

# Una misura dell'efficacia esterna della formazione universitaria basata sull'uso delle competenze acquisite

**Bruno Chiandotto, Silvia Bacci<sup>1</sup>**

*Dipartimento di Statistica "G. Parenti" - Università degli Studi di Firenze*

**Riassunto.** Nel presente contributo si studia l'utilizzo sul luogo di lavoro delle competenze acquisite all'università da parte dei laureati dell'Ateneo fiorentino nell'anno 2000, a un anno e mezzo o due anni e mezzo dal conseguimento del titolo. Lo scopo è l'individuazione delle possibili determinanti del fenomeno indagato, ponendo particolare attenzione alle differenze che possono emergere tra differenti corsi di laurea. In virtù di questa esigenza si svolgono due tipi di analisi: una con cui si individuano gruppi omogenei di corsi di laurea e si procede all'adattamento di un modello logistico ad odds proporzionali per ciascuno dei gruppi individuati e di un modello logistico ad odds parzialmente proporzionali per l'intero ateneo. Una seconda per stimare un modello logistico ordinale a intercetta casuale con due livelli di aggregazione in cui i corsi di laurea rappresentano le unità di secondo livello.

**Parole chiave:** Efficacia esterna; Competenze; Regressione logistica ordinale; Modelli a odds proporzionali; Modelli a odds proporzionali parziali; Modelli multilivello; Analisi dei gruppi.

## 1. Introduzione

Per valutare la qualità dei servizi formativi offerti dalle università si possono misurare l'efficienza e l'efficacia sia interna che esterna. Quando si parla di efficacia esterna si fa riferimento alla "capacità" del titolo universitario di soddisfare le

---

<sup>1</sup> Il presente lavoro è stato finanziato nell'ambito del PRIN 2002, cofinanziato dal MIUR "Transizioni Università-lavoro e valorizzazione delle competenze professionali dei laureati: modelli e metodi di analisi multidimensionali delle determinanti". Coordinatore nazionale è L. Fabbris, coordinatore del gruppo di Firenze è B. Chiandotto (titolo del progetto dell'unità di ricerca locale "Valutazione del processo formativo universitario, sbocchi professionali e pianificazione dei percorsi formativi: modelli e metodi"). L'idea iniziale e l'impostazione del lavoro sono dovuti al contributo di entrambi gli autori, le elaborazioni e la conseguente costruzione dei modelli sono state svolte da S. Bacci.

richieste provenienti dal mondo del lavoro. Indicatori di tale capacità possono essere il tasso di neo-occupazione (Chiandotto e Bacci, 2004), il tempo che intercorre dal conseguimento del titolo all'ingresso nel mondo del lavoro, l'effettiva utilità del titolo di studio per lo svolgimento del lavoro, il grado d'impiego sul luogo di lavoro delle competenze acquisite all'università da parte dei laureati occupati, ecc.

Nel presente contributo l'attenzione si concentra sul grado di utilizzo sul luogo di lavoro delle competenze acquisite all'università avendo come obiettivo l'individuazione delle possibili determinanti del fenomeno indagato e tenendo in particolare considerazione la diversa capacità di ciascun corso di laurea di creare competenze in linea con le esigenze espresse dal mondo del lavoro. A tal fine, individuate opportune aggregazioni dei corsi di laurea, si procede alla stima di un modello di regressione (logistica) sia per ciascuna di tali aggregazioni sia per l'intero contingente dei laureati occupati assegnando, in quest'ultimo caso, ai gruppi di corsi di laurea il ruolo di variabile esplicativa. I risultati così ottenuti sono, infine, posti a confronto con quanto dedotto dall'applicazione, al medesimo insieme di dati, di un modello di regressione (logistica) di tipo multilivello, in cui i corsi di laurea rappresentano le unità di secondo livello.

I dati impiegati nell'analisi provengono da apposite rilevazioni sui profili individuali e sugli sbocchi occupazionali di laureati e diplomati condotte sia dal consorzio inter-universitario Almalaurea (per quanto riguarda i laureati delle sessioni estive) sia dal Dipartimento di Statistica "G. Parenti" dell'Università degli Studi di Firenze (per quanto riguarda i laureati dell'Ateneo fiorentino nelle rimanenti sessioni dell'anno solare). I dati impiegati nell'analisi qui presentata sono relativi al contingente dei laureati presso l'Ateneo fiorentino nell'anno 2000, intervistato a un anno e mezzo/ due anni e mezzo dal conseguimento del titolo.

In merito alla struttura della nota, nel Par. 2 vengono presentati e commentati alcuni dati relativi al fenomeno oggetto di analisi e alle sue possibili interazioni sia con variabili individuali sia con variabili concernenti le diverse tipologie del lavoro svolto.

Nel Par. 3 vengono presentati i risultati relativi all'analisi dei gruppi, nei Paragrafi 4 e 5 si riportano le stime e i risultati derivanti dall'adattamento di un modello di regressione per ciascun gruppo di corsi di laurea e per l'intero Ateneo fiorentino al fine di individuare l'effetto netto delle possibili determinanti sull'utilizzo delle competenze universitarie da parte dei laureati occupati.

Nei Paragrafi 6 e 7 vengono presentati e commentati i risultati relativi alla stima di un modello di regressione multilivello e, a completamento dell'analisi, si procede al confronto tra i risultati conseguiti attraverso l'impiego delle due

metodologie e vengono tratte alcune conclusioni in merito ai vantaggi/svantaggi del ricorso a modelli multilivello per dati gerarchizzati (cioè aggregati in gruppi)<sup>2</sup>.

## 2. Le competenze acquisite nell'Ateneo fiorentino

Nell'anno 2000 hanno concluso gli studi presso l'Ateneo fiorentino 5245 studenti, di cui 4846 laureati e 399 diplomati. Tra i 3856 laureati intervistati, 2882 (pari al 75% circa) risultano occupati al momento dell'intervista: di questi, 1867 (64,8%) utilizzano intensamente le competenze acquisite all'università, 730 (25,3%) le utilizzano abbastanza, i restanti 285 (9,9%) non le utilizzano per niente.

La distribuzione delle risposte tra le tre possibili modalità (molto, abbastanza, per niente) è tutt'altro che omogenea sia a livello di corsi di laurea<sup>3</sup> che a livello di facoltà (Tab. 1). Si può osservare che i laureati della Facoltà di Lettere e Filosofia che dichiarano di impiegare molto le competenze universitarie variano dal 40,0% per Filosofia al 65,0% per Lingue e letterature straniere moderne, mentre a Scienze MFN la percentuale varia dal 45,8% per Scienze naturali all'81,3% per Fisica.

Per "utilizzo in misura elevata" si distinguono i laureati occupati di Medicina e Chirurgia (86,7%) e quelli di Farmacia (81,5%); per contro, coloro che dichiarano di non utilizzare per niente le competenze universitarie provengono per la maggior parte da Scienze Politiche (21,6%), Lettere e Filosofia (18,7%), Agraria (16,4%) e Scienze MFN (16,2%).

In linea di massima, i titoli di studio a cui corrisponde un maggior impiego delle competenze acquisite all'università sono quei titoli di natura prevalentemente "tecnica", volti, cioè, a fornire conoscenze specifiche direttamente spendibili in particolari settori lavorativi, laddove i laureati con titoli di studio a carattere prevalentemente teorico o comunque non specialistico (come Scienze politiche) hanno maggiore difficoltà a trovare un impiego ad hoc nel mondo del lavoro.

Tali risultati, sostenuti dai valori piuttosto elevati assunti dalle statistiche V di Cramer e Chi quadrato (con un p-value inferiore a 0,0001, Tab. 2), sono indicativi del sussistere di una differenza significativa tra corsi di studio in merito all'uso delle competenze in ambito lavorativo: ciò giustifica un'analisi del fenomeno d'interesse che consideri l'effetto esercitato dagli specifici corsi di laurea.

---

<sup>2</sup> Nella presente nota la trattazione del modello multilivello è limitata ad una presentazione sommaria delle stime finali ottenute. Per un approfondimento si rimanda a Chiandotto, Bacci e Bertaccini, 2004.

<sup>3</sup> A causa dello scarso numero di laureati, il corso di laurea in Economia politica è stato aggregato con Economia e commercio e i corsi di laurea in Agricoltura tropicale e subtropicale e in Scienze agrarie tropicali e subtropicali sono stati aggregati con Scienze agrarie. Anche i due corsi di laurea in Lingue e letterature straniere di Lettere e Filosofia e di Scienze della Formazione sono stati considerati come un unico corso.

**Tabella 1.** Utilizzo sul luogo di lavoro delle competenze acquisite all'università, per corso di laurea e facoltà

<i>Facoltà/Corso di studi</i>	<i>Molto</i>	<i>Abbastanza</i>	<i>Per niente</i>	<i>(n)</i>
<b>AGRARIA</b>	60,3	23,3	16,4	73
Scienze agrarie	55,0	35,0	10,0	20
Scienze forestali	63,0	22,2	14,8	27
Scienze forestali ed ambientali	61,1	16,7	22,2	18
Scienze e tecnologie agrarie	62,5	12,5	25,0	8
<b>ARCHITETTURA</b>	68,0	27,9	4,1	556
Architettura	68,0	27,9	4,1	556
<b>ECONOMIA</b>	72,6	24,3	3,0	497
Economia aziendale	83,3	16,7	-	18
Economia e commercio	72,3	24,8	2,9	447
Scienze statistiche ed attuariali	70,8	20,8	8,3	24
Scienze statistiche ed economiche	75,0	25,0	-	8
<b>FARMACIA</b>	81,5	15,4	3,1	65
Chimica e tecnol. farmaceutiche	89,3	10,7	-	28
Farmacia	75,7	18,9	5,4	37
<b>GIURISPRUDENZA</b>	63,2	22,6	14,2	190
Giurisprudenza	63,2	22,6	14,2	190
<b>INGEGNERIA</b>	71,8	23,1	5,1	312
Ingegneria civile	81,9	18,1	-	83
Ingegneria elettronica	62,0	29,3	8,7	92
Ingegneria meccanica	69,6	24,6	5,8	69
Ingegneria informatica	64,7	35,3	-	17
Ing. per ambiente e territorio	81,6	7,9	10,5	38
Ing. delle telecomunicazioni	69,2	30,8	-	13
<b>LETTERE e FILOSOFIA</b>	53,6	27,7	18,7	347
Filosofia	40,0	31,4	28,6	35
Lettere	52,4	25,9	21,6	185
Lingue e letterature straniere	55,1	34,8	10,1	69
Lingue e letter. straniere moderne	65,0	22,5	12,5	40
Storia	61,1	22,2	16,7	18
<b>MEDICINA e CHIRURGIA</b>	86,7	12,0	1,3	75
Medicina e chirurgia	84,8	13,0	2,2	46
Odontoiatria e protesi dentaria	89,7	10,3	-	29
<b>SCIENZE d. FORMAZIONE</b>	67,3	20,7	12,0	309
Lingue e letterature straniere	62,5	12,5	25,0	16
Lingue e letter. straniere europee	61,5	23,1	15,4	13
Materie letterarie	60,9	17,4	21,7	23

<i>Facoltà/Corso di studi</i>	<i>Molto</i>	<i>Abbastanza</i>	<i>Per niente</i>	<i>(n)</i>
Pedagogia	57,5	27,5	15,0	40
Psicologia	62,2	26,7	11,1	45
Scienze della educazione	72,7	18,6	8,7	172
<i>SCIENZE POLITICHE</i>	36,0	42,4	21,6	236
Scienze politiche	36,0	42,4	21,6	236
<i>SCIENZE MAT.FIS.NAT.</i>	64,4	19,4	16,2	222
Chimica	69,2	15,4	15,4	39
Fisica	81,3	12,5	6,3	16
Matematica	48,0	40,0	12,0	25
Scienze biologiche	64,9	16,2	18,9	74
Scienze della informazione	58,3	41,7	-	12
Scienze geologiche	78,1	12,5	9,4	32
Scienze naturali	45,8	16,7	37,5	24
<i>Totale</i>	64,8	25,3	9,9	2882

**Tabella 2.** *Misure di associazione relative all'utilizzo delle competenze universitarie per corso di laurea*

<i>Statistiche d'associazione per</i>	<i>Statistica</i>	<i>Valore</i>	<i>GdL</i>	<i>Prob</i>
<i>Utilizzo delle competenze universitarie</i>	Chi-quadro	299,3294	76	< 0,0001
<i>vs</i>				
<i>Corso di laurea seguito</i>	V di Cramer	0,2360		

Naturalmente, oltre al corso di laurea in cui si è conseguito il titolo, possono essere individuati altri caratteri che influenzano il grado di utilizzo sul luogo di lavoro delle competenze universitarie (Chiandotto, Bacci e Bertaccini, 2004). Le percentuali più elevate di laureati che dichiarano di impiegare in misura rilevante le competenze universitarie si osservano tra coloro che lavorano in piccole aziende, che hanno svolto almeno un'attività formativa post-laurea e che coprono posizioni dirigenziali o svolgono la libera professione. L'uso delle competenze risulta, inoltre, positivamente correlato sia con la soddisfazione per il proprio lavoro che con il grado di necessità del titolo conseguito ai fini delle mansioni svolte.

### 3. Variabile risposta e cluster analysis

La variabile risposta è il *grado di utilizzo sul luogo di lavoro delle competenze acquisite all'università* che può assumere tre diverse modalità: molto (Y=1),

abbastanza ( $Y=2$ ), per niente ( $Y=3$ ). Trattandosi di una variabile ordinale, si adotta un *modello di regressione logistica ordinale*.

Essendo interessati a valutare l'effetto sulla variabile risposta esercitato dal corso di studi in cui è stato conseguito il titolo, si aggregano i corsi di laurea in modo da stimare un modello *per ciascuno* dei gruppi individuati. Per rendere l'analisi più completa si stimato un modello per l'intero contingente dei laureati, impiegando i corsi di laurea aggregati in gruppi come variabile esplicativa<sup>4</sup> (*cluster*).

Stanti le differenze rilevate tra corsi di studio all'interno delle singole facoltà (Tab. 1), si procede ad una suddivisione in gruppi basandosi su una *cluster analysis* (Chiandotto, 1978) piuttosto che ricorrere alla naturale aggregazione per facoltà. In particolare, si adotta il metodo di raggruppamento gerarchico basato sulla minima varianza (*approccio di Ward*)<sup>5</sup>, adottando come variabili di aggregazione<sup>6</sup>:

- Età mediana alla laurea
- Voto mediano di laurea
- Voto mediano di maturità
- % di maschi
- % di laureati che hanno svolto un tirocinio per completare gli studi
- % di laureati che hanno avuto esperienze lavorative durante gli studi
- % di laureati provenienti dai licei.

La *cluster analysis* ha dato origine ad una suddivisione dei corsi di laurea in 4 gruppi<sup>7</sup>, la cui composizione è mostrata nella Tab. 3. In linea di massima, al Gruppo

---

<sup>4</sup> In linea teorica, tali analisi sarebbero potute essere condotte per corsi di laurea piuttosto che per gruppi di corsi di laurea; a causa dello scarso numero di laureati presente in molti corsi e a causa dell'elevata numerosità dei corsi stessi, si è preferita la seconda soluzione in modo da pervenire a risultati più significativi e di più facile lettura.

<sup>5</sup> La scelta è caduta sul metodo di Ward, in quanto gli altri metodi di raggruppamento utilizzati (metodo del legame completo, metodo della mediana e metodo del centroide) hanno dato risultati insoddisfacenti, in termini di distribuzione dei corsi di laurea tra i gruppi individuati: si è, infatti, notata la tendenza alla formazione di un gruppo principale con corsi di studio tra loro eterogenei e di diversi gruppi secondari costituiti da uno o due corsi soltanto.

<sup>6</sup> Essendo le variabili di raggruppamento grandezze espresse su scale differenti, sono state standardizzate in senso statistico (media 0 e varianza 1). Altre forme di standardizzazione hanno prodotto risultati meno soddisfacenti. Inoltre, la scelta delle variabili di raggruppamento è stata vincolata da quelle a disposizione: ideale sarebbe stato poter disporre di misure oggettive delle differenze tra corsi di laurea, che prescindessero dall'aggregazione di variabili individuali. L'utilizzo di quest'ultimo tipo di variabili può influenzare la significatività delle covariate nei singoli modelli stimati: infatti, essendo ogni gruppo omogeneo rispetto alle variabili di raggruppamento, è probabile che le stesse, considerate a livello individuale, non risultino significative. Infine, si ricorda che la scelta di un certo insieme di variabili influenza i risultati della cluster analysis in termini di composizione dei gruppi; quindi, i gruppi individuati non sono certamente l'unica soluzione possibile.

<sup>7</sup> La scelta di adottare una suddivisione in 4 gruppi è motivata dal fatto che una aggregazione più fine avrebbe dato origine (come è stato constatato provando ad impiegare 5 e 6 gruppi) a problemi di

1 appartengono prevalentemente corsi a carattere umanistico (facoltà di Lettere e Filosofia e Scienze della Formazione) o, comunque, a carattere teorico (Scienze biologiche e Scienze naturali). I Gruppi 2 e 3 si ripartiscono, invece, corsi di natura tecnica (a grandi linee, il Gruppo 2 può essere identificato con l'area economico-ingegneristica e il Gruppo 3 con l'area medico-sanitaria). Il Gruppo 4 si identifica quasi completamente con i corsi di laurea della facoltà di Agraria, anch'essi a carattere prevalentemente tecnico.

**Tabella 3.** *Composizione dei gruppi di corsi di laurea*

GRUPPO 1	GRUPPO 2	GRUPPO 3	GRUPPO 4
Architettura	Economia aziendale	Chimica	Scienze agrarie
Filosofia	Economia e commercio	Chimica e tecn. farm.	Scienze forestali
Lettere	Giurisprudenza	Farmacia	Scienze dell'educazione
Lingue e lett. str.	Ingegneria civile	Fisica	Scienze e tecn. agrarie
Lingue e lett. str. Europee	Ingegneria elettronica	Medicina e chirurgia	Sc. forestali e ambientali
Lingue e lett. str. moderne	Ingegneria informatica	Odont. e prot. dentaria	
Materie letterarie	Ingegneria meccanica	Ing. ambiente e territorio	
Pedagogia	Matematica	Ing. per le telecom.	
Psicologia	Scienze dell'informaz.		
Scienze biologiche	Scienze geologiche		
Scienze naturali	Scienze politiche		
Scienze stat. e attuariali			
Scienze stat. ed econom.			
Storia			

Se si procede all'analisi della distribuzione delle variabili di aggregazione tra i gruppi, emergono profili più specifici per ciascuno dei 4 raggruppamenti (Tab. 4). I corsi di laurea del gruppo 1 si caratterizzano per una netta prevalenza di laureati donne (62,4%), provenienti dai licei (76,6%), con voti di maturità generalmente piuttosto bassi (il 21,7% ha ottenuto una votazione compresa tra 36/60 e 40/60) e che hanno conseguito il titolo universitario in tempi decisamente lunghi (il 12,3% è andato fuori corso di oltre 6 anni e il 26,2% ha conseguito il titolo dopo i 30 anni).

Il gruppo 2 si distingue per essere l'unico raggruppamento di corsi a prevalenza maschile (55,9%) e con la più alta percentuale di diplomati tecnici (40,0%); inoltre, i laureati del gruppo 2 hanno riportato voti di maturità mediamente elevati (il 25,5% si è diplomato con almeno 55/60), mentre i voti di laurea sono i più bassi di ateneo (il 54,8% si è laureato con meno di 100/110 e soltanto il 14,2 ha

---

massimizzazione della funzione di verosimiglianza in fase di stima del modello, soprattutto per i gruppi costituiti da pochi individui.

ricevuto 110/110 e lode<sup>8</sup>). I tempi di conseguimento del titolo sono comunque lunghi (il 25,6% impiega più del doppio della durata legale).

**Tabella 4.** *Profilo dei gruppi di corsi di laurea (valori %)*

	Gruppo			
	1	2	3	4
Femmine	62,4	44,1	51,0	69,8
Liceali	76,6	60,0	82,4	61,2
Voto maturità >55/60	19,0	25,5	38,6	13,2
Laureati in corso	6,2	2,3	24,2	13,2
Voto laurea <100/110	11,0	54,8	21,6	8,5
Lavoro durante gli studi	74,0	73,2	45,1	83,0
Tirocinio	7,1	8,2	34,6	93,8
<i>Laureati</i>	<i>1170</i>	<i>1221</i>	<i>246</i>	<i>245</i>

I laureati del gruppo 3 sono in gran parte ex liceali (82,4%), hanno ottenuto voti di maturità elevati (il 38,6% ha riportato una votazione superiore a 55/60) e, durante gli studi universitari, sono coloro che hanno avuto meno esperienze lavorative rispetto ai colleghi degli altri gruppi (soltanto il 45,1% contro il 70-80%), ma, in compenso, il 34,6% di loro ha svolto un tirocinio per il completamento degli studi (contro il 7-8% dei gruppi 1 e 2). I tempi di conseguimento del titolo sono i più brevi di ateneo: quasi ¼ dei laureati, infatti, termina gli studi in corso.

Infine, per quanto riguarda il gruppo 4, esso si caratterizza per una prevalenza di femmine (69,8%) e di diplomati presso istituti tecnici (38,8%) con voti di maturità bassi (il 26,4% ha ottenuto una votazione inferiore a 41/60); ben il 93,8% ha svolto un tirocinio per il completamento degli studi e il 58,9% riesce a conseguire il titolo in meno di una volta e mezzo la durata legale.

#### 4. Variabili esplicative

I modelli volti a indagare sull'utilizzo delle competenze universitarie in ambito lavorativo (variabile risposta) sono stimati sui 2882 neo-laureati dell'Ateneo fiorentino nell'anno 2000 che al momento dell'intervista erano occupati. Le variabili esplicative sono:

<sup>8</sup> Per avere un termine di paragone, si consideri che negli altri tre gruppi le stesse percentuali, nel primo caso, non superano il 21,6% (gruppo 3) e nel secondo caso sono almeno pari al 34,6% (gruppo 3).



1. *Voto di laurea (votolau)*, espresso in 110-mi, varia tra 66 e 113 (corrispondente a 110 e lode). Considerato che il voto di laurea e il voto medio agli esami sono fortemente correlati ( $r=0,84$ ) si è preferito non inserirli entrambi nel modello per evitare la multicollinearità.

2. *Età alla laurea (etalau)*: varia da 21,9 a 67,6 anni.

3. *Voto di maturità (votodip)*, in 60-esimi, varia da 36/60 a 60/60.

4. *Numero di ore lavorative a settimana (num\_ore)*: assume valori tra 2 e 90, con frequenze più elevate in corrispondenza di 30, 36, 40, 45 e 50.

5. *Sesso (sesso)*: è una variabile binaria con modalità “femmina” (*femmina*) e “maschio” (*maschio*). Essendo le femmine laureate il 56,4%, come riferimento è stata assunta la modalità femmina<sup>9</sup>.

6. *Esperienze lavorative durante gli studi (lavI)*: è una variabile binaria con modalità “no” (*lavI\_no* – il laureato non ha avuto esperienze lavorative durante gli studi) e “sì” (*lavI\_si* – il laureato ha avuto esperienze lavorative durante gli studi). Benché la maggior parte dei laureati abbia avuto esperienze lavorative durante gli studi, si è deciso di assumere come riferimento la modalità “no”, al fine di una più immediata interpretazione del coefficiente di regressione stimato.

7. *Svolgimento di uno stage o tirocinio per il conseguimento del titolo (tirocin)*: è una variabile binaria con modalità: “tirocinio non svolto” (*tirocin\_no*) e “tirocinio svolto” (*tirocin\_si*); oltre l’83% dei laureati dichiara di non aver svolto nessun tirocinio

8. *Tipo di maturità conseguita (tipmat)*: le modalità assumibili da tale variabile sono: “classica” (*classica*), “scientifica” (*scientifica*), “tecnica” (*tecnica*), “altra maturità” (*altramat*). Come riferimento è stata adottata la modalità “scientifica”.

9. *Classe sociale di appartenenza (p\_socgen)*: le modalità assumibili da tale variabile sono: “borghesia” (*borghesia* – assunta come riferimento), “classe media impiegatizia” (*impiegatizia*), “piccola borghesia” (*picc\_borgh*), “classe operaia” (*operaia*).

10. *Conoscenza della lingua inglese (lingI\_gb)*, con modalità “nessuna/scarsa” (*ingl\_scarso*), “sufficiente” (*ingl\_suff*), “buona” (*ingl\_buono*), “ottima” (*ingl\_ottimo*). La maggior parte degli intervistati dichiara una buona conoscenza dell’inglese.

11. *Conoscenza di word processor (info3\_wp)*: le modalità considerate sono le stesse della variabile *lingI\_gb*; la modalità più frequente risulta essere quella relativa ad una buona conoscenza di word processor (*wp\_buono*).

12. *Zona geografica della sede di lavoro (sede\_lav)*: tale variabile assume le modalità “Firenze e provincia”, “Nord ed estero”, “Centro”, “Sud” ; come

---

<sup>9</sup> Come regola generale si è scelto di adottare come riferimento per le variabili esplicative discrete la modalità con la frequenza osservata più elevata.

riferimento è stata assunta la modalità “Firenze e provincia”. Si è ritenuto più interessante verificare l'eventuale effetto sull'impiego delle competenze della zona sede di lavoro piuttosto che della zona di residenza del laureato, dal momento che quest'ultima non sempre coincide con il luogo di lavoro.

13. *Condizione occupazionale alla laurea (conocclau)*: in tal caso si vuole verificare se il fatto di svolgere lo stesso lavoro che si svolgeva già al momento della laurea oppure di essere passati ad un nuovo lavoro si rifletta in qualche modo sull'uso delle competenze acquisite all'università. La variabile assume le seguenti modalità: lavorava alla laurea e svolge lo stesso lavoro, lavorava alla laurea e ha cambiato lavoro, non lavorava alla laurea (modalità questa assunta come riferimento).

14. *Tipo di lavoro (tip\_lav)*: tale variabile distingue il lavoro dipendente (modalità di riferimento – *lav\_dipendente*) dal lavoro autonomo (*lav\_autonomo*); l'obiettivo è verificare l'attendibilità dell'opinione diffusa che i lavoratori autonomi ricorrono ad un impiego significativamente più elevato delle competenze acquisite durante gli studi rispetto ai lavoratori dipendenti.

15. *Tipo di contratto (tip\_contr)*: tale variabile distingue il lavoro stabile (modalità adottata come riferimento) dal lavoro a termine, intendendo in questo secondo caso tutti i tipi di contratto a tempo determinato (quali contratti di collaborazione, contratti di formazione ecc.).

16. *Ramo di attività economica (ramoatt)*: i vari rami di attività economica sono stati raggruppati in modo da distinguere tra settore pubblico (*pubblico*) e settore privato<sup>10</sup> (*privato* - quest'ultimo è stato assunto come riferimento).

17. *Posizione professionale del laureato (pos\_prof)*: le modalità originarie sono “dirigente”, “impiegato”, “libero professionista”, “lavoratore autonomo” e “altro”. Dal momento che, in fase di stima del modello, tale variabile è risultata significativa nel complesso, ma non in corrispondenza di tutte le modalità, si è preferito dicotomizzare la variabile distinguendo tra impiegati e operai da una parte (modalità assunta come riferimento – *impieg/operaio*) e dirigenti, liberi professionisti e lavoratori autonomi dall'altra (*dirig/libprof*): la logica seguita è stata quella di tenere separati ruoli professionali più “bassi”, che comportano mansioni caratterizzate da minori responsabilità formali, da ruoli professionali più “elevati”, caratterizzati da maggiori responsabilità.

18. *Attività di qualificazione post-laurea conclusa (att\_form)*: si tratta di una variabile binaria volta a valutare l'effetto sulla qualità del lavoro svolto esercitato dal proseguimento degli studi dopo la laurea.

---

<sup>10</sup> La variabile originaria che distingue tra i due settori non è stata utilizzabile in quanto inficiata da troppe risposte mancanti.

19. *Dimensione dell'azienda*<sup>11</sup> (*dim\_azienza*): si è distinto tra le modalità “piccola/media azienda” (*piccola\_az* - massimo 50 dipendenti) e “grande azienda” (*grande\_az* - oltre 50 dipendenti).

20. *Necessità del titolo* (*necess*): si è distinto tra titolo di studio richiesto per legge (*tit\_perlegge*), titolo di fatto utile (anche se non obbligatorio per legge – *tit\_utile*) e titolo inutile ai fini dell'attività lavorativa svolta (*tit\_inutile*). Come riferimento è stata assunta la modalità “titolo richiesto per legge”.

21. *Soddisfazione per il lavoro svolto* (*soddisf*): tale variabile assume le modalità “molto/moltissimo” (*molto\_sodd*), “abbastanza” (*abbastanza\_sodd*), “poco o per niente” (*poco\_sodd*) e riguarda la soddisfazione per il lavoro nel suo complesso (ponderando i singoli aspetti dell'attività lavorativa). La maggior parte degli intervistati ha dichiarato di essere molto o moltissimo soddisfatto del proprio lavoro.

22. *Ricerca di un nuovo lavoro* (*cerc\_lav*): in tal caso si è distinto tra gli occupati che cercano un nuovo lavoro e gli occupati che, invece, non sono in cerca di un nuovo lavoro.

In sintesi, il *laureato occupato – base* è una femmina, ha conseguito la maturità scientifica, ha una buona conoscenza della lingua inglese e dell'uso del computer, non ha avuto esperienze lavorative durante gli studi e nemmeno lavorava al momento del conseguimento del titolo; ha svolto almeno un'attività formativa post-laurea e lavora come impiegata a Firenze o nella sua provincia alle dipendenze del settore privato con contratto di lavoro stabile. Inoltre, l'azienda in cui è impiegata è di piccole/medie dimensioni e il titolo di studio è obbligatorio per legge per lo svolgimento del lavoro. Infine, si dichiara molto soddisfatta per il lavoro svolto e non è in cerca di nessun nuovo impiego.

## 5. **Modello logistico a *odds* proporzionali e modello logistico a *odds* parzialmente proporzionali**

La natura della variabile risposta (politomica ordinale con 3 possibili modalità) e i risultati dello *score* test per valutare l'ipotesi di *odds* proporzionali hanno suggerito l'adozione di un *modello logistico ordinale a odds proporzionali* (McCullagh e Nelder, 1989: cap. 5) per ciascuno dei 4 gruppi di corsi di laurea, avente la struttura generale di seguito riportata espressa in termini di logit.

---

<sup>11</sup> Si osservi che con la denominazione di azienda si intende più genericamente il luogo di lavoro, quindi non soltanto l'azienda privata, ma anche l'azienda pubblica e lo studio professionale.

$$\begin{cases} \text{logit}(P_{1i}) = \log\left(\frac{P_{1i}}{1-P_{1i}}\right) = \log\left(\frac{P_{1i}}{P_{23i}}\right) = \alpha_1 + \sum_{j=1}^h \beta_j \cdot x_{ji} \\ \text{logit}(P_{12i}) = \log\left(\frac{P_{12i}}{1-P_{12i}}\right) = \log\left(\frac{P_{12i}}{P_{3i}}\right) = \alpha_2 + \sum_{j=1}^h \beta_j \cdot x_{ji} \end{cases}$$

dove:  $i$  indica un generico laureato,  $Pr(Y=1)=P_1$ ,  $Pr(Y=2)=P_2$ ,  $Pr(Y=3)=P_3$ ,  $P_1+P_2=P_{12}$ ,  $P_2+P_3=P_{23}$ ,  $\alpha_1$  e  $\alpha_2$  indicano le intercette dei due modelli logit,  $x_j$  è la  $j$ -esima variabile esplicativa e  $\beta_j$  il corrispondente coefficiente di regressione; nel caso trattato, il numero  $h$  di covariate è pari a 22.

In altri termini, un modello logistico per dati ordinali è dato dall'insieme di tanti modelli logistici per dati binari quante sono le modalità assumibili dalla variabile risposta  $Y$  meno una. In maniera analoga al caso dicotomico e sotto l'assunzione di *odds* proporzionali, il coefficiente di regressione  $\beta_j$  può essere interpretato come l'incremento (se positivo) o il decremento (se negativo) che subiscono *tutti* i vari *logit* in corrispondenza di un incremento unitario della variabile esplicativa  $x_j$ . Nel caso in cui la  $Y$  assume 3 modalità,  $e^{\beta_j}$  indica sia gli *odds* di  $P_1$  rispetto a  $P_{23}$  sia gli *odds* di  $P_{12}$  rispetto a  $P_3$ .

I quattro modelli sono stimati facendo ricorso alla PROC LOGISTIC del software SAS con impiego dell'algoritmo di massimizzazione Fisher-scoring; per la selezione delle variabili esplicative e di eventuali effetti d'interazione ed effetti quadratici è stata adottata la procedura *forward*<sup>12</sup>. In merito alla valutazione della bontà di adattamento dei modelli stimati si è fatto ricorso al test del rapporto di verosimiglianza e al test dello score, oltre che al calcolo delle statistiche del Pearson e della devianza e al calcolo dell'R-quadro<sup>13</sup>; in merito alla significatività dei coefficienti stimati, il giudizio è stato espresso sulla base del test di Wald multivariato e del test di Wald univariato<sup>14</sup>.

Nella Tab. 5 si riportano, per ciascun gruppo, l'elenco delle covariate selezionate, le stime dei corrispondenti coefficienti di regressione e delle intercette, con i relativi errori standard e con i risultati del test di Wald univariato (al livello di

<sup>12</sup> Sono state applicate anche le procedure di selezione *stepwise* e *backward* che, per i modelli relativi ai Gruppi 1, 2 e 4, hanno dato origine a risultati identici; viceversa, per il modello relativo al Gruppo 3, si sono rilevati problemi di bontà di adattamento (in relazione all'applicazione della procedura *stepwise*) e di convergenza dell'algoritmo di massimizzazione (in relazione all'applicazione della procedura *backward*).

<sup>13</sup> Si ricorda che per i modelli non lineari l'R-quadro difficilmente raggiunge valori elevati; pertanto, una quota di varianza spiegata pari al 20-25% può essere considerata soddisfacente.

<sup>14</sup> Il primo è stato applicato per valutare la significatività complessiva di ciascuna covariata discreta polinomica; il secondo, invece, per valutare la significatività dei singoli coefficienti di regressione stimati.

significatività del 10%). Nella seconda colonna si dà indicazione, per ciascuna covariata categorica, della modalità assunta come riferimento (modalità con frequenza più elevata), in modo tale che il punto in corrispondenza del quale tutte le covariate assumono valore 0 rappresenta il valore assunto dalla variabile risposta per “l’individuo-base”, cioè l’individuo con le caratteristiche più diffuse. Nell’ultima colonna della Tab. 5 sono riportate le stime degli odds ratio (che, si ricorda, indicano la variazione che subisce  $P_1$  rispetto  $P_{23}$  o, analogamente,  $P_{12}$  rispetto a  $P_3$ , in corrispondenza di un incremento unitario di ciascuna variabile esplicativa).

I risultati ottenuti evidenziano una particolare criticità di alcune covariate che risultano significative per tre dei quattro gruppi considerati: l’utilità del titolo di studio ai fini delle mansioni ricoperte (Gruppi 1, 2 e 4), il livello di soddisfazione per il lavoro svolto (Gruppi 1, 2 e 3), la conclusione di almeno un’attività formativa post-laurea (Gruppi 2, 3 e 4).

Si nota che, con effetto molto simile tra i tre gruppi di riferimento, la probabilità di utilizzo elevato delle competenze universitarie rispetto alla probabilità di utilizzo limitato è positivamente influenzata dalla necessità legale o, quantomeno, dall’utilità del titolo di studio, dalla soddisfazione per il lavoro svolto e dall’aver terminato almeno un’attività formativa post-laurea.

Prendendo in considerazione i singoli gruppi, si osserva che ciascun modello stimato si distingue per la presenza di specifiche covariate non presenti negli altri modelli. Relativamente al Gruppo 1, i laureati che sono in cerca di un nuovo lavoro hanno la metà di probabilità di usare molto le competenze universitarie (nell’impiego attuale) rispetto ai colleghi che al momento dell’intervista dichiarano di non voler cambiare attività; è, dunque, plausibile che la scarsa soddisfazione derivante da un impiego che rispecchia poco la propria preparazione universitaria sia uno dei motivi determinanti che spinge un giovane laureato a cercare una nuova occupazione.

Per i laureati provenienti dal Gruppo 2, la posizione professionale ricoperta, la dimensione dell’azienda e il voto di laurea rappresentano determinanti significative per il fenomeno studiato: coloro che sono occupati in piccole aziende, che coprono ruoli di maggiore responsabilità e che hanno conseguito voti di laurea elevati hanno probabilità superiori di utilizzare in maniera elevata le competenze universitarie rispetto ad utilizzarle scarsamente di quanto non accada per i colleghi impiegati in aziende grandi con ruolo di impiegato o operaio e con voti di laurea più bassi<sup>15</sup>.

Relativamente al Gruppo 3, i lavoratori autonomi e gli occupati nel settore pubblico presentano stime degli *odds ratio* nettamente superiori rispetto ai lavoratori dipendenti e agli occupati nel settore privato; inoltre, quanto più il titolo di studio

---

<sup>15</sup> L’effetto di tale covariata risulta molto debole, a conferma di quanto emerso in precedenti analisi descrittive in merito allo scarso potere discriminante delle votazioni finali, a causa dell’eccessiva concentrazione sui valori più elevati.

viene conseguito in età avanzata, tanto minore è la probabilità di utilizzare molto le competenze acquisite all'università rispetto alla probabilità di usarle poco o per niente.

**Tabella 5.** Modelli logistici ordinali a odds proporzionali, per ciascun gruppo

<i>Effetto</i>	<i>Riferimento</i>	<i>Stima</i>	<i>s.e.</i>	<i>t</i>	<i>p-value</i>	<i>Odds ratio</i>
<i>Gruppo 1 (R<sup>2</sup>=0,27)</i>						
Intercetta 1	-	1,9775	0,1885	110,0778	<,0001	-
Intercetta 2	-	3,8092	0,2607	213,5712	<,0001	-
<i>tit_utile</i>	<i>tit_perlegge</i>	-0,8117	0,2261	12,8918	0,0003	0,444
<i>tit_inutile</i>	<i>tit_perlegge</i>	-3,3951	0,3536	92,1649	<,0001	0,034
<i>cerc_lav_sì</i>	<i>cerc_lav_no</i>	-0,6415	0,2512	6,5232	0,0106	0,526
<i>abbastanza_sod</i>	<i>molto_sod</i>	-0,6897	0,2279	9,1625	0,0025	0,502
<i>d</i>						
<i>poco_sod</i>	<i>molto_sod</i>	-1,4907	0,3642	16,7531	<,0001	0,225
<i>Gruppo 2 (R<sup>2</sup>=0,22)</i>						
Intercetta 1	-	-2,4672	1,2882	3,6680	0,0555	-
Intercetta 2	-	-0,2560	1,2856	0,0397	0,8422	-
<i>att_form_no</i>	<i>att_form_sì</i>	-0,3506	0,1675	4,3810	0,0363	0,704
<i>dirig/libprof</i>	<i>Impieg/operaio</i>	0,4196	0,2061	4,1440	0,0418	1,521
<i>Grande_az</i>	<i>piccola_az</i>	-0,4888	0,1797	7,3975	0,0065	0,613
<i>tit_utile</i>	<i>tit_perlegge</i>	-1,1002	0,1932	32,4162	<,0001	0,333
<i>tit_inutile</i>	<i>tit_perlegge</i>	-3,0929	0,3504	77,9196	<,0001	0,045
<i>abbastanza_sod</i>	<i>molto_sod</i>	-0,6967	0,1831	14,4867	0,0001	0,498
<i>d</i>						
<i>poco_sod</i>	<i>molto_sod</i>	-1,3932	0,3213	18,8010	<,0001	0,248
<i>votolau</i>	-	0,0463	0,0124	13,8852	0,0002	1,047
<i>Gruppo 3 (R<sup>2</sup>=0,20)</i>						
Intercetta 1	-	8,3295	2,4155	11,8907	0,0006	-
Intercetta 2	-	9,8271	2,4746	15,7702	<,0001	-
<i>att_form_no</i>	<i>att_form_sì</i>	-1,0962	0,5044	4,7233	0,0298	0,334
<i>lav_autonomo</i>	<i>lav_dipendente</i>	2,7989	1,2444	5,0594	0,0245	16,427
<i>poco_sod</i>	<i>molto_sod</i>	-1,2445	0,5154	5,8305	0,0158	0,288
<i>pubblico</i>	<i>privato</i>	1,8586	0,5388	11,8975	0,0006	6,415
<i>etalau</i>		-0,2624	0,0896	8,5806	0,0034	0,769
<i>Gruppo 4 (R<sup>2</sup>=0,26)</i>						
Intercetta 1	-	2,2181	0,5576	15,8256	<,0001	-
Intercetta 2	-	3,3970	0,6229	29,7390	<,0001	-
<i>att_form_no</i>	<i>att_form_sì</i>	-1,3308	0,4809	7,6576	0,0057	0,264
<i>dirig/libprof</i>	<i>Impieg/operaio</i>	1,0200	0,4964	4,2222	0,0399	2,773
<i>tit_utile</i>	<i>tit_perlegge</i>	-0,8744	0,5325	2,6961	0,1006	0,417
<i>tit_inutile</i>	<i>tit_perlegge</i>	-3,7764	0,8838	18,2576	<,0001	0,023

In merito al Gruppo 4, l'unica variabile aggiuntiva rispetto allo svolgimento di attività formative post-laurea e all'utilità del titolo di studio è la posizione professionale ricoperta dal laureato: i dirigenti e i liberi professionisti presentano *odds ratio* più elevati rispetto agli impiegati e agli operai.

Nel caso di stima del modello per l'intero Ateneo fiorentino, i risultati dello *score test* hanno suggerito di non accettare l'ipotesi di *odds* proporzionali per la covariata rappresentata dai gruppi di corsi di laurea: il modello applicato è, dunque, un *modello logistico a odds parzialmente proporzionali*, la cui struttura generale è quella riportata di seguito.

$$\begin{cases} \text{logit}(P_{1i}) = \log\left(\frac{P_{1i}}{1-P_{1i}}\right) = \log\left(\frac{P_{1i}}{P_{23i}}\right) = \alpha_1 + \sum_{j=1}^h \beta_j \cdot x_{ji} + \beta_{1,h+1} \cdot x_{h+1,i} \\ \text{logit}(P_{12i}) = \log\left(\frac{P_{12i}}{1-P_{12i}}\right) = \log\left(\frac{P_{12i}}{P_{3i}}\right) = \alpha_2 + \sum_{j=1}^h \beta_j \cdot x_{ji} + \beta_{2,h+1} \cdot x_{h+1,i} \end{cases}$$

dove  $x_{h+1}$  indica la variabile esplicativa (nel caso trattato la variabile *cluster*) per la quale non vale l'ipotesi di proporzionalità degli odds e  $\beta_{1,h+1}$  e  $\beta_{2,h+1}$  sono i coefficienti di  $x_{h+1}$  relativi, rispettivamente, al primo logit e al secondo logit. In altri termini, il modello ad odds parzialmente proporzionali assume odds proporzionali soltanto per alcune variabili esplicative ma non per altre; di conseguenza i due logit che confrontano  $Y = 1$  con  $(Y = 2) \cup (Y = 3)$  nel primo caso e  $(Y = 1) \cup (Y = 2)$  con  $Y = 3$  nel secondo caso si distinguono non solo per l'intercetta ma anche per il valore assunto dai coefficienti di regressione per i quali non si assume l'ipotesi di proporzionalità degli odds. Poichè la  $Y$  assume 3 modalità,  $e^{\beta_{1,h+1}}$  indica la variazione di  $P_1$  rispetto a  $P_{23}$  in corrispondenza di un incremento unitario di  $x_{h+1}$ , mentre  $e^{\beta_{2,h+1}}$  indica l'analoga variazione subita da  $P_{12}$  rispetto a  $P_3$ .

Una volta selezionate le variabili significative<sup>16</sup> e verificata l'ipotesi di proporzionalità degli odds, il modello definitivo per l'intero Ateneo è stato stimato facendo ricorso alla PROC GENMOD del software SAS con impiego dell'algoritmo di massimizzazione di Newton-Raphson.

Nella Tab. 6 sono riportati l'elenco delle covariate selezionate con indicazione dell'effetto misurato rispetto alla modalità assunta come riferimento (prima e seconda colonna); seguono le stime dei coefficienti di regressione, i relativi errori standard e le stime degli odds ratio<sup>17</sup>. In particolare, per la variabile *cluster*,

<sup>16</sup> La procedura seguita è analoga a quella già descritta all'inizio del paragrafo in relazione ai singoli gruppi di corsi di laurea.

<sup>17</sup> Rispetto alla Tab. 4, non sono riportati i risultati (valore della statistica test e p-value) del test di Wald univariato. Comunque sia, le covariate per cui si è assunta la proporzionalità dei relativi odds

non valendo l'ipotesi di proporzionalità degli odds, le stime dei coefficienti, degli errori standard e degli odds ratio vengono riportate per entrambi i logit, cioè per  $P_{12}/P_3$  e per  $P_1/P_{23}$ .

Determinanti significative del grado d'utilizzo delle competenze acquisite all'università risultano essere il corso di studi in cui si è conseguito il titolo (che entra nel modello attraverso il gruppo di appartenenza), lo svolgimento di attività formative post-laurea, la posizione professionale ricoperta, la dimensione dell'azienda in cui si lavora, l'utilità del titolo di studio, la soddisfazione per il lavoro svolto e, anche se in misura minore, il voto di laurea. L'effetto netto esercitato dalle determinanti ad odds proporzionali è analogo a quanto già rilevato attraverso l'analisi per singoli gruppi; si osserva, infatti, che gli odds ratio mostrati nelle Tabelle 5 e 6 sono piuttosto simili, benché alcune delle covariate risultate significative per singole aggregazioni di corsi di studio (età alla laurea, settore lavorativo, tipo di lavoro, ricerca di un nuovo lavoro) non siano entrate a far parte del modello generale.

**Tabella 6.** Modello logistico ordinale a odds parzialmente proporzionali

Effetto	Riferimento	Stima	Stima	E.S.	E.S.	Odds	Odds
		$P_{12}/P_3$	$P_1/P_{23}$	$P_{12}/P_3$	$P_1/P_{23}$	ratio $P_{12}/P_3$	ratio $P_1/P_{23}$
Intercetta	-	-0,9734		1,0523		-	
<i>att_form_no</i>	<i>att_form_sì</i>	-0,2461		0,1075		0,782	
<i>dirig/libprof</i>	<i>impieg/operaio</i>	0,4661		0,1232		1,594	
<i>grande_az</i>	<i>piccola_az</i>	-0,3375		0,1126		0,714	
<i>tit_utile</i>	<i>tit_perlegge</i>	-3,2424		0,2072		0,353	
<i>tit_inutile</i>	<i>tit_perlegge</i>	-1,0422		0,1219		0,039	
<i>abbastanza_sod</i>	<i>molto_sod</i>					0,510	
<i>d</i>		-1,3017		0,1984			
<i>poco_sod</i>	<i>molto_sod</i>					0,272	
<i>votolau</i>	-	-0,6732		0,1155			
		0,0410		0,0091		1,042	
Gruppo 1	Gruppo 2	-0,7510	-0,3904	0,2415	0,2683	0,472	0,677
Gruppo 1	Gruppo 3	0,0710	-0,4736	0,4187	0,4879	1,074	0,623
Gruppo 1	Gruppo 4	0,3125	-0,4561	0,3511	0,4335	1,367	0,634
Gruppo 2	Gruppo 3	0,8220	-0,0833	0,4156	0,4782	2,275	0,920
Gruppo 2	Gruppo 4	1,0635	0,1249	0,3531	0,4228	2,896	1,133
Gruppo 3	Gruppo 4	0,2415	0,0935	0,4893	0,5854	1,273	1,098

presentano p-value non superiori al 2,2%. Per contro, benché la variabile *cluster*, con un p-value pari a 0,0019, risulti significativa nel suo complesso, alcuni dei confronti tra gruppi non sono significativi. A tale problema si sarebbe potuto ovviare aggregando ulteriormente i gruppi; però si è preferito presentare i risultati in maniera tale da consentire un confronto (in termini di odds ratio) tra i quattro gruppi su cui è stata sviluppata l'analisi fin dall'inizio.



Per quanto riguarda l'effetto netto connesso al gruppo di appartenenza dei corsi di laurea, un laureato proveniente dal gruppo 1 ha una probabilità di utilizzare molto le competenze universitarie, rispetto alla probabilità di utilizzarle in maniera limitata ( $P_1/P_{23}$ ), di circa il 33-37% inferiore in confronto ai colleghi provenienti dagli altri gruppi. Viceversa, per un laureato del gruppo 3 lo stesso *odds ratio* è di almeno l'8,0% superiore rispetto agli intervistati dei rimanenti gruppi. Relativamente a coloro che provengono dai gruppi 2 e 4, i primi sono un po' più avvantaggiati dei secondi ( $P_1/P_{23} = 1,133$ ) nel trovare un'occupazione in linea con le conoscenze acquisite all'università; d'altra parte, come già osservato, entrambi hanno odds ratio maggiori dei colleghi del gruppo 1 ed inferiori ai colleghi del gruppo 3<sup>18</sup>.

## 6. Modello a intercetta casuale a due livelli: un confronto

La struttura gerarchica dei dati analizzati induce ad adattare un modello di regressione logistica ordinale multilivello, in cui le unità di primo livello sono rappresentate dai 2882 laureati dell'Ateneo fiorentino nell'anno 2000 che al momento dell'intervista si sono dichiarati occupati, mentre le unità di secondo livello sono costituite dai 38 corsi di laurea in cui è risultato suddiviso tale contingente.

Le variabili esplicative della precedente analisi sono state impiegate come covariate di primo livello, mentre le variabili utilizzate per l'aggregazione dei corsi di laurea in gruppi hanno svolto il ruolo di covariate di secondo livello. Il modello è stato stimato facendo ricorso, attraverso la PROC NLMIXED del software SAS, alla massimizzazione della funzione di log-verosimiglianza con impiego dell'algoritmo di ottimizzazione Dual Quasi-Newton e del metodo di quadratura Gaussiana adattiva.

Per quanto riguarda la selezione del modello, si è proceduto alla stima preliminare del modello vuoto, cioè privo di variabili esplicative, in modo da

---

<sup>18</sup> Esaminando gli *odds ratio* di confronto tra  $(Y=1) \cup (Y=2)$  con  $Y=3$ , cioè  $P_{12}/P_3$ , l'interpretazione dell'effetto netto esercitato dai quattro gruppi sulla variabile risposta può apparire meno chiara. Il fatto di focalizzare l'attenzione su uno degli odds ratio piuttosto che sull'altro dipende dal significato che si attribuisce alla modalità di risposta intermedia  $Y=2$ : in un caso ( $P_1/P_{23}$ ) essa viene cumulata con la modalità  $Y=3$  e, quindi, viene interpretata in un'accezione negativa; nell'altro caso ( $P_{12}/P_3$ ), invece, viene cumulata con la modalità  $Y=1$ , inducendo ad un'interpretazione positiva. Nel caso specifico trattato, si ritiene più opportuno assegnare a  $Y=2$  un significato negativo, dal momento che tale modalità di risposta denota la mancanza di una presa di posizione netta da parte dell'intervistato in merito al quesito postogli, mentre l'obiettivo del presente lavoro è quello di dare rilievo a quei corsi di laurea che si distinguono in maniera particolare (secondo il giudizio degli intervistati che naturalmente può modificarsi nel corso degli anni e delle esperienze acquisite) per la capacità di creare competenze in linea con le esigenze del mondo del lavoro.

valutare la significatività della varianza dei residui di secondo livello (fattore indispensabile per giustificare il ricorso ad un modello multilivello); quindi, attraverso una procedura di selezione *stepwise*, si è passati all'individuazione delle covariate di primo e secondo livello significative, nonché di eventuali termini quadratici ed effetti d'interazione. Il modello risultante è logistico a intercetta casuale con due livelli di aggregazione (a odds proporzionali) (Fielding *et al.*, 2003; Goldstein, 2003; Hedeker, 2004); la struttura generale per il generico corso di laurea  $k$  è (la simbologia è analoga a quella del precedente paragrafo):

$$\begin{cases} \log\left(\frac{P_{1ik}}{P_{23ik}}\right) = \log\left(\frac{P(Y_{ik} \leq 1)}{P(Y_{ik} > 1)}\right) = \alpha_{1k} + \sum_{j=1}^k \beta_j \cdot x_{jik} \\ \log\left(\frac{P_{12ik}}{P_{3ik}}\right) = \log\left(\frac{P(Y_{ik} \leq 2)}{P(Y_{ik} > 2)}\right) = \alpha_{2k} + \sum_{j=1}^k \beta_j \cdot x_{jik} \end{cases}$$

Più precisamente, essendo le intercette  $\alpha_{1k}$  e  $\alpha_{2k}$  variabili in funzione del gruppo  $k$  (corso di laurea),

$$\alpha_{1k} = \gamma_1 + U_{0k} \text{ e } \alpha_{2k} = \gamma_2 + U_{0k},$$

il modello assume la seguente forma:

$$\begin{cases} \text{logit}(P_{1ik}) = \log\left(\frac{P_{1ik}}{1 - P_{1ik}}\right) = \log\left(\frac{P_{1ik}}{P_{23ik}}\right) = \gamma_1 + \sum_{j=1}^h \beta_j \cdot x_{jik} + U_{0k} \\ \text{logit}(P_{12ik}) = \log\left(\frac{P_{12ik}}{1 - P_{12ik}}\right) = \log\left(\frac{P_{12ik}}{P_{3ik}}\right) = \gamma_2 + \sum_{j=1}^h \beta_j \cdot x_{jik} + U_{0k} \end{cases}$$

dove  $U_{0k}$  indica la componente residua di secondo livello relativa al gruppo  $k$ -esimo. Si assume che i residui di secondo livello siano distribuiti normalmente con media pari a zero e varianza costante.

La mancata convergenza dell'algoritmo di massimizzazione della funzione di log-verosimiglianza, in alcuni casi, e la non significatività delle stime ottenute, in altri, determinano l'assenza di coefficienti casuali nel modello stimato<sup>19</sup>. Ciò comporta l'impossibilità di discernere l'effetto netto delle singole variabili esplicative sulla variabile risposta *per ogni corso di laurea*. Per contro, l'analisi presentata nel Par. 5 consente la stima di coefficienti diversi e la conseguente

<sup>19</sup> Se queste circostanze non si fossero verificate si sarebbe avuto un *modello a intercetta e coefficienti casuali a due livelli di aggregazione*.

individuazione di insiemi di covariate differenti almeno per ogni gruppo di corsi di laurea.

Il modello stimato (Tab. 7) presenta, in primo luogo, una variabilità dei residui di secondo livello significativa, fatto questo che giustifica l'adozione di un'analisi multilivello; in secondo luogo, si osserva che nessuna covariata di secondo livello è risultata significativa, probabilmente a causa del processo di costruzione di tali variabili tramite aggregazione per corso di studi di caratteristiche individuali. Per quanto riguarda le covariate di primo livello entrate a far parte del modello, si nota che cinque delle sette selezionate sono presenti anche nel modello logistico a odds parzialmente proporzionali e mostrano effetti netti molto simili tra loro. Invece, diversamente da quanto emerso dall'analisi precedente, la dimensione dell'azienda non risulta più una variabile significativa; d'altra parte, entrano a far parte del modello sia il sesso del laureato che il tipo di lavoro svolto (autonomo oppure dipendente). Per i maschi la probabilità di impiego elevato delle competenze acquisite all'università rispetto ad un impiego ridotto è di circa il 18% inferiore in confronto alle colleghe femmine; per i lavoratori autonomi lo stesso rapporto di probabilità è del 42% circa inferiore rispetto ai lavoratori dipendenti<sup>20</sup>.

**Tabella 7.** *Modello logistico ordinale a intercetta casuale con due livelli di aggregazione (odds proporzionali)*

<i>Effetto</i>	<i>Riferimento</i>	<i>Stima</i>	<i>s.e.</i>	<i>t</i>	<i>p-value</i>	<i>Odds ratio</i>
<i>EFFETTI FISSI</i>						
Intercetta 1	-	-0,0863	0,9022	0,10	0,9243	--
Soglia	-	-2,1287	0,0804	26,47	<0,0001	--
<i>maschio</i>	<i>femmina</i>	-0,1962	0,0948	2,07	0,0453	0,8218
<i>lav_autonomo</i>	<i>lav_dipendente</i>	-0,5353	0,1552	3,45	0,0014	0,5855
<i>att_form_no</i>	<i>att_form_sì</i>	-0,2008	0,0888	-2,26	0,0296	0,8181
<i>dirig/libprof</i>	<i>impieg/operaio</i>	0,7452	0,1460	-5,11	<0,0001	2,1069
<i>tit_utile</i>	<i>tit_perlegge</i>	-1,0146	0,0994	10,20	<0,0001	0,3625
<i>tit_inutile</i>	<i>tit_perlegge</i>	-3,1033	0,1726	17,98	<0,0001	0,0449
<i>abbastanza_sodd</i>	<i>molto_sodd</i>	-0,7012	0,0948	7,40	<0,0001	0,4960
<i>poco_sodd</i>	<i>molto_sodd</i>	-1,3428	0,1540	8,72	<0,0001	0,2611
<i>votolau</i>	-	0,0172	0,0086	-2,01	0,0521	1,0174
<i>EFFETTI CASUALI</i>						
<i>St.Dev.(U<sub>0k</sub>)</i>	-	0,2904	0,06473	4,49	<0,0001	--

<sup>20</sup> Si osservi che tale relazione è in contraddizione con quanto rilevato per la medesima variabile nel modello logistico relativo al gruppo 3 (Tab. 5).

## 7. Considerazioni conclusive

I due approcci metodologici impiegati per individuare le determinanti dell'uso delle competenze universitarie da parte dei giovani laureati e per misurarne l'effetto netto portano, in parte, a considerazioni analoghe. Si può, infatti, affermare che la probabilità di trovare un'occupazione adeguata al bagaglio di conoscenze derivanti dal percorso universitario svolto è tanto più elevata quanto più il laureato si ritiene soddisfatto del proprio lavoro, quanto più il titolo di studio si è rivelato utile per l'ottenimento dell'impiego e quanto più alto è il voto di laurea. Inoltre, la conclusione di almeno un'attività di formazione post-laurea e l'inquadramento in ruoli di maggiore responsabilità (in veste di dirigente o libero professionista) rendono più probabile l'assegnazione di mansioni il cui espletamento richiede un maggiore uso delle competenze universitarie.

Al di là di queste sicuramente importanti conclusioni comuni ai due approcci, l'analisi per gruppi, diversamente dall'analisi multilivello<sup>21</sup>, consente di pervenire a considerazioni più approfondite, tramite l'individuazione di determinanti peculiari soltanto di alcuni gruppi di corsi di laurea. Naturalmente, la presenza di una certa variabile in un certo gruppo è da ricondurre a particolari criticità di quest'ultimo. Ad esempio, la significatività della variabile relativa alla ricerca di un nuovo lavoro per il gruppo 1 può essere spiegata con l'uso particolarmente scarso delle competenze universitarie che si rileva per i laureati provenienti da tale gruppo, fatto questo che spesso determina la voglia di cambiare occupazione; d'altra parte, l'importanza della dimensione aziendale e della posizione professionale ricoperta risultano critiche per i laureati del gruppo 2, in virtù delle maggiori possibilità che ingegneri, laureati in economia e laureati in giurisprudenza hanno di scegliere tra libera professione e lavoro dipendente e, conseguentemente, tra studio professionale, piccola azienda o grande azienda. Sempre in relazione al gruppo 2, la distribuzione dei voti di laurea relativamente meno concentrata su valori elevati che si può osservare (Chiandotto, Bacci e Bertaccini, 2004) per i corsi di studio di Giurisprudenza, Economia, Ingegneria e Scienze Politiche può spiegare la significatività di tale variabile per il gruppo 2. Infine, la presenza di medici, dentisti e farmacisti nel gruppo 3, le cui attività lavorative si prestano sia ad un'occupazione autonoma che ad un'occupazione alle dipendenze nel settore pubblico o nel settore privato, è coerente con la criticità rilevata per il tipo di lavoro (autonomo o dipendente) e per il settore d'impiego (pubblico o privato) in relazione al terzo gruppo di corsi di laurea.

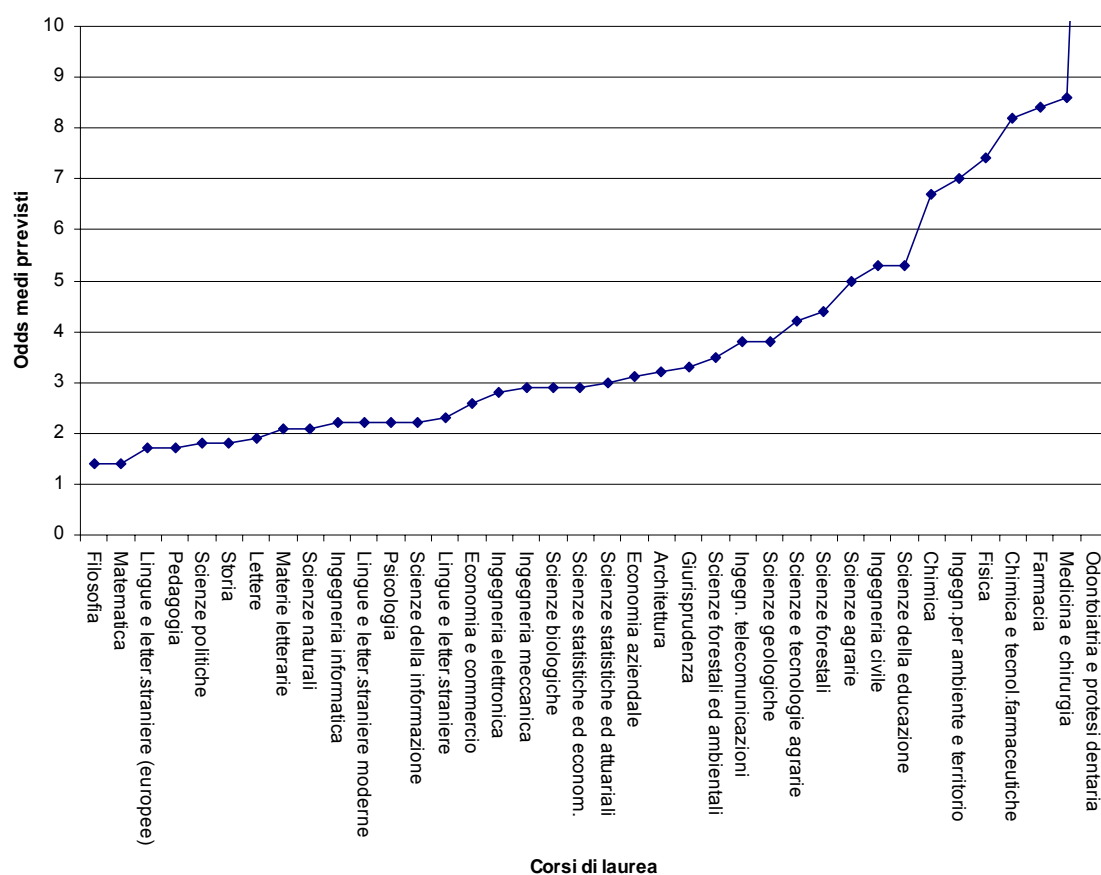
---

<sup>21</sup> In realtà, anche la regressione multilivello consentirebbe, teoricamente, di stimare effetti diversi per diversi corsi di laurea (unità di secondo livello), ma, come già sottolineato nel precedente paragrafo, problemi di convergenza degli algoritmi di ottimizzazione hanno di fatto impedito un approfondimento in tal senso dell'analisi.

L'analisi per gruppi consente di approfondire l'effetto sul fenomeno d'interesse esercitato dalle aggregazioni dei corsi di laurea tramite la stima degli *odds ratio*. In tale sede interessa evidenziare la coerenza di fondo tra i dati rilevati in sede di analisi descrittiva e le indicazioni a cui si perviene attraverso il confronto degli odds ratio per ogni coppia di gruppi: così se i corsi di laurea del gruppo 1 sono associati ad uno scarso uso delle competenze universitarie, i corsi del gruppo 3 si collocano all'estremo opposto, essendo i gruppi 4 e 2 in posizione intermedia.

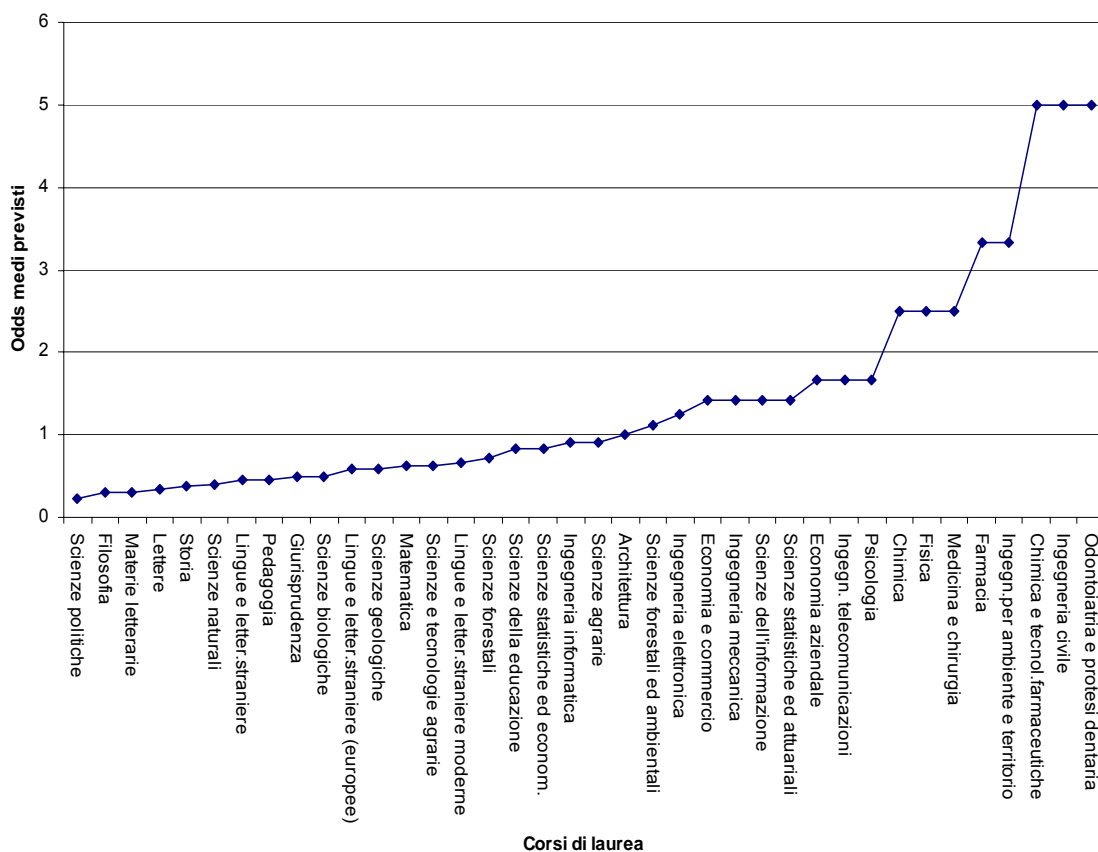
Ciò malgrado, sono evidenti alcune incoerenze di collocazione all'interno dei gruppi per certi corsi di studio: da una parte i corsi in Scienze statistiche e attuariali e Scienze statistiche ed economiche del gruppo 1 e in Scienze dell'educazione del gruppo 4 presentano percentuali di utilizzo elevato delle competenze nettamente maggiori di quanto farebbero presumere gli *odds ratio* stimati per i gruppi di appartenenza; il fenomeno opposto si registra, invece, per i corsi di laurea in Matematica, Scienze politiche e Scienze dell'informazione (appartenenti al gruppo 2) e in Chimica e Ingegneria delle telecomunicazioni (appartenenti al gruppo 3).

**Figura 1.** Odds medi previsti per ogni corso di laurea, stimati sulla base dei modelli logistici ad odds proporzionali



A conclusione dell'analisi svolta si è ritenuto interessante stimare gli odds per ogni singolo corso di laurea, sulla base del modello relativo al gruppo di appartenenza, in modo da creare una graduatoria dei corsi (Fig. 1). I risultati ottenuti sono soddisfacenti: non solo la graduatoria generale è coerente con i dati descrittivi di Tab. 1, ma anche le incoerenze sottolineate poco sopra e relative ad alcuni corsi di studio sono in gran parte eliminate. A titolo esemplificativo, si nota che Matematica scivola al penultimo posto, seguita a poca distanza da Scienze politiche; viceversa, i due corsi di laurea in Scienze statistiche e il corso in Scienze dell'educazione si collocano, rispettivamente, nella fascia media e nella fascia alta della graduatoria.

**Figura 2.** Odds medi previsti per ogni corso di laurea, stimati sulla base del modello logistico multilivello a intercetta casuale



Come ulteriore elemento di confronto, si riporta la medesima analisi svolta stavolta sulla base del modello a intercetta casuale (Fig. 2). Benché a livello generale la disposizione dei corsi di laurea sia analoga a quella rilevata in Fig. 1, si nota che

alcuni corsi occupano zone decisamente diverse nelle due graduatorie; si osservino a tal proposito le collocazioni di Matematica, Psicologia, Scienze dell'informazione, Giurisprudenza, Scienze geologiche e Scienze dell'educazione.

In definitiva, si può affermare che i due approcci adottati per lo studio sulle determinanti dell'uso delle competenze universitarie portano a risultati che soltanto in parte sono paragonabili. In particolare, l'approccio multilivello presenta sì una struttura teorica rigorosa che tiene esplicitamente in considerazione la naturale gerarchia delle unità oggetto d'interesse, ma tale fatto non sembra portare vantaggi sensibili per l'analisi del fenomeno; anzi, la complessità matematica del modello e, conseguentemente, degli algoritmi di ottimizzazione rendono particolarmente difficoltosa la stima di coefficienti casuali, elemento questo indispensabile per discernere in modo accurato tra i diversi effetti delle unità di secondo livello (i corsi di laurea) sulla variabile dipendente.

### Riferimenti bibliografici

- CHIANDOTTO B., BACCI S. (2004) Un modello multilivello per l'analisi della condizione occupazionale dei laureati. In: CROCETTA C. (a cura di) *Modelli statistici per l'analisi della transizione Università-lavoro*, Cleup, Padova: 211-234
- CHIANDOTTO B., BACCI S., BERTACCINI B. (2004) *I laureati e diplomati dell'Ateneo Fiorentino dell'anno 2000: profilo e sbocchi professionali*, Università degli Studi di Firenze, Firenze
- CHIANDOTTO B. (1978) L'analisi dei gruppi. Una metodologia per lo studio del comportamento elettorale (parte I), *Quaderni dell'Osservatorio Elettorale*, **4**, Firenze
- FIELDING A., YANG M., GOLDSTEIN H. (2003) Multilevel ordinal models for examination grades, *Statistical modelling*, **3(2)**: 127-153
- GOLDSTEIN H. (2003) *Multilevel Statistical Models*, Arnold Publishers, London.
- HEDEKER D. (2004) *Multilevel Models for Ordinal and Nominal Variables*, <http://tigger.uic.edu/~hedeker/ml.html>
- McCULLAGH P., NELDER J.A. (1989) *Generalized Linear Models*, Chapman and Hall, London