innanzitutto, che la variabile risposta "tempo di inserimento" è espressa in mesi a causa dei limiti di precisione della risposta in fase di rilevazione (avvenuta a distanza di almeno quattro anni dalla laurea); essa presenta valori interi da zero a 36 mesi, più una categoria finale comprendente tutti coloro che non avevano ancora trovato un'occupazione allo scadere dei tre anni dal conseguimento della laurea, e dunque è più correttamente misurabile su scala di misura ordinale che non su scala ad intervalli. Inoltre, poiché non si tratta di dati rivenienti da indagini longitudinali, ma solo di una ricostruzione "a posteriori" tramite interviste mirate, quindi priva di dati *censurati* nel senso stretto del termine⁵, non sembra opportuna l'applicazione di modelli di durata, più noti come *modelli di sopravvivenza*. O, per meglio dire, i vantaggi rivenienti dall'utilizzo dei modelli di durata, comunque utilizzabili an-

⁵ I dati censurati, com'è noto, sono le osservazioni relative ai soggetti che entrano a far parte dell'indagine dopo il suo inizio (dati censurati a sinistra) ed a quelli nei quali l'evento di riferimento non si verifica entro il termine della rilevazione (dati censurati a destra). Fra questi, sono compresi sia i casi che escono dalla rilevazione prima del termine dell'indagine (rappresentando, specularmente ai primi, le censure propriamente dette), sia quelli in cui, semplicemente, l'evento di riferimento non avviene prima del termine. Per riferimenti metodologici sull'analisi di sopravvivenza, cfr., ad es., Allison (1982), Collett (1994), Delvecchio (2000).

che in assenza di dati censurati nel senso stretto del termine, non sono tali da compensare alcune loro limitazioni metodologiche e/o computazionali⁶.

Se, invece, si definisce il problema come una analisi di dipendenza dei tempi di inserimento lavorativo, il modello lineare (con o senza tecniche multilivello) è inapplicabile in quanto la condizione di normalità delle distribuzioni multiple e singole, a partire dalla variabile obiettivo, è del tutto inesistente nel campione qui studiato e che, date le particolarità del fenomeno, non è presumibile possa verificarsi facilmente anche in altri ambiti territoriali o temporali.

Sarebbe, invece, perfettamente applicabile l'analisi di regressione per dati discreti ordinali, già altrove utilizzata per studiare fenomeni di durata o di sopravvivenza (cfr., per es., Delvecchio e d'Ovidio, 2002); in realtà, i tempi rilevati, formalmente espressi in mesi⁷, a causa del notevole lasso di tempo trascorso soffrono molto presumibilmente di una ridotta precisione, per cui, allo scopo di ridurre l'importanza di eventuali errori, andrebbero raggruppati in bimestri o trimestri.

Per non sacrificare troppo la precisione d'analisi, si è qui adottata la prima opzione, bimestralizzando i tempi di ingresso nel mondo del lavoro, ad esclusione del tempo nullo, che è relativo a coloro che hanno cominciato a lavorare poco dopo la laurea (escludendo comunque, così come sottolineato nel paragrafo precedente, coloro che già lavoravano nel medesimo ruolo professionale). Per quanto riguarda la metodologia, si è inizialmente fatto ricorso a metodi di analisi logit ordinale, tesaurizzando l'esperienza degli studi su citati ed utilizzando software appositi, come il MIXOR sviluppato da D. Hedeker, che effettua l'analisi di sopravvivenza ordinale secondo il modello logistico a odds proporzionali di McCullagh (1980), esteso da Bennett (1983) al contesto medico, con osservazioni anche censurate a sinistra.

Ci si è tuttavia scontrati con un difetto tipico dei modelli logit, legato alla struttura interna dei modelli stessi, basati, com'è noto, su tavole di contingenza multiple, le quali, al crescere del numero di variabili considerate o delle loro modalità,

⁶ Le condizioni di applicabilità dei modelli di durata sono varie: ad esempio, il modello semiparametrico di Cox prevede, oltre ad una sostanziale continuità della rilevazione, la proporzionalità dei rischi fra le diverse modalità di ciascuna covariata e rispetto alla modalità baseline, e ciò va determinato per ciascuna covariata; detta limitazione può essere parzialmente superata con alcuni artifici, come la definizione di covariate tempo-dipendenti, che però a volte implicano difficoltà computazionali che ne sconsigliano l'uso. I modelli parametrici, a loro volta, non sono sempre applicabili; il modello di Weibull, ad esempio, presenta talvolta problemi nella stima dei parametri (dipendendo dalla soluzione di un'equazione non lineare, generalmente con metodi iterativi come quello di Newton-Raphson), pur se appare meno vincolato dal punto di vista delle condizioni d'uso: esso, infatti, ipotizza solo che $\ln H(t) = \ln \lambda + \gamma \cdot \ln t$, ove $H(t)$ è la funzione di rischio cumulato al tempo t mentre λ e γ sono, rispettivamente, il parametro di scala ed il parametro forma della distribuzione di Weibull; in pratica, stimata $H(t)$ col metodo di Kaplan-Meier e riportati in un sistema di assi cartesiani i punti $(\ln \hat{H}(t), \ln t)$, detti punti devono disporsi all'incirca su una retta. (Delvecchio, 2000).

⁷ Unità di misura, peraltro, già poco adatta a soddisfare, nell'arco dei tre anni presi in considerazione, la condizione di sostanziale continuità sottostante all'applicazione dei modelli di Cox o di Weibull.

presentano sempre più frequentemente celle vuote, anche con il rischio di ottenere righe o colonne interamente vuote e dunque bloccando la procedura di calcolo per la presenza di matrici non singolari. Anche senza arrivare a questo estremo, l'eccessiva frammentazione della matrice multidimensionale dei dati porta talora a risultati poco esplicativi, in quanto molti effetti potenzialmente importanti si rivelano statisticamente non significativi.

È questo il caso presente: in base ai dati in oggetto, il modello di durata logit ordinale, che qui non si riporta per motivi di spazio, comprende pochi effetti significativi (e nessun riferimento alla Facoltà di laurea) e possiede una ridotta capacità esplicativa, essendo inferiore a 0,15 la statistica pseudo- R^2 di Cox-Schnell calcolata per detto modello. Per questo motivo, nel presente lavoro si preferisce fare riferimento ad un modello di regressione categoriale (ordinale) basato su tecniche di *Optimal Scaling*, e perciò non limitato né dalla condizione di normalità della distribuzione, né da problemi computazionali o di frammentazione delle variabili predittive.

4. Analisi multivariata dei tempi di inserimento

Il primo passo dell'analisi consiste nella stima del modello di dipendenza della variabile ordinale "numero di bimestri necessari per l'inserimento lavorativo post-laurea" a livello studente. La metodologia a cui si fa qui riferimento è quella di regressione categoriale sviluppata da Young, De Leeuw e Takane (1976), facente riferimento a tecniche di *Optimal Scaling*, che non prevede assunzioni a priori sulla distribuzione delle variabili indipendenti, che possono essere numeriche (normali o no), ordinali o categoriali; è possibile anche trattare variabili dicotomiche, pur se a volte la valutazione (e l'interpretazione) della posizione delle due categorie nello spazio fattoriale può essere difficoltosa⁸.

Per quanto riguarda i fondamenti teorici del metodo, va brevemente accennato che la j -ma variabile categoriale osservata viene specificata come prodotto di una matrice-indicatore \mathbf{G}_j (di dimensioni $n \times k_j$), data dall'osservazione se la h -ma categoria ($h=1,2,\dots,k_j$) della variabile stessa si sia verificata nell' i -mo individuo ($i=1,2,\dots,n$), e di un vettore $\boldsymbol{\omega}_j = [\omega_{j1} \dots \omega_{jh} \dots \omega_{jk_j}]'$ di parametri di scaling che, una volta stimati iterativamente con il metodo dei Minimi Quadrati Alternati, consentono di posizionare in modo ottimale le singole categorie nello spazio fattoriale. Maggiori informazioni sul metodo sono dati in Crocetta e d'Ovidio (2005), Vittadini (1999) e, particolarmente, in Lovaglio (1997). Una volta definito lo spazio fattoriale ottimale, sono ap-

⁸ La procedura utilizzata è la CATREG, inserita nel software SPSS in accordo con il gruppo di ricerca DTSS (Data Theory Scaling System) dell'Università di Leida (NL), diretto da Jan De Leeuw.

Tabella 3 - Modello di regressione categoriale 1 (variabile risposta=tempo di inserimento).

Variabili esplicative	b _h	Errore std	Z (Wald)	p-value
Facoltà	-0,348	0,0205	16,994	<0,0001
Possesso di partita IVA	-0,131	0,0345	3,791	0,0001
Attività lavorative pre-laurea	-0,121	0,0184	6,590	<0,0001
Qualità delle attività professionalizzanti	-0,106	0,0309	3,434	0,0003
Livello di conoscenza della lingua inglese	-0,096	0,0183	5,271	<0,0001
Livello delle conoscenze informatiche	-0,092	0,0205	4,506	<0,0001
Diploma di maturità	-0,072	0,0185	3,867	0,0001
Genere	-0,065	0,0187	3,495	0,0002
Lavoro adeguato alle aspirazioni	-0,060	0,0184	3,243	0,0006
Informazione sulle prospettive lavorative	0,042	0,0180	2,357	0,0092
Ramo di attività lavorativa	0,074	0,0179	4,139	<0,0001
Qualità degli insegnamenti specialistici	0,080	0,0308	2,609	0,0045
Carenze nelle competenze ottenute	0,101	0,0338	2,999	0,0014
Mancata continuazione attività dei genitori	0,108	0,0183	5,889	<0,0001
Abilitazione professionale/insegnamento	0,145	0,0196	7,376	<0,0001

plicabili alle variabili ottimamente trasformate, le cui modalità vengono considerate singole determinazioni di variabili normali standardizzate (cfr Young, De Leeuw e Takane, 1976), sia il modello lineare generalizzato di McCullagh-Nelder, (1989) con tecniche multilivello, sia metodi di analisi strutturale multigruppo (argomento trattato, nel campo d'indagine dell'analisi di durata, in Civardi e Zavarrone, 2002).

Nel modello sono state inserite inizialmente tutte le variabili presentate in precedenza, eliminando progressivamente quelle con coefficienti non statisticamente significativi al livello $\alpha=0,05$ (reinserendo eventualmente quelle che, eliminate al passo precedente, acquistano significatività), con la nota tecnica *Backward Stepwise*. Il modello risultante (altamente significativo, essendo il valore empirico del test ANOVA pari a $F=16,664 > f_{60;2144;0,01}=1,483$) è riportato in Tabella 3.

La capacità esplicativa del modello su riportato, non eccezionale anche a causa della nota e più volte riscontrata sconnessione fra mondo del lavoro ed esperienze formative (che costituiscono la grande maggioranza delle variabili prese in considerazione nella presente analisi), proprio in virtù di tale sconnessione appare sufficiente a farlo ritenere un accettabile modello di partenza, essendo $R^2=0,318$: quasi il 32% della variabilità della variabile risposta, in altri termini, risulta spiegato dalle variabili predittive selezionate.

Come si evince dalla tabella, l'influenza della Facoltà di laurea è (in valore assoluto) ben maggiore di qualunque altra variabile esplicativa, e ciò, insieme alla struttura evidentemente gerarchica dei dati (studenti entro Facoltà) giustifica la scelta di proseguire l'analisi dei tempi di inserimento dei laureati in un'ottica multilivello.

Per procedere in tale analisi, è tuttavia necessario ridefinire il modello di regressione categoriale senza tener conto dell'effetto della Facoltà, in modo da ottenere quantificazioni ottimali nello spazio k-dimensionale relativo soltanto agli studenti. Il modello risultante (anch'esso significativo, essendo il valore empirico dell'ANOVA pari a $F=14,561 > f_{54;2150;0,01}=1,511$, ma logicamente con una minore capacità esplicativa, essendo $R^2=0,268$) è riportato in Tabella 4.

Un breve sguardo ai principali risultati di quest'ultima regressione categoriale: le caratteristiche che sono maggiormente legate ad un abbreviamento dell'intervallo fra la laurea ed il primo lavoro sono l'*aver acquisito una partita IVA* ed aver compiuto, prima della laurea, *esperienze lavorative* (diverse da quella successivamente iniziata); immediatamente a ridosso di queste due caratteristiche, che si potrebbero porre per scontate, si posizionano il *livello delle conoscenze informatiche* e la *qualità delle attività professionalizzanti* fatte durante il corso di studi. All'altro estremo, le caratteristiche individuali che implicano un allungamento del tempo di inattività sono il *voto medio negli esami di profitto* (forse perché chi ha avuto una carriera universitaria brillante, a cui non sempre il voto di laurea è univocamente legato, decide più di sovente per una prosecuzione della propria formazione con dottorati, master ecc.), le *carenze rilevate nelle competenze*, il *conseguimento di abilitazione professionale o all'insegnamento* e, infine, la *mancata continuazione dell'attività dei genitori* (essendo questa l'interpretazione della variabile, come si evince dalla successiva Tabella 5).

Tabella 4 - Modello di regressione categoriale 2 (variabile risposta=tempo di inserimento).

Variabili esplicative	b_h	Errore std	Z (Wald)	p-value
Possesso di partita IVA	-0,179	0,0382	4,675	<0,0001
Voto di laurea	-0,171	0,0479	3,560	0,0002
Livello delle conoscenze informatiche	-0,155	0,0209	7,420	<0,0001
Attività lavorative pre-laurea	-0,151	0,0189	7,985	<0,0001
Qualità delle attività professionalizzanti	-0,139	0,0231	6,007	<0,0001
Diploma di maturità	-0,119	0,0192	6,226	<0,0001
Genere	-0,085	0,0198	4,310	<0,0001
Livello di conoscenza della lingua inglese	-0,058	0,0189	3,071	0,0011
Lavoro adeguato alle aspirazioni	-0,055	0,0196	2,793	0,0026
Informazione sulle prospettive lavorative	0,042	0,0186	2,279	0,0113
Ramo di attività	0,074	0,0186	4,003	<0,0001
Qualità degli insegnamenti specialistici	0,090	0,0225	4,001	<0,0001
Frequenza delle lezioni	0,121	0,0207	5,828	<0,0001
Carenze nelle competenze ottenute	0,142	0,0187	7,566	<0,0001
Mancata continuazione attività dei genitori	0,149	0,0376	3,971	<0,0001
Abilitazione professionale/insegnamento	0,160	0,0204	7,874	<0,0001
Voto medio negli esami di profitto	0,214	0,0481	4,446	<0,0001

Tabella 5 - Quantificazioni di categoria delle variabili categoriali ed ordinali.

Categoria	n	Quantif.	Categoria	N	Quantif.
<i>Genere</i>			<i>Possesso di partita IVA</i>		
F	1.439	-0,730	No	1660	-0,285
M	766	1,371	Sì, attualmente	393	-0,102
<i>Continuazione attività genitori</i>			Sì, in precedenza	48	1,208
Prosegue attività dei genitori	96	-2,154	N.R.	104	4,381
Non proseguimento attiv.	2.002	-0,112	<i>Ramo di attività</i>		
N.R.	107	4,065	Altro ramo	19	-4,073
<i>Diploma di maturità</i>			Non ha cercato personalmente	177	-1,857
Professionale Agrotecnico	6	-3,047	Agricoltura/allevamento	31	-1,246
Altro titolo	14	-2,022	Pubblica Amministrazione	606	-0,242
Maturità Classica	535	-1,358	Industria/artigianato	191	-0,228
Tecnico Geometri	32	-0,437	Commercio/pubblici esercizi	125	0,207
Maturità Linguistica	111	0,056	Servizi pubblici e privati	1003	0,373
Tecnico Commerciale	430	0,310	Credito, assicurazioni e finanza	53	4,431
Professionale Industriale	35	0,348	<i>Frequenza dei corsi</i>		
Maturità Scientifica	732	0,415	Regolarmente, quasi tutti i corsi	1174	-0,845
Tecnico Industriale	87	0,493	Regolarmente, pochi corsi	666	0,495
Maturità Magistrale	170	0,733	Saltuariamente, pochi corsi	264	1,815
Professionale Commercio	29	0,898	Occasionalmente, pochi corsi	101	1,815
Tecnico Agrario	24	5,741	<i>Liv. di conoscenza della lingua inglese</i>		
<i>Abilitazione professionale/insegnamento</i>			Nessuna conoscenza	157	-0,683
No	1032	-1,024	Molto limitata	451	-0,683
Sì, esercizio professione	732	0,605	Solo lingua scritta	198	-0,683
Sì, insegnamento	419	1,382	Parlata discretamente	1118	-0,145
Sì, professione e insegnam.	22	1,383	Parlata fluentemente	281	2,535
<i>Carenze nelle competenze ottenute</i>			<i>Livello delle conoscenze informatiche</i>		
Informatiche di base	18	-3,825	Livello nullo	25	-1,208
Giuridiche/normative	201	-1,516	Livello elementare	739	-1,208
Tecnico/professionali	561	-0,601	Livello abbastanza alto	1110	0,318
Informatiche avanzate	84	-0,354	Livello professionale	331	1,719
Economiche/finanziarie	137	-0,160	<i>Informazione sulle prospettive lavorative</i>		
Nessuna	779	0,159	No	294	-2,550
Altre	8	0,310	Sì, genericamente	909	0,392
Amministrative/contabili	160	0,566	Sì, abbastanza chiaramente	1002	0,392
Linguistiche	244	2,037	<i>Lavoro adeguato alle aspirazioni</i>		
Letterarie/culturali	13	3,708	Totalmente inadeguato	42	-2,382
<i>Attività lavorative pre-laurea</i>			Abbastanza inadeguato	160	-2,382
Nessun lavoro pre-laurea	1755	-0,491	Né adeguato né inadeguato	122	-2,119
Durante il corso	343	1,611	Abbastanza adeguato	969	0,084
Prima e durante il corso	12	2,859	Molto adeguato	808	0,658
Prima del corso	95	2,890	N.R.	104	1,220

Le differenze rispetto al modello precedente (comprensivo di Facoltà di laurea) sono minime, e tutte evidentemente legate alle caratteristiche interne delle Facoltà medesime: risultano significative, infatti, frequenza delle lezioni (praticamente obbligatoria in quasi tutti i corsi di Facoltà come Medicina e Chirurgia, ampiamente discrezionale in altre), media degli esami di profitto e voto di laurea (che, evidentemente, sono elementi importanti soltanto nel placement in alcune posizioni ben determinate, vincolate alla laurea conseguita).

I coefficienti presentati in tabella, benché consentano di stimare l'importanza delle singole variabili nella determinazione del tempo di inserimento lavorativo dei laureati, non sono però sufficienti a descrivere l'effetto delle variabili di tipo ordinale o categoriale sul tempo medesimo (espresso, a sua volta, in forma ordinale). Occorre, allo scopo, conoscere le quantificazioni di categoria contestualmente elaborate dalla procedura, riportate in Tabella 5 ordinate dal ramo negativo a quello positivo. Ovviamente, le variabili ivi descritte sono solo quelle categoriali ed ordinali, in quanto le variabili quantitative sono trasformate dalla procedura con una semplice standardizzazione e normalizzazione.

Come si può notare, la procedura ha portato a quantificare con il medesimo valore alcune modalità, in particolare di variabili espresse su scala ordinale, segno che il comportamento di dette modalità nei riguardi della variabile risposta (e, contestualmente, nei riguardi delle altre covariate) è globalmente uniforme: il che è in pieno accordo con la logica delle cose, posto che, ad esempio, la preparazione globale di chi ha seguito pochi corsi non cambia sensibilmente a seconda che la frequenza fosse saltuaria oppure occasionale; così come non vi è molta differenza, ai fini del placement, fra livelli nulli ed elementari di conoscenza dell'inglese o dell'informatica, né fra le motivazioni di base (generalmente, il bisogno di trovare un'occupazione in tempi brevi) che spingono ad accettare un lavoro inadeguato totalmente o solo parzialmente.

L'utilizzo delle quantificazioni di Tabella 5 nell'operazione di interpretazione del modello è intuitiva, trattandosi di "valori" che, moltiplicati per i coefficienti standardizzati di regressione, consentono di stimare il valore del tempo di inserimento (anch'esso quantificato e standardizzato, con valori da -1,269 per una attesa nulla a 1,750 per tutti i tempi superiori a 28 mesi) di un laureato con tali caratteristiche.

5. Analisi multilivello dei tempi di inserimento lavorativo

Come si evince dal confronto fra i modelli di regressione categoriale descritti nelle pagine precedenti, nell'ottica del modello di dipendenza ad effetti fissi l'effetto della Facoltà di laurea sembra molto limitato, posto che l'esclusione di tale fattore porta ad

una perdita di capacità esplicativa di meno del 10% della varianza complessiva (si tenga conto che le variabili esplicative risultate significative nel secondo modello contribuiscono per una quota ridottissima: se nel primo modello si ha un indice $R^2=0,318$ e nel secondo un $R^2=0,268$, il modello costruito con le medesime predittive del primo fuorché la Facoltà ha un indice $R^2=0,249$). Si vuole quindi verificare se un modello ad effetti casuali, nonostante il numero non elevato di Facoltà costituenti le unità di secondo livello, consenta un miglioramento della predittività statistica.

Per far ciò, partendo dai risultati della regressione categoriale si giunge a stimare, tramite le quantificazioni di categoria ottenute, un modello di regressione multilevel⁹ in cui la Facoltà di laurea viene considerata distintiva del livello di gruppo.

Allo scopo di verificare se sul fenomeno studiato esista realmente un effetto Facoltà, base essenziale per la costruzione di un modello multilevel, si è innanzitutto stimato il modello con sola intercetta casuale o “modello banale” (Delvecchio, 2001)

$$y_{ij} = \gamma_{00} + \eta_{0j} + \varepsilon_{ij},$$

ove ε_{ij} ed η_{0j} (che rappresentano, rispettivamente, gli errori di primo e di secondo livello) sono variabili casuali con le seguenti caratteristiche:

$$E(\varepsilon_{ij} \varepsilon_{i'j'})=0, \quad \varepsilon_{ij} \sim N(0, \sigma^2(\varepsilon)); \quad E(\eta_{0j_1} \eta_{0j_2})=0, \quad \eta_{0j} \sim N(0, \sigma^2(\eta)); \quad E(\eta_{0j} \varepsilon_{ij})=0.$$

Le stime dei parametri di covarianza per gli errori di primo livello sono pari, nel caso presente, a $\hat{\sigma}^2(\varepsilon)=0,777$; quelle relative agli errori di secondo livello sono invece pari a $\hat{\sigma}^2(\eta)=0,1904$.

Partendo da tali parametri, si è stimato il *coefficiente di correlazione intraclassa* $\hat{\rho}_w = \frac{\hat{\sigma}^2(\eta)}{\hat{\sigma}^2(\eta) + \hat{\sigma}^2(\varepsilon)} = 0,197$, il quale assicura che dipende dalla Facoltà di laurea poco meno del 20% della variabilità complessiva dei tempi di occupazione dopo la laurea: non moltissimo, ma una quota rilevante, tenendo conto del contesto.

La verifica dell'effettiva casualità intraclassa, effettuata tramite il test¹⁰

$$F = \frac{\tilde{n} \cdot \hat{\sigma}^2(\eta) + \hat{\sigma}^2(\varepsilon)}{\hat{\sigma}^2(\varepsilon)} \quad (\text{in cui } \tilde{n} = \bar{n} - \frac{\sum (n_j - \bar{n})^2}{J} / \sum n_j \text{ rappresenta l'ampiezza media corretta dei corsi di laurea nel campione})$$

risulta ampiamente positiva, essendo

⁹ Per la formalizzazione matematica dei modelli multilevel con variabili misurate su scala ordinale cfr., ad es., Fielding, Yang e Goldstein (2003). Nel presente studio, tuttavia, si è preferito utilizzare un modello multilevel di tipo lineare (costruito tramite il robusto modulo multilevel di LISREL, per il quale cfr. Du Toit, Du Toit e Cudeck, 1999) utilizzando in input le trasformate delle variabili selezionate dalla procedura CATREG (fra cui variabili categoriali ed ordinali ottimamente quantificate, le cui modalità trasformate, come già sottolineato in una precedente nota, si possono considerare determinazioni di variabili normali standardizzate)

¹⁰ Detto test, sotto l'ipotesi di base che non vi siano differenze fra i gruppi, ossia che sia nullo il coefficiente di correlazione residua interclasse, si distribuisce all'incirca come una F con J-1 e N-J gradi di libertà, per cui si respinge l'ipotesi di base se $F > f_{J-1; N-J; \alpha}$ (cfr., ad es., Snijders and Bosker, 1999).

$F=53,407 > f_{9,2195;0,05}=1,884$, per cui è definito statisticamente significativo l'effetto Facoltà sui tempi di inserimento lavorativo dei laureati, ed è dunque giustificato procedere con analisi più approfondite.

La prima analisi parte dalla supposizione che le h variabili esplicative (con $h=1,2,\dots,H$) osservate su ogni i -mo studente-unità di primo livello ($i=1,2,\dots,N$) non abbiano pendenze casuali rispetto al secondo livello, dato dalle j Facoltà considerate ($j=1,2,\dots,J$), ossia che l'effetto gruppo non agisca sui coefficienti di regressione (che quindi rimangono fissi nel passare da una Facoltà all'altra), ma soltanto sul parametro di intercetta, che può dunque essere considerato composto da una parte fissa e da un errore casuale, ossia $\beta_{0j} = \gamma_{00} + \eta_{0j}$. In tal caso, il modello di regressione può essere espresso come:

$$y_{ij} = \gamma_{00} + \beta_{10}x_{1ij} + \beta_{20}x_{2ij} + \dots + \beta_{H0}x_{Hij} + \eta_{0j} + \varepsilon_{ij}, \quad [1]$$

ove $\text{Cov}(X_{ij}, \varepsilon_{ij})=0$, ossia gli errori di primo livello sono indipendenti dalle variabili esplicative di primo livello (Goldstein, 1995; Snijders e Bosker, 1999).

Sotto le ipotesi previste per la [1], l'applicazione della metodologia multilevel alle variabili trasformate tramite la procedura CATREG ha portato (previa eliminazione progressiva delle variabili di primo livello divenute non significativamente

Tabella 6 - Modello di dipendenza multilevel a coefficienti fissi (variabile risposta = tempo di inserimento lavorativo post-laurea).

Parte fissa del modello	b_h	errore stand.	t	p-value
<i>Variabili predittive trasformate</i>				
Attività lavorative pre-laurea	-0,130	0,0187	6,965	<0,0001
Possesso di Partita IVA	-0,128	0,0419	3,056	0,0068
Livello delle conoscenze informatiche	-0,090	0,0227	3,961	0,0016
Livello di conoscenza della lingua inglese	-0,082	0,0235	3,479	0,0035
Genere	-0,075	0,0193	3,914	0,0018
Diploma di maturità	-0,066	0,0194	3,428	0,0038
Lavoro adeguato alle aspirazioni	-0,065	0,0189	3,435	0,0037
Qualità delle attività professionalizzanti	-0,049	0,0199	2,446	0,0185
Informazione sulle prospettive lavorative	0,039	0,0182	2,153	0,0299
Ramo di attività lavorativa	0,072	0,0179	4,025	0,0015
Carenze nelle competenze ottenute	0,104	0,0187	5,561	0,0002
Mancata continuazione attività dei genitori	0,131	0,0392	3,339	0,0043
Abilitazione professionale/insegnamento	0,136	0,0205	6,665	<0,0001
Parte casuale del modello	$\hat{\sigma}^2$	errore stand.	intervallo di confidenza 95%: lim.inferiore lim. superiore	
$\hat{\sigma}^2(\varepsilon)$	0,694	0,0210	0,658	0,741
$\hat{\sigma}^2(\eta)$	0,098	0,0457	0,038	0,256

connesse con il tempo di inserimento lavorativo introducendo il secondo livello, ossia la Facoltà di laurea, ed iterazione della regressione categoriale) a stimare il modello finale i cui coefficienti sono riportati nella Tabella 6.

Il modello risultante è molto più simile al primo modello di regressione categoriale (descritto in Tabella 3), senza la presenza della covariata “Facoltà”, che non al secondo (Tabella 4); rispetto a quest’ultimo, risultano escluse, non essendo statisticamente significativi i relativi coefficienti di regressione¹¹, soltanto alcune variabili formative (voto medio, frequenza lezioni, qualità insegnamenti specialistici, voto di laurea) il cui effetto viene assorbito dalla variabilità delle unità di secondo livello, le Facoltà di laurea.

Nel modello multilivello, rispetto al primo modello di semplice regressione categoriale, perde importanza (in termini di riduzione del tempo necessario per il *placement*) la *Qualità delle attività professionalizzanti* del corso di studi (variabile effettivamente alquanto “localizzata”, posto che per alcuni corsi di studi non vi sono molte attività riconosciute “professionalizzanti” dagli intervistati), mentre cresce l’influenza della *Mancata continuazione dell’attività dei genitori* sul ritardo nell’inserimento professionale, per ovvi motivi di “entrata” nel mondo del lavoro.

Le componenti di varianza non spiegata in questo modello sono ovviamente inferiori, a causa dell’introduzione delle covariate, che spiegano parte della variabilità di Y, ed anche il *coefficiente di correlazione intraclassa* stimato per questo modello è piuttosto ridotto ($\hat{\rho}_w = 0,124$), indicando che è dovuta alla Facoltà di laurea solo poco più del 12% della variabilità dei tempi di inserimento. Tuttavia, la verifica dell’effettiva casualità intraclassa, essendo $F = 31,477 > f_{9,2195;0,05} = 1,884$, consente di considerare ampiamente dimostrata la presenza dell’effetto gruppo anche in questo caso.

A conferma di questo risultato, il limite inferiore dell’intervallo di confidenza (al 99%) di $\hat{\sigma}^2(\eta)$, riportato nella tabella, è maggiore di zero in misura sufficiente da ritenere tale parametro statisticamente significativo, anche tenendo conto della nota ed evidente asimmetria dell’intervallo di confidenza dei parametri di covarianza.

La Tabella 7 evidenzia, nei residui di secondo livello delle stime bayesiane del modello, la maggiore o minore influenza della Facoltà di laurea sull’intercetta, ed in definitiva sul tempo di attesa prima dell’inserimento lavorativo. Anche alla luce di quanto già osservato nella Tabella 2, non stupisce constatare che i residui negativi più elevati (che indicano una riduzione sostanziale dei tempi di inserimento rispetto a quelli medi) competono ai laureati delle Facoltà di Farmacia e di Agraria (ai quali si

¹¹ Si ricorda che, come dimostrato da Brik e Raudenbush (1992), nell’ambito dei modelli multilevel la significatività dei coefficienti di regressione di primo livello è correttamente verificabile comparando il valore empirico del rapporto $t = (\text{stima coefficiente}) / (\text{errore standardizzato della stima})$ con il valore soglia, al livello α , di una distribuzione T con J-C-1 gradi di libertà (ove J rappresenta il numero di unità di secondo livello e C il numero di covariate osservate su tali unità: nel caso presente, dunque, si avrà g.d.l.=9); per l’analisi qui riportata, è stato prefissato $\alpha = 0,05$.

Tabella 7 - Residui bayesiani del modello [1] e varianza delle stime bayesiane.

Facoltà di laurea	Deviazione stime dai parametri di gruppo	Deviazione std. stime bayesiane
Agraria + Medicina Veterinaria	-0,405	0,100
Economia	-0,157	0,042
Farmacia	-0,475	0,100
Giurisprudenza	0,497	0,042
Scienze politiche	-0,126	0,065
Lettere e Filosofia	0,301	0,048
Lingue e letterature straniere	0,091	0,059
Medicina e chirurgia	-0,214	0,059
Scienze della formazione	-0,226	0,067
Scienze Matematiche, Fisiche e Naturali	-0,252	0,045

aggiunge, al fine di garantire una numerosità di gruppo sufficiente per l'analisi, la ridotta coorte di laureati in Medicina Veterinaria), mentre i residui positivi maggiori sono fatti rilevare dai laureati in Giurisprudenza, seguiti, a buona distanza, da quelli relativi a Lettere e Filosofia. Lo sbilanciamento numerico delle Facoltà con residui negativi rispetto alle altre è legato prevalentemente al peso rilevante assunto nel modello dalla Facoltà di Giurisprudenza (che, come mostrato nella Tab. 2, fa rilevare il maggiore sbilanciamento verso i tempi lunghi di *placement*), ovviamente a causa del cospicuo ammontare dei suoi laureati facenti parte del campione, ammontanti quasi a un quarto del totale. La variabilità delle stime bayesiane, infine, è generalmente abbastanza ridotta, anche in termini relativi.

Una volta accertato e quantificato l'effetto Facoltà sulla distribuzione dei tempi di accesso al lavoro dopo la laurea, non essendo significativamente connesse al tempo di *placement* le poche variabili esplicative di secondo livello disponibili (rapporto iscritti per docente e percentuale di laureati su iscritti), è stata tentata la costruzione e la stima di un modello multilevel a coefficienti casuali di primo livello, ponendo l'ipotesi che l'effetto gruppo influenzi anche alcune o tutte le variabili di primo livello, le quali, come si è più volte sottolineato, sono in buona parte logicamente legate al percorso formativo dell'individuo (e quindi, presumibilmente, appunto alla Facoltà in cui tale percorso si è svolto).

Supponendo che l'effetto gruppo agisca sui coefficienti di regressione delle h variabili esplicative di primo livello, tutti o alcuni di detti coefficienti possono essere considerati composti da una parte fissa e da un errore casuale, ossia $\beta_{hj} = \gamma_{h0} + \eta_{0j}$. Il modello di regressione [1] può essere dunque espresso come:

$$y_{ij} = \gamma_{00} + \sum_h \gamma_{h0} x_{hij} + \eta_{0j} + \sum_h \eta_{h0} x_{hij} + \varepsilon_{ij}, \quad [2]$$

ove, per ogni unità di secondo livello, si assume che gli errori di primo e di secondo livello siano indipendenti fra loro, ossia $\text{Cov}(\varepsilon_{ij}, \eta_{hj}) = 0$, ed inoltre che

$$\boldsymbol{\eta}_j = [\eta_{0j} \ \eta_{1j} \ \dots \ \eta_{Hj}]' \sim \mathcal{N}_H(\mathbf{0}, \boldsymbol{\Sigma}),$$

con

$$\boldsymbol{\Sigma} = E(\boldsymbol{\eta}_j \boldsymbol{\eta}_j') = \begin{bmatrix} \sigma_0^2(\eta) & \sigma_{01}(\eta) & \dots & \sigma_{0H}(\eta) \\ \sigma_{01}(\eta) & \sigma_1^2(\eta) & \dots & \sigma_{1H}(\eta) \\ \dots & \dots & \dots & \dots \\ \sigma_{0H}(\eta) & \sigma_{1H}(\eta) & \dots & \sigma_H^2(\eta) \end{bmatrix}, \quad \sigma_{rs}(\eta) = E(\eta_{rj} \cdot \eta_{sj}).$$

Sono date per acquisite tutte le altre assunzioni previste per il modello [1].

Ove solo alcune delle variabili di primo livello presentino pendenze casuali (e siano, per semplicità di notazione, le prime k variabili), il modello [2] diventa:

$$y_{ij} = \gamma_{00} + \sum_{h=1}^k \gamma_{h0} x_{hij} + \sum_{h=k+1}^H \beta_{hj} x_{hij} + \eta_{0j} + \sum_{h=1}^k \eta_{h0} x_{hij} + \varepsilon_{ij}. \quad [3]$$

In termini di procedura applicata, si è dapprima costruito un modello a soli coefficienti casuali, per poi riportare via via in forma fissa quei coefficienti la cui casualità non risultava statisticamente verificata. Una volta selezionati gli effetti casuali significativi (relativi, tra l'altro, a caratteristiche come "attività lavorative pre laurea", "carenze nelle competenze" e "abilitazione professionale o all'insegnamento"), si è verificato se l'accostamento del modello ai dati fosse migliorato.

Ciò non è avvenuto: anzi, tutti i criteri hanno fatto rilevare valori superiori, dal criterio di log-verosimiglianza ristretta agli indicatori AIC e BIC, ossia il criterio informativo di Akaike ed il criterio bayesiano di Schwartz; ciò implica, com'è noto, un accostamento peggiore; anche per i modelli in cui tutti o alcuni di detti coefficienti si presentano sia in forma fissa che casuale, l'accostamento ai dati è meno buono del modello con soli coefficienti fissi: nella migliore delle ipotesi, la differenza del valore di logverosimiglianza non è significativa rispetto al modello con soli coefficienti fissi, pur essendo il modello a coefficienti casuali più complesso e quindi meno chiaramente interpretabile.

Si può quindi ritenere che il modello più adeguato ad esprimere la relazione fra caratteristiche della formazione universitaria e tempo di inserimento lavorativo sia, per i laureati dell'Università di Bari, quello presentato nella Tabella 6 e che, dunque, la Facoltà di laurea influenzi in misura statisticamente significativa soltanto il parametro di posizione iniziale, mentre appare pressoché costante fra le varie Facoltà l'effetto delle covariate a livello individuale (comunque numericamente ridotte rispetto al modello non gerarchico da cui è partita la presente analisi). Invero, il minore intervallo di tempo fra la laurea ed il primo lavoro (qualsiasi sia la sua modalità, a tempo determinato o indeterminato, dipendente o autonomo) è legato soprattutto alle esperienze lavorative compiute prima della laurea¹², anche in ambiti e con ruoli

¹² Si ricorda, peraltro, che sono esclusi dalla presente analisi i laureati che lavoravano continuativamente prima del conseguimento del titolo ed immediatamente dopo, fra i quali, ovviamente, si trovano

differenti rispetto a quelli successivi. Tenendo conto del valore assunto, tramite la procedura di quantificazione, dalle modalità delle eventuali esperienze lavorative, è immediato concludere che i minori tempi vengono fatti registrare da chi ha lavorato prima degli studi (ed eventualmente anche durante), mentre chi non ha mai lavorato vede aumentare, a parità di altre condizioni, l'attesa per l'inserimento lavorativo post-laurea: ciò è in accordo con la generalizzata tendenza delle aziende a richiedere dai candidati all'assunzione un'esperienza lavorativa in ruoli più o meno consimili¹³. L'aver ottenuto una partita I.V.A. determina anch'essa una riduzione dei tempi.

La ricerca (e l'ottenimento) di un lavoro adeguato alle proprie aspirazioni risulta anch'essa una caratteristica abbastanza legata ai tempi di inserimento lavorativo, ma in senso inverso e difficile da interpretare: maggiore è l'adeguatezza, minore il tempo necessario per l'occupazione. Si tratta, spesso, di una relazione biunivoca (chi sta cercando lavoro da tempo è più facilmente portato ad accettare occupazioni meno confacenti alle proprie aspirazioni) e probabilmente spuria, dovuta all'influenza di altre caratteristiche logicamente connesse, come, ad esempio, il livello delle conoscenze informatiche e della lingua inglese (in particolare, per chi ha conoscenze informatiche a livello professionale o capacità di interpretariato, il mercato del lavoro presenta senz'altro caratteristiche più favorevoli ed il lavoro conseguito è certo più rispondente alle aspettative rispetto ad altri casi).

Vi è qualcosa da puntualizzare per quanto riguarda il genere: nel senso che, per noti motivi di resistenza delle imprese private (che considerano teoricamente sacrosanto ma praticamente antieconomico il diritto femminile ai congedi per gravidanze e problemi della figliolanza, che tradizionalmente sono competenza delle madri più che dei padri), per le laureate il conseguimento dell'obiettivo lavorativo è fin troppo spesso ritardato rispetto ai colleghi maschi, alla cui categoria corrisponde infatti la quantificazione migliore in termini di tempi di inserimento (cfr. Tab. 5).

Per quanto riguarda la relazione fra il diploma di maturità e la rapidità dell'inserimento lavorativo, il titolo maggiormente legato a quest'ultima è distintivo di una delle Facoltà che presentano di per sé una maggiore capacità di penetrazione del mercato del lavoro: il diploma Tecnico Agrario, la cui maggioranza di titolari si iscrive ad Agraria e inizia a lavorare in tempi abbastanza brevi dopo la laurea (sia per la particolare specializzazione conseguita, sia per l'esperienza pratica acquisita e sia, infine, per la numericamente ridotta concorrenza). A fronte di costoro, tuttavia, vi sono i laureati provenienti dal Professionale Agrotecnico, la cui esiguità numerica rende però critica l'assunzione di ipotesi sui motivi di tale difformità dai colleghi, pur conseguendo in maggioranza la stessa laurea ed avendone le stesse potenzialità.

tutti coloro che svolgevano il medesimo lavoro sia prima che dopo, anche se magari con un ruolo più qualificato ed adeguato al nuovo status di laureati.

¹³ Con l'osservazione, espressa spesso dai candidati, che è un po' arduo fare esperienze lavorative se tutti i possibili datori di lavoro, per l'assunzione, richiedono esperienze pregresse.

Non desta perplessità, invece, il cospicuo ritardo nel *placement* legato, mediamente, all'aver conseguito la Maturità classica (la cui caratteristica di fondo, oltre alla peraltro migliore preparazione generale, è l'orientamento della gran parte dei diplomati verso corsi di laurea aventi buon livello di coerenza con gli studi fatti ma minore efficienza in termini di placement, come Lettere e Filosofia oppure Giurisprudenza).

La qualità delle attività professionalizzanti è certamente un fattore che, fornendo competenze, tende a ridurre l'attesa prima dell'occupazione, pur se parte dell'effetto che tale caratteristica formativa fa rilevare in termini di semplice regressione categoriale (cfr. Tab. 4) risulta assorbito, com'è giusto che sia, dalla modellazione gerarchica per Facoltà: è noto, infatti, che in alcune Facoltà sono presenti corsi di laurea con discipline più teoriche ed in altre Facoltà vi sono corsi con discipline più pratiche ed orientate all'inserimento professionale.

Abbastanza significativa e sempre interessante è l'influenza, sui tempi di *placement*, dell'informazione sulle prospettive lavorative posseduta al momento dell'iscrizione universitaria: nel senso che la mancanza di informazioni (e quindi un minor livello di aspettative) risulta ridurre in modo cospicuo l'attesa prima del lavoro, presumibilmente in quanto implica un aumento di elasticità nella ricerca di lavoro.

D'altra parte, anche la ricerca del lavoro in un ramo di attività anziché in un altro influisce nello stesso senso: è parecchio più dispendioso in termini di tempo ottenere un lavoro nel settore finanziario e creditizio (spesso considerato un porto felice per chi sa impegnarsi, anche se alcune diseconomie trascinate negli anni hanno portato ultimamente ad un ridimensionamento del sistema bancario e quindi delle aspettative dei possibili candidati) che non in rami meno inflazionati, come industria, artigianato e agrozootecnica, senza contare casi diversi: il termine "Altro ramo" racchiude, infatti, sia il settore del tempo libero (laureati che fanno fruttare poco il proprio titolo lavorando, anche con successo, come animatori) che rami non specificati.

Eventuali carenze riscontrate nelle competenze conseguite tendono ad allungare i tempi di inserimento lavorativo in misura un po' meno consistente rispetto al modello non gerarchico, pur se ancora significativa grazie alla sostanziale omogeneità in ogni gruppo (ossia in ciascuna Facoltà). Le carenze più legate a tempi lunghi di attesa prima del lavoro sembrano essere di tipo letterario-culturale¹⁴, seguite da quelle linguistiche (come è ormai logico, in piena globalizzazione), mentre le carenze nelle competenze informatiche di base risultano connesse con tempi brevi di *placement*, così come, in misura meno consistente, le carenze giuridico-normative. Ciò non vuol dire che le conoscenze informatiche o giuridiche non servano: semplicemente, la loro mancanza (caratteristica di determinati percorsi formativi) non danno luogo ad una chiara e significativa perdita di tempo nel *placement* dei laureati.

¹⁴ Per trovar lavoro, evidentemente, conta ancora molto saper scrivere in italiano corretto e saper pensare con la propria testa, attività che richiede, com'è noto, un substrato culturale abbastanza articolato e non semplice nozionismo.

In ultimo, è logico che l'aver conseguito l'abilitazione all'esercizio della professione o all'insegnamento (o, peggio ancora, di entrambi) implica quasi sempre l'aver investito in tali attività formative tempo non dedicato alla ricerca occupazionale o ad attività in proprio, e tale effetto è maggiormente sentito per i laureati di quelle Facoltà in cui l'abilitazione è condizione essenziale per l'occupazione. Altrettanto ovvio (e meno soggetto alla struttura gerarchica Facoltà-laureato) è il motivo per cui intraprendere un'attività che non sia la prosecuzione dell'attività dei genitori influisca in misura significativa sulla tempistica di reperimento della prima occupazione, allungando i tempi rispetto a chi ha già una via di accesso professionale assicurata.

6. Osservazioni conclusive

I risultati statistici fin qui evidenziati definiscono un quadro abbastanza interessante, pur con il più volte menzionato limite della mancanza di informazioni sul lato della domanda di lavoro, ossia delle competenze richieste dalle aziende e sulle loro politiche del personale (argomento, peraltro, in costante evoluzione a causa dell'impatto dinamico della legge Biagi e delle sue interpretazioni più o meno corrette), e quindi con una minore capacità predittiva rispetto a quanto sarebbe stato auspicabile. L'effetto globale di tutte le caratteristiche personali selezionate nel presente studio ammonta, infatti, a poco più del 25% del totale: in altri termini, sono dovuti a fattori legati prettamente al mercato del lavoro ed alla situazione economica tre quarti dei casi di ritardo rispetto all'optimum, l'occupazione subito dopo la laurea¹⁵.

La constatazione che non vi sia, ai fini dei tempi di inserimento lavorativo, una significativa interazione di tipo gerarchico fra la Facoltà di laurea e le caratteristiche personali dei laureati concorrenti alla determinazione di detti tempi consente una semplificazione cospicua dell'analisi.

Resta, ovviamente, il punto fermo che, a causa delle caratteristiche professionali delle lauree che vi afferiscono, laurearsi in alcune Facoltà costituisce di per sé un vantaggio, così come altre Facoltà contano laureati con tempi di inserimento mediamente più lunghi (è il caso, ad esempio, della Facoltà di Giurisprudenza, perché la laurea omonima, che ne assorbe la quasi totalità dei casi, prevede lunghi tempi di tirocinio prima di dedicarsi ad una professione remunerativa, mentre l'alternativa di lavoro alle dipendenze è anch'essa soggetta ad attesa a causa del sovraffollamento in ambito privato e dei tempi lunghi dei pubblici concorsi). Questo vantaggio o svantaggio agisce, però, in modo strutturalmente omogeneo su tutti i laureati di una Fa-

¹⁵ Risultato alquanto utopistico nell'attuale congiuntura, ma che comunque, all'epoca, un centinaio di laureati ha effettivamente conseguito; a quanto dichiarato, la metà di costoro ha addirittura iniziato subito a lavorare a tempo indeterminato oppure in proprio (liberi professionisti/lavoratori autonomi).

coltà, indipendentemente dalle loro caratteristiche personali, formative o di approccio alla ricerca di lavoro. Ai fini della presente analisi, la struttura gerarchica Facoltà-laureati determina comunque una riduzione del numero di caratteristiche personali e formative che concorrono alla definizione di un più o meno lungo percorso di ricerca occupazionale, in quanto alcune di tali caratteristiche sono particolarmente legate alla tipologia di lauree afferenti ad alcune Facoltà e, quindi, il loro effetto viene assorbito dalla struttura gerarchica del modello.

In definitiva, lo studio qui condotto porta a confermare non poche evidenze pratiche, ma anche a confutare alcuni luoghi comuni: fra i quali, si noti, l'ipotesi che la disponibilità al trasferimento sia un fattore di primaria importanza, mentre la sua influenza nella presente ricerca appare surclassata da altri fattori legati alle inclinazioni individuali. Ciò appare di un certo valore nell'analisi di durata della condizione di non occupazione dei laureati, che fa parte, in fondo, dello studio dell'efficienza esterna della formazione universitaria.

Naturalmente, un quadro più esauriente sarà possibile solo se e quando saranno disponibili informazioni anche a riguardo delle richieste del mercato del lavoro, sulla base delle quali potrà assumere nuova importanza la mobilità territoriale.

Riferimenti bibliografici

- ALLISON P.D. (1982) Discrete-time methods for the analysis of event histories, in Leinhardt S. (ed.), *Sociological Methodology, 1982*, Jossey-Bass, San Francisco.
- BENNETT S. (1983) Analysis of survival data by the proportional odds model, *Statistics in Medicine, 2*: 273-277.
- BRYK A.S., RAUDENBUSH S.W. (1992) *Hierarchical Linear Models*, Sage, Newbury Park, CA.
- CIVARDI M., ZAVARRONE E. (2002) Modelli strutturali multigruppo per l'analisi dei tempi di permanenza nel sistema universitario. In G. Puggioni (a cura di), *Modelli e metodi per l'analisi di rischi sociali e sanitari*, CLEUP, Padova: 1-20.
- CLIFF N. (1996) *Ordinal Methods for Behavioral Data Analysis*, Lawrence Erlbaum.
- COLLETT D. (1994) *Modelling Survival Data in Medical Research*, Chapman & Hall, London.
- COX C. (1995) Location-scale cumulative odds models for ordinal data: a generalized non-linear model approach, *Statistics in Medicine, 14*: 1191-1203.

- CROCETTA C., D'OVIDIO F., TOMA E. (2005) Analisi multilivello del placement dei laureati presso l'Università di Bari. In L. Fabbris (a cura di), *Efficacia esterna della formazione universitaria: il progetto OUTCOMES*, CLEUP, Padova: 39-62.
- DE LEEUW J. (1977) *Canonical Analysis of categorical data*, University of Leiden, Leiden, The Netherlands.
- DELVECCHIO F. (1992) *Analisi statistica di dati multidimensionali*, Cacucci, Bari.
- DELVECCHIO F. (2000) *Analisi dei tempi di sopravvivenza - Appunti ad uso degli studenti*, Cattedra di Statistica Sociale, Dipartimento di Scienze Statistiche, Università di Bari, a.a. 1999-2000.
- DELVECCHIO F. (2001) *Cenni sui modelli multilevel - Appunti ad uso degli studenti*, Cattedra di Statistica Sociale, Dipartimento di Scienze Statistiche, Università di Bari, a.a. 2000-01.
- DELVECCHIO F., D'OVIDIO F. (2002) I tempi di permanenza nel sistema universitario. In: G. Puggioni (a cura di) *Modelli e metodi per l'analisi di rischi sociali e sanitari*, CLEUP, Padova: 105-128.
- DELVECCHIO G., D'OVIDIO F. (2004) Un modello con tecniche fuzzy neuro-adattative per la classificazione dei laureati dell'Università di Bari secondo l'occupazione a un anno dalla laurea. In C. Crocetta (a cura di), *Modelli statistici per l'analisi della transizione Università-lavoro*, CLEUP, Padova: 255-278.
- DU TOIT S., DU TOIT M., CUDECK R. (1999) *Introduction to the analysis of multilevel models with LISREL 8.30*, SSI Inc., Lincolnwood, IL.
- FABBRIS L. (1997) *Statistica multivariata. Analisi esplorativa dei dati*, McGraw-Hill, Milano.
- FIELDING A., YANG M., GOLDSTEIN H. (2003), Multilevel ordinal models for examination grades, *Statistical modelling*, vol. 3, 2: 127-153.
- GOLDSTEIN H. (1995) *Multilevel Statistical Models*, Arnold, London / Halsted, New York..
- LONG J.S. (1997) *Regression Models for Categorical and Limited Dependent Variables*, SAGE Publications, London-Thousand Oaks, CA.
- LOVAGLIO P.G., (1997) Un algoritmo di regressione multipla con dati categoriali, *Quaderni di Statistica e Matematica Applicata alle Scienze Economico-Sociali*, XIX, 3: 281-291.
- MCCULLAGH P. (1980) Regression models for ordinal data (with discussion), *Journal of the Royal Statistical Society, Series B*, 42: 109-142.
- MCCULLAGH P., NELDER J.A. (1989) *Generalized linear models*, Chapman & Hall, London.
- SNIJDERS T., BOSKER R. (1999) *Multilevel Analysis*, SAGE Publ., London.

- VITTADINI G. (1999) Analysis of Qualitative Variables in Structural Models with Unique Solutions. In: M. Vichi & O. Opitz (eds.), *Classification and Data Analysis: Theory and Application*, Series Classification Data Analysis and Knowledge Organization, Springer Verlag, Heidelberg.
- YOUNG F.W., DE LEEUW J., TAKANE Y. (1976) Regression with qualitative and quantitative variables: an Alternating Least Squares Method with Optimal Scaling Features, *Psychometrika*, **41**: 505-528.

Analysis of times of placement of graduates from the University of Bari

Summary. *The times of placement of graduates are presumably influenced not only by factors which are connected to the particular demands of the labour market (achieved experience, flexibility, etc), but also by characteristics of the graduates themselves, such as learning ability or work mobility, particularly in Southern Italy where the demand for graduates is still less than in other areas. In this paper, using multivariate statistical methods, we carry out a verification of the effective influence of such factors on the greater or smaller speed of placement of graduates from the University of Bari. We also used multilevel models, that took into account their Faculty of provenance as second level units. In this way we were able to highlight the different opportunities offered by different disciplinary fields.*

Keywords: *Duration analysis, Placement, University of Bari, Categorical Regression, Multilevel Analysis*