

**UNIVERSITA' DEGLI STUDI DI PADOVA**

**Facoltà di Scienze Statistiche**

**Corso di Laurea in Statistica e Gestione delle Imprese**

TESI DI LAUREA

**I FATTORI CONDIZIONANTI IL RISCHIO DI INGRESSO  
ALL'UNIVERSITA'**

**Relatore: Dott. Ugo Trivellato**

Dipartimento di Scienze Statistiche

**Laureando: Sartori Daniele**

**Matr.485252-GEI**

**ANNO ACCADEMICO 2005-2006**

*Puoi essere un dio in terra, ma se ti rubano la palla sono guai.*  
(Gianni Brera)

*Prendi la vita e difendila senza pietà.*  
(Timoria)

*Quando la gente è d'accordo con me, provo la strana  
sensazione di avere torto.*  
(Oscar Wilde)

## **Ringraziamenti**

Un grazie innanzitutto al Dott. Ugo Trivellato per avermi seguito ed appoggiato durante la stesura di questa tesi con disponibilità e professionalità.

Vorrei inoltre ringraziare la Dott.ssa Anna Giraldo per i preziosissimi consigli ed aiuti dati con le elaborazioni in SAS.

Grazie al mio amico Stefano con il quale ho condiviso in questi 3 splendidi anni universitari la mitica 504, facendo fuori una dopo l'altra, tutte le matricole; vorrei ringraziarlo soprattutto per avermi insegnato il valore della puntualità...

Grazie a Cesare e Loris, i miei 2 grandi amici di Roana, per essere cresciuti insieme a me dall'asilo fino ad oggi, passando per elementari, medie, superiori e mega-partite di calcio.

Grazie a Marco, per tutto il tempo trascorso insieme ai tempi delle medie e delle superiori, e per i ricordi delle migliori feste: forse però a Riccione avevamo leggermente esagerato!

Grazie a Bindi, che con me ha condiviso nelle ultime due vacanze a Lisbona e a Bressanone, non solo una camera costantemente disordinata ed il letto di Pedro, ma anche un'esperienza emozionante ed indimenticabile.

Grazie a Pez, il ragazzo che dopo i momenti di baldoria, è anche l'amico con cui spesso mi trovo a riflettere e a parlare degli argomenti più delicati.

Vorrei ringraziare in un elenco probabilmente ingeneroso tutti gli amici e le amiche sia di Roana che di Padova per le belle giornate, serate e nottate trascorse insieme. Prometto loro più spazio fra 2 anni (speriamo) se mi saranno ancora vicini! Grazie quindi a Elisabetta, Silvia, Cerchiuz, Daniela, Luca, Michele, Francesca, Simone, Piero (Barbofono), Cinzia, Marina, Federico, Claudia, Tigran, Alessandra, Paolo, Lisa, Penner, Piero (Pampa), Nicola, Nunzia, Marco (Gregge), Giulia.

Grazie ai compagni dell'A.C. Canove, con i quali condivido gioie e sofferenze durante allenamenti, partite e feste-salvezza, da 5 anni a questa parte.

Grazie a tutti i componenti della mia famiglia, che per tutti questi anni mi sono stati vicini, ricoprendomi di affetto: un grazie particolare ai nonni Giorgio e Mirella, ed agli Maurizio, Margherita e Luigi. Quest'ultimo, va necessariamente ringraziato anche per avermi trasmesso il valore della juventinità, salvandomi così dalla altrimenti inevitabile sofferenza nerazzurra.

Un ringraziamento particolarmente grande va a mia sorella Debora, che da 9 anni a questa parte ha dato un motivo ed un significato in più alla mia vita.

Grazie specialmente a mamma e papà, che giorno dopo giorno lavorano per donarmi una vita felice. Grazie per avermi insegnato i veri valori della vita, grazie per la fiducia in me riposta, grazie per il vostro amore, grazie per il vostro sostegno economico e morale durante tutti questi anni di scuola ed Università, perché senza non sarei sicuramente riuscito a raggiungere questo obiettivo. Per questo motivo sono entusiasta di potervi dire che questo traguardo è tanto mio quanto vostro.

Infine, un grazie ed un abbraccio va a nonna Rosina e alle bisnonne Carlina ed Erminia, che oggi sarebbero sicuramente fiere di questo traguardo.

# Indice

<b>Introduzione</b>	<b>6</b>
<b>1 La popolazione a rischio di ingresso all'Università per vincoli economici: il piano della ricerca</b>	<b>8</b>
1.1 L'obiettivo della nostra ricerca .....	8
1.2 Individuazione della popolazione a rischio di ingresso all'Università .....	9
1.3 Studio della distribuzione della popolazione rispetto al reddito .....	10
1.4 Studio della popolazione rispetto al titolo di studio dei genitori .....	11
1.5 Unità di aggregazione delle risorse .....	11
1.6 Scale di equivalenza .....	12
1.7 Il software utilizzato per le analisi .....	14
<b>2 Il sistema educativo italiano: alcuni richiami su ambiente culturale familiare e stratificazione sociale</b>	<b>15</b>
2.1 Introduzione .....	15
2.2 L'impronta culturale della famiglia .....	16
2.3 Evidenze .....	17
2.4 Aspettative sbagliate .....	17
2.5 Le altre cause .....	18
2.6 Conclusioni .....	20

<b>3 Determinanti della probabilità di abbandono dell'Università in Italia</b>	<b>22</b>
3.1 Introduzione .....	22
3.2 Il fenomeno dell'abbandono dell'Università .....	23
3.3 I motivi dell'interesse .....	24
3.4 Effetti dei differenti tipi di Scuola Superiore .....	24
3.5 Effetti familiari indiretti .....	25
3.6 Determinanti della probabilità di abbandono .....	26
<b>4 Elaborazione dei dati</b>	<b>28</b>
4.1 Operazioni preliminari .....	28
4.2 Ridefinizione dei pesi .....	23
4.3 Calcolo dei redditi familiari equivalenti .....	30
4.4 Selezione dei due campioni .....	31
4.5 Il caricamento delle basi di dati in DAD .....	32
4.6 Curve .....	34
<b>5 Risultati finali</b>	<b>35</b>
5.1 Considerazioni introduttive .....	35
5.2 La relazione tra reddito e livello di istruzione .....	36
5.3 Reddito .....	38
5.4 L'impronta culturale della famiglia .....	40
5.5 Osservazioni conclusive .....	43
<b>Appendice A: Indagine della Banca d'Italia</b>	<b>45</b>
A.1 Introduzione .....	45
A.2 Le indagini del 2000 e del 2002 .....	46
A.3 Il disegno campionario .....	46
A.4 Mancate interviste e attendibilità delle risposte .....	48
A.5 Calcolo dei pesi campionari delle famiglie .....	48
A.6 Ulteriori considerazioni sulla qualità dei dati .....	49

A.7 Descrizione delle basi di dati e delle variabili .....	50
<b>Appendice B: Il software DAD 4.3</b>	<b>54</b>
B.1 Introduzione al Software .....	54
B.2 I motivi della creazione del pacchetto DAD .....	55
B.3 Il linguaggio di programmazione .....	56
B.4 Presentazione di DAD .....	57
B.5 Le basi di dati in DAD .....	57
B.6 Le applicazioni in DAD .....	58
B.7 Grafici .....	59
B.8 Gli strumenti del pacchetto DAD .....	61
 <b>Bibliografia</b>	 <b>63</b>

## **Introduzione**

L'obiettivo di questa tesi è delimitare la popolazione a rischio di ingresso all'Università, qualificarla secondo la distribuzione del reddito, ed individuare in particolare il segmento a basso reddito per il quale vi possono essere vincoli di natura economica che ostacolano l'accesso all'Università.

Inoltre, analizzeremo il fattore socio-culturale (che siamo in grado di esaminare attraverso il titolo di studio dei genitori), che potrebbe essere un fattore rilevante tanto quanto il (se non di più del) reddito.

Questo lavoro è stato realizzato utilizzando i dati dell'indagine sui bilanci delle famiglie della Banca d'Italia degli anni 2000 e 2002, e avvalendosi di due programmi statistici: SAS 8.2 e DAD 4.3.

Nel primo capitolo si definiranno precisamente l'obiettivo della nostra ricerca e l'organizzazione del nostro lavoro; verranno inoltre presentati gli aspetti operativi che riguardano le scelte da effettuare prima di studiare la distribuzione della popolazione (soprattutto illustrando le scale di equivalenza).

Nel secondo capitolo si metterà in evidenza come l'Italia sia caratterizzata da un livello di istruzione inferiore rispetto agli altri Paesi sviluppati: si descriveranno brevemente le cause del fenomeno e si segnalerà l'importanza dell'impronta culturale della famiglia.



Nel terzo capitolo si evidenzierà come l'Italia sia caratterizzata da un tasso di abbandono dell'Università elevato e da una bassa mobilità intergenerazionale in termini di livello di istruzione e reddito: si richiameranno sinteticamente i motivi di questi fenomeni, in particolare le principali determinanti che portano all'abbandono dell'Università.

Nel quarto capitolo si provvederà a dare una spiegazione del metodo di elaborazione dei dati e delle principali operazioni svolte: l'illustrazione delle operazioni preliminari, la ridefinizione dei pesi campionari, l'individuazione dei due campioni, l'importazione delle basi di dati da SAS al foglio elettronico di DAD; si vedrà poi il modo di operare in DAD attraverso le applicazioni statistiche, le curve ed i grafici. Infine, nel quinto capitolo, si confronterà la distribuzione della popolazione totale con le distribuzioni dei due campioni rispetto al reddito ed al titolo di studio dei genitori. Si concluderà presentando i risultati finali e commentando sinteticamente.

Completano la tesi due appendici: nella prima si illustrano brevemente le indagini della Banca d'Italia sui bilanci delle famiglie italiane, le basi di dati e le variabili utilizzate in questo lavoro; nella seconda si introduce il software DAD 4.3, presentandone in breve le caratteristiche, il linguaggio di programmazione e gli strumenti principali.

# **Capitolo 1**

## **La popolazione a rischio di ingresso all'Università per vincoli economici: il piano della ricerca**

### **1.1 L'obiettivo della nostra ricerca**

In questo lavoro il nostro fine è rappresentato dall'individuazione della popolazione per la quale vi possono essere vincoli che ostacolano l'accesso all'Università.

Tali vincoli potrebbero essere di natura economica oppure potrebbero essere rappresentati da una scarsa cultura della famiglia.

Per svolgere questo lavoro, abbiamo utilizzato i dati dell'indagine sui bilanci delle famiglie della Banca d'Italia del 2000 e del 2002.

In questo primo capitolo cercheremo di definire nel modo più preciso possibile i passaggi necessari per raggiungere il nostro obiettivo e l'organizzazione del nostro

lavoro. Presenteremo inoltre le scale di equivalenza, che sono strumenti operativi necessari per confrontare famiglie di diversa numerosità e composizione.

## **1.2 Individuazione della popolazione a rischio di ingresso all'Università**

Di base, dovremo caratterizzare la distribuzione della popolazione totale rispetto al reddito, individuando in particolare il segmento a basso reddito (ad esempio il primo quintile), per il quale vi potrebbero essere vincoli di natura economica che ostacolano l'accesso all'Università. In secondo luogo, bisognerà studiare con grande attenzione l'impronta culturale della famiglia, che potrebbe rivelarsi un fattore importante tanto quanto il reddito familiare, se non addirittura di più.

Per lo svolgimento di questo compito, abbiamo pensato di individuare 2 campioni.

Il primo mira a individuare, la popolazione delle famiglie con figli che conseguono il diploma di scuola media superiore, e possono quindi proseguire gli studi all'Università. Operativamente, abbiamo approssimato questa sottopopolazione con le famiglie con persone in età compresa tra i 16 ed i 20 anni e con titolo di studio più elevato la licenza di scuola media (o il diploma professionale triennale) che siano studenti.

Il secondo campione mira invece, a individuare la sottopopolazione delle famiglie con figli che, avendone titolo, non proseguono gli studi all'Università.

Operativamente, questo campione è stato approssimato con le famiglie con persone in età compresa tra i 18 ed i 22 anni con titolo di studio più elevato il diploma di scuola media superiore che non siano studenti.

Per entrambi i campioni, ci siamo trovati di fronte ad una numerosità campionaria bassa: per risolvere questo problema abbiamo provveduto ad un pooling delle famiglie dalle indagini della Banca d'Italia del 2000 e del 2002, stando bene attenti ad includere una sola volta le famiglie panel. Operativamente, si uniscono le 8.001

famiglie del 2000 e le 8.011 famiglie del 2002 in un unico database e si eliminano da questo le famiglie (all'incirca 4.000) che sono presenti in entrambe le indagini. Partendo da questo campione pooled (composto all'incirca da 12.000 famiglie), si individuano poi i due sottocampioni di interesse.

### **1.3 Studio della distribuzione della popolazione rispetto al reddito**

Dopo aver individuato il campione pooled e i due sottocampioni, che rappresentano rispettivamente la popolazione totale ed i due campioni, il passo successivo sarà quello di studiare la distribuzione di queste tre popolazioni rispetto al reddito familiare (standardizzato con le scale di equivalenza come spiegato nel paragrafo 1.6), effettuando dei confronti tra la distribuzione di ciascuno dei due sottocampioni con quella del campione pooled (tramite distribuzioni per decili o quintili e rappresentazioni grafiche delle distribuzioni).

Questo dovrebbe servirci per riuscire ad individuare una soglia (o più) che ci consenta di selezionare nel primo sottocampione di famiglie quelle a basso reddito, con conseguente rischio di esclusione dei figli dal proseguimento all'Università.

Lo stesso procedimento dovrebbe farci trovare una soglia (o più) che ci porti ad individuare nel secondo sottocampione di famiglie con figli diplomati (scuola media superiore) quelli che non hanno proseguito gli studi all'Università ragionevolmente per vincoli economici, e non per scelte di altra natura.

## **1.4 Studio della popolazione rispetto al titolo di studio dei genitori**

L'altro carattere che analizzeremo, a fianco del reddito familiare, è l'ambiente socio-culturale, che operativamente studieremo attraverso il titolo di studio dei genitori. Come vedremo nei capitoli 2 e 3, quando presenteremo alcuni risultati sul livello di istruzione in Italia e sulle determinanti dell'abbandono dell'Università, secondo parte della letteratura l'impronta culturale della famiglia (ovvero, operativamente, il livello di istruzione dei genitori) potrebbe addirittura essere più importante dei fattori economici e finanziari, per quanto riguarda la decisione di accedere all'Università e/o di proseguire/abbandonare gli studi universitari.

Inoltre, reddito e titolo di studio sono variabili fortemente correlate tra di loro, come mostreremo nel seguito. Il nostro obiettivo finale, peraltro difficile da raggiungere e che realizzeremo solo in parte, è determinare quale delle due variabili incida di più sull'ingresso all'Università.

## **1.5 Unità di aggregazione delle risorse**

Per iniziare il lavoro, bisogna innanzitutto scegliere l'unità di aggregazione delle risorse (famiglia o persona): solitamente è utilizzata la famiglia.

La definizione di famiglia è diversificata. Nell'Indagine sui bilanci delle famiglie italiane la famiglia è intesa come "l'insieme di persone conviventi che, indipendentemente dai legami di parentela, provvede al soddisfacimento dei bisogni mediante la messa in comune di tutto o parte del reddito percepito dai suoi componenti".

## 1.6 Scale di equivalenza

In questo nostro lavoro, prima di studiare la distribuzione della popolazione rispetto al reddito, è necessario effettuare delle scelte per rendere confrontabili famiglie di ampiezza e composizione diverse.

Infatti, il benessere economico di una famiglia non dipende unicamente dalla quantità di risorse a sua disposizione, ma anche dalla dimensione e dalla composizione della famiglia stessa. Un introito mensile di 1.000 euro può consentire una vita dignitosa ad una famiglia formata da una sola persona, ma è quasi sicuramente insufficiente per una coppia con due figli. Le famiglie si differenziano prevalentemente per tre aspetti: il reddito, il numero di componenti e le caratteristiche personali dei singoli componenti, che ne determinano le necessità (quelle di un bambino sono diverse da quelle di una persona adulta o di uno studente universitario).

L'impiego di una scala di equivalenza nasce appunto dall'esigenza di eliminare, tramite opportuni parametri, le differenze esistenti tra famiglie di diversa ampiezza e composizione. Si tratta quindi di un insieme di coefficienti usati per rendere equivalente, cioè confrontabile, il reddito di ogni famiglia a quello di una generica famiglia di riferimento.

Vi sono numerose scale di equivalenza e, naturalmente, la scelta di quella da utilizzare può condizionare significativamente le successive analisi. Le scale di equivalenza possono essere raggruppate in 5 classi: scale econometriche, pragmatiche, soggettive, basate su bilanci standard e scale implicite nelle misure di assistenza sociale.

La tipologia forse più utilizzata è quella delle *scale econometriche*.

Stimando la relazione econometrica tra le variabili consumo, reddito e numero dei componenti, si può calcolare di quanto deve aumentare il reddito al crescere del numero dei componenti per ottenere la stessa quota di consumi e, quindi, lo stesso

tenore di vita. Una delle scale econometriche più utilizzate è la *scala Carbonaro*, la cui famiglia di riferimento è formata da due persone, alla quale è assegnato come coefficiente 1.

*Scala di equivalenza Carbonaro*

<b>Ampiezza della famiglia</b>	<b>Coefficienti</b>
1	0,60
2 (standard)	1,00
3	1,33
4	1,63
5	1,90
6	2,16
7 o più	2,40

La scala utilizzata nelle statistiche sulla povertà in Italia è appunto quella di Carbonaro, che, come si può vedere nella tabella, prende come riferimento le famiglie composte da due persone ed attribuisce, ad esempio, a quelle di quattro un coefficiente pari a 1,63. Una coppia con due figli, per avere lo stesso tenore di vita di una coppia senza figli con un reddito di 1.000 euro, deve dunque percepire un reddito di 1.630 euro.

## **1.7 Il software utilizzato per le analisi**

Le analisi svolte in questo lavoro sono state condotte utilizzando prevalentemente il software DAD (*Distributive analysis/Analyse Distributive*), che è brevemente presentato nell'Appendice B.



## **Capitolo 2**

### **Il sistema educativo italiano: alcuni richiami su ambiente culturale familiare e stratificazione sociale**

#### **2.1 Introduzione**

In questo capitolo riassumiamo alcune evidenze che documentano come l'Italia sia caratterizzata da un livello di istruzione inferiore rispetto a molti altri paesi sviluppati: c'è un gap elevato per quanto riguarda l'università, più ridotto nel caso delle scuole secondarie superiori. Checchi e Dardanoni (2002) hanno mostrato che, sebbene per le recenti generazioni aumenti sensibilmente il livello d'istruzione, la scelta del percorso scolastico-universitario in Italia è ancora fortemente legata all'ambiente culturale familiare, e quindi all'istruzione dei genitori.

Inoltre, le risorse pubbliche investite per l'istruzione (principalmente per gli stipendi degli insegnanti) non sono appropriatamente distribuite; se si effettua un confronto

con gli altri paesi, infatti, una notevole quota di risorse è assegnata alla scuola dell'obbligo e relativamente poche risorse sono invece destinate all'Università. La letteratura e le Indagini della Banca d'Italia (1993-95-98-2000) evidenziano che la frequenza universitaria non è condizionata tanto dal reddito familiare, quanto piuttosto dal grado di istruzione della famiglia. Gli effetti del reddito e del grado di istruzione dei genitori si riducono sensibilmente, invece, quando consideriamo la frequenza delle scuole secondarie superiori, eccezion fatta per l'influenza del diploma della madre.

## **2.2 L'impronta culturale della famiglia**

L'evidenza della stratificazione sociale nel percorso scolastico può essere schematizzata nel seguente modo: i genitori acculturati sollecitano maggiormente una crescita culturale dei figli, li motivano, li aiutano nei compiti per casa, ed i ragazzi alla fine della scuola dell'obbligo otterranno una valutazione positiva e saranno orientati a dei licei, non saranno bocciati e avranno una grande probabilità di proseguire all'università. Invece, i figli dei genitori poco istruiti finiranno le scuole dell'obbligo con voti bassi, saranno consigliati ad andare in Istituti Tecnici o Professionali; qui saranno poco entusiasti della scuola e del loro curriculum di studi, saranno bocciati più volte ed usciranno con bassa propensione a proseguire all'Università.

Come notato da Checchi (2002), l'accesso all'università è pesantemente condizionato dalla scuola secondaria superiore frequentata, ma l'istruzione della famiglia gioca comunque un ruolo rilevante. Al contrario, il reddito familiare, al netto dell'istruzione familiare, sembra non influenzare il proseguimento degli studi e il livello di istruzione dei figli, cosicché in Italia il fattore culturale sembra essere molto più

importante rispetto al fattore finanziario nell'influenzare l'istruzione universitaria dei figli.

### **2.3 Evidenze**

Con un'analisi un po' più approfondita condotta da Bernardi e Trivellato (1995) relativa all'anno accademico 1996/97 su una coorte di 1.000 giovani italiani (metodo per contemporanei), si nota che una parte di essi (3,6 %) non completa nemmeno la scuola dell'obbligo: essa è concentrata soprattutto nelle regioni del Sud, in particolare nelle aree di degrado. Un ulteriore 17% non va alle scuole medie superiori o le abbandona senza nessun diploma; il 12,8 % si ferma dopo il diploma triennale; il 21,4% raggiunge il diploma quinquennale ma non entra all'università, mentre il 28,1% si ritira da essa senza laurearsi: infine, solamente il 17,1 % riesce a raggiungere una laurea (triennale, quadriennale o quinquennale che sia).

### **2.4 Aspettative sbagliate**

La prima spiegazione riguardo il basso livello di istruzione concerne l'aspettativa di un ritorno economico dall'istruzione. Bisognerebbe riuscire a valutare se vale la pena andare all'Università da un punto di vista strettamente economico, ovvero se il valore attuale netto dei redditi marginali futuri (il surplus rispetto al reddito senza alcun titolo di studio) è maggiore rispetto al costo attuale sostenuto per gli anni di istruzione aggiuntiva. Tale valutazione è difficile da verificare, perché basata sull'aspettativa dei redditi futuri. Siccome i tassi di ritorno economico sono prezzi di

equilibrio, è richiesta la previsione del trend futuro della domanda ed offerta di capitale umano, compito molto difficile in un periodo di cambiamento tecnologico. Naturalmente quando le aspettative di reddito di una professione sono alte, è facile prevedere che crescerà il tasso di iscritti alle facoltà che portano a tali professioni. Le fluttuazioni di iscritti alle varie facoltà universitarie, segnalano la possibilità di aspettative sbagliate riguardo il futuro lavoro conseguente alla scelta del percorso scolastico-universitario.

Da uno studio condotto da Brunello, Lucifora e Winter-Ebner (2002) su un campione di 6.829 studenti universitari da oltre 50 facoltà di 10 stati europei, emerge che essi tendevano a sovrastimare il loro guadagno all'inizio della carriera. Gli studenti italiani invece, avevano un'aspettativa di un ritorno economico inferiore rispetto ai coetanei europei.

## **2.5 Le altre cause**

Un'altra componente estremamente importante da tenere in considerazione è la scarsità di risorse investite nell'istruzione. Da un'analisi un po' più approfondita svolta dall'OECD (2002) relativa all'anno 1999 (come illustrato nella seguente tabella), si nota come l'Italia spenda più degli altri paesi per la scuola dell'obbligo, sia ad un livello intermedio per quanto riguarda la scuola secondaria superiore (meno di U.S.A, Francia e Germania ma più di Giappone ed Inghilterra), e spenda, invece, nettamente meno rispetto a Francia, Germania, Inghilterra, U.S.A. e Giappone per quanto riguarda l'Università.

*Tabella- Spesa per l'Istruzione in base al livello di scuola – dati in milioni di \$*

	Scuola Elementare	Scuola Media	Scuole Superiori	Università
Francia	3.901	4.139	7.152	7.867
Germania	4.937	3.818	6.603	10.393
<b>Italia</b>	<b>5.133</b>	<b>5.354</b>	<b>6.518</b>	<b>7.552</b>
Giappone	3.154	5.240	6.039	10.278
Regno Unito	6.233	3.627	5.608	9.554
U.S.A	6.692	6.582	8.157	192.220
<i>Media OECD</i>	<i>3.847</i>	<i>4.148</i>	<i>5.465</i>	<i>9.210</i>

D'altra parte però, vanno riconosciuti i risultati ottenuti, ricordando che negli ultimi 50-60 anni lo sviluppo culturale medio dell'Italia è stato notevole. Se nel 1950 si arrivava molto spesso solo fino alla licenza elementare (a 10 anni), oggi è in vigore la scuola dell'obbligo fino ai 16 anni.

Il terzo punto molto importante riguardo il basso livello culturale in Italia è il modesto tasso di ritorno (economico) dall'istruzione. Qui è utile riferirsi allo studio di Brunello, Comi e Lucifora (2001) (vedi tabella seguente): essi hanno analizzato le differenze tra l'Italia e i paesi più sviluppati: l'Italia risulta essere il paese che premia maggiormente (sempre in termini economici) il raggiungimento del diploma di maturità, ed il secondo per quanto riguarda il conseguimento della laurea (risultati ottenuti facendo il rapporto sul reddito medio senza avere né diploma né laurea).

*Tabella- Aspettative di reddito in base a livello di istruzione e genere; non diplomati=100*

		Non diplomati	Diplomati	Laureati	Laureati/non diplomati
Francia	Maschi	100	124.08	204.80	2.05
	Femmine	100	136.43	216.53	2.17
Germania	Maschi	100	124.60	181.39	1.81
	Femmine	100	146.86	207.73	2.08
<b>Italia</b>	<b>Maschi</b>	<b>100</b>	<b>188.29</b>	<b>267.09</b>	<b>2.67</b>
	<b>Femmine</b>	<b>100</b>	<b>195.54</b>	<b>229.76</b>	<b>2.30</b>
Regno Unito	Maschi	100	162.81	250.20	2.50
	Femmine	100	140.88	264.02	2.64
U.S.A	Maschi	100	163.88	294.59	2.95
	Femmine	100	173.28	291.82	2.92

## 2.6 Conclusioni

Abbiamo parlato delle aspettative sbagliate, delle risorse limitate investite nell'istruzione, del basso ritorno economico dagli anni di istruzione: nessuna di queste però tiene conto del tipo di comportamento descritto prima, dove avevamo accennato come il tasso di abbandono durante le scuole superiori e l'Università e/o un basso tasso di accesso all'Università siano responsabili del basso livello d'istruzione. Sostanzialmente, la frequenza dell'Università sembra favorita dal reddito familiare solamente quando si esclude l'impronta culturale della famiglia; ma questo cambia non appena si introduce anch'essa nell'analisi. Se di primo acchito le famiglie italiane potrebbero apparire condizionate più che altro dal fattore economico, poichè le

famiglie ricche non hanno problemi a mandare i loro figli all'università, in realtà la scelta del percorso formativo è condizionata dall'istruzione dei genitori.

Il fatto che, in alcuni studi (come indicato dalle stime delle Indagini della Banca d'Italia condotte nel 1993-95-98-00), il reddito familiare abbia un segno negativo statisticamente significativo, potrebbe voler dire che genitori ricchi ma non istruiti sottovalutano l'istruzione dei loro figli e non li mandano all'Università.

## **Capitolo 3**

# **Determinanti della probabilità di abbandono dell'Università in Italia**

### **3.1 Introduzione**

L'Italia è caratterizzata da un tasso di abbandono dell'Università elevato. Per studiare questo fenomeno l'ISTAT ha svolto un'indagine (2001), utilizzando un campione di quasi 23.000 ragazzi diplomati (scuola secondaria superiore) nel 1998, e raccogliendo informazioni riguardo la loro attività fino al 2001. Grazie a questo campione, si sono studiate le determinanti della probabilità di abbandono dell'Università trovando nell'impronta culturale della famiglia (ed in particolare gli anni di formazione scolastica del padre) e nel precedente percorso di istruzione dello studente in questione, le più importanti determinanti della decisione di ritiro.



Discuteremo inoltre le possibili ragioni per cui i coefficienti familiari potrebbero sottostimare il vero effetto dell'impronta culturale della famiglia sulla decisione di continuare e parleremo della mobilità intergenerazionale bassa in termini di livello di istruzione e reddito.

### **3.2 Il fenomeno dell'abbandono dell'Università**

Grande attenzione è stata rivolta sia da politici che da economisti al fenomeno di abbandono dell'Università. L'interesse sull'argomento è principalmente guidato dalla sua importanza quantitativa nei paesi sviluppati: come dimostrato da Cingano e Cipollone (2003), nel 2000 almeno un terzo degli studenti dei Paesi OECD (Francia, Inghilterra, Italia, Germania, Giappone e USA) lasciano l'Università senza raggiungere nessun titolo (laurea quadriennale, laurea triennale, diploma di laurea...). Il tasso di non completamento mette in luce una grande varianza nei paesi sviluppati, con un intervallo che va dal quasi 20% in Irlanda, Giappone, Turchia e Regno Unito all'oltre 40 % in Austria, Francia, Italia e Svezia.

Ovviamente parte delle differenze osservate tra i paesi derivano da differenze nel sistema educativo. Per esempio, gli studenti iscritti in corsi di laurea brevi (2 o 3 anni) hanno una grande probabilità di completarli con successo, mentre quelli iscritti in paesi dove tali corsi non esistono corrono un grande rischio di non riuscire a completare la formazione universitaria. Anche se un fattore come la lunghezza dei corsi di laurea è in grado di spiegare parte delle differenze osservate fra i vari paesi, un'ulteriore motivazione potrebbe risiedere nella maggiore o minore capacità di un sistema educativo di ridurre le iniquità sociali in termini di opportunità educative. Per esempio, il fatto che due studenti altrimenti identici siano esposti ad una differenza significativa di rischio di ritiro che dipende dal reddito e/o dall'impronta culturale delle loro famiglie, può essere considerato un risultato molto indesiderato.

### **3.3 I motivi dell'interesse**

In Italia il fenomeno di abbandono dell'università ha meritato una tale attenzione per due ragioni.

Primo, il tasso di abbandono in Italia è il più alto fra i paesi dell'OECD da oltre 30 anni a questa parte. Dal 1950 ai primi anni '80 il tasso di abbandono ha infatti sfiorato il 70 % e solamente negli ultimi anni è in lieve declino. In Italia, l'elevato tasso di abbandono coesiste con un livello medio di istruzione basso ed una percentuale di laureati sulla popolazione totale particolarmente bassa (19 %).

Il secondo motivo è spiegato da Checchi (1999): l'Italia dimostra una grande persistenza intergenerazionale quanto a livello di istruzione e reddito familiare rispetto agli altri paesi sviluppati; ciò significa che le famiglie svantaggiate sono meno propense ad investire per l'istruzione dei loro ragazzi.

La recente riforma del sistema universitario, che sta introducendo dei corsi di laurea brevi per cercare di colmare il gap del tasso di abbandono tra l'Italia e gli altri paesi dell'OECD, potrebbe avere un impatto piuttosto limitato in termini di equalizzazione delle opportunità educative, se gli studenti lasciano l'Università non per la lunghezza del corso di laurea ma a causa della mancanza di risorse (finanziarie) o dell'insufficiente cultura (universitaria) della famiglia.

### **3.4 Effetti dei differenti tipi di scuola superiore**

Riguardo agli effetti dell'impronta culturale della famiglia, il presente paragrafo contribuisce a valutare gli esiti dei differenti tipi di scuola (superiore) sulla decisione di continuare. Evans e Schwab (1995), Neil (1997) e Altonji (2000) si sono

focalizzati in particolare sul ruolo giocato negli USA dalle scuole private cattoliche in contrapposizione alle scuole pubbliche, sulle performance accademiche. In Europa, avendo una grande varietà di tipi di scuole secondarie superiori, l'attenzione è stata rivolta alla divisione tra scuole tecniche e generali (rispettivamente Istituti Tecnici e Licei). Ovviamente le stime appropriate di tali effetti richiedono di tener conto della possibilità che gli studenti scelgano tra i differenti tipi di scuola sulla base di dimensioni non osservabili correlate con la loro abilità scolastica.

Dall'estensione di questo controllo (cioè catturando le precedenti performance scolastiche che sono in grado di controllare l'eterogeneità non osservata), i risultati di Bertola e Checchi (1999), Dustmann (2001), Margolis e Simonnet (2002) e Cappellari (2003) suggeriscono che i differenti tipi di scuola (in particolare Licei contrapposti agli Istituti Tecnici) hanno un impatto non trascurabile sulla successiva transizione (e carriera) all'Università.

I risultati puntualizzano l'esistenza di forti effetti dell'impronta culturale della famiglia sulla possibilità di abbandono dell'Università.

Questa analisi può non essere efficiente, se il ruolo della famiglia si limita a vincoli finanziari o al miglioramento delle condizioni in termini di motivazioni allo studio.

### **3.5 Effetti familiari indiretti**

La valutazione degli effetti marginali non tiene conto che le misure dell'impronta culturale (come il tipo di scuola superiore) potrebbero essere influenzate da altri regressori inclusi nella descrizione stessa (come l'impronta culturale della famiglia). Dal momento che il campione dell'ISTAT non include studenti che non sono iscritti alla scuola superiore o che l'hanno abbandonata, non si possono esprimere le stime degli effetti dell'impronta culturale della famiglia sulla scelta della scuola superiore.

Tuttavia, l'esplorazione informale del campione indica che i figli di genitori molto istruiti hanno generalmente una probabilità molto maggiore di iscriversi in scuole orientate all'università (Licei), che sono fortemente correlate sia con una maggiore frequenza universitaria, sia con un rischio minore di ritirarsi da essa.

Va messo in evidenza come, rispetto alle stime previste dallo studio di Cingano e Cipollone (2003) sugli effetti dell'impronta culturale della famiglia sull'abbandono dell'università, aumentino i coefficienti degli anni di istruzione di entrambi i genitori; in particolare, l'importanza degli anni di istruzione della madre aumentano più del doppio.

Questi risultati suggeriscono che gli indicatori della scelta dell'istruzione capta parte degli effetti dell'impronta culturale della famiglia sul ritiro, dal momento che i figli di famiglie più o meno avvantaggiate tendono ad iscriversi rispettivamente ai Licei o ad altre scuole.

Quindi, il vero effetto dell'impronta culturale della famiglia sulla decisione di proseguire gli studi all'università potrebbe essere maggiore rispetto ai coefficienti familiari stimati.

Infine, dallo studio condotto da Cingano e Cipollone, non è stato colto nessun cambiamento rilevante negli effetti stimati delle condizioni economiche locali (catturate mediante il tasso di attività della forza lavoro).

### **3.6 Determinanti della probabilità di abbandono**

In questo capitolo si è cercato di analizzare le determinanti della probabilità di abbandono dell'Università guardando a studi condotti su un campione rappresentativo di italiani diplomati. In particolare, tale studio si è focalizzato sulle seguenti variabili: la posizione socio-economica, l'impronta culturale della famiglia e

le condizioni locali (misurate attraverso il tasso di disoccupazione locale e la misura del grado di urbanizzazione), controllando le caratteristiche osservabili individuali.

I risultati evidenziano così fatti veramente interessanti.

Primo: controllando l'impronta culturale della famiglia e le condizioni locali, e tenendo conto dei costi della frequenza universitaria, sono stati notati forti effetti dell'impronta familiare sulla decisione di ritirarsi dall'università. Infatti, i figli di padri laureati sono soggetti ad un rischio di ritiro significativamente più basso rispetto ai ragazzi nati in famiglie svantaggiate.

Secondo, il percorso scolastico individuale, che agisce attraverso la scelta della scuola secondaria superiore, è anch'esso parte degli effetti dello sfondo culturale della famiglia.

Tenendo conto della già osservata mobilità intergenerazionale bassa in termini di livello di istruzione e di reddito familiare (reddito e titolo di studio sono variabili fortemente correlate), il percorso scolastico dovrebbe quindi essere imputato anche allo status socio-economico familiare, dal momento che i figli di famiglie più o meno avvantaggiate tenderanno ad iscriversi rispettivamente ai Licei o ad altre scuole.

## **Capitolo 4**

### **Elaborazione dei dati**

#### **4.1 Operazioni preliminari**

Dopo aver scaricato i dati dall'archivio della Banca d'Italia si è passati alla costruzione di due basi di dati, una per l'anno 2000 e una per il 2002.

Si è quindi passati a lavorare questi dati con SAS.

In un primo momento, il nostro lavoro è stato quello di riunire le variabili di interesse dei files familiari rfam, ricfam, confam e g (che descrivono rispettivamente il reddito familiare, la ricchezza familiare, i consumi familiari ed il godimento dell'abitazione) in un'unica tabella, e le variabili di interesse dei files individuali rper e carcom (rispettivamente redditi individuali e caratteristiche socio-demografiche personali) in un'altra. Ovviamente, abbiamo svolto queste operazioni sia con i files dell'indagine 2000 che con quelli dell'indagine 2002.

Questa operazione è stata necessaria perché sarebbe impossibile elaborare con SAS tabelle contenenti dati individuali (con 22.000 osservazioni) e tabelle contenenti dati familiari ( 8.011 osservazioni) contemporaneamente.

Per risolvere questo problema abbiamo perciò creato, grazie al comando RETAIN, un unico dataset che contenesse in un file familiare sia le variabili del file individuale sia le variabili del file familiare.

Per non trovarci di fronte ad una numerosità campionaria troppo bassa, abbiamo fatto un pooling di famiglie dalle indagini del 2000 e del 2002, contando una sola volta le famiglie che erano presenti in entrambe le indagini.

Per abbinare i dati è stata utilizzata la variabile chiave N\_QUEST che è il valore identificativo della famiglia

Abbiamo così unito le due tabelle in un unico database e con il comando NODUPKEY abbiamo eliminato le famiglie “doppie”.

## **4.2 Ridefinizione dei pesi**

Dopo aver creato il campione pooled della popolazione totale, combinando i dati dell'indagine 2000 e quelli del 2002, la successiva elaborazione è consistita nel riassegnare i pesi campionari: abbiamo già la variabile *\_peso*, ossia l'inverso della probabilità di inclusione nel campione, relativo alle 4396 famiglie del 2000 e alle 8011 del 2002.

Se vogliamo che anche il nostro campione di 12407 famiglie rappresenti correttamente i 20 milioni di famiglie italiane, è necessario riassegnare i pesi campionari.

Operativamente, dobbiamo fare in modo che la somma degli 8011 pesi delle famiglie del 2002, sia uguale alla somma dei 12407 pesi delle famiglie che derivano dall'unione dei dataset come descritto sopra. Questa operazione è stata svolta in SAS

semplicemente utilizzando gli operatori aritmetici, e la nuova variabile l'abbiamo chiamata *PESOFI*.

### 4.3 Calcolo dei redditi familiari equivalenti

L'operazione seguente da fare in SAS è standardizzare i redditi familiari (YFAM) del nostro dataset con una scala di equivalenza, per rendere equivalenti, e quindi confrontabili, i redditi tutte le famiglie a quello di una generica famiglia di riferimento (composta da due persone).

Abbiamo adottato la scala di equivalenza di Carbonaro, ed anche in questo caso, è stato sufficiente utilizzare gli operatori aritmetici di SAS. La nuova variabile è stata denominata *YFAM\_EQUI*.

La formula utilizzata per calcolare il reddito equivalente, *YFAM\_EQUI*, di ciascuna famiglia è la seguente:

$$YFAM\_EQUI_{N\_QUEST} = \frac{YFAM_{N\_QUEST}}{C(N\_COMP_{N\_QUEST})}$$

laddove *N\_QUEST* è l'identificatore della famiglia, *YFAM* è il reddito totale familiare, *N\_COMP* è il numero dei componenti e *C(.)* è il coefficiente della Scala Carbonaro che dipende dal numero dei componenti.

Arrivati a questo punto, dobbiamo esportare i dati del File SAS **Poptotale** in un file Excel (grazie al comando File – Esporta Dati... e seguendo il procedimento suggerito da SAS).



## 4.4 Selezione dei due campioni

Come avevamo già spiegato nel capitolo 1, in questa indagine abbiamo pensato di individuare 2 sottocampioni, rappresentativi di 2 sottopopolazioni.

Il primo sottocampione (composto da 1085 famiglie) è costruito individuando la popolazione delle famiglie con figli in età compresa tra i 16 ed i 20 anni e con titolo di studio più elevato la licenza di scuola media (o il diploma professionale triennale) che siano studenti.

In SAS, per fare tale selezione, abbiamo dovuto naturalmente lavorare con i dataset individuali, e dare i seguenti comandi:

```
data terry.passaggio;  
set terry.daniele;  
if studio=3;  
run;
```

```
data terry.passaggio;  
set terry.passaggio;  
if 16<= eta <=20;  
run;
```

```
data terry.passaggio;  
set terry.passaggio;  
if aponoc1=6;
```

Nel secondo sottocampione (costituito da 585 famiglie) abbiamo invece individuato il sottoinsieme delle famiglie con figli in età compresa tra i 18 ed i 22 anni con titolo di studio più elevato il diploma di scuola media Superiore che non siano studenti, individuando così i ragazzi che pur avendone titolo, non proseguono gli studi all'università.

Anche qui ovviamente abbiamo lavorato su dataset individuali, ed abbiamo dato i seguenti comandi:

```
data terry.cross;
set terry.daniele;
if studio=4;
run;
```

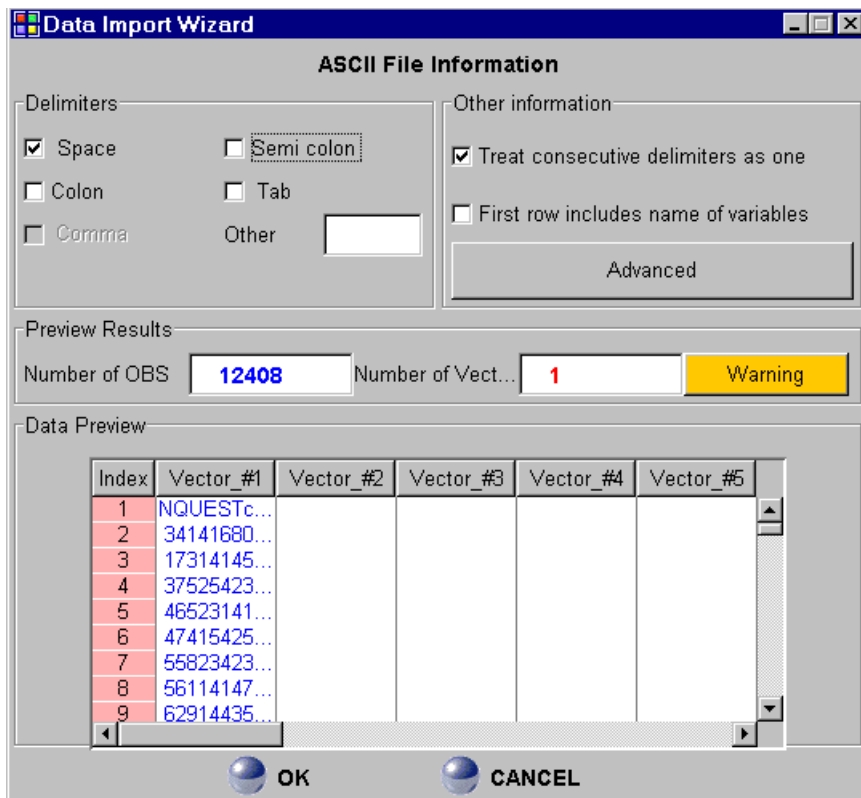
```
data terry.cross;
set terry.cross;
if 18<= eta <=22;
run;
```

```
data terry.cross;
set terry.cross;
if aponoc1 ^= 6;
run;
```

Anche in questo caso, dobbiamo esportare i dati dei files SAS **Sottocampione1** e **Sottocampione 2** in files Excel (grazie al comando File – Esporta Dati...).

## 4.5 Il caricamento delle basi di dati in DAD

Dopo averlo installato, eseguiamo il software DAD. Ci appare il foglio elettronico. Il primo passo consiste nel caricare le basi di dati di interesse, cioè quelle riguardanti i dati del campione pooled e dei due sottocampioni. Come sappiamo, i files da importare sono *Popolazione totale.txt*, *Sottocampione1.txt* e *Sottocampione2.txt*. Importiamo quindi il nostro file Excel in DAD selezionando dal menù principale l'opzione FILE/OPEN ed inserendo il nome del file da importare (nel nostro caso *Popolazione totale.txt*) e digitiamo INVIO: ci apparirà la finestra seguente Data Import Wizard:



Per caricare la base di dati, è di fondamentale importanza selezionare le opzioni che ci interessano perché essa sia caricata nel migliore dei modi e, dopo aver controllato l'esatto caricamento dei dati tramite il numero di osservazioni, il numero di variabili e la tabella di anteprima, clicchiamo sul pulsante OK.

Tale procedimento va seguito anche per importare i dati riguardanti i due campioni caricando i files *Sottocampione1.txt* e *Sottocampione2.txt*.

Abbiamo così caricato le tre basi di dati di interesse su DAD.

A questo punto si può selezionare (EDIT/SET SAMPLE DESIGN) il vettore dei pesi campionari (nel nostro caso PESOFL).

Ovviamente lo facciamo per i dati di tutti e tre i nostri files.

Dopo aver caricato le basi di dati e aver specificato lo schema di campionamento, si può procedere con il calcolo delle statistiche che ci interessano.

## 4.6 Curve

DAD ha incorporati strumenti che facilitano l'uso di curve per visualizzare informazioni sulle distribuzioni.

Ad esempio, per rappresentare graficamente la distribuzione del campione pooled rispetto al reddito, si deve selezionare dal menù principale, CURVES/QUANTILE (indicando il numero di distribuzioni: in questo caso 1).

Dopo aver scelto le variabili inerenti all'applicazione, cliccare sul pulsante "Graph". Per disegnare la curva di un'altra distribuzione (per esempio del sottocampione 1), bisogna ritornare alla finestra precedente, reinizializzare la variabile di interesse e cliccare nuovamente sul pulsante "Graph".

Quando appare la finestra del grafico, cliccare sul pulsante "Draw all" per disegnare tutte le curve.

## **Capitolo 5**

### **Risultati finali**

#### **5.1 Considerazioni introduttive**

Per svolgere questo lavoro, abbiamo utilizzato i dati dell'indagine sui bilanci delle famiglie della Banca d'Italia del 2000 e del 2002 e per analizzare i dati è stato utilizzato il pacchetto di programmi DAD 4.3.

L'obiettivo del nostro lavoro è quello di individuare la popolazione a rischio di ingresso all'Università, per la quale vi possono essere vincoli che ne ostacolano l'accesso.

Come già detto, il nocciolo sta nel capire se tali vincoli sono di natura prettamente economica oppure sono rappresentati da una scarsa cultura della famiglia.

## 5.2 La relazione tra reddito e livello di istruzione

Come già detto nei precedenti capitoli, la letteratura dice che l'Italia ha una grande persistenza intergenerazionale in quanto a livello di istruzione e reddito familiare rispetto agli altri paesi sviluppati: questo significa che le famiglie svantaggiate sono meno propense ad investire per l'istruzione dei loro ragazzi, i quali non arriveranno a raggiungere un titolo di studio elevato, e percepiranno (una volta entrati nel mondo del lavoro) un reddito modesto: saranno così a loro volta genitori non orientati ad investire molto per l'istruzione dei loro figli.

Infatti, come dimostrato dai seguenti risultati, reddito e titolo di studio sono variabili fortemente correlate tra di loro.

A questo proposito, specifichiamo bene la variabile Titolo di studio (che come tutte le altre variabili utilizzate in questo lavoro verrà spiegata più avanti nell'Appendice A):

STUDIO = Titolo di studio del capofamiglia

1 = nessun titolo di studio

2 = licenza elementare

3 = licenza media

4 = diploma

5 = laurea

6 = specializzazione post-laurea

### CAMPIONE POOLED

Titolo di studio	Media	Mediana	1. quintile
1	19.672	16.227	11.125
2	24.671	20.770	12.863
3	26.915	22.349	13.252
4	35.741	29.931	18.436
5	55.313	43.999	26.143
6	80.392	64.254	31.579

### SOTTOCAMPIONE 1

Titolo di studio	Media	Mediana	1. quintile
1	9.198	9.050	4.172
2	18.234	15.423	7.669
3	22.114	19.018	12.105
4	27.954	24.564	15.446
5	39.237	32.669	20.666
6	82.391	94.142	64.254

### SOTTOCAMPIONE 2

Titolo di studio	Media	Mediana	1. quintile
1	5.626	1.895	1.895
2	26.728	20.028	9.015
3	24.251	20.378	13.497
4	28.780	24.619	17.703
5	37.202	40.799	19.595
6	/	/	/

Sia per la media, per la mediana e per il primo quintile di tutte e tre le popolazioni, notiamo che il reddito cresce significativamente al crescere del livello di istruzione. Precisato questo, è immediato evidenziare una forte correlazione tra reddito e titolo di studio: notiamo in particolare un aumento considerevole del reddito al conseguimento della laurea (il reddito medio dei laureati è 55.313 Euro, +55% rispetto a quello dei diplomati, mentre il reddito mediano è 43.999 Euro, +47% rispetto a quello dei diplomati).

Inoltre, rileviamo un aumento significativo del reddito al conseguimento del diploma quinquennale delle scuole secondarie superiori: nel campione pooled il reddito mediano dei diplomati è 29.931 Euro, +34% rispetto al reddito mediano di quelli che hanno la licenza media (+29% nel sottocampione 1 e +20% nel sottocampione 2) .

Infine, è bene sottolineare come senza alcun titolo di studio la situazione sia gravosa: come si può notare soprattutto nei due sottocampioni, già la sola licenza elementare migliora notevolmente le cose.

In sintesi, possiamo concludere dicendo che come ci aspettavamo reddito e titolo di studio sono variabili fortemente correlate, e che come è immediato notare dalle tabelle, il diploma e soprattutto la laurea sono i titoli di studio più rilevanti per accrescere il reddito.

### **5.3 Reddito**

Dopo aver stabilito che reddito e titolo di studio sono variabili fortemente correlate tra di loro, dobbiamo tornare al nostro principale obiettivo: quale di queste due variabili è più importante per determinare l'ingresso all'università? I vincoli che ne ostacolano l'accesso sono prevalentemente di natura economica?



Cominciamo quindi studiando la distribuzione di queste tre popolazioni rispetto al reddito familiare standardizzato con le scale di equivalenza ed effettuando dei confronti tra la distribuzione di ciascuno dei due sottocampioni con quella del campione pooled.

Guardiamo la distribuzione delle 3 popolazioni rispetto al reddito in questa tabella.

	CAMPIONE POOLED	SOTTOCAMPIONE 1	SOTTOCAMPIONE 2
1. quintile	14.518	13.664	14.380
2. quintile	21.340	18.729	19.497
Mediana	24.827	21.653	22.540
3. quintile	28.970	24.795	25.991
4. quintile	42.617	35.145	38.543
5. quintile	1.084.367	223.040	177.444
Media	31.194	25.983	26.821

Subito si nota una cosa importante: delle 3 popolazioni, quella che ha il reddito medio (ma anche mediano e soprattutto il primo quintile) più basso è il primo sottocampione, cioè quello delle famiglie con figli che conseguono il diploma e possono quindi proseguire gli studi all'Università.

Il secondo sottocampione, quello delle famiglie con figli che avendone titolo, non proseguono gli studi all'Università, ha una distribuzione del reddito che sta sopra a quella del primo campione e soprattutto ha il primo quintile (vedi tabella) significativamente superiore a quello del primo sottocampione (+5.24%). Questa è un'indicazione per noi importante, poiché significa che le famiglie che non mandano i loro figli all'Università, seppure diplomati e quindi con la possibilità di accedervi, non sembrano farlo per motivi di reddito: esse hanno addirittura un reddito superiore al campione totale delle famiglie con figli diplomati.

Questa asserzione è confortata dal fatto che il primo quintile del secondo sottocampione è molto vicino al primo quintile del campione pooled (-0,96%): questo dovrebbe confermarci che il segmento a basso reddito del secondo campione coincide praticamente con quello della popolazione totale e quindi per esso, il campione delle famiglie con figli che pure avendone titolo, non accedono all'Università, i vincoli che ne ostacolano l'accesso non dovrebbero essere di natura economica.

#### **5.4 Impronta culturale della famiglia**

Per completare la nostra analisi studiamo ora il fattore socio-culturale (che siamo in grado di esaminare attraverso il titolo di studio del padre), che potrebbe rivelarsi il fattore che incide maggiormente sul proseguimento (o abbandono) all'università. Anche qui andiamo a studiare la caratterizzazione del campione pooled e dei 2 sottocampioni, stavolta però rispetto al titolo di studio del capofamiglia.

##### **CAMPIONE POOLED**

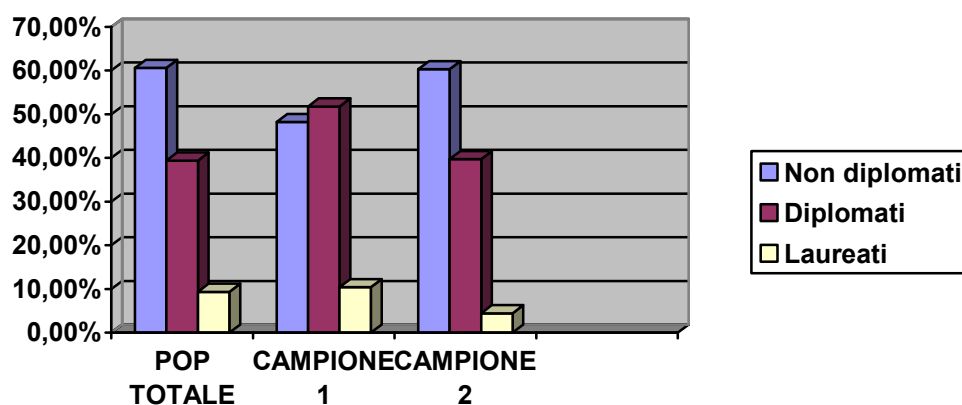
LAUREATI	9.36 %
DIPLOMATI	39.41 %
SENZA DIPLOMA	60.59 %

##### **SOTTOCAMPIONE 1**

LAUREATI	10.41 %
DIPLOMATI	51.79 %
SENZA DIPLOMA	48.21 %

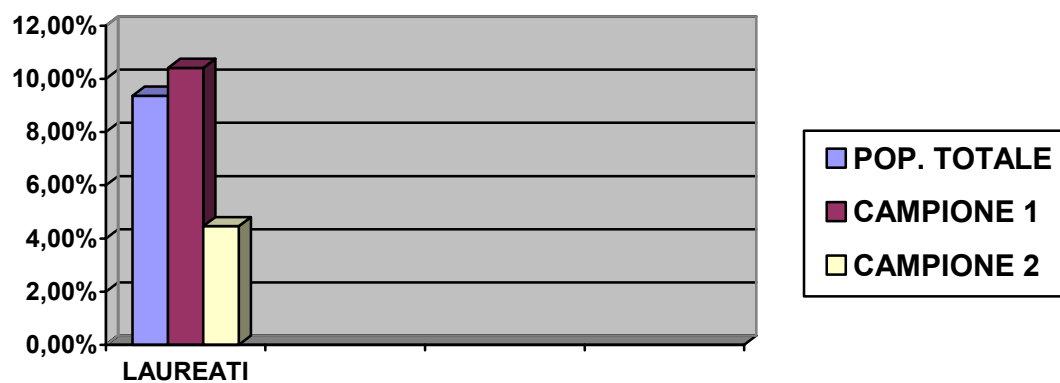
## SOTTOCAMPIONE 2

LAUREATI	4.46 %
DIPLOMATI	39.71 %
SENZA DIPLOMA	60.29 %



In particolare, ci focalizziamo sulle percentuali di capofamiglia laureati nelle nostre 3 popolazioni, mettendole a confronto grazie alla seguente tabella e all'istogramma.

	CAMPIONE POOLED	SOTTOCAMPIONE 1	SOTTOCAMPIONE 2
Laureati	9,36 %	10,41 %	4,46 %



Troviamo un risultato importante per la nostra ricerca: come risulta evidente dall'ultimo istogramma, la percentuale di laureati del secondo sottocampione (4,46%), molto bassa, è meno della metà di quella del primo sottocampione. Quindi, nel secondo campione, quello delle famiglie con figli diplomati (scuola secondaria superiore) che non hanno proseguito gli studi all'università, l'abbandono non è avvenuto per vincoli economici, ma per scelte di altra natura, e probabilmente per la scarsa cultura universitaria della famiglia.

Da evidenziare pure la percentuale di diplomati del primo sottocampione, nettamente superiore a quelle del secondo sottocampione e del campione pooled: questo risultato ci dà la conferma dell'importanza dell'impronta culturale della famiglia. Infatti, i ragazzi del primo campione sono quelli diplomati e con la possibilità di proseguire gli studi all'università: appare perciò evidente come anche per la frequenza delle scuole secondarie superiori e per il raggiungimento del diploma quinquennale incida molto l'impronta culturale delle famiglie.

Dall'analisi di questi dati emerge quindi che la scelta del percorso scolastico-universitario in Italia è ancora fortemente legata all'istruzione familiare.

L'importanza dell'impronta familiare può essere riassunta nel seguente modo: i genitori acculturati motivano maggiormente i figli durante la scuola dell'obbligo, i ragazzi saranno orientati a dei licei ed avranno una grande probabilità di andare all'Università; i figli di genitori poco istruiti invece, avranno scarso successo fin dalle scuole dell'obbligo, saranno consigliati ad andare in Istituti Tecnici o Professionali, dai quali usciranno con pochissima intenzione di andare all'Università.

## 5.5 Osservazioni conclusive

Avevamo già detto in precedenza che reddito e titolo di studio sono variabili correlate tra di loro, ma quello che ora ci è chiaro, è che l'impronta culturale della famiglia incide di più sull'ingresso all'Università.

Inoltre, come già affermato nel capitolo 3, l'accesso all'Università è influenzato fortemente dalla scuola secondaria superiore frequentata, ma l'istruzione della famiglia gioca un ruolo rilevante anche nel condizionare la scelta della scuola superiore.

L'indagine dell'ISTAT del 2001 fatta utilizzando un campione di 23.000 diplomati, indica infatti che i figli di genitori molto istruiti hanno generalmente una probabilità molto maggiore di iscriversi al Liceo e successivamente all'università, ed hanno inoltre una probabilità minore di ritirarsi da essa.

Il nostro principale risultato evidenzia quindi che la frequenza universitaria non è condizionata dal reddito familiare, quanto piuttosto dal grado di istruzione della famiglia. O meglio, la frequenza dell'Università sembra favorita dal reddito familiare solamente quando si esclude l'impronta culturale della famiglia; questo cambia però non appena si introduce anch'essa nell'analisi.

Infatti, se di primo acchito le famiglie italiane potrebbero apparire condizionate più che altro dal fattore economico, poiché le famiglie ricche non hanno problemi a pagare le tasse universitarie e mandare i loro figli al college, in realtà la scelta del percorso formativo è condizionata dall'istruzione dei genitori.

Il fatto che il reddito familiare abbia un segno negativo statisticamente significativo, potrebbe voler dire che genitori ricchi ma non istruiti sottovalutano l'istruzione dei loro figli e non li mandano all'Università.

Come affermato dalla letteratura, i figli di padri laureati sono soggetti ad un rischio di ritiro dall'università significativamente più basso rispetto ai ragazzi nati in famiglie

svantaggiate, e come abbiamo visto nel nostro studio, i ragazzi che pur avendone la possibilità non proseguono gli studi, provengono nella maggioranza dei casi da famiglie con una scarsa cultura universitaria.

## **Appendice A**

### **Indagine della Banca d'Italia**

#### **A.1 Introduzione**

Dal 1965 la Banca d'Italia conduce ogni due anni un'indagine sui bilanci delle famiglie italiane allo scopo di acquisire una più approfondita conoscenza dei comportamenti economici delle famiglie.

Le indagini campionarie prese in considerazione per sviluppare questa ricerca sono: *“l'indagine sui bilanci delle famiglie italiane nell'anno 2000”* e *“l'indagine sui bilanci delle famiglie italiane nell'anno 2002”*.

## **A.2 Le indagini del 2000 e del 2002**

Le interviste inerenti l'indagine del 2000 sono state condotte nel periodo compreso tra febbraio e luglio del 2001.

L'indagine ha riguardato 8.001 famiglie, di cui 3.873 sono famiglie *panel*, ovvero intervistate anche nella precedente rilevazione (o in indagini precedenti) e le rimanenti 4.128 sono famiglie intervistate per la prima volta.

Per quanto riguarda l'indagine 2002 le interviste si sono svolte nel periodo compreso tra febbraio e settembre del 2003.

L'indagine ha coinvolto 8.011 famiglie; la quota del campione costituita da famiglie *panel* è risultata pari al 45 %.

In accordo con il piano di campionamento, spiegato nei paragrafi successivi, a ciascuna famiglia viene attribuito un peso inversamente proporzionale alla sua probabilità di inclusione nel campione.

I pesi vengono successivamente modificati sia per aumentare la precisione degli stimatori sia per allineare la struttura del campione a quella della popolazione in base alle informazioni esogene sulla distribuzione per sesso, classe di età, ampiezza comunale e area geografica di residenza della popolazione note da fonte Istat.

## **A.3 Il disegno campionario**

Le indagini sui bilanci delle famiglie italiane condotte dalla Banca d'Italia sono state effettuate, fino al 1987, sulla base di rilevazioni indipendenti nel tempo.

A partire dall'indagine sul 1989, per favorire l'analisi dell'evoluzione dei fenomeni rilevati, è stato introdotto uno schema di rilevazione che prevede la presenza nel campione di una quota di unità già rilevate in occasioni di precedenti indagini



(famiglie panel). Una famiglia panel mantiene lo stesso numero di questionario (N\_QUEST) utilizzato nell'indagine (o nelle indagini) precedente. In questo modo è possibile studiare l'evoluzione dei fenomeni con maggiore accuratezza; si può ad esempio valutare la mobilità delle famiglie tra classi di reddito o ricchezza o degli individui tra i diversi stati occupazionali.

L'estrazione del campione è stata effettuata seguendo uno schema a due stadi (comuni e famiglie), con stratificazione delle unità di primo stadio (comuni) secondo la regione e la classe di ampiezza demografica del comune. All'interno di ogni strato, vengono dapprima individuati i comuni nei quali effettuare interviste, includendo tutti quelli con popolazione superiore ai 40.000 abitanti ed estraendo casualmente quelli di dimensione inferiore e successivamente vengono estratte casualmente le famiglie da intervistare.

Per quanto riguarda il 2002, per la parte panel del campione, i comuni sono stati estratti tra quelli già oggetto di rilevazione nell'indagine sul 2000 (comuni panel); le famiglie residenti in questi comuni che avevano partecipato a due o più rilevazioni sono state tutte reinserite nel campione; le 34 restanti famiglie panel da intervistare sono state estratte casualmente tra quelle rilevate solo nella precedente indagine.

Le famiglie non panel sono state invece estratte casualmente dalle liste anagrafiche, sia nei comuni panel che in quelli non panel.

**Tav. 1**

Area geografica	Comuni di rilevazione		Totale
	<i>Panel</i>	<i>Non panel</i>	
Nord .....	139	13	152
Centro .....	66	5	71
Sud e Isole .....	108	13	121
<b>Totale .....</b>	<b>313</b>	<b>31</b>	<b>344</b>

#### **A.4 Mancate interviste e attendibilità delle risposte**

Le mancate interviste possono costituire un problema nelle indagini statistiche in quanto possono condurre a campioni nei quali i segmenti di popolazione meno disposti a collaborare possono essere sottorappresentati, producendo stime distorte (*selectivity bias*).

Un'indicazione in proposito si ricava valutando il numero dei contatti necessari per ottenere l'intervista: relativamente all'anno 2002, per effettuare le 8.011 interviste gli intervistatori hanno compiuto 15.068 azioni di contatto.

La difficoltà a ottenere l'intervista è crescente al crescere del reddito e del titolo di studio del capofamiglia; minori difficoltà si incontrano con le famiglie con un ridotto numero di componenti, con capofamiglia pensionato e di sesso femminile.

Per limitare i possibili effetti negativi della mancata partecipazione sono state adottate alcune misure. In primo luogo, le famiglie che non è stato possibile intervistare sono state sostituite con altre estratte con criteri casuali negli stessi comuni.

In secondo luogo, al termine della rilevazione, è stata effettuata una stratificazione a posteriori sulla base di alcune caratteristiche individuali dei soggetti intervistati, che consente di riequilibrare, all'interno del campione, il peso dei diversi segmenti di popolazione.

#### **A.5 Calcolo dei pesi campionari delle famiglie**

Il campione non è autoponderante, cioè è necessario l'uso di pesi per restituire la struttura della popolazione (e ottenere stime non distorte). Utilizzando lo stimatore di *Horwitz-Thompson (HT)*, il peso campionario di ciascun individuo della famiglia è

dato dall'inverso della probabilità di inclusione nel campione. Ad esempio: se in una popolazione di 1.000 maschi e 1.000 femmine ho estratto un campione dai rispettivi strati di 10 maschi e 20 femmine, allora a ciascuna unità sarà attribuito il peso  $w_m$  o  $w_f$  (a seconda se maschio o femmina):

$$w_m = 1000/10 \qquad \text{oppure} \qquad w_f = 1000/20$$

In generale, si scrive:  $w_h = N_h/n_h$ .

Come già spiegato nel paragrafo A.3, nell'indagine la selezione del campione avviene in due stadi. Prima si selezionano i comuni e poi, casualmente al suo interno, le famiglie. I comuni con oltre 40.000 sono selezionati (quasi) tutti. I comuni con meno di 40.000 abitanti sono selezionati secondo il metodo *PPS* (Probability Proportional to Size), cioè più è grande il comune (in termini di popolazione) e maggiore è la sua probabilità di essere incluso nel campione:  $\text{prob}I = m_h \Phi_h / \Phi$ . In entrambi i casi, una volta estratto il comune, la probabilità di ciascuna famiglia è inversamente proporzionale alla dimensione del comune:  $\text{prob}II = n_{ih} / \Phi_h$ .

## **A.6 Ulteriori considerazioni sulla qualità dei dati**

Sebbene per un'indagine sul reddito e la ricchezza la mancata risposta e l'*underreporting*, di cui si è già detto, costituiscano senz'altro i fenomeni che maggiormente possono determinare problemi di qualità delle stime, numerosi altri fattori sono indicati in letteratura come potenziali elementi di disturbo.

Il questionario, in primo luogo, non è uno strumento neutro: il modo in cui le domande sono poste, come sono ordinate e poste in relazione tra di loro, le modalità di risposta previste, sono tutti elementi in grado di influire sulle risposte degli intervistati, in particolare – ma non solo – nella rilevazione di opinioni, aspettative e altri elementi di natura soggettiva.

Un ulteriore fattore importante è costituito dagli intervistatori, il cui atteggiamento o modo di illustrare le domande può influire sulla risposta degli intervistati.

Altri problemi possono derivare dai processi cognitivi del rispondente; al di là degli elementi intenzionali riconducibili ai fenomeni di *under-reporting*, le risposte a determinate domande sono influenzate dalla capacità di elaborazione (ad esempio per le domande ipotetiche) o di memoria (per le domande retrospettive) del rispondente. Vanno inoltre considerati gli elementi di conoscenza vera e propria, soprattutto quando, per questioni pratiche, è talvolta consentito a un soggetto di rispondere per conto di un altro.

Aspetti più generali, come la motivazione del rispondente, il tempo e l'impegno a lui richiesto sono ulteriori elementi che sono in grado di influire sulla qualità delle risposte fornite.

## **A.7 Descrizione delle basi di dati e delle variabili**

Gli archivi di dati utilizzati, ricavati dalle indagini della Banca d'Italia, sono di dominio pubblico: si può facilmente disporre di questi documenti scaricandoli dal sito ufficiale della Banca d'Italia all'indirizzo [www.bancaditalia.it](http://www.bancaditalia.it) alla voce "statistiche-consultazione dei dati".

Le informazioni così scaricate, non sono però collegate tra di loro: per farlo c'è stato quindi bisogno di interrogare i dati ed incrociarne i risultati utilizzando come chiave la variabile N\_QUEST che identifica univocamente la famiglia. Dalle basi di dati ricavate sono state selezionate le variabili di interesse creando tre nuove tabelle esportate sui files di testo *Popolazione totale.txt*, *Sottocampione1.txt* e *Sottocampione2.txt*. Successivamente sono state importate dal software utilizzato.

Le variabili, seguite dalla loro descrizione, presenti nelle tre basi di dati utilizzate in questo lavoro sono le seguenti:

NQUEST	numero del questionario, che identifica univocamente la famiglia
ANNO	anno della rilevazione
YL	reddito da lavoro dipendente
YT	reddito da pensioni e da altri trasferimenti
YM	reddito da lavoro autonomo e impresa
YC	reddito da capitale (reddito da fabbricati, interessi su titoli di stato, interessi su depositi bancari e postali)
CN	consumi non durevoli
CD	spese per l'acquisto di beni durevoli
AR	attività reali (immobili, aziende, oggetti di valore)
AF	attività finanziarie (titoli di stato, azioni, obbligazioni, depositi bancari e postali)
PF	passività finanziarie
PESOFI	peso campionario: è l'inverso della probabilità di inclusione nel campione
GODABIT	titolo di godimento dell'abitazione di residenza 1 = di proprietà 2 = in affitto 3 = a riscatto 4 = in usufrutto o in uso gratuito
IREG	codice ISTAT regione
AREA5	ripartizione del territorio in 5 aree geografiche 1 = Nord-Ovest 2 = Nord-Est 3 = Centro 4 = Sud 5 = Isole

AREA3	ripartizione del territorio in 3 aree geografiche
	1 = Nord
	2 = Centro
	3 = Sud e Isole
ACOM4C	ripartizione dei comuni per popolazione residente (in 4 classi)
	1 = meno di 20.000 abitanti
	2 = tra 20.001 e 40.000 abitanti
	3 = tra 40.001 e 500.000 abitanti
	4 = più di 500.000 abitanti
NCOMP	numero dei componenti della famiglia
NPERC	numero dei percettori di reddito della famiglia
ANASC	anno di nascita del componente della famiglia
SESSO	sexso del componente della famiglia
	1 = maschile
	2 = femminile
STACIV	stato civile del componente della famiglia
	1 = coniugato
	2 = celibe/nubile
	3 = separato/divorziato
	4 = vedovo/a
STUDIO	titolo di studio del componente della famiglia
	1 = nessuno
	2 = licenza elementare
	3 = licenza media
	4 = diploma
	5 = laurea
	6 = specializzazione post-laurea

SETTP7            ripartizione per settore di attività economica (occupazione principale)

1 = agricoltura

2 = industria e costruzioni

3 = commercio, riparazioni, alberghi e ristoranti

4 = trasporti e comunicazioni

5 = intermediazione monetaria, finanziaria e assicurazioni

6 = P.A., attività immobiliari, servizi alle imprese e altri servizi

7 = condizione non professionale

QUALP10        ripartizione per condizione professionale (occupazione principale)

1 = operaio o posizione simile

2 = impiegato o insegnante

3 = impiegato direttivo / quadro

4 = dirigente

5 = libero professionista

6 = imprenditore individuale

7 = lavoratore autonomo

8 = titolare o coadiuvante di impresa familiare

9 = socio o gestore di società

10 = condizione non professionale

## **Appendice B**

### **Il Software DAD 4.3**

#### **B.1 Introduzione al Software**

Il DAD, che sta per “*Distributive analysis/Analyse Distributive*”, è stato progettato per facilitare l’analisi ed i confronti di benessere sociale, disuguaglianza, povertà ed uguaglianza utilizzando micro-dati (o dati disaggregati). E’ un software distribuito gratuitamente. Il programma include la stima di un largo numero di indici e curve che sono utili per confronti tra le distribuzioni. Prevede anche diversi strumenti statistici, tra cui statistiche descrittive di base, regressioni e stime non parametriche semplici di funzioni di densità.

La prima versione di DAD è stata lanciata nel Settembre 1998. Inizialmente essa era nata da una richiesta del Centro di Ricerca e Sviluppo Internazionale del Canada (IDRC) fatta all’Università di Laval (Quebec, Canada) per una ricerca sulle micro-conseguenze di politiche macroeconomiche e correttive (MIMAP – Micro Impacts of



Macroeconomic and Adjustment Policies) allora effettuata in Africa. La prima versione è stata poi rivista e corretta da successive versioni che hanno aumentato il numero delle operazioni, oltre a rendere il programma più stabile e flessibile. La versione utilizzata per questa ricerca è la 4.3.

Il software DAD è stato progettato da Jean-Yves Duclos e Araar Abdelkrim e programmato in linguaggio Java da Araar Abdelkrim e Carl Fortin.

Per poter disporre del programma è necessario accedere al sito internet ufficiale all'indirizzo <http://www.mimap.ecn.ulaval.ca> e registrarsi; successivamente si può scaricare il file ed eseguire l'installazione.

## **B.2 I motivi della creazione del pacchetto DAD**

Diversi fattori hanno motivato i progettisti alla creazione di DAD. In primo luogo il crescente bisogno degli analisti dei Paesi in via di sviluppo di mettere a fuoco dei “profili” di povertà e disuguaglianza. Molte politiche di sviluppo sono stabilite utilizzando criteri di povertà e sono effettuate, tra l'altro, attraverso l'elaborazione di stime di povertà, documenti strategici di riduzione della povertà (PRSP), analisi dell'impatto sociale delle azioni contro la povertà etc...

Molte di queste valutazioni erano fatte da consulenti stranieri e da personale tecnico di organizzazioni internazionali, mentre i ricercatori e gli analisti locali spesso non ne erano coinvolti. Innanzitutto per rompere questa netta divisione tra esperti stranieri ed analisti locali, ai ricercatori dell'Università di Laval è sembrato utile introdurre degli strumenti che gli studiosi dei Paesi in via di sviluppo potessero utilizzare agevolmente.

In secondo luogo, il problema di una certa difficoltà nell'accesso ai dati per gli analisti dei Paesi in via di sviluppo (in parte a causa dello scetticismo e del comportamento reticente delle agenzie statistiche e del personale delle organizzazioni

internazionali di ricerca) era diventato molto meno grave dalla fine degli anni '90 grazie ad una forte pressione esterna. Di conseguenza, dopo aver superato questo primo ostacolo, la nuova barriera da infrangere era diventata analizzare questi dati. In terzo luogo, la maggior parte delle analisi era (e in parte è ancora) maneggiata come se non fosse soggetta ad imprecisione statistica. Una quantità considerevole di energie e risorse viene sprecata in discussioni su “risultati” statisticamente non rilevanti. Variazioni nelle stime della povertà anche del 4-5% sono spesso statisticamente non significative alla luce degli usuali criteri di precisione statistici. Inutile dire come gli sforzi degli analisti e “policy makers” atti a spiegare variazioni dell'1 o 2% nei tassi di povertà (come spesso accade) siano una pura perdita di risorse. Questa situazione sfavorevole doveva essere risolta da un miglior uso di appropriate tecniche statistiche. Il DAD è stato progettato in gran parte proprio per aiutare a superare questo tipo di ostacoli. Realizzare ciò ha significato sgombrare il campo dai risultati statisticamente non significativi e lasciare più tempo e risorse all'interpretazione di quei risultati distributivi statisticamente significativi.

### **B.3 Il linguaggio di programmazione**

DAD è stato programmato in *Java*. Linguaggio orientato agli oggetti, il Java è indipendente: una volta scritte, le applicazioni in Java possono essere eseguite su qualunque sistema operativo così come su Internet. Per fare in modo che DAD fosse completamente gratuito, i progettisti hanno deciso di non legare il suo utilizzo ai programmi statistici commerciali come ad esempio, *Excel*, *Stata* o *SPSS*.

## **B.4 Presentazione di DAD**

Obiettivo primario è stato rendere l'ambiente di DAD intuitivo e facile da usare. Per fare questo usa i menu per selezionare le variabili e le opzioni che servono per tutte le applicazioni e vengono sfruttate finestre pop-up e fogli elettronici come principali strumenti di lavoro. Questo permette agli utenti di visualizzare molte delle informazioni con un'occhiata e controllarle facilmente. L'insieme dei dati è facilmente importabile da files esterni (da hard disk, floppy disk o cd-rom) che possono essere di tipo ASCII, ma anche salvati in formato testo o Excel. Piccoli insiemi di dati possono essere facilmente immessi manualmente. Le basi di dati possono essere modificate (cambiamenti ai valori delle singole celle possono essere eseguiti direttamente sul foglio di lavoro elettronico), nuove osservazioni possono essere aggiunte, e nuovi vettori di dati possono essere generati usando operatori aritmetici o logici. DAD ha inoltre la possibilità di caricare due basi di dati simultaneamente, visualizzandole alternativamente sul foglio elettronico. Questa struttura rende molto facile eseguire operazioni su ciascuna o due basi di dati, ed inoltre permette a DAD di valutare se esse sono indipendenti o meno quando si tratta di calcolare gli standard errors degli stimatori che utilizzano informazioni da entrambi i campioni.

## **B.5 Le basi di dati in DAD**

Le basi di dati in DAD sono visualizzate su fogli elettronici simili a quelli di *SPSS*, *STATA* o *Microsoft Excel*.

Una base di dati usata in DAD è quindi una tabella composta da colonne (massimo 20) che rappresentano le variabili di interesse e righe (massimo 150.000) che rappresentano le singole unità statistiche del campione.

I valori delle variabili possono essere interi (per esempio la categoria di reddito a cui appartiene una famiglia) o decimali (p.e. il reddito della famiglia stessa).

## **B.6 Le applicazioni in DAD**

Per ogni applicazione disponibile in DAD c'è una specifica finestra che facilita la specificazione delle variabili, dei parametri e delle opzioni che servono a generare le statistiche distributive volute, come ad esempio la finestra specifica dell'applicazione per il calcolo degli indici di povertà proposti da *Foster, Greer e Thorbecke (FGT)* con una distribuzione (c'è una finestra specifica per il caso di due distribuzioni).

Molte finestre delle applicazioni sono divise in tre riquadri. Il primo riquadro serve a specificare le variabili rilevanti per la stima da calcolare (nel nostro caso YFAM\_EQUI o CF\_STUD).

Il secondo riquadro, generalmente nella parte bassa della finestra dell'applicazione, specifica i valori dei parametri e le opzioni dello stimatore.

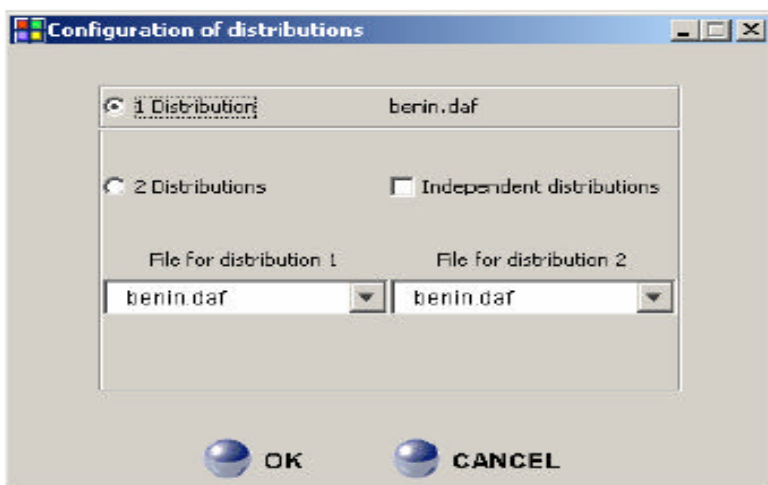
Il terzo riquadro attiva i pulsanti riguardanti i vari tipi di risultati che possono essere generati.

Alcune finestre delle applicazioni possono generare a loro volta eventuali finestre di dialogo pop-up. Ad esempio, cliccando sul pulsante "Compute line" nella finestra dell'applicazione *POVERTY|FGT* si specifica il range all'interno del quale la linea di povertà dovrebbe essere stimata.

Le seguenti variabili sono quelle tipicamente richieste per il calcolo delle statistiche in DAD:

- *Variabile di interesse*, “Variable of interest”
- *Variabile del gruppo*, “Group variable”
- *Numero del gruppo*, “Group number”
- *Peso del Campione*, “Sampling Weight”

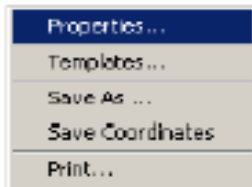
Le applicazioni di DAD con due distribuzioni possono essere lanciate solo dopo aver caricato due basi di dati. Ogni volta che si lancia un’applicazione che può supportare due distribuzioni, la finestra di dialogo (presentata di seguito), si apre per permettere all’utente di specificare il numero voluto di distribuzioni da usare e, soprattutto, quali. La finestra di applicazione per due distribuzioni è molto simile a quella per una. La differenza sostanziale è l’aggiunta di un secondo riquadro per la specificazione delle variabili da utilizzare per la seconda distribuzione.



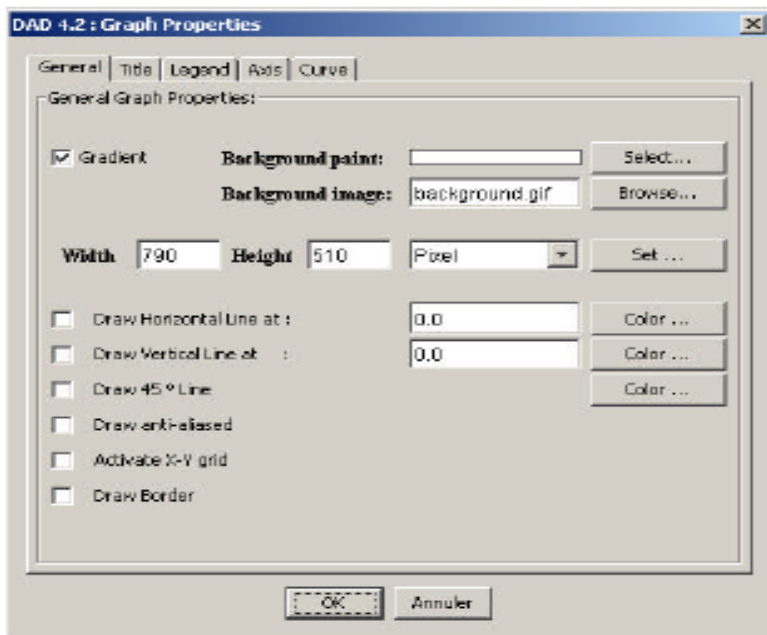
## B.7 Grafici

Le recenti versioni di DAD sono piuttosto flessibili in termini di modifica, salvataggio e stampa di grafici. Sulla maggior parte delle finestre di applicazione è presente un pulsante “Graph” per poter disegnare i grafici istantaneamente. Molte

delle opzioni per la costruzione dei grafici sono raggiungibili dalla finestra di dialogo “Graph Properties”.



The Popup Menu



I grafici tracciati con DAD possono anche essere salvati in diversi formati riconoscibili dalle loro estensioni: *png* (Portable Network Graphic), *pmb* (Bitmap Image File), *tif* (Tag Image File Format), *jpg* (JPEG File Interchange Format), *pdf* (Portable Document Format) e *ps* (Postscript).

## B.8 Gli strumenti del pacchetto DAD

Gli strumenti principali del DAD sono:

1. Stima degli indici di:

- povertà (*Watts, CHU, FGT, S-Gini, Sen*): normalizzati e non normalizzati (o indici di povertà assoluta e relativa), con le linee di povertà assoluta e relativa;
- benessere sociale (*Atkinson, S-Gini, Atkinson-Gini*);
- disuguaglianza (*S-Gini, Atkinson, Entropia, Atkinson-Gini ed altri*);
- redistribuzione, uguaglianza verticale, riordinamento e disuguaglianza orizzontale.

2. Scomposizione della povertà in sottogruppi della popolazione.

3. Calcolo di indici di disuguaglianza in sottogruppi della popolazione.

4. Evoluzione della disuguaglianza attraverso imposte e/o trasferimenti e sussidi (sovvenzioni).

5. Studio della povertà attraverso gli effetti di redistribuzione del reddito e di sviluppo (crescita, espansione).

6. Test per la verifica della robustezza dei confronti distributivi.

7. Stima delle curve di dominanza stocastica di primo e di secondo ordine, per povertà, benessere sociale e disuguaglianza.

8. Robustezza delle scomposizioni nei sottogruppi della popolazione.

9. Stima delle principali curve di secondo ordine: curve di Lorenz ordinarie e generalizzate, curve cumulative del poverty gap, curve dei quantili normalizzate e non normalizzate, curve di concentrazione ordinarie e generalizzate, etc...

10. Stima delle principali curve di primo ordine: funzioni di distribuzione cumulative, curve del divario di povertà (poverty deficit), curve di intensità della povertà, etc...

11. Stima delle differenze nelle curve e negli indici.
12. Stima delle linee “critiche” di povertà per i confronti di povertà assoluta e relativa.
13. Stima degli scarti quadratici medi asintotici di tutte le stime degli indici, punti sulle curve, linee critiche di povertà, punti di incrocio, etc..., tenendo conto della dipendenza o indipendenza dei campioni che sono stati confrontati.
14. DAD tiene conto degli errori di campionamento nelle linee di povertà specificate per calcolare gli indici di povertà assoluta e relativa.



## Bibliografia

Agostinelli C. e Sartorelli S., *Introduzione al linguaggio di SAS System*, Facoltà di Scienze Statistiche, Padova, 2002.

Bertola G. e Checchi D., *Sorting and Private Education in Italy*, Lavoro e Relazioni Industriali, 2001, 2, pp. 87-124.

Biancotti C., D'Alessio G., Faiella I. e Neri A., *I bilanci delle famiglie italiane nell'anno 2002*, Supplementi al Bollettino Statistico Anno XIV n°12, Banca d'Italia, Roma, 2004.

Brunello G., Comi S. e Lucifora C., Italia, 2001. in Harmon C., Walker I. e Wetergaard Nielsen N., *Education and Earnings in Europe*. Edward Edgar, Cheltenham, Gran Bretagna, 2001.

Brunello G., Lucifora C. e Winter-Ebner R., *Wage Expectations of European college students*. Cesifo working paper, 2002.

Cappellari L., *The effects of high school choices on academic performance and early labour market outcomes*, Quaderni dell'Istituto di economia dell'impresa e del lavoro, Università Cattolica, Milano, 2003.

Cecchi D., *The Italian educational system: family background and social stratification*, University of Milan, 2003.

Cecchi D. e Dardanoni V., *Mobility Comparisons: Does using different measures matter?*, 2002, forthcoming in *Research on Inequality 2002*.

Cingano F. e Cipollone P., *Determinants of University drop-out probability in Italy*, Bank of Italy, Research Department, 2003.

D'Alessio G. e Faiella I., *I bilanci delle famiglie italiane nell'anno 2000*, Supplementi al Bollettino Statistico Anno XII n°6, Banca d'Italia, Roma, 2002.

Duclos J.Y., Abdelkrim A. e Fortin C., *DAD: Distributive Analysis/Analyse Distributive User's Manual*, Université Laval, Montreal, 2004.

Dustmann C., *Parental Background, Primary to Secondary School Transition and Wages*, (Discussion Paper No.367).IZA, 2001.

Evans W.N. e Schwab R.M., *Finishing High School and Starting College: Do Catholic Schools Make a Difference?* Quarterly Journal of Economics, 1995, 110(4), pp. 941-974.

Guiso L., Jappelli T. e Terlizzese D., *Earnings uncertainty and precautionary saving*, Journal of Monetary Economics, 1992, vol. 30, pp. 307-337.

Neil D., *The effects of Catholic Secondary Schooling on Educational Achievement*, Journal of Labor Economics, 1997, 15(1), pp. 98-123.

OECD, *Education at Glance OECD Indicators 2002*, Paris, 2003.

Rovati G., *Rapporto sulle politiche contro la povertà e l'esclusione sociale*, Ministero del Lavoro, Roma, 2005.

Trivellato U., *Appunti di Statistica Economica*, Facoltà di Scienze Statistiche, Padova, 2002.

Trivellato U. e Bernardi L., *Scolarità e formazione professionale nel Mezzogiorno: nuove evidenze da un'analisi dei flussi*, Economia e Lavoro, 1995, 28: 3-34.

Siti Internet:

[www.bancaditalia.it](http://www.bancaditalia.it)

[www.istat.it](http://www.istat.it)

[www.mimap.ecn.ulaval.ca](http://www.mimap.ecn.ulaval.ca)