



Università degli Studi di Ferrara

DIPARTIMENTO DI ECONOMIA, ISTITUZIONI, TERRITORIO

Corso Ercole I D'Este n.44, 44100 Ferrara

Quaderni del Dipartimento

n.11/2002

Aprile 2002

**DISUGUAGLIANZE DI RETRIBUZIONE TRA LAUREATI:
UN'INDAGINE STATISTICA**

DI ROBERTO IORIO

Aprile 2002

**DISUGUAGLIANZE DI RETRIBUZIONE TRA
LAUREATI:
UN'INDAGINE STATISTICA**

DI ROBERTO IORIO*

Abstract

Il presente lavoro indaga, in un campione di laureati, il tema delle disuguaglianze salariali, analizzandone le fonti, le caratteristiche e le possibili motivazioni della loro persistenza. A tale scopo si conduce un'indagine esplorativa sulle determinanti del reddito, si indaga sull'esistenza di differenziali salariali tra lauree diverse, sul differente effetto delle determinanti del reddito tra laureati in facoltà diverse e sulle motivazioni delle scelte in istruzione. Le indagini vengono condotte mediante una pluralità di tecniche statistiche (osservazioni di statistica descrittiva, regressione lineare, logistica binaria, logistica multinomiale, alberi di classificazione con CART).

*Dottorando in Economia presso l'Università di Ferrara, Facoltà di Economia, Dipartimento di Economia, Istituzioni, Territorio. E-mail: iorio@economia.unife.it

Premessa

L'origine di questo lavoro fa riferimento alla Facoltà di Economia dell'Università degli Studi di Trento. Esso è infatti la sintesi e rielaborazione di un rapporto per una ricerca finanziata dal C.N.R. (nell'ambito del progetto "Qualificazione della forza lavoro, terzo settore e opportunità di creazione di posti di lavoro nel contesto europeo"), la cui curatrice è stata la prof.ssa Maria Luigia Segnana (Dipartimento di Economia). A sua volta questa ricerca trae origine dalla mia tesi di laurea, la cui relatrice è stata sempre la prof.ssa Segnana.

I dati analizzati sono stati raccolti e rielaborati da un team coordinato dal prof. Carlo Borzaga (Dipartimento di Economia).

L'utilizzazione della tecnica statistica CART mi è stata suggerita dal prof. Giuseppe Espa (Istituto di Statistica e Ricerca Operativa).

La revisione formale di questo articolo, nonché alcune elaborazioni e suggerimenti bibliografici relativi al paragrafo 6, sono dovuti all'opera preziosissima della dott.ssa Beatrice Valline.

Colgo naturalmente l'occasione per ringraziare vivamente queste persone, senza le quali questo lavoro non sarebbe potuto nascere o almeno non sarebbe stato ciò che è.

Un ringraziamento va anche alla prof.ssa Daniela Del Boca (Università di Torino, Dipartimento di Economia S. Cagnetti de Martiis), che ha avuto modo di leggere questo scritto, apprezzandolo ed invitandomi alla pubblicazione, e soprattutto al prof. Patrizio Bianchi (Università di Ferrara, Dipartimento di Economia, Istituzioni, Territorio) che ha confermato la sua stima nei miei confronti consentendone appunto la pubblicazione.

Parte del contenuto di questo articolo si ritrova in "Accesso al lavoro e salari dei laureati in un contesto ad elevata occupazione: il caso della provincia di Trento" di Carlo Borzaga – Roberto Iorio – Beatrice Valline, pubblicato su *Rivista di Politica Economica*, luglio-agosto 2001, pagg. 245-281.

1. Introduzione

Una cospicua letteratura ha posto in evidenza le considerevoli dimensioni di un fenomeno accaduto negli anni '80 e attenuatosi solo parzialmente negli anni '90: l'aumentato divario salariale tra soggetti con livelli di istruzione differenti, ma anche all'interno di gruppi omogenei rispetto a questo fattore.¹

L'obiettivo di questo lavoro è approfondire il tema delle disuguaglianze salariali mediante un'indagine statistica effettuata su un campione di laureati a Trento e/o residenti a Trento².

Gli obiettivi di questa indagine sono molteplici; l'indagine riguarda infatti:

- le determinanti del reddito;
- i differenziali salariali tra laureati in differenti facoltà;
- il diverso comportamento delle determinanti del reddito in differenti facoltà;
- le motivazioni delle scelte in istruzione.

Il comune denominatore "economico" tra questi temi è rappresentato appunto da un'analisi delle disuguaglianze, con un tentativo di spiegazione del loro mancato rientro, fondato sulle motivazioni delle scelte in istruzione.

Il tema delle disuguaglianze di reddito, tornato, come abbiamo detto, di stringente attualità, è altresì uno dei temi "classici" dell'economia: una domanda di cui si è spesso cercata la risposta è dunque quali ne siano le fonti. La teoria del capitale umano fornisce una spiegazione: il salario è posto in relazione alla produttività, la produttività deriva dal capitale umano accumulato dai soggetti, cioè dall'istruzione ricevuta e dall'esperienza maturata sul posto di lavoro³. Altre teorie hanno messo in discussione il legame tra istruzione e produttività, ma non quello tra istruzione e salari (teoria dello *screening*; *teoria dei segnali*).⁴ Naturalmente non mancano le evidenze empiriche che confermano questa relazione. Ma cosa dire di soggetti che hanno lo stesso livello di istruzione e di esperienza lavorativa? Che queste siano due determinanti del reddito, è ipotesi che non ha bisogno di ulteriori conferme, ma da questo non si può concludere che siano le sole. Infatti, solo se esistono altre determinanti del reddito, possono darsi disuguaglianze anche all'interno di gruppi omogenei per istruzione ed esperienza lavorativa: le disuguaglianze tra gli individui si spiegherebbero con un diverso livello di tali ulteriori determinanti.

Il campione preso in considerazione in questo studio è appunto omogeneo per livello di istruzione, essendo costituito interamente da laureati. Si tratta tuttavia di un gruppo certamente eterogeneo dal punto di vista del reddito conseguito, anche tenuto conto dell'effetto dell'esperienza lavorativa. È dunque necessario individuare le altre fonti della disuguaglianza.

La prima domanda a cui bisogna rispondere è dunque: quali variabili hanno effetto sul livello del reddito post-laurea e in particolare esistono fattori "altri" rispetto all'istruzione in

¹ Tra la vasta letteratura sull'argomento segnaliamo a titolo indicativo Murphy - Welch (1992) sull'incremento dei differenziali tra livelli di istruzione diversa, Karoly (1993) e Juhn - Murphy - Pierce (1993) sull'aumento dei differenziali tra gruppi omogenei.

² Si rinvia all'Appendice A1 per una descrizione completa del campione oggetto di indagine.

³ Il riferimento classico è Becker (1962).

⁴ Si veda, ad esempio, Spence (1974).

grado di influenzarlo? Per rispondere a questa domanda abbiamo suddiviso le determinanti del reddito in categorie concettuali, al fine di favorire una visione d'insieme del problema. Come vedremo, l'indagine condotta, oltre a confermare il ruolo cruciale dell'esperienza lavorativa, conferma l'esistenza di una discriminazione per sesso, rivela l'importanza delle variabili legate alla situazione lavorativa e geografica, mentre un minor rilievo sembrano assumere il capitale familiare e quello formativo extra-universitario.

Abbiamo affermato che il campione analizzato è omogeneo per livello di istruzione, essendo tutti i soggetti laureati. Tuttavia un'indagine sulle disuguaglianze impone di verificare l'ipotesi semplificatrice della teoria del capitale umano, secondo cui *ceteris paribus* a parità di anni di studio (ovvero di livello di istruzione) corrispondono redditi uguali: l'indagine da noi condotta mira invece a verificare se esistono delle differenze salariali derivanti dalla differenziazione qualitativa dell'istruzione universitaria. Cioè, la domanda a cui cerchiamo di rispondere è: laurearsi in una disciplina piuttosto che in un'altra può essere di per sé una fonte di disuguaglianza salariale? Avremo modo di vedere che la risposta a questa domanda è affermativa, con una differenza cruciale tra facoltà "scientifiche" ed "umanistiche".

Il tema dei differenziali salariali per facoltà si presta ad alcuni approfondimenti interessanti. È infatti possibile verificare se le differenti determinanti del reddito hanno un effetto diverso in differenti facoltà o gruppi di facoltà. La nostra indagine individua in particolare un differente effetto dell'esperienza lavorativa nelle facoltà "scientifiche" e "umanistiche". Di ciò cerchiamo di individuare brevemente le ragioni.

I dati raccolti consentono di indirizzare l'analisi in un'ulteriore direzione, verso le motivazioni delle scelte in istruzione. Infatti, i redditi individuali, e dunque i differenziali di reddito, sono generati dall'intreccio delle condizioni di domanda e offerta; secondo la teoria i differenziali tenderebbero a rientrare, tramite un riorientamento delle scelte dei soggetti, interessati alla massimizzazione dei propri guadagni e quindi attratti dalle scelte in istruzione che si prospettano più remunerative. Un'indagine sulle motivazioni delle scelte mostra tuttavia la dominanza delle motivazioni "non economiche" nella scelta della facoltà, una dominanza che però progressivamente diminuisce nelle leve più giovani, molto più orientate da specifiche motivazioni "economiche". Dall'altro lato, permane un vincolo familiare nelle scelte in istruzione degli studenti che, più che realizzarsi attraverso la specifica scelta della facoltà, si effettua in modo indiretto, ovvero sulla scelta della scuola superiore. Infatti, è attraverso quest'ultima scelta e la diffusa applicazione da parte degli studenti della continuità curricolare tra scuole superiori e università che si ritrova, anche nelle scelte universitarie, il ruolo del capitale familiare.

2. Le metodologie di analisi

Una particolarità di questo studio è costituita dall'utilizzazione di tecniche di analisi molteplici. L'osservazione dei dati di statistica descrittiva (medie, mediane, correlazioni) è effettuata spesso a fini interpretativi. Quindi si utilizza una metodologia statistica innovativa, CART, fondata sul concetto degli alberi di classificazione. Questa tecnica di indagine è largamente integrata con la più lungamente collaudata analisi di regressione (lineare, logistica binaria e logistica multinomiale). Soprattutto lo studio sulle determinanti del reddito è condotto mediante un'interazione di queste tecniche.

Analizziamo in dettaglio tali tecniche ed il modo in cui sono state utilizzate.

2.1 La tecnica CART

Nell'analizzare i dati raccolti abbiamo utilizzato una tecnica di recente introduzione: CART. Il termine è un acronimo di *Classification and Regression Trees*, una procedura statistica introdotta da Leo Breiman, Jerome Friedman, Richard Olshen e Charles Stone nel 1984⁵. Si tratta di una procedura non parametrica che genera alberi di classificazione o di regressione. Nel nostro caso, la variabile dipendente, suddivisa in classi, imponeva l'utilizzazione di alberi di classificazione. Si tratta di una tecnica particolarmente indicata per l'analisi esplorativa dei dati e dunque per la selezione delle variabili in vista della costruzione di modelli più tradizionali. Poiché uno degli intenti del nostro lavoro è considerare quali sono le variabili decisive nella determinazione del reddito, da individuarsi tra le numerosissime variabili rilevate, l'utilizzazione di questa tecnica ci è parsa particolarmente indicata.

Il funzionamento di questo tipo di analisi è concettualmente semplice, per quanto si basi su algoritmi matematici complessi.⁶ Un semplice esempio, simile al caso di studio considerato nel presente lavoro, può chiarire quanto detto.

Si supponga che il reddito sia diviso in due classi (basso ed alto). Il *dataset* conterrà un certo numero di individui a reddito basso e i rimanenti a reddito alto: questo è il punto di partenza dell'albero, il cosiddetto nodo - radice. Si consideri ora una delle variabili esplicative del reddito, ad esempio le ore di lavoro. Si prenda un certo valore di questa variabile (ad es. 35 ore settimanali). Utilizzando questa determinazione della variabile il nodo-radice viene diviso (si effettua uno *split*) in due nodi: quello sinistro comprende i soggetti con meno di 35 ore di lavoro settimanali, quello destro i soggetti con più di 35 ore di lavoro. È lecito attendersi che il nodo sinistro veda una prevalenza di soggetti a reddito basso, quello destro una prevalenza di soggetti a reddito alto. I due nodi saranno cioè meno "impuri" di quello di partenza. A loro volta i due nodi ottenuti possono essere ulteriormente divisi in modo analogo, a partire dalla stessa o da altre variabili, finché non si ottengono nodi che non possono più essere ulteriormente divisi, perché tutti i soggetti presenti in tali nodi appartengono ad una stessa classe di reddito o comunque, pur non essendo il nodo omogeneo, non esiste alcun

⁵ Per le nostre analisi abbiamo utilizzato il *software* STATISTICA, modulo *classification tree*, opzione *CART-style univariate split selection method*.

⁶ Per una trattazione sistematica si rimanda a Breiman *et alii* (1984).

variabile in grado di dividerlo in modo da aumentarne l'omogeneità. Ad ogni passo va dunque individuata la variabile, e il valore della variabile, che meglio è in grado di dividere i singoli nodi, cioè che maggiormente riduce l'“impurità” dei singoli nodi, dove l'impurità viene rigorosamente definita con un'opportuna formula matematica.

In modo più formale, possiamo dire che l'impurità è una funzione così definita:

$$(1) \quad i(t) = [p(1|t), p(2|t), \dots, p(j|t), \dots, p(J|t)]$$

essendo:

$i(t)$ = impurità del nodo t

$p(j|t)$ = proporzione di casi che afferiscono alla classe j .

La funzione di decremento dell'impurità del nodo t data dallo *split* s è così definita:

$$(2) \quad i(s,t) = i(t) - p_s \cdot i(t_s) - p_d \cdot i(t_d)$$

essendo:

p_s = proporzione di casi destinati a sinistra

p_d = proporzione di casi destinati a destra

$i(t_s)$ = impurità del nodo di sinistra

$i(t_d)$ = impurità del nodo di destra

Si dice che s^* è lo *split* ottimale del nodo t nei due nodi t_s e t_d se per s^* la (2) è massimizzata, cioè se è:

$$(3) \quad i(s^*,t) = \max_i i(s,t).$$

La definizione di funzione di decremento dell'impurità consente altresì di comprendere meglio il concetto di punto di arresto dell'albero. Quelle individuate sono condizioni - limite. Di fatto viene generalmente definita una regola di arresto. Ad esempio, il nodo t sarà terminale se:

$$(4) \quad \max_i i(s,t) \quad .$$

I concetti fin qui esposti riguardano in generale la creazione di alberi di classificazione. Vi sono poi differenti algoritmi matematici applicabili a questa struttura logica di base. Ogni tecnica è caratterizzata dall'applicazione di determinati algoritmi: CART è una delle tecniche possibili di sviluppo di alberi di classificazione. Una importante caratteristica distintiva di CART è la seguente: abbiamo detto che lo *split* prosegue finché non si raggiungono nodi terminali, cioè quei nodi che non possono essere ulteriormente divisi. In effetti un albero così ottenuto rischierebbe di essere troppo complesso; pertanto CART effettua una “potatura” dei rami dell'albero principale, ottenendo così dei sotto - alberi che vengono confrontati tra loro, al fine di scegliere il migliore in base al criterio detto del minimo costo - complessità.

Formalmente, la misura del costo - complessità è data da:

$$(5) \quad R_c|T| = R|T| + \lambda |T|$$

Essendo:

$R|T|$ = costo di errata classificazione

$|T|$ = complessità dell'albero (numero dei nodi terminali)

Dunque λ è un parametro prefissato di penalizzazione della dimensione dell'albero.

Quello prescelto dev'essere cioè un albero che abbia le qualità, tra cui esiste un *trade-off*, della "snellezza" e della capacità previsiva. Il secondo concetto va chiarito. Un albero di classificazione individua una regola di classificazione, data dalla sequenza delle variabili individuate e dalla corrispondente suddivisione dei casi in esame. Tale regola di classificazione ha potere conoscitivo solo se può essere estesa dai dati in esame a tutti i casi analoghi: applicando la stessa regola di classificazione a casi analoghi si dovrebbe ottenere, a limite, la stessa classificazione. Quanto più le classificazioni ottenute applicando ad altri casi la regola ottenuta nel campione di partenza si distanziano da quella prevista, tanto più debole sarà la capacità previsiva dell'albero. Ora, CART consente di testare questa capacità previsiva dei sotto-alberi, estraendo più sotto-campioni da quello dato; le procedure utilizzate da CART a questo scopo sono quella del campione di apprendimento (*learning sample*) e di *cross-validation*.

La creazione di un albero di classificazione permette dunque di individuare le variabili che maggiormente influenzano la variabile dipendente. È evidente, infatti, che le variabili che riescono a dividere i nodi sono rilevanti nella determinazione della variabile dipendente (queste variabili meglio delle altre dividono i soggetti a reddito basso da quelli a reddito alto).

È dunque possibile effettuare una "classifica" di importanza delle variabili, fondata appunto sulla loro capacità di dividere i nodi. Tuttavia una classifica fondata unicamente sulle variabili prescelte ad ogni passo per l'effettuazione dello *split* potrebbe essere ingannevole: vi possono essere variabili che ad ogni passo sono di poco superate da un'altra e dunque, pur essendo variabili rilevanti, non sono mai privilegiate per l'effettuazione degli *split*. Nel determinare l'importanza delle variabili CART tiene dunque conto anche delle cosiddette variabili "surrogato", cioè di quelle variabili che, se fossero prescelte per lo *split* in luogo di quelle effettivamente scelte, genererebbero un albero di classificazione accurato quasi quanto l'originale.⁷

Esposte brevemente le caratteristiche della tecnica CART possiamo dunque apprezzarne pregi e limiti.

Innanzitutto, come detto, si tratta di una tecnica non parametrica, che dunque non fa nessuna assunzione sulla distribuzione della popolazione, a differenza delle ipotesi di normalità che sono proprie dell'analisi di regressione per poter validamente effettuare inferenze dal campione alla popolazione. L'indagine con CART è inoltre particolarmente indicata per indagini "esplorative", in cui cioè non si abbia ben chiara la natura teorica del fenomeno in analisi e quindi quali siano le variabili che influenzano la variabile dipendente e quali poco o per nulla. Dunque, poiché questo studio si propone, tra l'altro, di stabilire quali siano le determinanti del reddito (almeno tra le variabili rilevate) e quali hanno invece uno scarso effetto, la tecnica CART appare senz'altro preziosa. Una semplice osservazione della "classifica" delle variabili è sufficiente ad effettuare questa indagine. È teoricamente possibile fare girare il programma inserendo tutte le variabili rilevate e fare quindi, in base alla "classifica" ottenuta, una scrematura, salvo poi approfondirla con l'ausilio di nozioni teoriche, individuando, per esempio, quelle variabili che hanno sì una forte relazione con la variabile dipendente, ma con una relazione causale inversa rispetto a quella voluta.

⁷ Vedi Breiman *et alii* (1984), pag. 146 segg.

La stessa cosa è decisamente più problematica con una regressione: il confronto dei coefficienti standardizzati è tecnica grezza e approssimativa, il contributo all'aumento dell'R-quadrato può essere fatto solo per inserimenti successivi di variabili e dunque risente dell'ordine di inserimento.

Un altro vantaggio di CART rispetto alla regressione è quello di non ipotizzare alcuna forma funzionale di relazione tra le variabili; l'arbitrarietà è semmai spostata sulla forma funzionale degli algoritmi matematici.

Un limite di CART rispetto alla regressione può essere considerata la mancata valutazione quantitativa dell'effetto di ogni singola variabile sulla variabile dipendente, cosa che la regressione effettua mediante la stima dei parametri. La "classifica" delle variabili è una valutazione di importanza relativa delle variabili, confrontate tra loro, non un'indicazione sull'effetto che queste hanno sulla variabile dipendente.

Un limite tecnico di CART consiste nel fatto che gli algoritmi che ne sono alla base tendono a sovrastimare le variabili con molte categorie.

Più che un limite, una caratteristica diversa della tecnica CART rispetto alla regressione è che quest'ultima valuta l'effetto di ogni variabile sulla dipendente *ceteris paribus*, cioè essendo tutte le altre variabili costanti; questa condizione non si pone con la tecnica degli alberi di classificazione. Dunque, una regressione si rivela tecnica "superiore" nella misura in cui siamo interessati esclusivamente agli effetti diretti delle singole variabili sulla dipendente.

In effetti, in un *dataset* ricco come quello considerato le variabili correlate e gli effetti indiretti sono numerosissimi e dunque un'indagine che prescindesse da questi elementi risulterebbe avere un debole potere conoscitivo. Perciò ci è parso necessario affiancare un'analisi di regressione a quella basata sugli alberi di classificazione.

2.2 L'analisi di regressione

La regressione è dunque lo strumento che consente di valutare l'effetto di ogni variabile indipendente sulla dipendente tenendo costante l'effetto delle altre variabili. Nelle indagini economiche il tipo di regressione tradizionalmente utilizzata è quella lineare. L'utilizzazione di questa tecnica sarebbe resa problematica dal modo in cui è stata rilevata la retribuzione: essa è stata rilevata per classi, per di più non perfettamente omogenee in ampiezza, specie in riferimento alle classi estreme, rendendo quindi non-normale la distribuzione dei suoi errori, contravvenendo pertanto ad un'ipotesi fondamentale della regressione per la correttezza dell'inferenza statistica.

Tuttavia ci è sembrato che si potesse ugualmente far ricorso a questo tipo di analisi in virtù della considerazione che il numero di classi in cui la rilevazione del reddito è divisa è piuttosto grande (14), tanto da non rendere particolarmente grave lo scostamento dall'ipotesi di normalità degli errori. Una tecnica statistica i cui presupposti non vengono in alcun modo inficiati da questa strutturazione della variabile indipendente è la regressione logistica. Essa presuppone infatti una variabile dipendente dicotomica: in questo caso è agevole "dicotomizzare" la variabile dipendente, assegnando valore 0 a chi è al di sotto di una certa

classe di reddito e 1 a chi è al di sopra; secondo la consuetudine il valore prescelto per la “dicotomizzazione” è stato quello della classe mediana.

Un tipo di indagine che consente di ovviare ai principali limiti delle regressioni lineare e logistica è la regressione logistica multinomiale (o politomica) ordinata, che, affine nei principi alla regressione logistica binaria⁸, ha come variabile dipendente una variabile categorica ordinata, qual è appunto il reddito diviso in classi.

In questo modo è possibile aggirare i principali limiti della regressione lineare (la violazione del principio rigoroso della continuità della variabile dipendente) e della regressione logistica (la “dicotomizzazione” di una variabile che non è stata rilevata in maniera dicotomica).

2.3 L'utilizzazione delle diverse tecniche

In base a quanto detto, si comprende che un'opportuna tecnica di analisi può consistere nell'effettuazione di un'indagine “esplorativa” con CART, che consente di analizzare contemporaneamente una grande quantità di variabili, scartando le variabili irrilevanti, individuando quindi il modello “ottimo” con la regressione, che consente invece di apprezzare gli effetti *ceteris paribus* delle variabili e di stimarne l'effetto dal punto di vista quantitativo.

In questo lavoro abbiamo effettuato un allargamento di questa metodologia di ricerca. Questa è infatti particolarmente adatta quando lo scopo della ricerca è la verifica di particolari aspetti della teoria, quando cioè si tratta di stimare l'effetto di alcune specifiche variabili. Invece lo scopo della nostra indagine, se da un lato è la verifica di alcuni nessi causali individuati in letteratura, dall'altro risiede nell'individuazione *tout-court* delle determinanti del reddito. Per quest'ultimo scopo ci è parso opportuno “rinforzare” la parte esplorativa, utilizzando anche regressione lineare e logistica a tale fine.⁹

Un'interazione delle differenti tecniche è stata utilizzata anche nell'analisi di un altro problema: la valutazione del differente effetto di alcune variabili-chiave in differenti facoltà o gruppi di facoltà. A tale scopo CART è stato utilizzato confrontando il coefficiente di importanza relativa ottenuto da tali variabili nei differenti gruppi in esame, analogamente a

⁸ In tutto il testo, quando parliamo semplicemente di regressione logistica ci riferiamo alla regressione logistica binaria.

⁹ La tecnica con cui abbiamo proceduto è stata la seguente: nella regressione lineare abbiamo valutato l'effetto delle singole variabili sull'*R-quadrato* corretto, nella regressione logistica si è tenuta sotto osservazione la significatività della variazione del *chi-quadrato*. Poichè queste tecniche risentono dell'ordine di introduzione delle variabili, va chiarito l'ordine seguito: dopo aver creato un modello con quelle variabili per cui la teoria, precedenti indagini ed una forte correlazione parziale con la variabile dipendente indicavano una sicura relazione col reddito, si sono creati modelli alternativi, caratterizzati dall'introduzione di una variabile per volta tra quelle rimanenti. Alla fine sono state considerate determinanti del reddito (e si è costruito il modello “finale”) le variabili in grado di alzare l'*R-quadrato* corretto, nella regressione lineare, avendo accettabile significatività, e di comportare una variazione del *chi-quadrato* significativa al 95% nella regressione logistica. Quindi si è effettuato un confronto tra i risultati ottenuti con le diverse tecniche.

Va precisato che è stata effettuata una scrematura “a priori” delle variabili, eliminando quelle troppo soggettive (ad esempio: le motivazioni della scelta della facoltà, rilevate chiedendo ai soggetti di indicare un punteggio a queste associate) e le variabili per cui era palese un'assenza di legame logico col reddito. Questo anche per non appesantire eccessivamente dal punto di vista computazionale il modello.

quanto fatto con la regressione lineare col confronto dei coefficienti. Questa tecnica è stata poi impiegata per un'analisi più sofisticata del problema, con l'utilizzazione di variabili-interazione.

3. Le variabili considerate

3.1 *La variabile dipendente: la retribuzione mensile*

La variabile dipendente presa in considerazione nelle analisi di regressione e logistica e in CART è la *retribuzione mensile*, considerando, come abbiamo visto precedentemente, il valore centrale delle classi in cui è stata rilevata per la regressione lineare, una sua “dicotomizzazione” per la regressione logistica, ed una sua divisione in quattro classi per l'analisi con CART. Ci siamo dunque discostati dalla prassi consueta, che considera la retribuzione oraria. Le ragioni di questa scelta risiedono soprattutto nel modo in cui è stata rilevata la retribuzione, per classi, come abbiamo detto, per cui la retribuzione oraria media può essere ottenuta unicamente dividendo il valore centrale della classe per il numero di ore di lavoro; in tal modo, però, non solo soggetti con valori di reddito diversi possono essere associati a un valore numerico uguale, come accade per la variabile retribuzione a causa della sua divisione in classi, ma anche soggetti con reddito orario uguale possono essere associati a valori diversi, generando quindi una proliferazione degli errori.

3.2 *Le variabili indipendenti*

Gli studi empirici sulle determinanti del reddito sono innumerevoli; innumerevoli sono cioè i tentativi di stimare una “funzione di guadagno”, sia allo scopo di valutare il rendimento dell'istruzione (è in questo contesto che Mincer introduce il termine di “funzione di guadagno”), sia al fine di individuare quali siano gli argomenti di tale funzione e quale sia il loro effetto: è in questa seconda direzione che si muove la nostra analisi.

Psacharopoulos (1987) fornisce un'utile catalogazione delle variabili individuate come rilevanti nella determinazione del reddito.

Egli distingue innanzitutto tra le due macrocategorie delle caratteristiche personali e di quelle ambientali/istituzionali. Le prime a loro volta si dividono in caratteristiche non soggette a cambiamento e quelle per la maggior parte soggette a cambiamento.

Una possibile riclassificazione, sostanzialmente fedele a quella di Psacharopoulos, ma al tempo stesso funzionale ai dati a nostra disposizione, può essere così formulata:

- (1) *caratteristiche individuali*
- (2) *capitale umano*
- (3) *situazione lavorativa*
- (4) *situazione ambientale*

A sua volta la categoria (2) *capitale umano* può suddividersi in sottocategorie:

- (2a) *capitale familiare*
- (2b) *capitale formativo pre-universitario*
- (2c) *capitale formativo universitario*
- (2d) *capitale formativo post-universitario ed extra-universitario*
- (2e) *capitale lavorativo*

L'individuazione di tali determinanti è frutto di analisi empiriche, ma esse non possono non essere filtrate dalla teoria. Senza di essa è impossibile determinare una direzione causale tra reddito e fattori ad esso connessi. Nella previsione degli effetti attesi della nostra analisi si congiungono dunque, necessariamente, i risultati di indagini precedenti e previsioni di tipo più squisitamente teorico.

(1) Caratteristiche individuali (*Sesso - età - età al momento della laurea - stato civile - numero di figli*).

Esse svolgono un importante ruolo di segnalazione.

È stato sovente rilevato un effetto discriminante per *Sesso*, per cui le donne hanno un reddito inferiore a quello degli uomini, effetto distorsivo di una corretta allocazione delle risorse, perché soggetti con pari abilità vengono remunerati diversamente senza nessun'altra ragione che una posizione pregiudiziale. La minore retribuzione femminile può altresì derivare da certi fattori che caratterizzano la carriera lavorativa femminile, soggetta, ad esempio, ad interruzioni per maternità, ecc.¹⁰ La "problematicità" di questa variabile sta nelle sue complesse interazioni con una serie di altri fattori, che si rivelano decisivi nella determinazione del reddito (facoltà scelta, ore di lavoro effettuate) per cui la difficoltà consiste nel valutarne l'effetto *ceteris paribus*: solo quando si accertasse questo si potrebbe concludere, in termini economicamente rigorosi, per l'esistenza di una discriminazione¹¹.

Il ruolo dell'*età* è esplicitamente rilevato dalla teoria del capitale umano: l'effetto sul reddito si struttura secondo una parabola, prima ascendente, poi discendente (Mincer, 1974, Antonelli, 1985); questo andamento deriva in effetti dall'esperienza lavorativa, che, come vedremo, aumenta la produttività dell'individuo; oltre una certa soglia di età, però, il logoramento del capitale umano è superiore al suo incremento dato dall'esperienza. Esperienza lavorativa ed età sono dunque due variabili altamente correlate in un campione di soggetti attivi nel mondo del lavoro, come è quello in analisi, tanto da rendere impossibile isolare gli effetti di una variabile rispetto all'altra (quindi, se l'esperienza lavorativa è esclusa dall'analisi, l'età ne costituisce una *proxy*). Va aggiunto che è presumibile che un'età più alta, a parità di esperienza lavorativa, svolga un effetto-segnalazione negativo, essendo intesa come minore possibilità di "formare" l'individuo. Questo effetto dell'età "in sé", tenendo l'esperienza lavorativa costante, può essere approssimato dall'*età al momento della laurea*, per cui ci si attende dunque un effetto negativo sul reddito.

Ambiguo è il rapporto causale con il reddito dello *stato civile* e del *numero dei figli*: da un lato la presenza di sussidi alla famiglia può determinare un reddito più alto per soggetti coniugati o con molta prole, dall'altro si tratta di variabili correlate con l'età e dunque può essere che l'effetto sul reddito sia unicamente indiretto.

¹⁰ Abburà (1997) parla di rendimenti inferiori per le donne, mentre diversa è la stima di Brunello - Comi-Lucifora (1999)

¹¹ Sul ruolo del sesso come elemento di disuguaglianza nei redditi si veda Oaxaca (1987).

(2) **Capitale umano**

(2a) **Capitale familiare** (*titolo di studio del padre - titolo di studio della madre*)

Numerose indagini hanno messo in rilievo il ruolo dell'ambiente familiare: alcune delle critiche alla teoria del capitale umano si sono basate proprio sul sostenere l'importanza decisiva di questo fattore, ridimensionando quindi il ruolo dell'istruzione. Delle interessanti indagini empiriche relative alla realtà italiana hanno messo in luce i meccanismi con cui questo fattore agisce sui redditi. Lo studio di Checchi - Ichino - Rustichini (1996) mostra che l'effetto della scolarità paterna sulla mobilità economica dei figli è positivo, ma passa attraverso la scolarità dei figli; cioè, più è elevato il titolo di studio dei genitori, più lo sarà quello dei figli e più saranno alti i loro salari. Non è invece significativo l'effetto diretto della scolarità dei padri sui guadagni dei figli, a parità di titolo di studio di questi. Dunque su un campione di soggetti aventi tutti lo stesso titolo di studio ci si attende uno scarso effetto del titolo di studio dei genitori sul reddito. Viceversa, in conformità con la letteratura, ci si attende un forte effetto di questo fattore sulle scelte formative pre-universitarie.

(2b) **Capitale formativo pre-universitario** (*tipo di diploma superiore conseguito - voto di diploma - anno di iscrizione all'università*)

Un'abbondante letteratura¹² illumina gli effetti del tipo e della qualità della formazione pre-universitaria sulle scelte universitarie: i percorsi formativi si pongono generalmente su una linea di continuità tra *tipo di diploma superiore conseguito* e facoltà scelta e solo un elevato *voto di diploma* può convincere il giovane delle possibilità di successo se si discosta da questa linea. Lo stesso voto di diploma è indice di un'abilità che tende quindi a manifestarsi in un voto di laurea più alto. Dunque l'effetto di questo aspetto del capitale formativo sul reddito di un soggetto laureato passa attraverso le scelte e i risultati universitari. Più difficile è prevederne un effetto diretto.

Infine, può essere interessante valutare se ha qualche effetto sul reddito la decisione di rimandare *l'iscrizione all'università*: vi può essere una sorta di "dispersione" del capitale scolastico, ma anche un arricchimento derivante magari da contatti con il mondo del lavoro. È chiaro dunque che l'effetto di questa variabile è fortemente interrelato con altre.

(2c) **Capitale formativo universitario** (*facoltà - gruppo di facoltà - voto di laurea - tempi di laurea - università frequentata - frequenza delle lezioni - periodo di studio all'estero - possesso di altre lauree*).

Abbiamo già accennato alla necessità di verificare la correttezza dell'ipotesi di base della teoria del capitale umano, per cui pari livelli di istruzione, a parità di esperienza lavorativa e di fattori inosservabili o residuali, danno pari livelli di reddito; date invece le differenze di *skills* generate dai diversi corsi di studi è lecito attendersi differenze retributive, connesse alle differenti facoltà presso cui si è conseguita la laurea. Un riferimento empirico in questo senso, relativo alla realtà italiana, è dato da Checchi *et alii* (1999), i quali calcolano i tassi di

¹² Si veda ad es. Checchi *et alii* (1999).

rendimento delle singole facoltà, che risultano considerevolmente diversi gli uni dagli altri. Alla stessa conclusione giungono Brunello - Comi - Lucifora (1999).

Il problema è piuttosto complesso e su di esso ritorneremo ampiamente in seguito. Ora accenniamo soltanto ad una questione. Esiste certamente un grado di insostituibilità tra laureati in diverse discipline, per cui in alcuni rami lavorativi e per alcune professioni vi è una maggiore presenza di laureati in certe facoltà piuttosto che in altre; ora, rami diversi e professioni diverse sono caratterizzate da diverse condizioni di mercato e dunque anche da remunerazioni diverse. Il problema è dunque: le differenze retributive tra laureati in facoltà diverse sono riconducibili unicamente ad una differente presenza in rami e professioni diverse oppure le differenze permangono anche tenendo costanti queste componenti? L'analisi di regressione può consentire di rispondere a questa domanda.

Per quanto riguarda il *voto di laurea* e i *tempi di studio*, in essi si intrecciano concetti legati al capitale umano (voto più alto e tempi più brevi come conseguenza di una maggiore abilità che genererà quindi maggiore produttività: sarebbero questi gli unici indicatori probanti di quell'abilità, appunto, che sfugge a tante rilevazioni) e alla teoria dei "segnali" (si presume che chi ha avuto prestazioni universitarie migliori sia più abile, quindi viene meglio remunerato). In ogni caso, qualunque sia la base teorica alla luce della quale interpretare queste variabili, ci si attende dunque un segno positivo per voti di laurea alti e tempi di laurea brevi¹³.

Molti autori hanno sottolineato l'importanza della qualità della scuola¹⁴. Questo concetto può essere "catturato" dalla variabile *università frequentata*: a fronte di insegnamenti formalmente uguali vi possono essere fattori che "fanno la differenza", come la qualità degli insegnanti o l'eccellenza delle strutture. In termini di qualità dell'istruzione vi può essere anche un aspetto "soggettivo", dato dalla capacità di sfruttare appieno il servizio-istruzione che viene offerto. Un aspetto importante in questo senso è dato dalla *frequenza o meno delle lezioni*: è lecito attendersi che, a parità di tutte le altre condizioni, chi frequenta i corsi abbia una qualità più alta del servizio-istruzione.

Con gli attuali programmi di scambio internazionale, agli studenti è concessa la possibilità di effettuare dei periodi di *studio all'estero*. Dal punto di vista del capitale umano questo rappresenta, in linea di principio, una sicura acquisizione: aumenta la conoscenza della lingua, aumentano le conoscenze generali dell'individuo (una diversa cultura, un diverso sistema scolastico); difficile dire, però, se un capitale umano così "generico" può avere effetto sul reddito.

(2d) Capitale formativo post-universitario ed extra-universitario (*corsi post-laurea - conoscenze linguistiche - conoscenze informatiche*).

Si tratta di conoscenze che possono formarsi dentro o al di fuori delle strutture universitarie e possono essere anche di formazione familiare (ad esempio, madrelingua di lingue diverse dall'Italiano) e che certamente arricchiscono gli *skills* dell'individuo e perciò ci si attende un loro effetto positivo. In effetti è presumibile che alcune di queste conoscenze siano diventate così diffuse (soprattutto la conoscenza dell'Inglese e dell'informatica) da non costituire più un fattore distintivo; cioè bisogna domandarsi fino a che punto siano richiesti

¹³ Si veda Crawford - Johnson - Summers (1997), sia pure nella scuola superiore.

¹⁴ Si veda ad esempio Card - Krueger (1992).

standard veramente alti di conoscenze e non si richieda piuttosto unicamente una soglia minima, trovarsi al di sotto della quale è penalizzante, ma non è fonte di un premio rilevante trovarsi al di sopra.

(2e) Capitale lavorativo (*esperienza lavorativa - periodi di formazione effettuati dall'impresa - stage durante la tesi - lavoro durante gli studi - numero di lavori svolti - continuità lavorativa col lavoro pre-laurea*).

Dell'*esperienza lavorativa* abbiamo già detto parlando dell'età: le due variabili sono così fortemente correlate che è impossibile separarne gli effetti. Il ruolo riconosciuto all'esperienza lavorativa dalla teoria del capitale umano è quella di fornire essenzialmente capitale umano "specifico", cioè prettamente spendibile all'interno dell'impresa e quindi così prezioso per l'impresa stessa da spingerla ad aumentare il salario dei dipendenti esperti; la causa di questo incremento salariale può essere individuata sia nella crescita della produttività, sia nell'aumentare gli incentivi a restare con l'impresa.

Esiste una copiosa letteratura che dibatte gli effetti della formazione effettuata dall'impresa e discute della differenza tra capitale umano "generico" e "specifico". Nell'indagare la relazione tra queste due forme di capitale umano, può essere interessante considerare il *numero di lavori svolti*: confrontando due soggetti che abbiano pari esperienza lavorativa, chi ha svolto un maggior numero di lavori avrà un minore capitale specifico: se il ruolo del capitale specifico è dunque preponderante, ci si potrebbe attendere un effetto negativo sul reddito di questa variabile (in presenza della variabile *esperienza lavorativa*). D'altro canto, un suo segno positivo potrebbe stare ad indicare la trasferibilità del capitale umano fornito dalle imprese.

Infine, l'aver "dimestichezza" col mondo del lavoro può rappresentare un vantaggio: è dunque interessante rilevare l'effetto di tutte quelle variabili che indicano l'esistenza o meno di *contatto col mondo del lavoro* o di un vero e proprio *rapporto di lavoro preesistente* al momento della laurea.¹⁵

(3) Situazione lavorativa (*numero di ore di lavoro - ramo di attività - settore, pubblico o privato, di attività - dimensioni dell'impresa - tipo di professione - numero di ore di lavoro effettuate - tipo di contratto - coerenza del lavoro svolto con gli studi effettuati - utilità della laurea nello svolgimento della professione - modalità con cui si è trovato un lavoro - tempi di ricerca del lavoro*).

Si tratta, come si vede, di una categoria composita, in cui forti sono le correlazioni.

Vi sono *rami* di attività, tipi di attività in cui la remunerazione è mediamente più alta; è da verificare, tuttavia, se questo effetto permane *ceteris paribus*, cioè a prescindere dal tipo di laurea posseduta e dal livello professionale.

Allo stesso tempo, *professioni* ad alto grado di autonomia decisionale (mansioni dirigenziali e lavoro autonomo) implicano un'alta "produttività" del lavoro e pertanto ci si attendono retribuzioni più alte.

¹⁵ L'effetto positivo sul reddito del lavoro nel periodo scolastico, sia pure negli Stati Uniti ed a livello di *High School* è rilevato da Crawford - Johnson - Summers (1997)

Un'altra variabile importante è rappresentata dal tipo di *contratto di lavoro*: essa consente di gettare uno sguardo sugli effetti salariali delle nuove forme di lavoro "flessibile" (ad esempio a tempo indeterminato, contratto interinale, di formazione lavoro) o addirittura illegale (assenza di contratto).

Parimenti ci si attende un effetto sul reddito delle *dimensioni dell'impresa*: una migliore organizzazione del lavoro, che dovrebbe essere propria delle grandi imprese, un *management* più professionale dovrebbero garantire risultati più alti. La *coerenza* del lavoro con gli studi svolti e *l'utilità* della laurea per le mansioni svolte dovrebbero garantire una maggiore produttività del lavoro, quindi salari più alti: si tratta di una variabile con forti implicazioni di politica dell'istruzione¹⁶.

Interessante può essere altresì indagare sugli eventuali effetti sul reddito delle *modalità con cui si è trovato lavoro* e dei *tempi di ricerca del lavoro*. Da un lato, infatti, si può valutare l'importanza dei canali informativi per trovare lavoro, dall'altro è possibile tentare una verifica empirica della teoria del *job search* (un maggiore tempo di ricerca è legato ad un maggiore salario?) Va sottolineato però che, in base a questa teoria, non c'è un legame causale che va dai tempi di ricerca al salario conseguito.

In conclusione, di questa categoria concettuale va detto che essa può assumere un significato interessante se si ipotizza che il trovarsi in una situazione lavorativa favorevole, in senso reddituale, non è dovuto a fattori casuali ma è frutto di abilità individuale (soggetti più abili verranno richiesti da imprese più grandi, svolgeranno un'attività "liberale", svolgeranno mansioni più qualificate, quindi più coerenti con la laurea conseguita).¹⁷

(4) Situazione geografica (residenza - sede del lavoro)

Un contesto economico favorevole significa mediamente salari più alti. Numerosi studi¹⁸ sottolineano l'effetto del fattore geografico in un contesto come quello italiano, caratterizzato da forti squilibri¹⁹.

¹⁶ Variabili di tipo "lavorativo" sono incluse da Bartel - Sicherman (1999) nella loro equazione di rendimenti.

¹⁷ Sul concetto di "abilità lavorativa" e della sua specificità si veda Becker (1962).

¹⁸ Si vedano, ad esempio, Brunello - Comi - Lucifora (1999).

¹⁹ Si noti che il fattore rilevante dovrebbe essere la sede del lavoro, piuttosto che la residenza. Tuttavia la presenza di molti dati mancanti per la prima di queste due variabili ci ha spinto a considerare, nelle regressioni, la seconda, che comunque, nella maggior parte dei casi, coincide con la prima.

4. I risultati dell'indagine: le determinanti del reddito

4.1 I risultati di CART

L'analisi svolta con CART è stata effettuata riclassificando la variabile dipendente in quattro categorie al fine di non appesantire troppo l'elaborazione del modello.²⁰

Si è inoltre tenuto presente che la variabile dipendente è una categoria ordinata e pertanto un errore di previsione per cui, ad esempio, un soggetto appartenente alla prima classe di reddito viene assegnato alla quarta, è più grave dell'attribuzione dello stesso soggetto alla seconda classe. Pertanto si sono introdotti dei costi di errata classificazione.²¹ Questa procedura non è però senza costi aggiuntivi, ed infatti la capacità previsiva di un modello siffatto si è rivelata più bassa rispetto a quella del modello con costi di errata classificazione unitari. In ogni caso i due alberi non sono molto differenti.

Nella Tavola 1 esponiamo l'albero di classificazione ottenuto imponendo costi di errata classificazione unitari. Mostriamo anche il *ranking* di importanza delle variabili, la matrice dei costi di classificazione e la matrice di classificazione.

Leggiamo ora l'albero di classificazione ottenuto (Tavola 1). Le caselle rettangolari rappresentano i nodi dell'albero; i numeri al di sopra di esse indicano il numero dei soggetti inclusi in quel nodo; al di sotto di ogni nodo è indicata la *condizione di split*; essa individua il nodo di sinistra; all'interno di ogni casella il numero in alto a sinistra indica il numero di ogni nodo, quello in alto a destra la classe di reddito predetta.

Considerando, ad esempio, il nodo 5, la regola di classificazione individuata da CART è la seguente: i soggetti che svolgono le professioni 2, 3, 4 o 8 e lavorano più di 34,5 ore alla settimana apparterranno alla terza classe di reddito. Questa "predizione" va poi confrontata con l'osservazione dei dati.

Un altro esempio: soltanto nel nodo 10 la classe prevista è 1. Dunque solo i 94 soggetti appartenenti a questo nodo appartengono, secondo CART, alla prima classe di reddito; cioè tutti e solo i soggetti che svolgono le professioni 1, 5, 6, 7 o 9, lavorano nei rami 1 o 9 e lavorano meno di 24,5 ore settimanali appartengono alla prima classe di reddito. Osserviamo ora la matrice di classificazione (Tavola 2): la somma del primo rigo è appunto 94, tuttavia la prima casella riporta il numero 47: ciò vuol dire che solo 47 dei soggetti con le caratteristiche sopra indicate appartengono effettivamente alla prima classe di reddito.

²⁰ Un modello con due sole classi di reddito ha mostrato errori previsivi troppo alti.

²¹ All'errore di una classe (un soggetto appartenente alla classe n viene classificato come appartenente alla classe $n-1$ o $n+1$) è stato assegnato peso 1,5; all'errore di due classi (un soggetto appartenente alla classe n viene classificato come appartenente alla classe $n-2$ o $n+2$) è stato assegnato peso 5; all'errore di tre classi (un soggetto appartenente alla classe n viene classificato come appartenente alla classe $n-3$ o $n+3$) è stato assegnato peso 15.

TAVOLA 1
Albero di classificazione per tutti i soggetti del campione

Classification Tree for RETRIB4

Number of splits = 6; Number of terminal nodes = 7

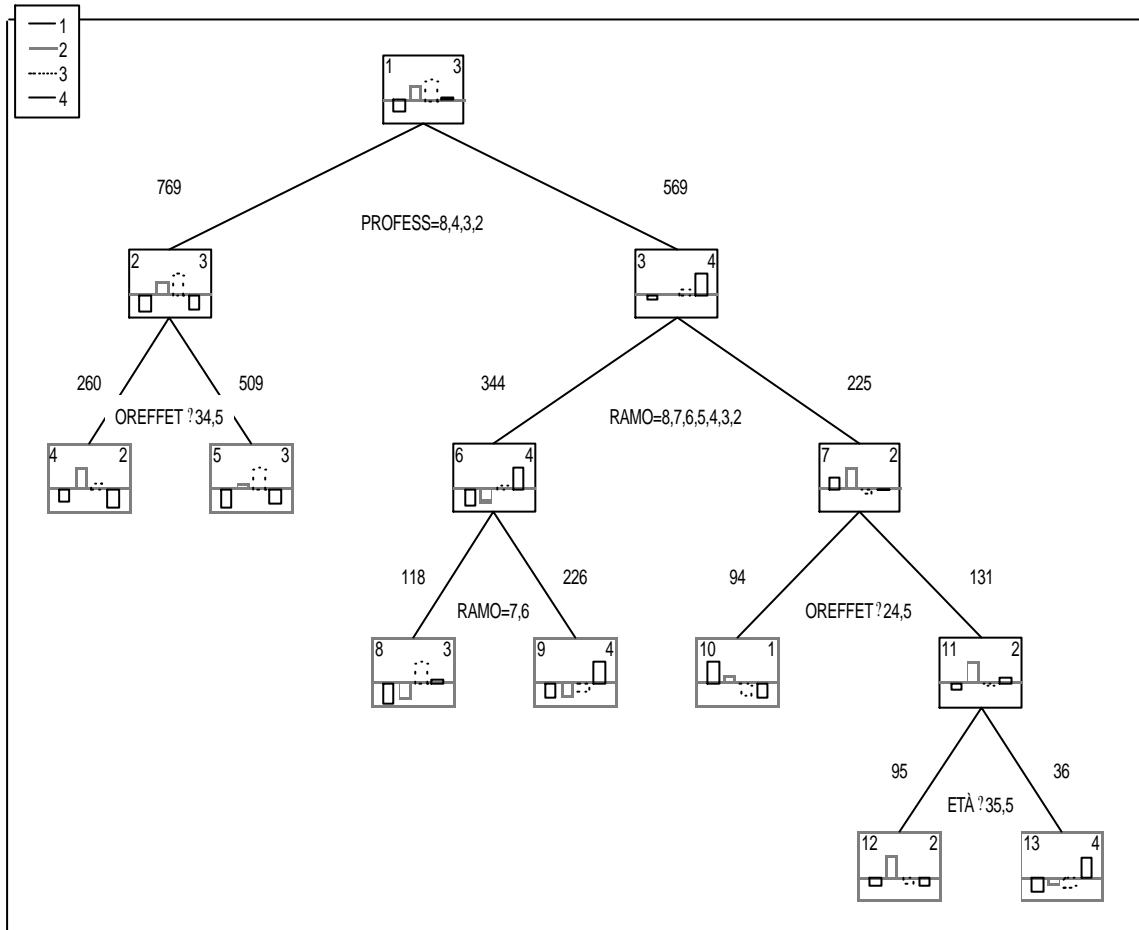


TAVOLA 2
Matrice di classificazione

Campione N = 1338	CLASSI OSSERVATE				
	1	2	3	4	
CLASSI PREDETTE	1	47	30	10	7
2	42	185	104	24	
3	16	173	350	88	
4	23	31	46	162	

TAVOLA 3
Ranking dell'importanza delle variabili indipendenti basato su *split* univariati²²
 0=scarsa importanza; 100=elevata importanza

Numero di ore settimanali effettivamente lavorate (oreffet)	100	Svolgimento di un secondo lavoro (altrolav)	11
Ramo di attività economica (ramo)	94	Conoscenza del francese (francese)	11
Età al momento dell'intervista (età)	86	Attività lavorativa durante l'università (lav_uni)	11
Professione (prof)	85	Studi all'estero (stud_est)	10
Forma del contratto di lavoro (contratto)	73	Conoscenza dell'inglese (inglese)	10
Anni di esperienza lavorativa (esplav)	62	Conoscenza del tedesco (tedesco)	10
Dimensione dell'azienda in cui si lavora (dimens)	54	Periodi di preparazione organizzati dall'azienda (prep_lav)	10
Sesso (sesso)	50	Età alla laurea (etalau)	10
Università dove si è conseguita la laurea (università)	35	Situazione lavorativa attuale (sit_lav)	9
Modalità con cui si è trovato lavoro (trovo)	35	Conoscenza dello spagnolo (spagnolo)	9
Diploma conseguito (diploma)	32	Conoscenza di una 5ª lingua (lingua 5)	8
Facoltà frequentata (facoltà)	28	Voto di maturità (votodipl)	8
Numero di figli (numfigli)	28	Tempo impiegato per conseguire la laurea (tempolau)	8
Stato civile (staciv)	27	Frequenza di un dottorato di ricerca o Phd (postlau3)	7
Sede del lavoro (sede_lav)	26	Frequenza corso post-laurea o master universitari (postlau1)	6
Mesi di ricerca attiva del lavoro (mesi_ric)	25	Superamento di un esame di stato (postlau5)	6
Titolo di studio del padre (tit_padre)	18	Effettuazione di altri studi post-laurea (postlau6)	6
Titolo di studio della madre (tit_madre)	15	Anno di iscrizione all'università rispetto all'anno di diploma (dipl_uni)	5
Frequenza di stage aziendale durante la tesi (lau_stage)	14	Frequenza di corso post-laurea o master non universitari (postlau2)	5
Frequenza con cui si sono seguite le lezioni (freq_lez)	14	Conoscenza di una 6ª lingua (lingua 6)	5
Utilità della laurea per il lavoro svolto (util_lau)	14	Situazione lavorativa attuale rispetto alla precedente (lav_att)	5
Voto di laurea (votolau)	14	Conseguimento della lode (lode)	4
Settore pubblico o privato di attività economica (pub_priv)	13	Effettuazione praticantato o tirocinio per accedere ad ordini professionali (postlau4)	4
Conoscenze informatiche (informatica)	12	Eventuali altre lauree possedute (altrelau)	3
Numero di lavori svolti dopo la laurea (num_lav)	12		

²² Le variabili sono descritte in dettaglio in Appendice A3, disposte in ordine alfabetico secondo la sigla tra parentesi.

Valutiamo ora l'importanza delle variabili nel determinare il reddito (Tavola 3). L'albero ottenuto vede come variabile che effettua il primo *split* il tipo di professione svolta e tuttavia la variabile complessivamente più importante risultano essere le ore di lavoro effettuate.

Le variabili "lavorative" risultano nel complesso molto rilevanti: oltre alle già citate ore di lavoro e professione, sono rilevanti anche esperienza lavorativa, ramo lavorativo, dimensioni dell'impresa, tipo di contratto e modo in cui si trova lavoro.

Tra le caratteristiche individuali risultano assai rilevanti sesso ed età, un po' meno stato civile e numero dei figli.

Nella categoria relativa al "capitale umano" sono piuttosto importanti sia la facoltà che l'università frequentata (ma non ha quasi effetto la riclassificazione di questa variabile per gruppo di facoltà). Anche il tipo di diploma conseguito riveste una notevole importanza.

Rilevante è poi il peso del fattore geografico, mentre, tra le variabili a cui la letteratura o ragionevoli previsioni sembrerebbero attribuire un certo peso, questo non è riscontrato da CART né per il titolo di studio dei genitori né tantomeno per le conoscenze linguistiche e informatiche e per gli studi post-laurea; anche il voto di laurea e i tempi di studio non sembrano variabili che possiedono un grande valore esplicativo. Debole è il peso di tutte le altre variabili prese in considerazione.

L'analisi discriminante è stata utilizzata anche con un'altra finalità: confermare quanto è più rigorosamente riscontrato, come vedremo, con l'analisi di regressione, cioè che alcune variabili rilevanti, particolarmente l'età e l'esperienza lavorativa, hanno un differente effetto nelle diverse facoltà e nei due gruppi di facoltà, umanistiche e scientifiche. A tale scopo si sono creati due alberi di classificazione distinti per i due gruppi. Si è quindi proceduto ad un confronto dei *ranking* delle variabili nei due gruppi. Quanto riscontrato con questa analisi conferma quanto vedremo in seguito: età ed esperienza lavorativa hanno un'importanza relativa inequivocabilmente superiore nelle facoltà scientifiche. Abbiamo condotto anche un'analisi per facoltà, con modelli includenti un minor numero di variabili, a causa della ridotta numerosità dei campioni; anche questa analisi è coerente con la regressione e con quanto detto sopra: in facoltà come Economia e Medicina l'effetto dell'esperienza lavorativa è più forte che a Lettere o Sociologia; unico risultato "anomalo" il debole effetto dell'esperienza lavorativa ad Ingegneria.

4.2 I risultati dell'analisi di regressione

Passiamo ora ad una sintetica disamina dei risultati in base a quanto ottenuto con l'analisi di regressione lineare e logistica, rimandando per una valutazione più approfondita all'esame diretto delle Tavole 4, A2.1, A2.2, A2.3. Ad eccezione di alcuni aspetti secondari, i risultati ottenuti con i due tipi di regressione sono sostanzialmente concordi tra loro. Per quanto riguarda un confronto con quanto previsto dalla letteratura teorica ed empirica, i risultati sono per lo più in linea con questa, soprattutto per quanto riguarda le variabili più indagate, come l'esperienza lavorativa e le caratteristiche individuali. Altre variabili, meno indagate in letteratura, presentano qui un significativo effetto, il cui ruolo, peso e direzione può

essere oggetto di riflessione più che di conclusioni consolidate. Interessante è altresì il non rilevante effetto riscontrato per alcune variabili di cui ci si poteva attendere un effetto importante, come le conoscenze linguistiche ed informatiche e la formazione post-universitaria.

I modelli stimati con la regressione lineare e logistica si rivelano soddisfacenti dal punto di vista della capacità interpretativa: l'R-quadrato della regressione lineare è vicino a 0,60, il che vuol dire che si è riusciti a spiegare più della metà della variabilità dei redditi con le variabili a nostra disposizione. La quota non spiegata va ricondotta alla presenza di fattori non rilevati (ad esempio il livello economico della famiglia di provenienza, che è, come abbiamo visto, una variabile considerata rilevante in molti studi empirici e di cui il livello culturale, rilevato, non è una *proxy* molto precisa), o addirittura non rilevabili (ad esempio l'abilità individuale non spiegata da altri indicatori, come l'abilità negli studi) o casuali (la "fortuna", che qualche autore considera fattore di primaria importanza).²³

La bontà della stima è confermata dalla distribuzione dei residui, che approssima abbastanza bene una distribuzione normale.

In sintesi, le variabili che, secondo la regressione lineare, influenzano significativamente il reddito sono: il sesso, l'età (o l'esperienza lavorativa), lo stato civile, la facoltà frequentata, il voto di laurea, l'età al momento della laurea, le ore di lavoro effettive, la professione svolta, le dimensioni dell'impresa, l'utilità della laurea, la residenza. Con la regressione logistica le variabili individuate sono le stesse, ad eccezione del voto di laurea e dello stato civile, che sono escluse dal modello finale mentre il ramo lavorativo supera le soglie poste. Tutte queste variabili risultano significative in un modello di regressione logistica multinomiale ordinata che comprenda tutte e solo queste.

Esponiamo di seguito (Tavola 4) il risultato della regressione lineare.

TAVOLA 4
Risultati della regressione lineare aggregata²⁴
 Variabile dipendente: *Retribuzione mensile netta (retribuz)*

<i>Variabili</i>	<i>Coefficienti non standardizzati</i>	<i>Errore standard</i>	<i>Coefficienti standardizzati</i>	<i>T</i>	<i>Sig.</i>
(costante)	-249,213	612,572		-,407	,684
SESSO	-286,569	56,832	-,123	-5,042	,000
STATO CIVILE	184,959	54,961	,079	3,365	,001
ETA'	76,776	7,067	,369	10,864	,000
ARCHITETTURA	-683,987	153,485	-,113	-4,456	,000
LETTERE	-184,117	106,293	-,056	-1,732	,084
LINGUE	-226,036	179,473	-,029	-1,259	,208
SCIENZE	-290,898	108,618	-,079	-2,678	,008
AGRARIA	-231,114	190,421	-,028	-1,214	,225
FARMACIA	173,460	248,495	,016	,698	,485
SOCIOLOGIA	-248,901	122,790	-,064	-2,027	,043
INGEGNERIA	-85,465	114,149	-,022	-,749	,454
ECONOMIA	157,513	91,773	,061	1,716	,086
MEDICINA	562,560	134,066	,119	4,196	,000

²³ Si veda Jencks *et alii* (1972)

²⁴ La facoltà di confronto è Giurisprudenza, la residenza di confronto è il Trentino.

4. I risultati dell'indagine: le determinanti del reddito

<i>Variabili</i>	<i>Coefficienti non standardizzati</i>	<i>Errore standard</i>	<i>Coefficienti standardizzati</i>	<i>T</i>	<i>Sig.</i>
ETA' ALLA LAUREA	-59,927	9,148	-,193	-6,551	,000
VOTO DI LAUREA	15,618	4,294	,087	3,637	,000
ORE SETTIM. EFFETT. LAVORATE	34,055	2,831	,294	12,030	,000
PROFESSIONE 1	-2141,108	814,780	-,059	-2,628	,009
PROFESSIONE 2	-1689,063	323,666	-,205	-5,219	,000
PROFESSIONE 3	-1587,785	523,035	-,075	-3,036	,002
PROFESSIONE 4	-1255,134	287,566	-,306	-4,365	,000
PROFESSIONE 5	-968,524	355,638	-,088	-2,723	,007
PROFESSIONE 6	-1300,939	276,821	-,552	-4,700	,000
PROFESSIONE 7	-552,663	283,005	-,138	-1,953	,051
PROFESSIONE 8	-607,119	278,085	-,207	-2,183	,029
DIMENSIONE 1	-123,532	67,110	-,047	-1,841	,066
DIMENSIONE 2	-15,802	61,747	-,006	-,256	,798
RESIDENZA ESTERO	930,906	183,659	,107	5,069	,000
RESIDENZA ALTO ADIGE	182,807	79,164	,050	2,309	,021
RESIDENZA VENETO	-165,871	97,025	-,037	-1,710	,088
RESIDENZA ALTRO NORD	96,521	106,070	,020	,910	,363
RESIDENZA CENTRO-SUD	-725,857	345,233	-,044	-2,103	,036

Numero di casi inclusi nell'analisi: 986

F:44,552

R-quadrato: 0,591

g.d.l. 31

R-quadrato corretto: 0,578

Sig.:0,000

Tra i risultati dell'analisi di regressione alcuni ci sembrano di particolare rilievo.

- 1) Esiste una discriminazione per sesso: anche scontando l'effetto delle ore di lavoro, della laurea conseguita e di tutte le altre variabili rilevanti che possono essere inegualmente distribuite tra i due sessi, la nostra indagine conferma l'esistenza di retribuzioni più basse per le donne.
- 2) L'effetto dell'esperienza lavorativa è conforme a quanto previsto dalla teoria, cioè al suo aumentare il reddito cresce in maniera "parabolica". La regressione esposta nella Tavola 4 include l'età, che risulta avere maggiore potere esplicativo rispetto all'esperienza lavorativa. Tuttavia il termine quadratico relativo all'età non è significativo, probabilmente a causa dell'elevata correlazione col termine di primo grado. Viceversa, una regressione costruita con l'esperienza lavorativa in luogo dell'età (si veda la tavola A2.3) mostra segno negativo, con elevata significatività, del termine quadratico.
- 3) L'effetto delle ore di lavoro non è lineare. Infatti, se si costruisce una regressione avente la retribuzione oraria media come variabile dipendente, la variabile ore di lavoro risulta significativa, con segno negativo. Dunque la retribuzione oraria non è costante al variare delle ore di lavoro e pertanto la relazione tra retribuzione mensile e ore di lavoro non è lineare, ma è concava. Tuttavia né la funzione quadratica né quella logaritmica rappresentano correttamente la relazione tra queste due variabili, che dunque ha verosimilmente una forma complessa.

- 4) Di grande importanza teorica è la presenza della variabile facoltà tra le determinanti del reddito, a dimostrazione di un'eterogeneità dell'offerta formativa. Sull'argomento torneremo più estesamente nel successivo paragrafo.

4.3 Riepilogo e confronto dei risultati

Abbiamo analizzato gli effetti sul reddito delle differenti variabili secondo quanto ottenuto con le due differenti tecniche: analisi di classificazione ed analisi di regressione. È dunque opportuno vedere quali sono i punti in comune e le differenze dei risultati delle due analisi.

C'è un gruppo di variabili su cui entrambi i tipi di analisi convergono nell'assegnare un effetto sul reddito:

- il sesso
- l'età
- la laurea conseguita
- l'esperienza lavorativa
- la professione
- il ramo lavorativo
- il numero di ore di lavoro effettuate
- la residenza

Numerose sono poi le variabili su cui le indagini concordano riguardo all'assenza (o scarsità) di effetto sul reddito. Tra queste si segnala il titolo di studio dei genitori, le conoscenze linguistiche e informatiche, gli studi post-laurea.

Vi sono poi variabili a cui un tipo di analisi attribuisce peso e l'altro no: ad esempio, le regressioni lineare e logistica individuano come variabili significative il voto di laurea, l'età al momento della laurea, l'utilità della laurea per lo svolgimento del lavoro; queste variabili sono valutate come poco rilevanti dall'analisi CART. Il contrario si verifica per il modo in cui si è trovato lavoro, l'università frequentata, il tipo di contratto di lavoro, una delle variabili chiave, quest'ultima, per CART ed invece al di sotto, sia pure di poco, dei valori di soglia si accettabilità nella regressione lineare e logistica.

Come giustificare queste differenze? Abbiamo sottolineato le differenti "filosofie" di analisi delle due tecniche, in particolare l'assenza, nell'analisi CART, della clausola del *ceteris paribus* che è invece l'essenza dell'analisi di regressione. Vi sono dunque variabili che hanno una forte relazione col reddito, ma questa relazione è mediata dall'effetto che queste variabili hanno su altre che hanno un più diretto legame col reddito: l'analisi di classificazione darà peso a queste variabili, cosa che non farà l'analisi di regressione. È quanto accade, ad esempio, con la variabile che esprime le modalità con cui si trova lavoro: queste sono collegate al tipo di lavoro svolto e questo ha un'influenza diretta sul reddito; così l'università prescelta è legata al fattore geografico ed alle facoltà presenti nelle università scelte, il tipo di contratto di lavoro è fortemente legato al numero di ore svolte ed al tipo di professione, variabili che hanno un legame più diretto con il reddito: a parità di queste il tipo di contratto di lavoro cessa di avere un peso rilevante.

Viceversa, ci sono variabili che, pur non avendo una forte correlazione col reddito, una volta scontati tutti gli effetti derivanti dall'intreccio delle correlazioni, risultano avere un peso effettivo sul reddito: è forse in questo modo che si può spiegare la significatività attribuita, mentre CART le considera poco rilevanti, a voto di laurea, età al momento della laurea e utilità della laurea.

5. I risultati dell'indagine: i differenziali salariali

5.1 I differenziali salariali per facoltà e per gruppo di facoltà

Nelle funzioni di guadagno il fattore istruzione è solitamente espresso dal numero di anni di studio, effettivo o “legale” che sia. Ci si attende dunque che due soggetti con lo stesso livello di istruzione formale abbiano lo stesso reddito, fatte salve le differenze derivanti da una diversa quantità degli altri “fattori”. Il nostro obiettivo è analizzare in profondità la veridicità di questa ipotesi. La questione può essere analizzata mediante un’analisi di regressione, ma anche con l’ausilio di osservazioni di statistica descrittiva, che aiutino a rinforzare e corroborare i risultati raggiunti con la regressione. In questa analisi sarà particolarmente utile effettuare un confronto tra facoltà scientifiche ed umanistiche, che ha il pregio di permettere il confronto tra campioni adeguatamente numerosi, vantaggio che si perde nei più analitici e disaggregati confronti tra singole facoltà, che pure effettueremo.

Partiamo dunque da un’osservazione di statistica descrittiva. Se conduciamo, nel nostro *database*, un’analisi differenziata per facoltà, vediamo che esistono notevoli differenze nelle retribuzioni medie: vediamo nella Tavola 5 il valore medio della retribuzione per ogni facoltà.

Tavola 5
Retribuzione mensile ed oraria media e mediana - per facoltà²⁵

FACOLTA'		Retribuzione mensile netta (classe)(retrib)	Retribuzione mensile netta (valore centrale) (retribuz)	Indicatore 2 Retribuzione oraria (ritzora)²⁶
Architettura	Media	6,27	1764,4231	47,7622
	N	52	52	46
	Deviaz. stand.	2,67	811,3978	19,2785
	Mediana	6,00	1650,0000	50,6667
Lettere e Filosofia	Media	6,68	1842,8571	69,6448
	N	189	189	159
	Deviaz. stand.	2,23	663,6571	42,4596
	Mediana	7,00	1900,0000	59,3750
Lingue	Media	6,90	1931,6667	58,5546
	N	30	30	25
	Deviaz. stand.	2,06	695,4883	19,1521
	Mediana	7,00	1900,0000	55,0000
Scienze	Media	7,43	2051,0345	61,1537
	N	145	145	122
	Deviaz. stand.	2,25	656,7351	20,6403
	Mediana	8,00	2150,0000	58,0183

²⁵ Per la precisione, alla facoltà di Economia è stata aggregata la facoltà di Scienze Statistiche, alla facoltà di Sociologia quella di Scienze Politiche, in base dunque a criteri di affinità. La stessa procedura è stata seguita nell’effettuazione delle regressioni, in modo da ridurre il numero delle categorie, eliminando quelle troppo esigue (vedi Tavola A1.4 per la consistenza numerica delle singole facoltà).

²⁶ Questo indicatore (Retribuzione mensile/ore settimanali di lavoro) fornisce una misura approssimata del reddito orario se moltiplicato per 7/30. Si è preferito lasciarlo in questa forma perchè per questa indagine siamo interessati non ai valori assoluti ma ai confronti.

5. I risultati dell'indagine: i differenziali salariali

FACOLTA'		Retribuzione mensile netta (classe)(retrib)	Retribuzione mensile netta (valore centrale) (retribuz)	Indicatore 2 Retribuzione oraria (rtzora)²⁶
Giurisprudenza	Media	7,61	2254,8611	62,8456
	N	144	144	120
	Deviaz. stand.	3,42	1273,1125	48,7136
	Mediana	8,00	2150,0000	55,8536
Agraria	Media	8,56	2568,0000	63,2036
	N	25	25	21
	Deviaz. stand.	2,72	1255,6439	25,0734
	Mediana	8,00	2150,0000	60,0000
Farmacia	Media	8,56	2518,7500	62,8475
	N	16	16	12
	Deviaz. stand.	2,28	1108,7343	32,6480
	Mediana	8,00	2150,0000	51,3889
Sociologia	Media	8,73	2553,9683	72,9029
	N	126	126	109
	Deviaz. stand.	2,48	1048,5343	39,9449
	Mediana	8,00	2150,0000	63,3333
Ingegneria	Media	8,88	2655,8824	60,1523
	N	119	119	103
	Deviaz. stand.	2,60	1207,1148	20,8838
	Mediana	8,00	2150,0000	53,7500
Economia	Media	8,93	2697,4286	65,2754
	N	350	350	303
	Deviaz. stand.	2,64	1263,7470	30,5430
	Mediana	8,00	2150,0000	57,1429
Medicina	Media	11,83	4006,8421	95,8536
	N	95	95	80
	Deviaz. stand.	2,30	1145,4956	31,2023
	Mediana	13,00	4500,0000	100,0000
Totale	Media	8,31	2468,7839	66,7336
	N	1291	1291	1100
	Deviaz. stand.	2,91	1200,5964	35,0066
	Mediana	8,00	2150,0000	59,7222

Vediamo sinteticamente la differenza del reddito medio tra i due gruppi di facoltà nella Tavola 6.

Il problema è dunque: queste differenze possono essere attribuite unicamente ad una differente distribuzione o peso dei fattori rilevanti per la determinazione del reddito? Ad esempio, accertato che l'età è un fattore determinante del reddito, nel campione i laureati in Medicina potrebbero essere mediamente più anziani e la loro superiorità reddituale potrebbe derivare unicamente da questa circostanza. Per smentire questa ipotesi è necessario costruire una regressione, che includa tutte le variabili determinanti del reddito, verificando il livello di significatività della variabile facoltà. È esattamente quanto abbiamo fatto nel paragrafo precedente. Come abbiamo visto, la facoltà risulta una variabile nel complesso significativa. Anche confrontando singole facoltà, è possibile determinare la superiorità di alcune facoltà su altre *ceteris paribus*.

Tavola 6
Retribuzione mensile ed oraria media e mediana - per gruppo di facoltà

<i>FACOLTA'</i>		<i>Retribuzione mensile netta (classe)</i>	<i>Retribuzione mensile netta (valore centrale)</i>	<i>Indicatore 2 retribuzione oraria</i>
Umanistiche	Media	7,49	2152,8689	67,9012
	N	488	488	412
	Dev. Standard	2,80	1020,7659	42,9027
	Mediana	8,00	2150,0000	59,7222
Scientifiche	Media	8,80	2660,7721	66,0344
	N	803	803	688
	Dev. Standard	2,86	1260,2247	29,2901
	Mediana	8,00	2150,0000	58,4350
Totale	Media	8,31	2468,7839	66,7336
	N	1291	1291	1100
	Dev. Standard	2,91	1200,5964	35,0066
	Mediana	8,00	2150,0000	59,7222

È interessante rilevare che l'analisi di regressione sostanzialmente conferma il primato della facoltà di Medicina: tutte le facoltà hanno segno negativo (anche se non sempre statisticamente significativo) se paragonate a questa. Ciò sembrerebbe in realtà coerente con la teoria del capitale umano che individua nel numero di anni di studio il fattore decisivo nella determinazione del reddito: in effetti la facoltà di Medicina è quella che presenta il corso di studi legale più lungo, sei anni, contro Ingegneria che ne prevede cinque e tutte le altre che ne prevedono quattro. Ora, come è noto, in Italia sono una piccolissima percentuale di studenti si laurea entro i termini legali. Si può pensare tuttavia che le facoltà che prevedono tempi legali di studio più lunghi prevedano anche tempi effettivi più lunghi. Invece la Tavola A2.7, che indica i tempi medi e mediani di studio per ogni facoltà, mostra chiaramente che Medicina non è la facoltà per cui si richiedono i tempi di studio più lunghi. Anche i tempi mediamente brevi di Economia, facoltà "ricca", dimostrano che non c'è una relazione inequivoca tra anni di studio e rendimento lavorativo.

Rimane dunque in piedi l'ipotesi da noi formulata, di una differenziazione qualitativa e dunque di produttività, delle differenti lauree che generano *ceteris paribus* reddito diverso.

L'analisi di regressione dovrebbe togliere ogni dubbio in proposito e tuttavia ci sembra opportuno rafforzarne l'analisi e corroborarne le conclusioni. A tale scopo cerchiamo di individuare i fattori che, inegualmente distribuiti, potrebbero essere alla base dei differenziali, cercando di dimostrare che i differenziali persistono anche tenendo conto di questa ineguale distribuzione.

Dall'effetto sull'*R-quadrato*, nella regressione lineare, e dal *ranking* delle variabili, con CART, è possibile individuare le determinanti che maggiormente influiscono sul reddito. Qui ne prendiamo in analisi quattro che risultano particolarmente influenti e che appaiono inegualmente distribuite tra le facoltà:

- a) *il sesso*
- b) *il ramo lavorativo*

- c) *la professione svolta*
 d) *le ore di lavoro*

a) *Il sesso*

Nelle facoltà scientifiche c'è una maggior percentuale di maschi rispetto alle facoltà umanistiche (questo dato è mostrato nelle tavole successive). Proprio perché il sesso è stato individuato dalle regressioni come fattore discriminante, e dunque determinante del reddito, si può "sospettare" che questa differente composizione percentuale per sesso sia alla base dei differenziali salariali tra i due gruppi di facoltà. In effetti se la differenza tra il reddito medio dei due gruppi di facoltà (Tavola 6) è superiore alla differenza del reddito dei soli maschi (Tavola 7), tuttavia tale differenza rimane e per le donne tale differenza rimane consistente (Tavola 8) Lo stesso tipo di dato emerge generalmente dal confronto per singole facoltà: le differenze tra un gruppo di facoltà e l'altro persistono anche prendendo in considerazione soggetti dello stesso sesso.

Tavola 7
Retribuzione mensile ed oraria media e mediana dei soggetti di sesso maschile - per gruppo di facoltà

<i>FACOLTA'</i>		<i>Retribuzione mensile netta (classe)</i>	<i>Retribuzione mensile netta (valore centrale)</i>	<i>Indicatore 2 retribuzione oraria</i>
Umanistiche	Media	8,83	2664,7590	77,3736
	N	166	166	138
	Dev. Standard	2,94	1273,7355	48,9672
	Mediana	9,00	2400,0000	64,4767
Scientifiche	Media	9,33	2876,9309	67,5337
	N	492	492	428
	Dev. Standard	2,74	1327,2298	30,0763
	Mediana	9,00	2400,0000	59,8611
Totale	Media	9,20	2823,4043	69,9328
	N	658	658	566
	Dev. Standard	2,80	1316,2141	35,8181
	Mediana	9,00	2400,0000	60,5556

Tavola 8
Retribuzione mensile ed oraria media e mediana dei soggetti di sesso femminile - per gruppo di facoltà

<i>FACOLTA'</i>		<i>Retribuzione mensile netta (classe)</i>	<i>Retribuzione mensile netta (valore centrale)</i>	<i>Indicatore 2 retribuzione oraria</i>
Umanistiche	Media	6,80	1888,9752	63,1304
	N	322	322	274
	Dev. Standard	2,47	735,8602	38,7245
	Mediana	7,00	1900,0000	56,5789
Scientifiche	Media	7,97	2318,8103	63,5664
	N	311	311	260
	Dev. Standard	2,87	1061,7050	27,8287
	Mediana	8,00	2150,0000	56,5789

FACOLTA'		Retribuzione mensile netta (classe)	Retribuzione mensile netta (valore centrale)	Indicatore 2 retribuzione oraria
Totale	Media	7,38	2100,1580	63,3427
	N	633	633	534
	Dev. Standard	2,73	934,9787	33,8297
	Mediana	7,00	1900,0000	56,5789

b) *Il ramo lavorativo*

Come abbiamo affermato, in questo studio abbiamo stimato una funzione di guadagno. Essa deriva, sia pure in maniera complessa, come avverte Psacharopoulos (1987), dalle condizioni di domanda ed offerta. Consideriamo ad esempio il mercato dei laureati in una data facoltà in uno schema standard di domanda ed offerta: quest'ultima mette in relazione il salario desiderato con il numero di tali laureati disposti a lavorare (nel breve periodo tale numero è fisso), la domanda dipende dalla produttività marginale di tali laureati e dal prezzo del bene venduto. Le teorie sull'istruzione si concentrano sull'offerta e sulla produttività dei lavoratori; il fattore prezzo è invece, in questo contesto, un fattore esogeno e dipende dal mercato in cui l'impresa opera. Dunque questi laureati, se anche avessero gusti uniformi e produttività uguale, guadagnerebbero diversamente a seconda del ramo in cui lavorano. Dunque, le differenze di salario tra diversi laureati potrebbero trovare una spiegazione unicamente in una diseguale distribuzione dei laureati tra i diversi rami. Confrontando "umanisti" e "scienziati", i primi potrebbero essere relativamente più presenti nei rami più "poveri", i secondi nei rami più "ricchi". Anche se i laureati delle due categorie fossero remunerati in modo analogo in ciascun ramo, dalla diseguale distribuzione deriverebbe un vantaggio reddituale complessivo per gli "scienziati".²⁷ L'analisi dei dati farebbe concludere per una compresenza delle due spiegazioni. Si nota che i rami più "ricchi" sono quelli in cui è più forte la presenza di laureati in facoltà scientifiche. Tuttavia all'interno di quasi tutti i rami, e soprattutto in quelli più "poveri" permangono le differenze salariali, soprattutto ai danni dei laureati in Lettere. Si può dunque concludere per una produttività diversa in ciascun ramo delle differenti lauree, con un generale vantaggio per le lauree di tipo "scientifico".

c) *La professione*

Un discorso analogo a quello fatto per il ramo lavorativo può essere fatto per la professione svolta. Da un lato vi sono percorsi di continuità tra facoltà e professione (ad esempio, le professioni autonome possono essere svolte solo da laureati in determinate facoltà), dall'altro differenti professioni sono caratterizzate da differenti livelli di produttività e da condizioni di mercato differenti. È dunque evidente che esistono differenziali tra tipi diversi di professione ma, come è stato detto per il ramo lavorativo, le differenze tra facoltà

²⁷ Naturalmente rimarrebbero da spiegare le ragioni di una tale diversa distribuzione: scelte o non perfetta sostituibilità di laureati in discipline diverse?

potrebbero dipendere unicamente dalla distribuzione dei laureati tra le varie professioni, mentre, analizzando all'interno di un'unica professione, le retribuzioni potrebbero essere simili.

Anche in questo caso l'analisi empirica smentisce questa ipotesi: la distribuzione delle lauree non varia molto di professione in professione, mentre tra coloro che esercitano una medesima professione permane il divario tra le varie lauree.

d) *Le ore di lavoro*

Come si vede nella Tavola 5, la graduatoria delle facoltà secondo la retribuzione mensile è assai diversa, se non per gli estremi, da quella della retribuzione oraria. La stessa differenza tra i due gruppi di facoltà è ridotta moltissimo se si considera la retribuzione oraria. Se si effettua una regressione in cui si include una variabile *dummy* indicante il tipo di facoltà (umanistiche e scientifiche), si può constatare il vantaggio reddituale *ceteris paribus* delle facoltà scientifiche; tuttavia la significatività di tale variabile risulta inferiore al 90%; la significatività ed il peso di questa variabile aumentano in misura considerevole se si escludono le ore di lavoro. Queste considerazioni sembrerebbero dimostrare che buona parte della differenza reddituale dei due gruppi si spiegherebbe col differente numero di ore di lavoro svolte dai laureati nei due gruppi. Questa considerazione sembrerebbe confermata dal fatto che le retribuzioni orarie tra i due gruppi sono molto vicine (vedi Tavola 6), mentre il numero delle ore di lavoro mediamente lavorate nei due gruppi è decisamente diverso (vedi Tavola 9).

Tavola 9
Numero medio di ore settimanali effettivamente lavorate - per gruppo di facoltà

<i>Facoltà</i>	<i>Media</i>	<i>N</i>	<i>Dev. Standard</i>	<i>Mediana</i>
Umanistiche	35,43	417	10,83	38
Scientifiche	41,26	701	9,59	42
Totale	39,09	1118	10,45	40

Si può pensare, cioè, che per tutti i laureati valga un'unica funzione lineare che lega reddito mensile e ore di lavoro; l'unica differenza sarebbe che la posizione media degli "umanisti" si troverebbe su un punto più basso di questa retta rispetto agli "scienziati", essendo più bassa la media di ore di lavoro. La retribuzione sarebbe dunque uguale a parità di ore di lavoro. La retribuzione oraria sarebbe dunque costante rispetto al numero di ore di lavoro.

Tuttavia abbiamo visto in precedenza (par. 4.2) che la retribuzione mensile cresce meno che proporzionalmente rispetto alle ore di lavoro, ovvero che la relazione tra retribuzione oraria e ore di lavoro non è costante, ma decrescente. Se la funzione che lega la retribuzione mensile alle ore di lavoro fosse una sola per tutti i soggetti, una sola sarebbe la relazione tra retribuzione oraria ed ore di lavoro. Di conseguenza, essendo quest'ultima relazione decrescente, a due livelli diversi di ore di lavoro dovrebbero corrispondere due livelli diversi di retribuzione oraria; ma accade esattamente il contrario: gli "scienziati" lavorano mediamente di più degli "umanisti", ma guadagnano all'ora quanto loro. L'unica ipotesi

compatibile con le osservazioni diventa allora la seguente: i due gruppi di facoltà hanno due distinte funzioni, concave, che legano retribuzione mensile ed ore di lavoro; di queste due funzioni, quella degli "scienziati" è più alta rispetto a quella degli "umanisti". Ne consegue che esistono due funzioni, decrescenti, che legano retribuzione oraria ed ore di lavoro, quella degli "scienziati" più alta rispetto a quella degli "umanisti". In tal modo è possibile che, per livelli diversi di ore di lavoro, la retribuzione oraria sia uguale.

Dunque l'ipotesi formulata è che, a parità di ore di lavoro, gli "scienziati" guadagnino di più degli "umanisti" sia mensilmente che all'ora. Il maggiore reddito mensile medio degli "scienziati" deriverebbe quindi in parte dal maggiore reddito orario medio, in parte dal più alto numero medio di ore di lavoro. Il numero di ore di lavoro, dunque, spiega le differenze di retribuzione, ma solo parzialmente.

I dati riportati nelle Tavole 10-11-12 sembrano confermare l'ipotesi formulata. Esse, infatti, mostrano il numero medio delle ore di lavoro al crescere della classe di retribuzione oraria, in aggregato e nei due gruppi di facoltà: come si vede, in tutti e tre i casi sussiste una relazione inversa tra le due variabili, con valori più alti per le facoltà scientifiche.

Tavola 10

Numero medio di ore settimanali effettivamente lavorate per classe di retribuzione oraria - tutti

<i>Retribuzione oraria netta (Rtrzora4: 4 classi)</i>	<i>Media</i>	<i>N</i>	<i>Dev. Standard</i>	<i>Mediana</i>
1	41,78	267	9,41	41,00
2	39,47	281	7,70	40,00
3	38,93	280	8,94	40,00
4	35,70	271	13,65	40,00
Totale	38,97	1099	10,36	40,00

Tavola 11

Numero medio di ore settimanali effettivamente lavorate per classe di retribuzione oraria - facoltà scientifiche

<i>Retribuzione oraria netta (Rtrzora4: 4 classi)</i>	<i>Media</i>	<i>N</i>	<i>Dev. Standard</i>	<i>Mediana</i>
1	43,74	165	7,99	43,00
2	40,84	182	7,58	40,50
3	41,12	164	8,79	40,00
4	39,09	176	12,04	42,00
Totale	41,15	687	9,41	42,00

Tavola 12

Numero medio di ore settimanali effettivamente lavorate per classe di retribuzione oraria - facoltà umanistiche

Retribuzione oraria netta (Rtrzora4: 4 classi)	Media	N	Dev. Standard	Mediana
1	38,62	102	10,63	40,00
2	36,95	99	7,31	38,00
3	35,84	116	8,23	36,00
4	29,44	95	14,30	30,00
Totale	35,32	412	10,84	38,00

Il discorso fin qui svolto può ovviamente ripetersi nell'effettuazione di un'analisi per singole facoltà: la "classifica" della retribuzione oraria risulta assai diversa rispetto a quella della retribuzione mensile, ma questo non deve far pensare che tutte le differenze si spieghino unicamente col differente numero di ore di lavoro.

5.2 Un approfondimento sui differenziali per facoltà

Abbiamo visto che frequentare una facoltà piuttosto che un'altra ha di per sé conseguenze sul reddito guadagnato, indipendentemente dal valore delle altre variabili. Siamo giunti a questa conclusione effettuando un'analisi di regressione applicata all'intero campione a nostra disposizione, constatando che la variabile *facoltà* è significativa anche in presenza delle altre variabili determinanti del reddito.

L'equazione stimata è stata dunque:

$$(1) \quad Y = k + a_1F_1 + a_2F_2 + \dots + a_{n-1}F_{n-1} + b_1X_1 + b_2X_2 + \dots + b_mX_m + u$$

essendo

Y = reddito

k = costante

F_1, F_2, \dots, F_{n-1} = variabili *dummy* indicante la facoltà frequentata (essendo n il numero delle facoltà)

X_1, X_2, \dots, X_m = variabili significative

u = residui.

Un'equazione di questo tipo stima l'effetto del frequentare una facoltà piuttosto che un'altra (tale effetto è espresso dai diversi a_i), ipotizzando che le altre variabili (le diverse X_j) abbiano un effetto identico nelle diverse facoltà.

È possibile verificare l'ipotesi alternativa che le diverse variabili abbiano effetto differente in ogni facoltà. Per fare questo è necessario creare delle "interazioni" tra le *dummy* che indicano la facoltà e le variabili in analisi.

L'analisi condotta in questa direzione è stata a livello più aggregato: abbiamo infatti effettuato il confronto tra i due gruppi di facoltà, umanistico e scientifico.

L'equazione stimata è stata dunque:

$$(2) \quad Y = k + aG + b_1X_1 + b_2X_2 + \dots + b_mX_m + c_1X_1G + c_2X_2G + \dots + c_mX_mG + u$$

essendo:

G = variabile *dummy* indicante il gruppo di facoltà (0 = facoltà umanistica, 1 = facoltà scientifica; per la suddivisione vedi Nota alla Tavola A1.4 in Appendice).

In tal modo ogni variabile ha una quota di effetto “costante” al variare della facoltà (espresso dai diversi b_i) ed una quota variabile (espressa dai diversi c_i ²⁸). Infatti, stimando la (2), se un generico c_i risulta significativamente diverso da 0, allora si può concludere che la variabile generica X_i ha effetto diverso nei due gruppi di facoltà. Se invece c_i è nullo, l'effetto della variabile non è sistematicamente diverso nei due gruppi.

Alternativamente è possibile effettuare questo confronto stimando le due equazioni:

$$(3) \quad Y_1 = k_1 + d_1X_1 + d_2X_2 + \dots + d_mX_m + u$$

$$(4) \quad Y_2 = k_2 + e_1X_1 + e_2X_2 + \dots + e_mX_m + u$$

In tal caso, considerando una generica variabile X_i , se d_i è significativamente diverso da e_i vuol dire che la variabile generica X_i ha un coefficiente diverso nei due gruppi.

Delle due tecniche statistiche, la prima è preferibile, perché è volta a stimare un solo parametro; nel secondo caso è necessario un confronto tra due parametri, più problematico.

Nella stima della (2) un problema statistico è dato dalla compresenza di tutte le variabili “interazione” che sono fortemente correlate tra loro. Appare dunque opportuno stimare equazioni in cui non siano presenti contemporaneamente tutte le interazioni.

Le variabili di cui abbiamo stimato il differente effetto nei due gruppi sono state: l'esperienza lavorativa (o l'età), il sesso, il voto di laurea e l'età al momento della laurea. L'unica variabile per cui il coefficiente della variabile interazione è apparso significativo è stata l'esperienza lavorativa (o l'età): il coefficiente della variabile interazione è positivo, la qual cosa vuol dire che l'impatto dell'esperienza lavorativa è più consistente nelle facoltà scientifiche che in quelle umanistiche e che il reddito è più alto *ceteris paribus* nelle facoltà scientifiche. La Tavola A2.6 mostra i risultati della regressione lineare costruita inserendo la sola interazione dell'esperienza lavorativa (che abbiamo trattato con una funzione logaritmica) con la variabile indicante il gruppo di facoltà. È interessante rilevare che quest'ultima variabile presenta segno negativo: questo starebbe ad indicare una superiorità reddituale delle facoltà umanistiche per livelli nulli di esperienza²⁹, rapporto che poi si rovescia al crescere dell'esperienza lavorativa, proprio a causa del maggiore impatto che questo fattore ha nelle facoltà scientifiche.

Per quanto riguarda le altre variabili, di nessuna si può concludere con certezza se abbia o no un differente impatto nei due gruppi: nessun'altra variabile “interazione” presenta significatività superiori al 95%.

Il confronto dei coefficienti delle regressioni lineari condotte nei due gruppi di facoltà (Tavole A2.4 - A2.5) mostra un risultato coerente con quanto visto in precedenza: il

²⁸ Si ha infatti $c = \partial^2 Y / \partial X \partial G$, cioè c indica il variare dell'effetto di X su Y al variare di G .

²⁹ La variabile *facesplav* è pari a $\ln(\text{anni esperienza lavorativa} + 1)$: abbiamo costruito così questa variabile in modo da non avere casi in cui ci fosse il logaritmo di zero. Inoltre, in tal modo, il coefficiente della variabile indicante il tipo di facoltà indica esattamente la derivata parziale del reddito rispetto a tale variabile quando l'esperienza lavorativa è pari a zero.

coefficiente dell'esperienza lavorativa è più elevato nel gruppo scientifico, anche se i due intervalli di confidenza al 95% hanno una piccola parte in comune.

Come detto, delle altre variabili non si è potuto accertare con sufficiente attendibilità statistica il differente effetto nei due gruppi di facoltà.

Come abbiamo già accennato nel par. 4.1, il differente effetto dell'esperienza lavorativa e dell'età nei due gruppi sembra confermata anche da CART. A tale scopo abbiamo costruito due distinti alberi di classificazione per i due gruppi di facoltà ed abbiamo quindi confrontato il differente effetto delle variabili nei due gruppi. Questo tipo di indagine conferma pienamente i risultati dell'analisi di regressione: età ed esperienza lavorativa hanno un peso molto più consistente nel gruppo scientifico, mentre le differenze non sono così rilevanti per le altre variabili.

Un livello più disaggregato di analisi, che confronti le singole facoltà (almeno le più numerose) conferma il differente effetto dell'età nei due gruppi, con la vistosa eccezione di Giurisprudenza. Lo stesso risultato di massima si ottiene anche con CART, che però registra un risultato anomalo per Ingegneria, dove l'età risulterebbe pochissimo rilevante.

La discriminazione per sesso sembra più consistente nelle facoltà umanistiche. L'elevata età di laurea ha un forte effetto negativo sul reddito particolarmente per i laureati in Ingegneria e Giurisprudenza, mentre il voto di laurea sembra avere molto peso a Sociologia e a Giurisprudenza.

Spostando il livello dell'analisi possiamo ora dare un'indicazione del differente impatto delle determinanti del reddito nei due gruppi, esaminando il reddito medio al variare di queste variabili. Si tratta naturalmente di una tecnica più grossolana dell'analisi di regressione, perché non si valuta il differente impatto di ciascuna variabile *ceteris paribus*, poiché al variare della variabile in osservazione c'è il contestuale (e non necessariamente omogeneo nei due gruppi) variare di tutte le altri determinanti del reddito.

Può trattarsi comunque di un'utile indicazione; nel caso dell'esperienza lavorativa questa analisi conferma i risultati delle tecniche più sofisticate: considerando il numero di anni "mediano" nel campione (6 anni di esperienza, al di sotto dei quali si trova il 49,42% dei soggetti del campione) e confrontando i due sottogruppi, si conferma una più forte crescita per le facoltà scientifiche che per quelle umanistiche.

Nel sottogruppo con esperienza lavorativa da 0 a 6 anni avremo i risultati riportati nella Tavola 13.

Tavola 13
Retribuzione mensile (valore centrale) media e mediana dei soggetti con
esperienza lavorativa da 0 a 6 anni – per gruppo di facoltà

<i>Facoltà</i>	<i>Media</i>	<i>N</i>	<i>Dev. Standard</i>	<i>Mediana</i>
Umanistiche	1923,5023	217	974,8341	1900
Scientifiche	2128,3715	393	803,2796	2150
Totale	2055,4918	610	872,9413	1900

Nel sottogruppo con esperienza lavorativa oltre i 6 anni avremo i risultati riportati nella Tavola 14.

Tavola 14
Retribuzione mensile (valore centrale) media e mediana dei soggetti con
esperienza lavorativa oltre i 6 anni – per gruppo di facoltà

<i>Facoltà</i>	<i>Media</i>	<i>N</i>	<i>Dev. Standard</i>	<i>Mediana</i>
Umanistiche	2407,6419	229	1026,4573	2150
Scientifiche	3225,1412	354	1386,9446	2750
Totale	2904,0309	583	1318,7701	2400

Come si vede agevolmente, la differenza tra i due gruppi di facoltà è molto più marcata tra i soggetti con maggiore anzianità lavorativa e l'ascesa retributiva degli "scienziati" è molto più forte che per gli "umanisti" (per i primi si ha un aumento del 25,17%, nei secondi del 51,55%)

Un'analisi disaggregata per facoltà mostra, riguardo all'influenza dell'età, dinamiche più veloci per Ingegneria ed Economia, per il versante scientifico, e per Giurisprudenza nel settore umanistico; lente invece le dinamiche di Lettere ma anche di Scienze. Se si considera la media delle classi di reddito per le quattro leve nelle sette facoltà più numerose del campione (si vedano le Tavole. A2.8), si nota un forte aumento passando dalla leva più giovane a quella precedente, con un "balzo" la cui entità non si ripete più per le leve successive: questo dato può essere interpretato con l'andamento "a parabola" dell'effetto dell'esperienza lavorativa, forte nei primi anni di lavoro e più debole successivamente. Si noti che in alcune facoltà, come Giurisprudenza e Medicina, il "balzo" è particolarmente drastico; non a caso per i laureati in queste due facoltà sono particolarmente lunghi i periodi di tirocinio, in cui i guadagni sono bassi: le *performance* reddituali della leva più giovane risentono probabilmente di questo effetto. Sono dunque i successivi passaggi (tra seconda e terza e tra terza e quarta leva) a dare probabilmente un quadro più veritiero delle dinamiche retributive "a regime".

Il differente effetto dell'esperienza lavorativa a seconda del tipo di laurea può trovare una spiegazione con il settore in cui i laureati nelle diverse facoltà sono occupati in percentuale maggiore; infatti, come è noto, nel settore pubblico le dinamiche retributive sono più piatte che nel settore privato e questo spiegherebbe i bassi premi all'esperienza per i laureati in Lettere e Scienze, per lo più impegnati nell'insegnamento pubblico.

Ma soprattutto le ragioni del differente effetto vanno cercate in un più forte aumento della produttività per effetto dell'esperienza nelle facoltà scientifiche, e dunque con un processo di *learning-by-doing* più intenso per i laureati in queste facoltà; ovvero, con un più forte calo del livello di sostituibilità con l'aumentare dell'esperienza in facoltà come Ingegneria o Economia: la tendenza alla formazione di figure iperspecializzate, particolarmente accentuata in questi settori, fa sì che le imprese investano fortemente nella formazione al loro interno, in modo tale che un dipendente, al trascorrere degli anni di lavoro, acquista competenze sempre più specialistiche e diversificate rispetto agli altri, tanto da renderlo sempre meno sostituibile e quindi sempre più "prezioso" per l'impresa.

La stessa esigenza di competenze elevate in alcuni settori è forse alla base anche della minore discriminazione per sesso esistente: in un campo come quello ingegneristico o medico in cui si richiedono competenze molto elevate per emergere e segnalarsi agli utenti o alle

imprese, donne altamente competenti subiranno una minore discriminazione rispetto a quanto non accada in settori in cui non c'è uno stesso premio per la competenza, perché le conoscenze richieste sono più standardizzate.

Infatti, anche se, come abbiamo affermato in precedenza, per nessun'altra variabile oltre all'esperienza lavorativa è possibile concludere con certezza che abbia un effetto diverso nei due gruppi di facoltà; tuttavia un confronto del semplice reddito medio per i due gruppi sembra mostrare una discriminazione per sesso decisamente minore nelle facoltà scientifiche, soprattutto per effetto, appunto, di facoltà come Medicina e Ingegneria.

Riportiamo ora i valori medi della retribuzione mensile e oraria per sesso nei due gruppi di facoltà (Tavole 15 e 16); l'osservazione è stata effettuata solo tra i soggetti occupati a tempo pieno, in modo da annullare l'eventuale effetto del part-time maggiormente diffuso tra le donne e che quindi inciderebbe maggiormente nelle facoltà umanistiche, a più forte componente femminile.

Tavola 15

Retribuzione mensile e oraria media e mediana nelle facoltà scientifiche - per sesso

<i>SESSO</i>		<i>Retribuzione mensile netta (valore centrale)</i>	<i>Indicatore 2 retribuzione oraria</i>
Maschi	Media	2876,9309	67,5337
	N	492	428
	Dev. Standard	1327,2298	30,0763
	Mediana	2400,0000	59,8611
Femmine	Media	2318,8103	63,5664
	N	311	260
	Dev. Standard	1061,7050	27,8287
	Mediana	2150,0000	56,5789
Totale	Media	2660,7721	66,0344
	N	803	688
	Dev. Standard	1260,2247	29,2901
	Mediana	2150,0000	58,435

Tavola 16

Retribuzione mensile e oraria media e mediana nelle facoltà umanistiche - per sesso

<i>SESSO</i>		<i>Retribuzione mensile netta (valore centrale)</i>	<i>Indicatore 2 retribuzione oraria</i>
Maschi	Media	2664,7590	77,3736
	N	166	138
	Dev. Standard	1273,7355	48,9672
	Mediana	2400,0000	64,4767
Femmine	Media	1888,9752	63,1304
	N	322	274
	Dev. Standard	735,8602	38,7245
	Mediana	1900,0000	56,5789
Totale	Media	2152,8689	67,9012
	N	488	412
	Dev. Standard	1020,7659	42,9027
	Mediana	2150,0000	59,7222

In questo sottogruppo di soggetti (gli occupati a tempo pieno), nelle facoltà umanistiche il reddito maschile supera quello femminile del 41,07%, nelle facoltà scientifiche solo del 24,07%.

Di grande interesse sarebbe stato rilevare l'andamento nel tempo dei differenziali salariali tra laureati in facoltà o gruppi di facoltà diverse. Purtroppo il reddito è stato rilevato in un unico momento. Tuttavia la presenza di un profilo temporale nei dati, dato dalla compresenza di quattro leve diverse di laureati, ci ha consentito di tentare un esperimento in questa direzione, come illustrato in Appendice B.

6. Le motivazioni delle scelte in istruzione

In un'analisi sulla determinazione dei salari naturalmente non si può non tener conto del lato dell'offerta. Se i redditi individuali, e dunque i differenziali di reddito, sono generati dall'intreccio delle condizioni di domanda e offerta, un modello di scelta razionale prevederebbe un rientro dei differenziali, tramite un riorientamento delle scelte dei soggetti, interessati alla massimizzazione dei propri guadagni e quindi attratti dalle scelte in istruzione che si prospettano più remunerative. Invece la distribuzione dei laureati per percorso formativo e l'evoluzione nel tempo di tale distribuzione non sembrano seguire una strategia economicamente razionale, nel senso che non sembra che le scelte si stiano sistematicamente spostando verso le facoltà che garantiscono redditi più elevati. Un'analisi di questi squilibri è proposta da Martini (2000): confrontando l'offerta e la domanda di laureati per gli anni 1992 e 1998, l'autore rileva squilibri particolarmente marcati (superiori alla media) per i gruppi giuridico, letterario-linguistico-psicologico, politico-sociale e scientifico-farmaceutico. Ciononostante, i laureati di quasi tutti questi gruppi (soprattutto quelli del gruppo giuridico), hanno continuato a crescere più che proporzionalmente tra il 1978 e il 1996. Sono cresciuti invece meno di quanto ci si poteva attendere i laureati in Ingegneria e in Agraria. L'unico gruppo con un tasso di squilibrio basso e un tasso di crescita dei laureati elevato, è quello economico-statistico. Questi comportamenti non sembrano neppure imputabili a carenze informative degli studenti al momento della scelta del percorso formativo (Brunello – Lucifora - Winter-Ebmer, 2001). Per comprendere le ragioni di questi squilibri tra domanda e offerta è necessario approfondire i meccanismi che determinano le scelte formative. Il primo aspetto da considerare è se e fino a che punto le scelte in istruzione siano effettivamente, come ipotizza Becker (1962), scelte di investimento nel reddito futuro: diversi studi empirici hanno concluso per un'effettiva rilevanza sulle scelte in istruzione degli incentivi economici provenienti dal mercato del lavoro, per cui i redditi attesi hanno un effettivo potere di orientamento nella scelte (Mattila, 1982 e Grogger - Eide, 1995, mentre più problematico è il risultato di Arai, 1995), tuttavia la scelta in istruzione presenta evidentemente molte caratteristiche comuni alle scelte di consumo, guidate cioè dal "piacere", ovvero dall'interesse per gli studi. La predominanza di queste motivazioni può fornire una valida spiegazione ai mancati riorientamenti delle scelte in istruzione in una direzione riequilibratrice dei differenziali salariali.

Un'altra ipotesi esplicativa, che trova negli studi empirici effettuati un considerevole supporto, è che la decisione di iscriversi ad una certa facoltà sia non una scelta isolata ma sequenziale, influenzata innanzitutto dal *background* familiare e dalle esperienze formative pre-universitarie, in particolare quelle della scuola superiore³⁰. Più in generale va sottolineata la complessità delle motivazioni sottostanti le scelte in istruzione, che vanno inquadrare nella problematica generale delle scelte umane, che evidentemente dipendono da ragioni più complesse di quelle ipotizzate dalla teoria economica tradizionale (si veda a questo proposito Gambetta, 1996).

³⁰ Checchi (2000) sostiene che i giovani provenienti da famiglie con livelli di istruzione elevati, avendo maggior probabilità di frequentare una scuola superiore di tipo "generalista", hanno successivamente maggiori probabilità di iscriversi all'università.

In questo paragrafo questi temi saranno utilizzati per una verifica empirica sul campione di laureati, utilizzando una metodologia di indagine solitamente estranea alle ricerche economiche, basata cioè su dichiarazioni individuali delle motivazioni di scelta. Ai soggetti intervistati è stato chiesto infatti di assegnare un punteggio (scala di Likert), da 1 a 5, a 14 differenti motivazioni della scelta della facoltà. Dai risultati (esposti nella Tavola 17) si può evincere che quella in istruzione è essenzialmente una scelta “di consumo”: la media dei punteggi assegnati alla motivazione di scelta fondata sull’interesse per gli studi è 4,43, mentre le motivazioni di tipo “economico”³¹ (cioè scelta motivata da prospettive di alti guadagni o dalla maggiore probabilità di trovare lavoro in breve tempo), registrano punteggi decisamente inferiori, rispettivamente 2,35 e 2,83, dimostrando, tra l’altro, che le performance occupazionali di una facoltà sono più apprezzate di quelle reddituali.

È interessante tuttavia rilevare la differenza dei valori riportati da questo tipo di motivazioni nei due differenti gruppi di facoltà: le motivazioni di tipo “lavorativo” risultano sensibilmente più forti nelle facoltà scientifiche, che in effetti garantiscono performance lavorative migliori. Fermo restando dunque l’interesse come ragione primaria della scelta, sembra possibile concludere, pur con tutte le cautele imposte da un siffatto tipo di rilevazione, che le motivazioni “economiche” hanno comunque un loro peso nel processo di scelta: chi è più sensibile a queste motivazioni sceglie tendenzialmente le facoltà più remunerative, (come dimostrerebbe la maggiore concentrazione di soggetti animati da queste motivazioni nelle facoltà scientifiche); *ergo*, c’è almeno una quota di soggetti che presumibilmente sceglie (anche) in base a queste ragioni.

Un’analisi disaggregata per facoltà conferma quanto fin qui detto: le motivazioni “lavorative” sono decisamente più forti nelle facoltà “ricche”; di converso, le motivazioni culturali sono più alte nelle facoltà “povere” (da segnalare l’eccezione di Medicina): chi sceglie Lettere, probabilmente, lo fa nonostante conosca gli scarsi risultati reddituali e occupazionali di questa facoltà, ma il suo livello di interesse è tale da prevalere su altri tipi di motivazione.

Tra le motivazioni di scelta “non economiche” un peso consistente (media 2,73) ha la coerenza degli studi universitari con quelli superiori, un dato che è abbastanza omogeneo tra tutte le facoltà: è dunque improbabile che incentivi occupazionali o retributivi possano condurre a scelte che contraddicano percorsi lineari di istruzione.

Sempre in tema di “vincoli” alla scelta, secondo le dichiarazioni dei soggetti, il volere della famiglia avrebbe un peso non molto forte ma neppure trascurabile (1,94).

Da segnalare, infine, il peso discretamente consistente di una variabile “geografica”, costituita dalla vicinanza della sede della facoltà prescelta (2,31).

Un’osservazione interessante riguarda la dinamica delle motivazioni nelle quattro differenti leve di laureati: a fronte di una generale sostanziale costanza dei valori assegnati alle differenti motivazioni, si nota una crescita piuttosto marcata, andando verso le leve più giovani, delle motivazioni di tipo “lavorativo”: si può forse concludere per una crescita dell’elasticità della domanda di istruzione universitaria agli incentivi economici. Ovvero, questo risultato può essere letto come una probabile conseguenza dell’aumentata dispersione delle prestazioni

³¹ L’uso di questo termine è, in questo ambito, affine al linguaggio comune; sarebbe più corretto parlare, con Becker (1962), di investimento nel reddito futuro. Nello stesso senso si è parlato di motivazioni lavorative.

reddituale e occupazionali tra facoltà: quanto più forti sono i divari, tanto maggior peso hanno le motivazioni economiche nella scelta di una facoltà piuttosto che di un'altra³².

La rilevazione condotta sulle motivazioni soggettive della scelta sembra far concludere per un peso mediamente rilevante dei fattori di continuità nei percorsi di istruzione e per un peso non molto grande del "volere familiare"; resta da verificare se i dati oggettivi sono coerenti con queste dichiarazioni.

A tal fine abbiamo analizzato in primo luogo i percorsi scolastici (tipo di scuola superiore – facoltà universitaria) seguiti dai soggetti del campione. Se la preoccupazione della coerenza fosse poco rilevante, si potrebbero attendere percentuali di scelta universitaria abbastanza analoghi tra i vari tipi di scuola superiore; i dati smentiscono clamorosamente questa ipotesi: la "coerenza" sembra un aspetto estremamente rilevante nella scelta. Si vede quindi che i diplomati in scuole in cui è prevalente la formazione umanistica prediligono scelte universitarie di questo tipo e l'inverso dicasi per scuole in cui prevale una formazione tecnico-scientifica (vedi Tavola 18). Di particolare interesse è la divaricazione delle scelte tra i diplomati al Liceo classico e scientifico, due scuole i cui *curricula* sono in effetti poco distanti, e comunque entrambi tendenti ad una formazione "generalista": ebbene, i primi scelgono in prevalenza (58,7%) facoltà umanistiche, i secondi in più netta prevalenza (74,2%) facoltà scientifiche. Tra i fenomeni notevoli si segnala l'80% di scelta delle facoltà umanistiche per i diplomati all'Istituto magistrale, il 50% di scelta di Architettura dal Liceo artistico, il 55% di scelta di Lingue dal Liceo linguistico, e l'equilibrio nelle scelte per i diplomati dell'Istituto tecnico industriale (in cui il 21,9% dei diplomati ha scelto Sociologia, percentuale assai superiore alla media generale, che è 8,8%).

Sembra di poter concludere, dunque, che i fattori di continuità sono molto rilevanti e quindi la scelta della scuola superiore si rivela di grande peso nella scelta universitaria.³³ Questo ovviamente riduce la flessibilità delle scelte e quindi il possibile effetto di motivazioni economiche. D'altro canto, scelte del tutto scollegate da criteri di continuità con i *curricula* formativi della scuola superiore possono derivare da sfiducia verso la formazione ricevuta. Naturalmente al criterio della continuità sottostanno sia considerazioni legate ad un'autovalutazione della propria abilità (che si ritiene superiore in materie in cui si è ricevuta una formazione migliore) sia all'interesse per gli studi, o perché si sviluppa un interesse

³² In conclusione di questa sezione, è forse opportuna una considerazione sulla validità di questi risultati: oltre a possibili autocensure e difficoltà di autovalutazione delle motivazioni, l'aver richiesto il peso di tutte le motivazioni proposte e non l'indicazione della motivazione maggiormente determinante, genera risultati non univoci: la situazione in cui un soggetto assegna punteggi a due motivazioni (5,1) ed un altro (3,5) genera una media dei punteggi (4,3), che è la stessa che si ha nella situazione in cui un soggetto assegna punteggi (5,4) e l'altro (3,2); però, nel primo caso per un soggetto è più forte la prima motivazione e per l'altro la seconda; nel secondo caso entrambi pongono la prima motivazione come principale. Non è dunque possibile stabilire con certezza, osservando i soli valori medi, per quale percentuale di soggetti una certa motivazione sia la motivazione principale. Queste ed altre difficoltà sono aggirabili con tecniche opportune (ad esempio costruendo indici di "scalabilità") che tuttavia non ci è apparso il caso di utilizzare in questo tipo di indagine, poiché ci pare che il livello di non-smentita dell'ipotesi principale formulata (le motivazioni "non-economiche" nella scelta della facoltà hanno un peso molto consistente) sia comunque sufficientemente elevato.

³³ uno studio dell'Osservatorio del mercato del lavoro (1989) in Trentino l'affinità della scelta scolastica è stata decisiva nella scelta universitaria nel 66,3% dei casi.

maggiore per le discipline più approfondite o perché questo orientamento di interessi già preesisteva alla scelta della scuola superiore.

Un altro tema assai dibattuto nella letteratura economica è l'influenza familiare sulle scelte scolastiche.³⁴ Come abbiamo già avuto modo di dire, il *database* non include le condizioni economiche della famiglia di provenienza, ma solo il titolo di studio dei genitori. Abbiamo visto che questo si dimostra sostanzialmente ininfluenza sulle prestazioni reddituali dei laureati; è da verificare, invece, in che modo influisce sulle scelte in istruzione: dalle dichiarazioni dei soggetti la volontà dei genitori sembra avere poca influenza. Forse questo è vero se ci si riferisce ad una volontà "esplicita", ma se ci riferiamo ai condizionamenti derivanti dall'ambiente familiare, inteso almeno in termini di condizioni culturali, ebbene sembra innegabile una forma di influenza. Se il livello culturale dei genitori fosse ininfluenza sulla scelta della facoltà, ci aspetteremmo una distribuzione dei titoli di studio dei genitori più o meno analoga per le diverse facoltà; invece le differenze sono consistenti, tanto da non apparire puramente casuali: così, il livello culturale medio dei genitori più alto lo si riscontra a Giurisprudenza, Farmacia e Medicina e Chirurgia, il più basso a Scienze Politiche, Psicologia, Sociologia e Medicina Veterinaria (vedi Tavola 19). Non è un caso che le prime tre facoltà siano forse quelle per cui i laureati maggiormente traggono vantaggio da un'attività paterna già avviata, la qual cosa costituisce quindi un sicuro incentivo per la scelta di questa facoltà, un vantaggio che invece è sostanzialmente escluso per Scienze Politiche, Psicologia e Sociologia, facoltà che solo da poco tempo stanno conoscendo una certa fortuna professionale e per cui quindi difficilmente si può contare su tradizioni familiari consolidate. Si è parlato di attività paterna, e non anche materna, non a caso, ma perché il titolo di studio materno è distribuito più uniformemente tra le facoltà: se la spiegazione fornita del titolo di studio paterno mediamente più alto per le facoltà sopra citate è corretta, questa minore influenza del titolo di studio materno si spiega con la maggiore rarità di figure professionali femminili nei campi suddetti.

Ancora più clamorosa è l'influenza del titolo di studio dei genitori sulla scelta della scuola superiore: qui lo scostamento dall'equidistribuzione è relevantissimo ed anzi esiste una vera e propria "gerarchia" dei tipi di scuola, con posizioni sensibilmente staccate tra loro. Il livello del titolo di studio paterno più alto lo si riscontra tra i diplomati al Liceo classico, quindi seguono gli altri Licei (linguistico e scientifico), l'Istituto professionale, l'Istituto magistrale, quindi gli istituti tecnici; una gerarchia quindi coerente con quella che è la percezione media del "livello culturale" delle varie scuole (per cui la scuola "generalista" è superiore a quella "specialistica" e quella "umanistica" superiore a quella "scientifica"). In questo caso l'andamento del titolo di studio materno segue abbastanza fedelmente quello del titolo paterno (vedi Tavola 20).

³⁴ Per dei delucidanti riferimenti su questo tema si vedano Abburrà (1997) e Giannini (1998). Il primo rivela il peso della famiglia nella scelta già della scuola superiore, che è una scelta spesso già orientata ad un proseguimento o meno della carriera scolastica, il secondo sottolinea l'esistenza di elementi di continuità generazionale nelle scelte di capitale umano.

Al termine di questa analisi, si può concludere che esiste una sicura influenza del capitale culturale familiare sulle scelte di istruzione universitaria, non solo o non tanto in via diretta, quanto soprattutto in via indiretta, tramite i condizionamenti sulla scelta della scuola superiore e la relativa rigidità dei percorsi tra scuola superiore e scelta universitaria. In questo contesto, unitamente al dato soggettivo sull'importanza delle motivazioni "di consumo" della scelta universitaria (che a loro volta non sono indipendenti dal tipo di scuola superiore prescelta e dunque dall'influenza familiare, che agisce ovviamente non necessariamente imponendo delle scelte per ragioni di vero o presunto prestigio culturale, ma anche se non soprattutto stimolando interessi autentici nei figli, ad esempio verso una cultura "generalista" piuttosto che specialistica) si comprende come le motivazioni "economiche" non siano prioritarie o, quantomeno, non esclusive.

Tavola 17

**Grado di importanza delle motivazioni nella scelta della facoltà. Valori medi
(1=non importante – 5=molto importante)**

MOTIVAZIONI	Media
Interesse personale per l'oggetto degli studi	4,43
Buone prospettive di trovare facilmente lavoro	2,83
Coerenza con il ciclo di studi precedenti	2,73
Opportunità di trovare un lavoro redditizio e prestigioso	2,35
Presenza della facoltà nella sede universitaria più vicina al luogo di abitazione	2,31
E' stata consigliata come la facoltà più adatta alle capacità	1,94
Volere della famiglia	1,47
Prospettive di carriera consentite dalla laurea per il lavoro svolto in precedenza	1,44
I compagni di scuola/amici erano intenzionati ad iscriversi/ erano già iscritti	1,3
In attesa di trovare occupazione	1,22
Proseguire l'attività già svolta in famiglia	1,21
In attesa di decidere cosa fare	1,16
È la stessa facoltà frequentata dai genitori/fratelli	1,12
In attesa di partire per il servizio civile/militare	1,11

Tavola 18

Tipo di facoltà scelta rispetto alla scuola superiore frequentata (valori %)

	Facoltà umanistiche	Facoltà scientifiche
Liceo classico	58,7	41,3
Liceo scientifico	25,8	74,2
Liceo linguistico	80,0	20,0
Liceo artistico/istituto d'arte	42,9	57,1
Istituto magistrale/scuola magistrale	71,8	28,2
Istituto professionale (4-5 anni)	62,5	37,5
Istituto tecnico commerciale	34,6	65,4
Istituto tecnico per geometri	12,7	87,3
Istituto tecnico industriale	24,4	75,6
10. Altro istituto tecnico	50,0	50,0
Totale	39,0	61,0

Tavola 19
Titolo di studio dei genitori per facoltà di provenienza degli intervistati.
Valore medio (1=nessun titolo – 9=specializzazione post-laurea)

	Titolo di studio del padre	Titolo di studio della madre
Giurisprudenza	5,30	4,50
Lingua e letteratura straniera	5,26	4,81
Farmacia	5,24	4,24
Medicina e chirurgia	5,11	4,15
Ingegneria	4,76	4,15
Scienze matematiche, fisiche e naturali	4,64	3,85
Economia	4,63	3,98
Lettere e filosofia	4,54	3,95
Architettura	4,46	3,91
Agraria	4,41	3,89
Psicologia	4,13	3,63
Sociologia	3,93	3,53
Scienze politiche	3,78	3,44
Totale	4,69	4,03

Tavola 20
Titolo di studio dei genitori per scuola superiore di provenienza degli intervistati.
Valore medio (1=nessun titolo – 9=specializzazione post-laurea)

	Titolo di studio del padre	Titolo di studio della madre
Liceo classico	5,57	4,65
Liceo linguistico	5,10	4,83
Liceo scientifico	5,02	4,29
Istituto tecnico commerciale	4,15	3,74
Istituto magistrale/scuola magistrale	4,11	3,31
Istituto tecnico per geometri	3,85	3,25
Liceo artistico/istituto d'arte	3,57	3,79
Altro istituto tecnico	3,44	3,38
Istituto tecnico industriale	3,40	2,95
Istituto professionale (4-5 anni)	3,38	2,81
Totale	4,69	4,03

7. Conclusioni

L'obiettivo di questa indagine statistica, condotta con una pluralità di metodologie (statistica descrittiva, regressione lineare, logistica binaria, logistica multinomiale, alberi di classificazione con CART), era verificare le fonti delle disuguaglianze di retribuzione in un gruppo di laureati. Cerchiamo di esporre sinteticamente le principali conclusioni a cui siamo giunti.

In primo luogo, le indagini effettuate confermano l'esistenza di molteplici ed eterogenee determinanti del reddito, che sono altrettante fonti di disuguaglianza, al di là del semplice titolo di studio. Tali determinanti possono includersi nelle categorie delle caratteristiche individuali (sesso, età), del capitale umano (laurea conseguita, voto di laurea, età al momento della laurea), del capitale lavorativo (esperienza lavorativa) e della situazione lavorativa (ramo di attività, tipo di professione svolta, dimensioni dell'impresa, ore di lavoro effettuate, grado di utilità della laurea). Di altri fattori, considerati rilevanti da altri studi o dal "sentire comune" (come il titolo di studio dei genitori, la conoscenza della lingua inglese e dell'informatica) non è stato invece possibile accertare l'effetto positivo sul reddito.

Un altro importante risultato raggiunto nella nostra analisi rivela la parzialità dell'ipotesi tradizionale della teoria del capitale umano, secondo cui ciò che conta, riguardo all'istruzione, per la determinazione del reddito è il numero di anni di studio ovvero il livello del titolo raggiunto, indipendentemente dal tipo di studi effettuati. Il risultato ottenuto mostra che frequentare una facoltà piuttosto che un'altra, ovvero un tipo di facoltà (umanistica o scientifica) piuttosto che l'altro, ha di per sé conseguenze sul reddito conseguito, indipendentemente dal livello delle altre variabili.

Questo risultato conferma l'osservazione, effettuata in numerosi studi, di differenze salariali all'interno di gruppi formalmente omogenei per istruzione, rivelando dunque un processo in atto di differenziazione qualitativa dell'istruzione.

Un ulteriore risultato rivela che le differenze reddituali tra i due gruppi sembrano derivare almeno in parte dal diverso effetto che le determinanti del reddito hanno nei due gruppi di facoltà; ad un grado accettabile di attendibilità statistica è l'effetto dell'esperienza lavorativa che comporta, al suo aumentare, una crescita reddituale più forte nelle facoltà scientifiche rispetto a quelle umanistiche.

Per quel che riguarda infine le motivazioni sottostanti le scelte in istruzione, un altro fondamentale aspetto della teoria del capitale umano, i risultati ottenuti mostrano che le motivazioni di tipo "economico", per cui le scelte in istruzione sono una forma di investimento, non sembrano essere le più rilevanti nella scelta universitaria: le motivazioni "non economiche", cioè una scelta guidata dall'interesse per gli studi, e i vincoli derivanti dalla continuità della carriera scolastica svolgono un ruolo determinante. Rilevante è pure l'influenza sulle scelte scolastiche del "capitale culturale" familiare.

APPENDICE

A1. Descrizione del campione

I dati utilizzati per la ricerca empirica, i cui risultati sono riportati nel Capitolo Quinto, sono tratti dagli esiti di un questionario postale effettuato dall'Università di Trento nel 1998. L'indagine ha interessato le seguenti categorie di soggetti:

1. i laureati presso l'ateneo trentino residenti in provincia di Trento;
2. i laureati presso gli Atenei di Verona, Padova, Venezia, Milano e Bologna residenti in provincia di Trento;
3. i laureati presso l'ateneo trentino non residenti in provincia di Trento.

All'interno di queste categorie è stato selezionato un campione costituito dai laureati in quattro anni accademici: 1982/83, 1986/87, 1990/91, 1994/95.

Il numero di questionari spediti è stato 2468. Di questi sono stati restituiti compilati 1402, pari ad un tasso di risposta del 58,5%.

Descriviamo ora la composizione del campione utilizzato (i soggetti che hanno risposto) secondo le variabili più significative.

Tavola A1.1
Composizione del campione per sesso

<i>Sesso</i>	<i>N</i>	<i>%</i>
Maschi	702	50,1
Femmine	700	49,9
Totale	1402	100

Tavola A1.2
Composizione del campione per leva

<i>Leva</i>	<i>N</i>	<i>%</i>
82/83	203	14,5
86/87	233	16,6
90/91	360	25,7
94/95	606	43,2
Totale	1402	100

Tavola A1.3
Composizione del campione per residenza e ateneo

Ateneo e residenza	N	%
Trentini a TN	570	40,7
Non trentini a TN	301	21,5
Trentini non a TN	530	37,8
Totale	1401	100

Tavola A1.4
Composizione del campione per facoltà

Facoltà	N	%
Architettura	46	3,3%
Lettere e Filosofia	196	14,0%
Lingue e Letterature straniere	31	2,2%
Scienze matematiche, fisiche e naturali	153	10,9%
Scienze statistiche	5	0,4%
Giurisprudenza	180	12,8%
Scienze politiche	9	0,6%
Psicologia	8	0,6%
Agraria	27	1,9%
Farmacia	17	1,2%
Sociologia	123	8,8%
Ingegneria	124	8,8%
Economia	372	26,5%
Medicina veterinaria	4	0,3%
Medicina e chirurgia	107	7,6%
Totale	1402	100,0%

Tavola A1.5
Composizione del campione per gruppo di facoltà

Gruppo facoltà	N	%
Umanistiche	547	39
Scientifiche	855	61
Totale	1402	100

Nota: con il termine di *facoltà umanistiche* abbiamo indicato: Lettere e Filosofia, Lingua e Letterature straniere, Giurisprudenza, Scienze Politiche, Sociologia;
con il termine di *facoltà scientifiche* abbiamo indicato: Architettura, Scienze matematiche, fisiche e naturali, Scienze statistiche, Agraria, Farmacia, Ingegneria, Economia, Medicina veterinaria, Medicina e chirurgia.

Tavola A1.6
Valore medio, mediano e deviazione standard delle principali variabili considerate, in aggregato e nei due gruppi di facoltà

<i>Facoltà</i>		<i>Retribuzione mensile netta (classi da 1 a 14) (retrib)</i>	<i>Retribuzione mensile netta (valore centrale) (retribuz)</i>	<i>Indicatore 1 retribuzione oraria (retrora)</i>	<i>Indicatore 2 retribuzione oraria (rtrzora)</i>
Umanistiche	Media	7,49	2152,8689	2,3578	67,9012
	N	488	488	412	412
	Deviaz. stand.	2,80	1020,7659	1,4054	42,9027
	Mediana	8,00	2150,0000	2,2000	59,7222
Scientifiche	Media	8,80	2659,5149	2,2214	66,0184
	N	804	804	689	689
	Deviaz. stand.	2,86	1259,9441	,8083	29,2718
	Mediana	8,00	2150,0000	2,1429	58,3333
Totale	Media	8,31	2468,1502	2,2724	66,7230
	N	1292	1292	1101	1101
	Deviaz. stand.	2,91	1200,3475	1,0729	34,9925
	Mediana	8,00	2150,0000	2,1429	59,7222

<i>Facoltà</i>		<i>Età</i>	<i>Voto di laurea</i>	<i>Età alla laurea</i>	<i>Anni per conseguire la laurea</i>	<i>Ore settimanali effettiv. lavorate</i>	<i>Anni di esperienza lavorativa</i>
Umanistiche	Media	34,85	105,66	28,38	7,54	35,43	9,37
	N	542	535	542	516	417	464
	Deviaz. stand.	7,06	5,70	5,63	3,28	10,83	8,53
	Mediana	33,00	108,00	26,00	7,00	38,00	7,00
Scientifiche	Media	33,64	102,79	26,67	7,06	41,25	7,38
	N	858	850	858	838	702	775
	Deviaz. stand.	5,16	6,98	2,91	2,29	9,59	6,27
	Mediana	33,00	104,00	26,00	7,00	42,00	6,00
Totale	Media	34,11	103,90	27,33	7,24	39,08	8,13
	N	1400	1385	1400	1354	1119	1239
	Deviaz. stand.	6,00	6,66	4,26	2,71	10,45	7,26
	Mediana	33,00	106,00	26,00	7,00	40,00	6,00

Tavola A1.7
Correlazioni tra le variabili principali (vedi Legenda variabili in A3)

		<i>RETRIBUZ</i>	<i>RETRDIC</i>	<i>SESSO</i>	<i>ETÀ</i>	<i>VOTOLAU</i>	<i>ETÀLAU</i>	<i>ESPLAV</i>	<i>OREFFET</i>
RETRIBUZ	Corr. di Pearson	1,000	,622	-,301	,367	-,022	,036	,316	,435
	Sig. (2 code)	,	,000	,000	,000	,425	,202	,000	,000
	N	1292	1292	1292	1290	1276	1290	1194	1100
RETRDIC	Corr. di Pearson	,622	1,000	-,232	,270	-,027	,018	,236	,332
	Sig. (2 code)	,000	,	,000	,000	,338	,512	,000	,000
	N	1292	1292	1292	1290	1276	1290	1194	1100
SESSO	Corr. di Pearson	-,301	-,232	1,000	-,126	,209	-,087	-,074	-,295
	Sig. (2 code)	,000	,000	,	,000	,000	,001	,009	,000
	N	1292	1292	1402	1400	1385	1400	1239	1119
ETÀ	Corr. di Pearson	,367	,270	-,126	1,000	-,056	,679	,765	-,034
	Sig. (2 code)	,000	,000	,000	,	,036	,000	,000	,262
	N	1290	1290	1400	1400	1383	1400	1238	1117
VOTOLAU	Corr. di Pearson	-,022	-,027	,209	-,056	1,000	-,154	,018	-,108
	Sig. (2 code)	,425	,338	,000	,036	,	,000	,527	,000
	N	1276	1276	1385	1383	1385	1383	1224	1105
ETÀLAU	Corr. di Pearson	,036	,018	-,087	,679	-,154	1,000	,488	-,076
	Sig. (2 code)	,202	,512	,001	,000	,000	,	,000	,011
	N	1290	1290	1400	1400	1383	1400	1238	1117
ESPLAV	Corr. di Pearson	,316	,236	-,074	,765	,018	,488	1,000	-,006
	Sig. (2 code)	,000	,000	,009	,000	,527	,000	,	,859
	N	1194	1194	1239	1238	1224	1238	1239	1043
OREFFET	Corr. di Pearson	,435	,332	-,295	-,034	-,108	-,076	-,006	1,000
	Sig. (2 code)	,000	,000	,000	,262	,000	,011	,859	,
	N	1100	1100	1119	1117	1105	1117	1043	1119

A2. Tavole fuori testo

Tavola A2.1
Risultati della regressione logistica binaria aggregata³⁵
 Variabile dipendente: classi di retribuzione «dicotomizzate» (retdic)

Variabile	B	S.E.	Wald	gdl	Sig	R	Exp(B)
(Costante)	-,8931	2,0800	,1844	1	,6676		
Sesso	-,6177	,2079	8,8265	1	,0030	-,0725	,5392
Età	,2328	,0283	67,6640	1	,0000	,2247	1,2621
FACOLTA'			22,4642	10	,0129	,0435	
Architettura	-1,6156	,9361	2,9786	1	,0844	-,0274	,1988
Lettere	-,6413	,8586	,9579	1	,4551	,0000	,5266
Lingue	-2,1851	1,0299	4,5017	1	,0339	-,0439	,1125
Scienze	-,7893	,8471	,8684	1	,3514	,0000	,4541
Giurispr.	-1,1495	,8516	1,8222	1	,1771	,0000	,3168
Agraria	,0456	1,1274	,0016	1	,9678	,0000	1,0466
Farmacia	-1,1639	1,1191	1,0816	1	,2983	,0000	,3123
Sociologia	-,2563	,8567	,0895	1	,7648	,0000	,7739
Ingegneria	-,7587	,8640	,7711	1	,3799	,0000	,4683
Economia	-,1943	,8353	,0541	1	,8160	,0000	,8234
Etàlau	-,1639	,0380	18,5819	1	,0000	-,1129	,8488
Oreffet	,0842	,0124	46,0168	1	,0000	,1840	1,0879
RAMO			40,7788	8	,0000	,1380	
Ramo 1	-1,1892	,4644	6,5565	1	,0105	-,0592	,3045
Ramo 2	-,8048	,5328	2,2818	1	,1309	-,0147	,4472
Ramo 3	-,7063	,9799	,5196	1	,4710	,0000	,4935
Ramo 4	-,9090	,5942	2,3403	1	,1261	-,0162	,4029
Ramo 5	,6757	,4791	1,9890	1	,1584	,0000	1,9654
Ramo 6	-,3264	,4984	,4290	1	,5125	,0000	,7215
Ramo 7	-,1246	,4980	,0626	1	,8024	,0000	,8828
Ramo 8	-,0311	,8062	,0015	1	,9692	,0000	,9694
PROFESSIONE			29,4196	8	,0003	,1016	
Prof 1	-6,7966	22,2891	,0930	1	,7604	,0000	,0011
Prof 2	-3,4796	1,5622	4,9612	1	,0259	-,0477	,0308
Prof 3	-1,7740	2,0595	,7420	1	,3890	,0000	,1696
Prof 4	-,8169	1,3238	,3808	1	,5372	,0000	,4418
Prof 5	-1,3208	1,4901	,7856	1	,3754	,0000	,2669
Prof 6	-,6613	1,2974	,2598	1	,6102	,0000	,5162
Prof 7	-,1552	1,3239	,0137	1	,9067	,0000	,8562
Prof 8	,9035	1,3576	,4428	1	,5058	,0000	2,4681
DIMENSIONE			11,0227	2	,0040	,0735	
Dimens 1	-,8170	,2595	9,9120	1	,0016	-,0780	,4417
Dimens 2	-,4744	,2277	4,3411	1	,0372	-,0424	,6223
UTILITA' LAUREA			15,1154	2	,0005	,0925	
Util_lau 1	-1,3466	,4170	10,4279	1	,0012	-,0805	,2601
Util_lau 2	-,6488	,2122	9,3507	1	,0022	-,0752	,5227
RESIDENZA			23,5038	5	,0003	,1019	
Trentino	-2,5790	,9506	7,3609	1	,0067	-,0642	,0759
Alto Adige	-1,6413	,9854	2,7743	1	,0958	-,0244	,1937
Veneto	-3,1970	1,0036	10,1479	1	,0014	-,0792	,0409
Altrnord	-2,7763	1,0213	7,3888	1	,0066	-,0644	,0623
Censud	-4,5551	1,6559	7,5671	1	,0059	-,0654	,0105

Numero di soggetti considerati: 998

Chi-quadrato: 502,955

³⁵ In questa analisi il termine di confronto per la facoltà è Medicina, per la residenza è l'estero. Nelle successive regressioni il termine di confronto per la facoltà è Giurisprudenza, per la residenza è il Trentino.

Indice correct: 80,86%
Indice sensitivity: 87,97%

g.d.l.: 39
Sign.: 0,0000

Tavola A2.2 - Risultati della regressione logistica multinomiale ordinata aggregata
Variabile dipendente: Retribuzione mensile netta (valore centrale)(retribuz)

	Coef.	Err. Stan.	z	P> z	[95% Int. di confid.]	
secco	-.680513	.1400323	-4.860	0.000	-.9549714	-.4060546
statciv	.3548829	.1295903	2.738	0.006	.1008906	.6088752
età	.216395	.0182336	11.868	0.000	.1806578	.2521321
architeta	-1.6216	.3828626	-4.235	0.000	-2.371997	-.8712031
lettere	-.4712015	.2721143	-1.732	0.083	-1.004536	.0621326
lingue	-1.006676	.4227162	-2.381	0.017	-1.835185	-.1781676
scienze	-.6092509	.2704233	-2.253	0.024	-1.139271	-.079231
agraria	-.2940725	.5124672	-0.574	0.566	-1.29849	.7103447
farmacia	.117983	.6435153	0.183	0.855	-1.143284	1.37925
sociol	-.5344531	.3018471	-1.771	0.077	-1.126063	.0571564
ingegn	-.2974153	.289505	-1.027	0.304	-.8648347	.270004
economia	.3179647	.2246496	1.415	0.157	-.1223405	.7582699
medicina	1.023422	.4105432	2.493	0.013	.2187723	1.828072
votolau	.0345601	.0102659	3.366	0.001	.0144393	.0546809
etàlau	-.1479019	.022767	-6.496	0.000	-.1925245	-.1032793
oreffet	.1049557	.008371	12.538	0.000	.0885488	.1213627
ramo 1	-.7721605	.3082394	-2.505	0.012	-1.376299	-.1680224
ramo 2	.1370895	.2930242	0.468	0.640	-.4372273	.7114063
ramo 3	-.4633241	.6105931	-0.759	0.448	-1.660065	.7334164
ramo 4	-.5408691	.397354	-1.361	0.173	-1.319669	.2379303
ramo 5	-.1184628	.2947742	-0.402	0.688	-.6962096	.459284
ramo 6	-.3295658	.3265018	-1.009	0.313	-.9694975	.310366
ramo 7	-.0194948	.3231996	-0.060	0.952	-.6529545	.6139648
ramo 8	.4016884	.5205114	0.772	0.440	-.6184951	1.421872
prof 1	-5.846704	1.684791	-3.470	0.001	-9.148833	-2.544575
prof 2	-3.707847	.8047229	-4.608	0.000	-5.285075	-2.130619
prof 3	-3.148412	1.49955	-2.100	0.036	-6.087476	-.2093475
prof 4	-2.637278	.7424229	-3.552	0.000	-4.0924	-1.182156
prof 5	-2.120773	1.012001	-2.096	0.036	-4.104258	-.1372883
prof 6	-2.284662	.7130362	-3.204	0.001	-3.682187	-.8871363
prof 7	-1.009956	.7309121	-1.382	0.167	-2.442517	.422606
prof 8	-.9768769	.7131933	-1.370	0.171	-2.37471	.4209563
dimens 1	-.5763956	.1649095	-3.495	0.000	-.8996124	-.2531788
dimens 2	-.2476781	.1439989	-1.720	0.085	-.5299107	.0345545
utillau1	-.3938731	.2692288	-1.463	0.143	-.921552	.1338057
utillau2	-.3267587	.1388853	-2.353	0.019	-.598969	-.0545485
estero	2.553007	.4508849	5.662	0.000	1.669289	3.436726
altoadig	.7579312	.1911599	3.965	0.000	.3832648	1.132598
veneto	-.7277243	.2300743	-3.163	0.002	-1.178662	-.276787
altrnord	.1949034	.2666958	0.731	0.465	-.3278108	.7176176
censud	-2.980491	1.012585	-2.943	0.003	-4.965122	-.9958597

Numero di soggetti considerati: 982

Chi-quadrato: 1018,87

g.d.l.: 41 Sign.: 0,0000

Pseudo R-quadrato: 0,2193

Tavola A2.3
Risultati della regressione lineare con effetti dell'esperienza lavorativa
 Variabile dipendente: *Retribuzione mensile netta(valore centrale) (retribuz)*

<i>Variabili</i>	<i>Coefficienti non standardizzati</i>	<i>Errore standard</i>	<i>Coefficienti standardizzati</i>	<i>T</i>	<i>Sig.</i>
(Costante)	1540,433	651,539		2,364	,018
SESSO	-315,423	59,982	-,136	-5,259	,000
STATCIV	228,393	57,295	,098	3,986	,000
ARCHITETTURA	-690,887	163,268	-,115	-4,232	,000
LETTERE	-149,851	113,332	-,046	-1,322	,186
LINGUE	-294,870	193,214	-,038	-1,526	,127
SCIENZE	-217,940	115,488	-,060	-1,887	,059
AGRARIA	-92,966	200,691	-,011	-,463	,643
FARMACIA	229,007	255,861	,021	,895	,371
SOCIOLOGIA	-228,236	130,902	-,059	-1,744	,082
INGEGNERIA	-46,128	122,456	-,012	-,377	,706
ECONOMIA	164,589	97,744	,064	1,684	,093
MEDICINA	777,937	139,752	,162	5,567	,000
VOTOLAU	12,853	4,564	,071	2,816	,005
ETÀLAU	-28,955	8,432	-,093	-3,434	,001
OREFFET	31,341	2,979	,270	10,522	,000
ESPLAV	71,724	7,893	,447	9,087	,000
ESPQUAD	-,920	,162	-,239	-5,665	,000
PROFESSIONE 1	-1781,782	844,399	-,050	-2,110	,035
PROFESSIONE 2	-1672,002	357,207	-,193	-4,681	,000
PROFESSIONE 3	-1572,971	547,042	-,077	-2,875	,004
PROFESSIONE 4	-1361,667	313,365	-,331	-4,345	,000
PROFESSIONE 5	-1105,246	379,578	-,103	-2,912	,004
PROFESSIONE 6	-1445,398	302,483	-,614	-4,778	,000
PROFESSIONE 7	-660,470	309,053	-,167	-2,137	,033
PROFESSIONE 8	-681,758	303,844	-,233	-2,244	,025
DIMENS 1	-174,886	70,630	-,066	-2,476	,013
DIMENS 2	-17,975	65,322	-,007	-,275	,783
ESTERO	951,513	206,945	,103	4,598	,000
ALTO ADIGE	132,425	83,085	,036	1,594	,111
VENETO	-222,153	102,506	-,050	-2,167	,030
ALTRO NORD	134,991	111,444	,028	1,211	,226
CENTRO-SUD	-555,685	353,970	-,035	-1,570	,117

Numero di casi inclusi nell'analisi: 928

R-quadrato: 0,571

R-quadrato corretto: 0,556

F: 37,242

g.d.l.: 32

Sig.:0,000

Tavola A2.4
Risultati della regressione lineare relativa alle facoltà scientifiche
 Variabile dipendente: *Retribuzione mensile netta (valore centrale) (retribuz)*

<i>Variabili</i>	<i>Coefficienti non standardizz.</i>	<i>Errore standard</i>	<i>Coefficienti standardizz.</i>	<i>t</i>	<i>Sig.</i>	<i>Intervallo di</i>	<i>Confidenza al 95%</i>
(Costante)	2558,530	878,974		2,911	,004	832,033	4285,026
SESSO	-311,957	77,506	-,125	-4,025	,000	-464,196	-159,719
STATO CIVILE	270,286	76,054	,113	3,554	,000	120,899	419,673
VOTO LAUREA	4,952	5,365	,028	,923	,356	-5,586	15,490
ETÀ LAU	-38,187	14,932	-,078	-2,557	,011	-67,516	-8,858
OREFFET	30,607	4,337	,235	7,057	,000	22,089	39,126
LNESP_1	487,238	54,228	,305	8,985	,000	380,723	593,754
PROFESSIONE 2	-1614,399	409,050	-,198	-3,947	,000	-2417,861	-810,937
PROFESSIONE 3	-1340,201	667,782	-,065	-2,007	,045	-2651,869	-28,533
PROFESSIONE 4	-1046,481	361,836	-,243	-2,892	,004	-1757,205	-335,758
PROFESSIONE 5	-785,625	442,269	-,076	-1,776	,076	-1654,338	83,088
PROFESSIONE 6	-1264,372	345,669	-,523	-3,658	,000	-1943,340	-585,404
PROFESSIONE 7	-485,301	350,999	-,130	-1,383	,167	-1174,739	204,137
PROFESSIONE 8	-419,856	348,946	-,137	-1,203	,229	-1105,262	265,549
RAMO 1	-838,604	134,471	-,295	-6,236	,000	-1102,735	-574,474
RAMO 2	375,474	171,556	,080	2,189	,029	38,501	712,447
RAMO 3	-739,163	258,468	-,087	-2,860	,004	-1246,851	-231,475
RAMO 4	-518,866	175,546	-,101	-2,956	,003	-863,676	-174,056
RAMO 5	-563,226	135,066	-,166	-4,170	,000	-828,525	-297,926
RAMO 6	-539,638	131,076	-,174	-4,117	,000	-797,100	-282,175
RAMO 7	-522,499	129,722	-,174	-4,028	,000	-777,301	-267,696
RAMO 8	-764,121	257,391	-,090	-2,969	,003	-1269,692	-258,549
DIMENSIONE 1	-327,251	93,860	-,121	-3,487	,001	-511,612	-142,890
DIMENSIONE 2	-13,437	87,510	-,005	-,154	,878	-185,326	158,452
ESTERO	858,090	295,368	,083	2,905	,004	277,923	1438,257
ALTO ADIGE	164,122	108,597	,043	1,511	,131	-49,186	377,430
VENETO	-236,977	134,430	-,050	-1,763	,078	-501,026	27,072
ALTRO NORD	-9,734	164,665	-,002	-,059	,953	-333,171	313,704
CENTRO-SUD	-1688,912	823,301	-,058	-2,051	,041	-3306,053	-71,771

Numero di casi inclusi nell'analisi: 588

R-quadrato: 0,573

R-quadrato corretto: 0,552

F: 26,823

g.d.l.: 28

Sig.:0,000

Tavola A2.5
Risultati della regressione lineare relativa alle facoltà umanistiche
 Variabile dipendente: *Retribuzione mensile netta (valore centrale) (retribuz)*

<i>Variabili</i>	<i>Coefficienti non standardizz.</i>	<i>Errore standard</i>	<i>Coefficienti standardizz.</i>	<i>t</i>	<i>Sig.</i>	<i>Intervallo di</i>	<i>Confidenza al 95%</i>
(Costante)	3021,587	1281,679		2,358	,019	499,697	5543,477
SESSO	-337,721	96,162	-,151	-3,512	,001	-526,934	-148,508
STATO CIVILE	95,095	93,007	,045	1,022	,307	-87,909	278,099
VOTO LAUREA	10,031	8,161	,051	1,229	,220	-6,028	26,089
ETÀ LAU	-6,626	9,807	-,032	-,676	,500	-25,923	12,670
OREFFET	30,084	4,649	,298	6,471	,000	20,937	39,231
LNE SP_1	203,208	64,576	,162	3,147	,002	76,145	330,271
PROFESSIONE 1	-3820,202	1116,754	-,198	-3,421	,001	-6017,578	-1622,826
PROFESSIONE 2	-4467,849	878,790	-,461	-5,084	,000	-6196,997	-2738,702
PROFESSIONE 3	-5424,065	1113,234	-,281	-4,872	,000	-7614,516	-3233,615
PROFESSIONE 4	-3509,532	796,549	-,966	-4,406	,000	-5076,859	-1942,205
PROFESSIONE 5	-3974,591	908,409	-,355	-4,375	,000	-5762,017	-2187,164
PROFESSIONE 6	-3275,535	789,643	-1,526	-4,148	,000	-4829,272	-1721,798
PROFESSIONE 7	-2511,657	801,101	-,551	-3,135	,002	-4087,940	-935,374
PROFESSIONE 8	-2615,751	785,655	-1,022	-3,329	,001	-4161,642	-1069,860
RAMO 1	44,809	193,667	,021	,231	,817	-336,259	425,876
RAMO 2	-128,304	183,620	-,035	-,699	,485	-489,604	232,996
RAMO 3	767,337	777,185	,040	,987	,324	-761,889	2296,563
RAMO 4	753,873	272,924	,157	2,762	,006	216,856	1290,890
RAMO 5	284,704	192,457	,125	1,479	,140	-93,983	663,390
RAMO 6	42,305	240,807	,010	,176	,861	-431,517	516,128
RAMO 7	344,069	263,813	,070	1,304	,193	-175,020	863,159
RAMO 8	540,190	429,209	,056	1,259	,209	-304,341	1384,721
DIMENSIONE 1	-112,102	115,987	-,046	-,967	,335	-340,322	116,119
DIMENSIONE 2	-97,503	105,232	-,040	-,927	,355	-304,562	109,555
ESTERO	1102,442	298,828	,150	3,689	,000	514,455	1690,430
ALTO ADIGE	200,605	135,824	,062	1,477	,141	-66,647	467,858
VENETO	2,521	163,488	,001	,015	,988	-319,165	324,207
ALTRO NORD	352,419	150,395	,098	2,343	,020	56,495	648,344
CENTRO-SUD	-446,753	396,199	-,046	-1,128	,260	-1226,332	332,826

Numero di casi inclusi nell'analisi: 340

R-quadrato: 0,538

R-quadrato corretto: 0,495

F: 12,468

g.d.l.: 29

Sig.:0,000

Tavola A2.6
Risultati della regressione lineare con effetto dell'interazione tra gruppo di facoltà ed esperienza lavorativa

Variabile dipendente: *Retribuzione mensile netta (valore centrale) (retribuz)*

<i>Variabili</i>	<i>Coefficienti non standardizzati</i>	<i>Errore standard</i>	<i>Coefficienti standardizzati</i>	<i>t</i>	<i>Sig.</i>
(Costante)	2284,700	657,296		3,476	,001
SESSO	-300,131	60,389	-,129	-4,970	,000
STATO CIVILE	228,473	59,226	,098	3,858	,000
VOTOLAU	8,164	4,426	,045	1,844	,065
ETÀLAU	-16,239	8,382	-,052	-1,937	,053
OREFFET	30,867	3,187	,266	9,686	,000
LNE _{SP} _1	215,991	61,113	,146	3,534	,000
TIPO_FAC	-434,225	153,431	-,180	-2,830	,005
FAC _{LNE} _{SP}	278,897	72,770	,258	3,833	,000
PROFESSIONE 1	-2641,451	872,214	-,075	-3,028	,003
PROFESSIONE 2	-1998,691	355,952	-,256	-5,615	,000
PROFESSIONE 3	-2027,422	561,562	-,099	-3,610	,000
PROFESSIONE 4	-1523,092	323,736	-,370	-4,705	,000
PROFESSIONE 5	-1214,992	389,616	-,113	-3,118	,002
PROFESSIONE 6	-1544,952	312,493	-,658	-4,944	,000
PROFESSIONE 7	-809,587	318,979	-,204	-2,538	,011
PROFESSIONE 8	-794,840	315,243	-,272	-2,521	,012
RAMO 1	-643,673	107,087	-,257	-6,011	,000
RAMO 2	152,669	124,775	,035	1,224	,221
RAMO 3	-567,068	240,777	-,057	-2,355	,019
RAMO 4	-254,922	149,114	-,050	-1,710	,088
RAMO 5	-395,545	106,458	-,137	-3,715	,000
RAMO 6	-385,699	117,941	-,115	-3,270	,001
RAMO 7	-388,927	114,632	-,119	-3,393	,001
RAMO 8	-474,061	221,177	-,053	-2,143	,032
DIMENSIONE 1	-254,715	73,195	-,097	-3,480	,001
DIMENSIONE 2	-32,946	67,693	-,012	-,487	,627
UTIL_LAU 1	-28,706	122,307	-,006	-,235	,814
UTIL_LAU 2	-128,573	62,662	-,053	-2,052	,040
ESTERO	942,642	212,014	,102	4,446	,000
ALTO ADIGE	148,025	84,527	,041	1,751	,080
VENETO	-188,310	104,926	-,042	-1,795	,073
ALTRONORD	109,138	111,818	,023	,976	,329
CENTRO-SUD	-554,298	364,002	-,035	-1,523	,128

Numero di casi inclusi nell'analisi: 929

R-quadrato: 0,550

R-quadrato corretto: 0,534

F: 33,163

g.d.l.: 33

Sig.:0,000

Tavola A2.7
Anni di studio per conseguire la laurea - per facoltà

<i>Facoltà</i>	<i>Media</i>	<i>N</i>	<i>Dev. St.</i>	<i>Mediana</i>
1 Architettura	8,05	56	2,29	8
2 Lettere	7,05	190	2,49	6
3 Lingue	7,38	29	1,99	7
4 Scienze	6,56	149	2,04	6
5 Giurisprudenza	6,96	170	2,11	6,50
6 Agraria	7,48	27	1,50	7
7 Farmacia	6,53	15	1,41	6
8 Sociologia	9,05	128	4,94	7
9 Ingegneria	7,31	123	1,67	7
10 Economia	6,77	360	2,15	6
11 Medicina	7,91	107	3,31	7
Totale	7,24	1354	2,71	7

Tavole A2.8
Classe di retribuzione mensile (retrib) media per leva -
in aggregato, nei due gruppi di facoltà, nelle facoltà più numerose

a) Tutti

<i>Leva</i>	<i>Media</i>	<i>N</i>	<i>Deviaz. Stand.</i>
82/83	10,51	186	2,79
86/87	9,53	215	2,79
90/91	8,80	322	2,43
94/95	7,05	489	2,38
Totale	8,48	1212	2,85

b) Facoltà scientifiche

<i>Leva</i>	<i>Media</i>	<i>N</i>	<i>Deviaz. Stand.</i>
82/83	11,17	126	2,61
86/87	10,22	137	2,74
90/91	9,14	188	2,47
94/95	7,45	305	2,14
Totale	8,99	756	2,80

c) Facoltà umanistiche

<i>Leva</i>	<i>Media</i>	<i>N</i>	<i>Deviaz. Stand.</i>
82/83	9,12	60	2,68
86/87	8,32	78	2,47
90/91	8,31	134	2,29
94/95	6,38	184	2,61
Totale	7,64	456	2,72

d) Ingegneria

<i>Leva</i>	<i>Media</i>	<i>N</i>	<i>Deviaz. Stand.</i>
82/83	12,33	6	2,25
86/87	10,61	23	2,66
90/91	9,88	26	1,93
94/95	7,63	62	1,89
Totale	8,96	117	2,55

e) Economia

<i>Leva</i>	<i>Media</i>	<i>N</i>	<i>Deviaz. Stand.</i>
82/83	11,70	37	2,65
86/87	10,25	55	2,54
90/91	9,37	86	2,19
94/95	7,85	156	2,06
Totale	9,06	334	2,59

f) Medicina e Chirurgia

<i>Leva</i>	<i>Media</i>	<i>N</i>	<i>Deviaz. Stand.</i>
82/83	12,45	44	1,59
86/87	12,43	21	1,36
90/91	12,13	16	1,50
94/95	9,20	5	2,95
Totale	12,20	86	1,76

g) Scienze

<i>Leva</i>	<i>Media</i>	<i>N</i>	<i>Deviaz. Stand.</i>
82/83	8,44	27	1,76
86/87	8,24	21	2,62
90/91	7,88	40	2,07
94/95	6,73	45	1,91
Totale	7,66	133	2,15

h) Lettere e Filosofia

<i>Leva</i>	<i>Media</i>	<i>N</i>	<i>Deviaz. Stand.</i>
82/83	7,86	14	1,70
86/87	7,22	27	1,87
90/91	7,36	59	1,73
94/95	6,41	70	2,15
Totale	6,99	170	1,98

i) Giurisprudenza

<i>Leva</i>	<i>Media</i>	<i>N</i>	<i>Deviaz. Stand.</i>
82/83	11,75	4	2,06
86/87	10,14	14	2,35
90/91	9,42	45	2,72
94/95	5,97	74	3,13
Totale	7,70	137	3,46

j) Sociologia

<i>Leva</i>	<i>Media</i>	<i>N</i>	<i>Deviaz. Stand.</i>
82/83	11,00	3	3,00
86/87	7,75	4	2,36
90/91	9,25	4	1,89
94/95	7,20	5	0,45
Totale	8,56	16	2,28

A3. Legenda e descrizione delle variabili ³⁶

NOME VARIABILE	DESCRIZIONE	VALORI
ALTRELAU	Eventuali altre lauree possedute	1. No 2. Sì
ALTROLAV	Eventuale svolgimento di una seconda attività lavorativa	1. No 2. Sì
COER_LAU	Grado di coerenza tra mansione svolta e formazione teorica	Da 1 a 5
CONTRATTO	Forma del contratto di lavoro	1. Contratto di formazione e lavoro 2. A tempo determinato 3. A tempo indeterminato 4. Collaboratore professionale occasionale 5. Collaborazione professionale coordinata e continuativa 6. Contratto di lavoro interinale 7. Nessun contratto
DIMENS	Dimensioni dell'azienda per cui si lavora, in relazione al numero degli addetti	1. Piccola azienda (meno di 15 addetti) 2. Media azienda (da 15 a 50 addetti) 3. Grande azienda (più di 50 addetti)
DIPL_UNI	Anno di iscrizione all'Università rispetto all'anno di diploma	1. Non nello stesso anno 2. Nello stesso anno
DIPLOMA	Titolo di scuola media superiore conseguito	1. Liceo classico 2. Liceo scientifico 3. Liceo linguistico 4. Liceo artistico/istituto d'arte 5. Istituto magistrale/Scuola magistrale 6. Istituto professionale (4-5 anni) 7. Istituto tecnico commerciale 8. Istituto tecnico per geometri 9. Istituto tecnico industriale 10. Altro istituto tecnico
ESPLAV	Anni di esperienza lavorativa	
ESPQUAD	(esplav) ²	
ETÀ	Età al momento della rilevazione	
ETÀ_LAU	Età al momento della laurea	
FACESPLAV	tipofac * ln (esplav + 1)	

³⁶ Le variabili binarie utilizzate nelle versioni "finali" (quelle riportate) delle regressioni sono codificate come 0-1, come vuole la prassi nell'uso delle *dummies*, quelle utilizzate solo nel modello finale di CART sono codificate come 1-2. E' chiaro, però, che nei tentativi di regressione effettuati, ma non riportati, tutte le variabili binarie sono state trattate come 0-1.

E' altresì chiaro che le variabili multicategoriali sono state trattate, nelle regressioni lineari, come altrettante variabili *dummy* codificate come 0-1.

Appendice A3. Legenda e descrizione delle variabili

NOME VARIABILE	DESCRIZIONE	VALORI
FACOLTA'	Facoltà presso cui si è conseguita la laurea (ordinate secondo la retribuzione mensile media)	<ol style="list-style-type: none"> 1. Architettura 2. Lettere e filosofia 3. Lingue e letterature straniere 4. Scienze matematiche, chimiche, fisiche e naturali 5. Giurisprudenza 6. Agraria 7. Farmacia 8. Sociologia e Scienze politiche 9. Ingegneria 10. Economia 11. Medicina e chirurgia e Medicina Veterinaria
FRANCESE	Conoscenza del Francese (autovalutazione)	Come Inglese
FREQ_LEZ	Frequenza con cui si sono seguite le lezioni	<ol style="list-style-type: none"> 1. Regolarmente, tutti i corsi o quasi 2. Regolarmente, solo alcuni i corsi 3. Saltuariamente, tutti i corsi 4. Saltuariamente, solo alcuni corsi 5. Regolarmente in alcuni anni, saltuariamente (o senza nessuna frequenza) in altri 6. Nessuna frequenza
INFORMATICA	Conoscenze informatiche (autovalutazione)	Da 1 a 5
INGLESE	Conoscenza dell'Inglese (autovalutazione)	<ol style="list-style-type: none"> 1. Nessuna 2. Scolastica 3. Discreta 4. Buona 5. Ottima 6. Madrelingua
LAU_STAGE	Frequenza di uno stage in azienda durante la preparazione della tesi di laurea	<ol style="list-style-type: none"> 1. No 2. Sì, in Italia 3. Sì. All'estero
LAV_ATT	Situazione lavorativa attuale rispetto alla situazione precedente	<ol style="list-style-type: none"> 1. L'attuale occupazione non è la stessa intrapresa dopo la laurea o che già era svolta al momento della laurea 2. L'attuale occupazione è la stessa intrapresa dopo la laurea o che già era svolta al momento della laurea
LAV_UNI	Eventuale attività lavorativa remunerata svolta durante gli studi universitari	<ol style="list-style-type: none"> 1. Studente a tempo pieno 2. Svolgimento di lavori saltuari 3. Prima studente lavoratore poi studente a tempo pieno. 4. Prima studente a tempo pieno poi lavoratore. 5. Sempre studente lavoratore
LINGUA 5	Conoscenza di una quinta lingua (autovalutazione)	Come Inglese
LINGUA 6	Conoscenza di una sesta lingua (autovalutazione)	Come Inglese

Appendice A3. Legenda e descrizione delle variabili

NOME VARIABILE	DESCRIZIONE	VALORI
LNESP_1	In (esplav + 1)	
LODE	Conseguimento della lode	1.No 2.Sì
MESI_RIC	Mesi dedicati alla ricerca attiva del lavoro	
NUM_LAV	Numero di lavori effettuato dopo la laurea	
NUMFIGLI	Numero di figli	
OREFFET	Numero di ore settimanali effettivamente lavorate	
POSTLAU 1	Frequenza (terminata o in atto) di un corso di specializzazione post-laurea o master organizzato da università	1.No 2.Sì, in Italia 3.Sì, all'estero 4.Sì, in Italia e all'estero
POSTLAU 2	Frequenza (terminata o in atto) di un corso di specializzazione post-laurea o master organizzato da strutture non universitarie	Come sopra
POSTLAU 3	Frequenza (terminata o in atto) di un dottorato di ricerca o Phd	Come sopra
POSTLAU 4	Effettuazione (terminata o in atto) di un praticantato o tirocinio per accedere ad ordini professionali	Come sopra
POSTLAU 5	Superamento di un esame di stato	Come sopra
POSTLAU 6	Effettuazione di altri studi post-laurea	Come sopra
PREP_LAV	Eventuali periodi di preparazione supplementare organizzati dall'azienda prima di iniziare il lavoro e/o durante l'attività lavorativa	1. No 2. Sì, di durata inferiore a un mese 3. Sì, di durata superiore a un mese
PROF	Posizione professionale (ordinate per retribuzione mensile media)	1.Lavoratore a domicilio per conto d'impresa 2. Operaio, subalterno e assimilati 3. Coadiuvante familiare 4. Impiegato o intermedio 5.Specializzando, dottorando, praticante, tirocinante 6. Lavoratore in proprio 7. Libero professionista 8. Dirigente o quadro direttivo 9. Imprenditore
PUB_PRIV	Settore (pubblico o privato) di attività economica	1. Pubblico 2. Privato
RAMO	Ramo di attività economica (ordinati per retribuzione mensile media)	1. Istruzione, ricerca e sviluppo 2. Altri campi rispetto a quelli nominati 3. Agricoltura, foreste, caccia e pesca 4. Commercio, alberghi e pubblici esercizi 5. Pubblica Amministrazione 7. Credito, assicurazione e servizi alle imprese 7. Industria (comprese le costruzioni) 8. Trasporti e telecomunicazioni 9. Sanità

Appendice A3. Legenda e descrizione delle variabili

NOME VARIABILE	DESCRIZIONE	VALORI
RESIDENZA	Residenza attuale	1. Trentino 2. Alto Adige 3. Veneto 4. Altra regione del nord (Altro Nord) 5. Centro-sud (Censud) 6. Estero
RETDIC	Retribuzione mensile netta percepita negli ultimi mesi (suddivisione di retrib in due classi)	0. Classi 1-7 1. Classi 8-14
RETRIB	Retribuzione mensile netta percepita negli ultimi mesi	14 classi (in lire) 1. < 500.000 2. 500.000 – 800.000 3. 800.000 – 1.000.000 4. 1.000.000 – 1.300.000 5. 1.300.000 – 1.500.000 6. 1.500.000 – 1.800.000 7. 1.800.000 – 2.000.000 8. 2.000.000 – 2.300.000 9. 2.300.000 – 2.500.000 10. 2.500.000 – 3.000.000 11. 3.000.000 – 3.500.000 12. 3.500.000 – 4.000.000 13. 4.000.000 – 5.000.000 14. > 5.000.000
RETRIBUZ	Retribuzione mensile netta percepita negli ultimi mesi (valore centrale delle classi di retrib)	250.000 650.000 900.000 1.150.000 1.400.000 1.650.000 1.900.000 2.150.000 2.400.000 2.750.000 3.250.000 3.750.000 4.500.000 6.000.000
RETRORA	Indicatore 1 della retribuzione oraria media percepita negli ultimi mesi	retrib/oreffet
RTRZORA	Indicatore 2 della retribuzione oraria media percepita negli ultimi mesi	retribuz/oreffet
RTRZORA 4	Suddivisione in quartili di rtrzora	
SEDE_LAV	Sede del lavoro attuale	1. Trentino 2. Alto Adige 3. Altra regione italiana 4. Estero
SESSO		0. Maschio 1. Femmina
SIT_LAV	Situazione lavorativa dopo la laurea rispetto a prima della laurea	1. Non ancora iniziato a lavorare dopo la laurea 2. Continuato il lavoro svolto prima della laurea 3. Iniziato un nuovo lavoro

Appendice A3. Legenda e descrizione delle variabili

NOME VARIABILE	DESCRIZIONE	VALORI
SPAGNOLO	Conoscenza dello Spagnolo (autovalutazione)	Come Inglese
STATCIV	Stato civile	1. Celibe/nubile 2. Coniugato/a o convivente o separato/a o divorziato/a o vedovo/a
STUD_EST	Periodi di studio all'estero durante l'Università	1.No 2.Sì, in Italia 3.Sì, all'estero
TEDESCO	Conoscenza del Tedesco (autovalutazione)	Come Inglese
TEMPOLA 2	Anni impiegati a conseguire la laurea (attorno alla mediana)	1.Fino a 6 anni 2.Da 7 anni in poi
TIPOFAC	Tipo di facoltà presso cui si è conseguita la laurea	0. Umanistica 1. Scientifica
TTT_MADRE	Titolo di studio materno	1. Nessuno 2. Licenza elementare 3. Avviamento professionale 4. Licenza media 5. Diploma di qualifica professionale 6. Diploma di scuola media superiore 7. Diploma post-secondario 8. Diploma di laurea 9. Specializzazione post-laurea
TTT_PADRE	Titolo di studio paterno	Come sopra

Appendice A3. Legenda e descrizione delle variabili

NOME VARIABILE	DESCRIZIONE	VALORI
TROVO	Modalità con cui si è trovato il lavoro	<ol style="list-style-type: none"> 1. Inserzione e/o risposta ad un annuncio pubblicato su un giornale/rivista specializzata 2. Inserzione presso un ufficio pubblico di collocamento 3. Partecipazione a concorsi pubblici 4. Domanda di assunzione ad impresa/ente pubblico 5. Chiamata diretta da parte del datore di lavoro 6. Assunzione a seguito di stage 7. Tramite contatti maturati nel precedente lavoro 8. Con l'aiuto di un servizio di consulenza/agenzia di selezione del personale 9. Attraverso parenti, amici o conoscenti 10. Su segnalazione ricevuta da altre persone 11. Proseguendo l'attività dell'impresa familiare 12. Iniziando un'attività in proprio 13. Altro
UNIVERSITA'	Università presso cui si è conseguita la laurea	<ol style="list-style-type: none"> 1. Università degli studi di Bologna 2. Università degli studi di Milano 3. Politecnico di Milano 4. Università "Cattolica del Sacro Cuore di Milano 5. Università Commerciale "Luigi Bocconi" di Milano 6. Università degli studi di Padova 7. Università degli studi di Trento 8. Università degli studi di "Ca' Foscari" di Venezia 9. Istituto universitario di architettura di Venezia 10. Università degli studi di Verona
UTIL_LAU	Utilità della laurea per il lavoro svolto	<ol style="list-style-type: none"> 1. Poco utile o inutile 2. Utile, ma non indispensabile 3. Indispensabile
VOTODIPLOMA	Voto conseguito all'esame di diploma superiore	In sessantesimi
VOTOLAU	Voto di laurea in centodecimi	

B. Un tentativo di analisi dinamica

Come detto nell'introduzione di questo lavoro, lo stimolo a questo studio statistico è stato dato dal ritorno all'attenzione generale del tema dei differenziali salariali, dovuto al loro incremento negli anni '80; gli anni '90 hanno però visto un arresto di questa tendenza all'aumento. Sarebbe stato pertanto di grande interesse verificare nel nostro campione l'andamento dei differenziali.

Purtroppo il *database* da noi analizzato non consente una valutazione di questa evoluzione salariale, essendo la retribuzione rilevata in un solo momento, quello dell'intervista. Tuttavia l'esistenza, all'interno di questo campione, di quattro leve di laureati consente un tentativo di analisi temporale.

E' chiaro infatti che, confrontando i laureati delle leve più recenti con quelle più remote, le differenze di reddito possono dipendere:

- a) da una differente esperienza lavorativa (maggiore naturalmente per i laureati da più tempo);
- b) da un'evoluzione salariale, per cui *ceteris paribus* (quindi anche a parità di esperienza lavorativa) le lauree più recenti sono favorite rispetto alle più remote (o viceversa);
- c) da un mutato effetto nel tempo di una o più variabili determinanti del reddito (ad esempio il premio per l'esperienza lavorativa è più forte nelle leve più recenti).
- d) da una differente distribuzione, nelle diverse leve, nel campione o nella popolazione, dei fattori determinanti del reddito.

La componenti che ci interessano sono dunque la seconda e la terza, ed in particolar modo ci interessa verificare se questa evoluzione salariale, se esiste, ha toccato in maniera differente le diverse facoltà o i due gruppi di facoltà (anche in questo caso condurremo un'indagine aggregata). E' dunque possibile effettuare una regressione che, "controllando" per l'esperienza lavorativa e per altre variabili determinanti³⁷, ci permetta di verificare se il secondo e il terzo effetto ci sono stati e se sono stati diversi nei due gruppi. Se, per quanto riguarda il punto 3, stimiamo solo l'effetto dell'esperienza lavorativa, avremo l'equazione:

$$(B1) \quad Y = ax + b_1z_1 + b_2z_2 + \dots + b_nz_n + cL + dG + eLG + fxL + gxG + hxLG$$

essendo:

x = esperienza lavorativa o funzione della stessa;

z_1, z_2, \dots, z_n = altre variabili significative;

G = variabile *dummy* indicante il gruppo di facoltà;

L = variabile indicante la leva di laurea.

³⁷ "Controllando" per le altre variabili determinanti eliminiamo l'effetto distorsivo dato dalla diseguale distribuzione di tali variabili nel campione; tenendo tuttavia costanti tali variabili, non cogliamo l'eventuale mutamento di tali fattori nella popolazione.

In tal modo sarà possibile stimare gli effetti di cui sopra:

- l'effetto b) è misurato da $c + eG$ (variazione non dipendente dal premio per l'esperienza lavorativa);
- l'effetto c) è misurato da $fx + hGx$ (variazione dipendente dal solo premio per l'esperienza lavorativa);
- il totale dell'effetto-leva (variazione del reddito al variare della leva) è dato da: $c + eG + fx + hxG$. Un segno positivo di questo valore significa che i laureati più recenti guadagnano di più dei laureati più remoti, a parità di esperienza lavorativa e di altre variabili determinanti.

La variazione dell'effetto-leva al variare del gruppo di facoltà sarà dato da: $e + hx$; dunque un segno positivo di questo valore significa che la crescita salariale nel tempo è stata più forte nelle facoltà scientifiche che in quelle umanistiche (o, se l'effetto leva è negativo, la diminuzione è stata più forte nelle facoltà umanistiche che in quelle scientifiche). Quindi un segno positivo di questo valore significa differenziali per facoltà crescenti nel tempo.

Come è facile immaginare, stimare questa equazione è problematico a causa delle forti correlazioni presenti tra le variabili "interazione". Dunque abbiamo stimato due modelli semplificati, uno comprendente il solo effetto b) (incremento salariale non dovuto al premio per l'esperienza), uno comprendente il solo effetto c) (incremento salariale dovuto unicamente al premio per l'esperienza). Per l'esperienza lavorativa abbiamo usato una funzione logaritmica (per la precisione $\ln(\text{esplav} + 1)$, che indicheremo ancora con x), le altre variabili considerate sono state il sesso e le ore di lavoro effettive. La *dummy* indicante il gruppo di facoltà è stata codificata, come nelle altre nostre analisi, con 0 = facoltà umanistiche, 1 = facoltà scientifiche. La variabile relativa alla leva è stata codificata con:

- 0 = leva 82/83;
- 4 = leva 86/87;
- 8 = leva 90/91;
- 12 = leva 94/95.

Esponiamo qui le stime relative a questi modelli (tra parentesi il valore della t di Student) (la variabile dipendente è il valore centrale delle classi di reddito):

$$(B2) \quad Y = 1212374,3 - 392413,6 \text{ sesso} + 43757,522 \text{ oreffet} + 270303,039 x + \\ (3,161) \quad (-6,226) \quad (14,558) \quad (2,325) \\ + 593742,371 G - 43948,322 L - 63050,003 LG \\ (4,559) \quad (-2,065) \quad (-4,408)$$

$$(B3) \quad Y = 501464,179 - 386438 \text{ sesso} + 43503 \text{ oreffet} + 522414,597 x - 7034,717 xL + \\ (2,464) \quad (-6,152) \quad (14,487) \quad (9,265) \quad (-0,869) \\ + 255534,444 xG - 29730,722 xLG \\ (4,798) \quad (-3,581)$$

Come si vede, secondo entrambi i modelli c'è un effetto-leve negativo e, con alta significatività, un segno negativo della variabile che esprime l'andamento nel tempo dei differenziali tra i due gruppi di facoltà, che starebbero quindi rientrando negli anni recenti. Questo sembrerebbe confermare quanto detto in apertura, che la tendenza all'aumento dei differenziali, quantomeno all'interno di gruppi "omogenei" per istruzione, si sarebbe almeno arrestata. Addirittura il prevalere di un effetto negativo complessivo nelle nostre stime, considerando un arco temporale che va appunto degli anni '80 fino a quasi tutti gli anni '90, può essere dovuto semplicemente ad una netta sovrarappresentanza delle due leve relative agli anni '90 rispetto alle leve degli anni '80. Se dunque non si può concludere che la tendenza all'aumento dei differenziali si è invertita, i risultati di queste indagini sono incompatibili con un loro costante aumento.

Questo risultato è altresì coerente con quanto rilevato nell'indagine sulle motivazioni delle scelte in istruzione: se le motivazioni "non economiche" continuano a prevalere su quelle "economiche", queste ultime sono apparse più forti nelle leve più recenti: una maggiore sensibilità ai crescenti differenziali può aver provocato un riequilibrio della domanda di formazione, riducendo i differenziali stessi.

Al termine di questa indagine va aggiunto che i risultati qui esposti vanno considerati con estrema cautela: l'estrapolazione di un'indagine dinamica da dati "statici" presenta molti rischi ed arbitri (per esempio è fondamentale una scelta corretta della funzione relativa all'esperienza lavorativa, che tuttavia non può che essere in larga misura una scelta per approssimazione). Questa indagine va considerata pertanto alla stregua di un esperimento. A risultati maggiormente affidabili si potrà giungere solo se (come sembra avere in programma il Dipartimento di Economia dell'Università di Trento) verrà riefettuata un'analogha raccolta di dati, in modo da avere un vero profilo temporale.

BIBLIOGRAFIA

- ABBURRA', L. (1997), "Proseguire o smettere: da cosa dipendono le scelte scolastiche individuali negli anni novanta?", *Polis*, XI, n. 3, dicembre, pp. 367-389.
- ALTONIJ, J. G. (1993), "The Demand for and Return to Education When Education Outcomes Are Uncertain", in *Journal of Labor Economics*, n. 1, pp. 48-84.
- ANTONELLI, G. (1985), *Risorse umane e redditi da lavoro*, Milano, Franco Angeli.
- ARAI, K. (1995), *The Economics of Education - An Analysis of College-Going Behavior*, Sprinter.
- BARTEL, A. P. - SICHERMAN, N. (1999), "Technological Change and Wages: an Interindustry Analysis", in *Journal of Political Economy*, April, pp. 285-325.
- BECKER, G. (1962), "Investment in Human Capital: a Theoretical Analysis" in *Journal of Political Economy*, 70, pp. 9-49.
- BORZAGA, C. - CONTRINI, E. - GABURRI, A. (1999), *Indagine sugli sbocchi occupazionali dei laureati dell'Università di Trento e dei laureati residenti in Provincia di Trento*, Università degli Studi di Trento, Dipartimento di Economia, mimeo.
- BREIMAN, L. - FRIEDMAN, J.H. - OLSHEN, R.A. - STONE, C.J. (1984), *Classification and Regression Trees*, Belmont (California), Wodsworth International Group.
- BRUNELLO, G. - COMI, S. - LUCIFORA, C. (1999), *The Returns to Education in Italy: a New Look at the Evidence*, mimeo.
- BRUNELLO G. – LUCIFORA C. – WINTER-EBMER R. (2001), *The wage expectations of european college students*, n. 30, Quaderni dell'Istituto di Economia dell'Impresa e del Lavoro, Milano, Università Cattolica del Sacro Cuore.
- CARD, D. - KRUEGER, A. (1992), "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States", in *Journal of Political Economy*, n. 1, pp 1-40.
- CHECCHI D. (2000), "University Education in Italy", in *International Journal of Manpower*, vol. 21, n. 3/4 pp. 177- 205.
- CHECCHI, D. - ICHINO, A. - FRANZONI, F. - RUSTICHINI, A. (1999), "Background familiare e carriera universitaria", paper presentato alla conferenza «Politiche pubbliche per il lavoro», Pavia, 17-18 giugno 1999.

- CHECCHI, D. - ICHINO, A. - RUSTICHINI, A. (1996), 'Immobili perché uguali? Mobilità occupazionale e scolastica tra generazioni in Italia e Stati Uniti', in CONFINDUSTRIA - CENTRO STUDI, "La mobilità della società italiana. Le persone, le imprese, le istituzioni" (a cura di G. Galli), v. I, parte A, pp. 511-568.
- CRAWFORD, D. L. - JOHNSON, A. W. - SUMMERS, A. A. (1997), "Schools and Labor Market Outcomes", in *Economics of Education Review*, n. 3, pp. 255-265.
- GAMBETTA, D. (1996), *Were They Pushed or Did They Jump? Individual Decision Mechanisms in Education*, Boulder – Oxford, Westview Press (1a edizione: 1987).
- GIANNINI, M., (1998) *Il capitale umano: misurazione, determinanti socio-economiche ed effetti sul reddito in Italia*, Università di Roma "La Sapienza", mimeo.
- GROGGER, J. - EIDE, E. (1995), "Changes in College Skills and the Rise in the College Wage Premium", in *The Journal of Human Resources*, Spring, pp. 280-310.
- JENCKS, C.S. - SMITH, M. - ACLAND, H. - BANE, M.J. - COHEN, D. - GINTIS, H. - HEYNS, B. - MICHELSON, S. (1972): "Inequality: A Reassessment of the Effect of Family and Schooling" in *America*, New York, Basic Books.
- JUHN, C. - MURPHY, K. M. - PIERCE, B. (1993), "Wage Inequality and the Rise in Returns to Skill", in *Journal of Political Economy*, Chicago, The University Chicago Press, v. 101. n. 3, pp. 410-442.
- KAROLY, L. A. (1993), "The Trend Inequality Among Families, Individuals, and Workers in the United States: A twenty-five Year Perspective, in DANZINGER, S. - GOTTSCHALCK, P, *Uneven tides*, New York, Russel Sage Foundation, pp. 19-97.
- MARTINI, M.(2000), 'Le linee di tendenza evolutive nel campo del lavoro intellettuale' in FONDAZIONE ARISTEIA (aa.vv.) (2000), *Viaggio tra i perché della disoccupazione intellettuale in Italia*, Milano, Giuffrè Editore, pp. 41-70.
- MATTILA, J.P. (1982), "Determinants of Male School Enrollments: a Time-Series Analysis", in *The Review of Economics and Statistics*, n. 64, pp. 242-251.
- MINCER, J. (1974), *Schooling, Experience and Earnings*, Columbia University Press.
- MURPHY, K. - WELCH, F. (1992), "The Structure of Wages", in *Quarterly Journal of Economics*, February, pp. 285-325.

OAXACE, R.L. (1987), "Sex Earnings Differentials" in Psacharopoulos, G. (edited by): *Economics of Education. Research and Studies*, Oxford, Pergamon Press.

PRAUSSELLO, F. - MARENCO, M. (1996), *Economia dell'istruzione e del mercato del lavoro*, Bari, Laterza.

PROVINCIA AUTONOMA DI TRENTO - OSSERVATORIO DEL MERCATO DEL LAVORO
Istruzione e mercato del lavoro in provincia di Trento. Esiti occupazionali dei laureati e dispersione scolastica universitaria, aprile 1989.

PSACHAROPOULOS, G. (1987), "Earnings Functions" in Psacharopoulos, G. (edited by): *Economics of Education. Research and Studies*, Oxford, Pergamon Press.

SPENCE, M. (1974), *Market Signalling*, Cambridge, University Press Harvard.

INDICE DELLE TAVOLE

1.	Albero di classificazione per tutti i soggetti del campione	17
2.	Matrice di classificazione	17
3.	Ranking dell'importanza delle variabili indipendenti basato su split univariati	18
4.	Risultati della regressione lineare aggregata	20
5.	Retribuzione mensile ed oraria (indicatore) media e mediana - per facoltà	24
6.	Retribuzione mensile ed oraria (indicatore) media e mediana - per gruppo di facoltà	26
7.	Retribuzione mensile ed oraria (indicatore) media e mediana dei soggetti di sesso maschile - per gruppo di facoltà	27
8.	Retribuzione mensile ed oraria (indicatore) media e mediana dei soggetti di sesso femminile - per gruppo di facoltà	27
9.	Numero medio di ore settimanali effettivamente lavorate - per gruppo di facoltà	29
10.	Numero medio di ore settimanali effettivamente lavorata per classe di retribuzione oraria - tutti i soggetti attivi	30
11.	Numero medio di ore settimanali effettivamente lavorate per classe di retribuzione oraria - facoltà scientifiche	30
12.	Numero medio di ore settimanali effettivamente lavorate per classe di retribuzione oraria - facoltà umanistiche	31
13.	Retribuzione mensile media e mediana dei soggetti con esperienza lavorativa da 0 a 6 anni - per gruppo di facoltà	33
14.	Retribuzione mensile media e mediana dei soggetti con esperienza lavorativa oltre i 6 anni - per gruppo di facoltà	34
15.	Retribuzione mensile media e mediana nelle facoltà scientifiche - per sesso	35
16.	Retribuzione mensile media e mediana nelle facoltà umanistiche - per sesso	35
17.	Grado di importanza delle motivazioni nella scelta della facoltà – valori medi	41
18.	Tipo di facoltà scelta rispetto alla scuola superiore frequentata – valori percentuali	41
19.	Titolo di studio dei genitori per facoltà di provenienza degli intervistati	42
20.	Titolo di studio dei genitori per scuola superiore di provenienza degli intervistati	42
A1.1	Composizione del campione per sesso	44
A1.2	Composizione del campione per leva	44
A1.3	Composizione del campione per residenza e ateneo	45
A1.4	Composizione del campione per facoltà	45
A1.5	Composizione del campione per gruppo di facoltà	45
A1.6	Valore medio, mediano e deviazione standard delle principali variabili considerate- in aggregato e nei due gruppi di facoltà	46
A1.7	Correlazioni tra le variabili principali	47

A2.1	Risultati della regressione logistica binaria aggregata	48
A2.2	Risultati della regressione logistica multinomiale ordinata aggregata	49
A2.3	Risultati della regressione lineare con effetti dell'esperienza lavorativa	50
A2.4	Risultati della regressione lineare relativa alle facoltà scientifiche	51
A2.5	Risultati della regressione lineare relativa alle facoltà umanistiche	52
A2.6	Risultati della regressione lineare con effetto dell'interazione tra gruppo di facoltà ed esperienza lavorativa	53
A2.7	Tempo medio e mediano per conseguire la laurea - per facoltà	54
A2.8	(a-j) Classe di retribuzione mensile media per leva- in aggregato, nei due gruppi di facoltà, nelle facoltà più numerose	54

INDICE GENERALE

Premessa	pag. 1
1. Introduzione	pag. 2
2. Le metodologie di analisi	pag. 4
2.1 <i>La tecnica CART</i>	<i>pag. 4</i>
2.2 <i>L'analisi di regressione</i>	<i>pag. 7</i>
2.3 <i>L'utilizzazione delle diverse tecniche</i>	<i>pag. 8</i>
3. Le variabili considerate	pag. 10
3.1 <i>La variabile dipendente</i>	<i>pag. 10</i>
3.2 <i>Le variabili indipendenti</i>	<i>pag. 10</i>
4. I risultati dell'indagine: le determinanti del reddito	pag. 16
4.1 <i>I risultati di CART</i>	<i>pag. 16</i>
4.2 <i>I risultati dell'analisi di regressione</i>	<i>pag. 19</i>
4.3 <i>Riepilogo e confronto dei risultati</i>	<i>pag. 22</i>
5. I risultati dell'indagine: i differenziali salariali	pag. 24
5.1 <i>I differenziali salariali per facoltà e per gruppo di facoltà</i>	<i>pag. 24</i>
5.2 <i>Un approfondimento sui differenziali per facoltà</i>	<i>pag. 31</i>
6. Le motivazioni delle scelte in istruzione	pag. 36
7. Conclusioni	pag. 43
Appendice	pag. 44
A1. <i>Descrizione del campione</i>	<i>pag. 44</i>
A2. <i>Tavole fuori testo</i>	<i>pag. 48</i>
A3. <i>Legenda e descrizione delle variabili</i>	<i>pag. 57</i>
B. <i>Un tentativo di analisi dinamica</i>	<i>pag. 63</i>
Bibliografia	pag. 66
Indice delle tavole	pag. 69