

**UNIVERSITÀ DEGLI STUDI DI PADOVA**  
**Facoltà di Scienze Statistiche**



**CORSO DI LAUREA SPECIALISTICA IN SCIENZE STATISTICHE, ECONOMICHE,  
FINANZIARIE ED AZIENDALI**

**TESI DI LAUREA**

**LA VALUTAZIONE DEGLI EFFETTI  
DELL'ESTENSIONE DEL PERIODO DI AMMISSIBILITÀ  
ALLE LISTE DI MOBILITÀ NEL VENETO.  
UN'ANALISI SUI SALARI MEDIANTE  
“PRINCIPAL STRATIFICATION”**

**Relatore: Prof. Enrico Rettore**

**Correlatore: Prof. Erich Battistin**

**Laureanda: Roberta De Stefani**

**Anno accademico 2008-2009**



# Indice

|                                                                                        |           |
|----------------------------------------------------------------------------------------|-----------|
| <b>Indice .....</b>                                                                    | <b>3</b>  |
| <b>Introduzione .....</b>                                                              | <b>5</b>  |
| <br><b>Capitolo 1</b>                                                                  |           |
| <b>Le politiche del lavoro.....</b>                                                    | <b>9</b>  |
| 1.1 Le politiche del lavoro attive in Italia .....                                     | 11        |
| 1. 2 Un confronto con l'Europa .....                                                   | 13        |
| <br><b>Capitolo 2</b>                                                                  |           |
| <b>Le liste di Mobilità .....</b>                                                      | <b>15</b> |
| 2.1 Soggetti interessati e criteri di ammissibilità .....                              | 16        |
| 2.1.1 Le imprese .....                                                                 | 16        |
| 2.1.2 I lavoratori .....                                                               | 17        |
| 2.2 Le diverse componenti della politica .....                                         | 18        |
| 2.2.1 La componente passiva (indennità di mobilità) .....                              | 18        |
| 2.2.2 La componente attiva .....                                                       | 19        |
| 2.3 Modalità di distribuzione dei benefici .....                                       | 20        |
| 2.3.1 Periodo di permanenza nelle liste .....                                          | 20        |
| 2.3.2 La dimensione dell'azienda: i due sotto-programmi.....                           | 22        |
| 2.4 La valutazione del programma: obiettivi e problemi di identificazione .....        | 23        |
| 2.5 Ricerche precedenti.....                                                           | 25        |
| <br><b>Capitolo 3</b>                                                                  |           |
| <b>Breve panoramica sui metodi per la valutazione degli effetti di politiche .....</b> | <b>29</b> |
| 3.1 Il problema fondamentale dell'inferenza causale .....                              | 30        |
| 3.2 Notazione .....                                                                    | 31        |
| 3.3 Regression Discontinuity Design.....                                               | 34        |

|                                                                                           |            |
|-------------------------------------------------------------------------------------------|------------|
| <b>Capitolo 4</b>                                                                         |            |
| <b>Principal Stratification.....</b>                                                      | <b>37</b>  |
| 4.1 Effetti Causali e Variabili Intermedie.....                                           | 40         |
| 4.2 Gli effetti principali .....                                                          | 43         |
| 4.3 Il problema del ‘Censoring by death’ .....                                            | 45         |
| 4.4 L’approccio della Principal Stratification nei problemi di ‘Censoring by Death’ ..... | 46         |
| 4.4.1 Metodi non parametrici: intervalli di confidenza .....                              | 50         |
| 4.4.2 Metodi parametrici: il modello di mistura.....                                      | 50         |
| <br><b>Capitolo 5</b>                                                                     |            |
| <b>I dati.....</b>                                                                        | <b>53</b>  |
| 5.1 Le Fonti Amministrative .....                                                         | 54         |
| 5.2 Il campione di riferimento .....                                                      | 56         |
| <br><b>Capitolo 6</b>                                                                     |            |
| <b>La stima dell’impatto dell’anno aggiuntivo .....</b>                                   | <b>61</b>  |
| 6.1 Il disegno di valutazione.....                                                        | 63         |
| 6.1.1 Le fonti di distorsione.....                                                        | 64         |
| 6.1.2 Il controllo dell’effetto età tramite <i>regression discontinuity design</i> .....  | 65         |
| 6.1.3 Identificazione degli strati principali e degli effetti casuali .....               | 68         |
| 6.2 Analisi e commento dei risultati.....                                                 | 76         |
| 6.2.1 Uomini che percepiscono l’indennità di mobilità .....                               | 77         |
| 6.2.2 Uomini senza indennità di mobilità.....                                             | 82         |
| 6.2.3 Donne con indennità di mobilità .....                                               | 85         |
| 6.2.4 Donne senza indennità di mobilità .....                                             | 89         |
| <br><b>Conclusioni e possibili sviluppi.....</b>                                          | <b>93</b>  |
| <br><b>APPENDICE .....</b>                                                                | <b>97</b>  |
| <br><b>Bibliografia.....</b>                                                              | <b>103</b> |

# Introduzione

Nel 1991 il Parlamento italiano emana la legge 223 che riorganizza la legislazione vigente in materia di mercato del lavoro e sostegno all'occupazione. Tra le innovazioni introdotte, il decreto legge dispone la reintroduzione dei licenziamenti collettivi, definiti tali laddove siano attuati da parte di imprese che per riduzione o trasformazione di attività si trovino nella necessità di licenziare almeno cinque lavoratori nell'arco di centoventi giorni. Come previsto anche dalla legislazione precedente, le imprese che attuano una ristrutturazione, riorganizzazione o conversione aziendale o in caso di fallimento, possono operare una riduzione del personale ed usufruire della Cassa Integrazione Guadagni Straordinaria<sup>1</sup> per un periodo di tempo limitato<sup>2</sup>. Al fine di provvedere ad un sostegno dell'occupazione anche in un momento successivo alla CIGS, la nuova legislazione sancisce che *“l'impresa che sia stata ammessa al trattamento straordinario di integrazione salariale, qualora nel corso di attuazione del programma ... ritenga di non essere in grado di garantire il reimpiego a tutti i lavoratori sospesi e di non poter ricorrere a misure alternative, ha facoltà di avviare le procedure di mobilità”*<sup>3</sup>

Il nuovo ammortizzatore sociale introdotto dal decreto legge, denominato “Liste di Mobilità”, è destinato all'insieme dei lavoratori coinvolti da licenziamenti collettivi o licenziati dopo un periodo di CIGS da aziende medie o

---

<sup>1</sup> Nel seguito CIGS

<sup>2</sup> Fino ad un anno se la Cigs è concessa a causa di crisi aziendale, fino a due anni negli altri casi

<sup>3</sup> Cfr. art.1, co.1, legge 223/1991

grandi<sup>4</sup>. Il programma agisce su due diversi aspetti, da un lato la componente attiva offre la possibilità ai lavoratori di iscriversi a particolari liste<sup>5</sup>, gestite dalle agenzie regionali per l'impiego (competenza poi trasferita ai Centri per l'Impiego dal 1997) che si occupano di agevolare il reinserimento nel mercato del lavoro attraverso notevoli sgravi contributivi offerti alle aziende che assumono da tali liste; dall'altro la componente passiva assegna ai lavoratori un'indennità, chiamata appunto di mobilità. Nel 1993 la legge 236 allarga il diritto di partecipare alle LM (su base volontaria) anche a tutti i lavoratori soggetti a licenziamento individuale o provenienti da imprese con meno di 15 dipendenti, per i quali tuttavia non è prevista nessuna indennità.

In base alla dimensione dell'impresa di provenienza, la politica si articola dunque in due sottoprogrammi: il primo rivolto ai lavoratori provenienti da medie e grandi imprese che beneficino del sussidio monetario; all'impresa che assume tali lavoratori viene offerta una riduzione del carico fiscale sui contributi e il trasferimento dei benefici inizialmente destinati al lavoratore. Il secondo sottoprogramma prevede solo sgravi contributivi per i datori di lavoro che assumono ed è rivolto ai lavoratori licenziati dalle piccole imprese. Entrambi i sottoprogrammi prevedono l'erogazione dei diversi trattamenti per una durata massima che varia in funzione dell'età dei lavoratori al momento dell'iscrizione in lista, pari ad un anno per coloro che hanno un'età inferiore a 39 anni, pari invece a due anni per coloro che hanno età compresa tra 40 e 49 anni. Non verranno considerati nell'analisi i lavoratori di età superiore per i quali sono previsti tre anni di permanenza in lista ed un programma aggiuntivo, detto "mobilità lunga", pensato come ponte verso la pensione.

Obiettivo di questo studio è valutare l'effetto della politica sulle retribuzioni successive all'uscita dalle Liste. Poiché le differenze nella durata di permanenza nelle LM sono una funzione dell'età, possiamo concentrare la valutazione sull'impatto differenziale di un anno aggiuntivo di ammissibilità alle liste sul livello medio dei salari a tre anni dall'iscrizione.

---

<sup>4</sup> Aziende con più di 15 addetti

<sup>5</sup> Nel seguito anche LM

La letteratura relativa alla valutazione d'impatto della mobilità è molto ampia. L'interesse degli analisti per questa politica è stato fin da subito elevato in quanto la combinazione delle due componenti che operano in direzioni opposte non permette di stabilire a priori il segno dell'impatto. I dati di cui disponiamo ci permettono di proseguire e sviluppare le analisi precedenti; l'abbinamento tra le fonti amministrative e gli archivi Inps ha infatti reso possibile l'allargamento dell'insieme informativo anche ai dati sui salari e nel contempo di ampliare la finestra di osservazione negli anni sia precedenti sia successivi al periodo di permanenza in lista. I dati utilizzati nello studio sono relativi ai lavoratori entrati in mobilità dal 1995 al 1998 nel Veneto.

L'esercizio di valutazione, per la complessità del disegno della politica, presenta delle forti limitazioni alla costruzione di un adeguato gruppo di confronto. Tutti i metodi utilizzati sono finalizzati ad annullare la distorsione causata dalla selezione non casuale sia per dimensione aziendale ed età che definiscono i pacchetti di benefici sia per stato occupazionale<sup>6</sup> del lavoratore che determina l'osservabilità del salario. Per eliminare qualsiasi fonte di eterogeneità tra i soggetti considerati, le analisi verranno svolte distintamente per i quattro gruppi definiti da genere e da trattamenti ricevuti (ammissibilità o meno a ricevere il sussidio). Al fine di rendere nullo l'effetto età, dato che risulta l'unica variabile che determina la durata della permanenza nelle liste, il metodo più adatto è il *regression discontinuity design*, che permette di effettuare una valutazione d'impatto per i soggetti al margine tra periodi di ammissibilità di durata diversa. Un'ultima fonte di distorsione è causata dalla selezione non casuale del gruppo di soggetti occupati, il solo gruppo sul quale è possibile effettuare la valutazione d'impatto sui salari. Lo studio si concentra sull'applicazione di un metodo sviluppato nell'ambito della 'principal stratification', elaborato da Rubin e Frangakis (2002). Tale metodo propone di utilizzare lo stato occupazionale per stratificare i gruppi di controllo e di trattamento in sottogruppi, chiamati strati principali, che sintetizzano importanti caratteristiche non osservabili dei

---

<sup>6</sup> Per stato occupazionale intendiamo la condizione di disoccupazione o occupazione. In quest'ultimo caso non è preso in considerazione l'eventuale tipo di contratto.

lavoratori. La validità del metodo si basa sull'assunzione di indipendenza tra gli strati e l'assegnazione al trattamento, ipotesi che assicura di poter effettuare inferenza valida, condizionatamente agli strati stessi. Nel nostro studio l'approccio della "principal stratification" ci permette di stimare l'impatto solo per un particolare strato, costituito dai soggetti che risulterebbero comunque occupati a prescindere dal trattamento al quale sono assegnati. Non è possibile tuttavia estendere i risultati all'intera popolazione.

L'evidenza empirica mostra come l'estensione del periodo di ammissibilità alle liste da uno a due anni non ha alcun effetto significativo sulle retribuzioni a medio termine, confermando i risultati ottenuti in Paggiaro et al. (2005).

Il seguito della tesi è così organizzato. Nel capitolo 1 si dà una definizione generale delle politiche del lavoro distinguendo tra i diversi obiettivi che le caratterizzano. Si passerà successivamente a descrivere le principali politiche attive in Italia e ad un confronto tra i sistemi di welfare italiano e degli altri paesi europei sulla base del livello di protezione dei lavoratori e del livello di spesa. Nel capitolo 2 verranno descritte le linee guida del programma Liste di Mobilità, con particolare attenzione agli aspetti problematici per la sua valutazione. Inoltre si fornirà una breve panoramica sugli studi effettuati in passato. Il capitolo 3 presenta i principali metodi sviluppati per la valutazione d'impatto e descrive il metodo del *regression discontinuity design* utilizzato nella prima parte del nostro studio. Nel capitolo successivo viene introdotto il concetto di 'principal stratification' prima in ambito generale, poi in un'applicazione specifica nel problema di selezione non casuale definito come 'censoring by death'. Dopo una breve presentazione delle fonti utilizzate e del campione sul quale viene svolta l'analisi, il capitolo 6 presenta le analisi effettuate. Infine nelle conclusioni viene fornito un riepilogo dei principali risultati ottenuti.

# Capitolo 1

## Le politiche del lavoro

Le politiche del lavoro sono un insieme di iniziative volte ad affrontare cause e conseguenze della disoccupazione. Una loro classificazione sistematica è presente nel *Labour Market Policies*, un modulo definito da Eurostat, che suddivide gli interventi in nove differenti categorie:

1. assistenza alla ricerca di lavoro
2. formazione professionale
3. schemi di suddivisione del lavoro
4. incentivi all'occupazione
5. politiche d'inserimento al lavoro dei disabili
6. creazione di lavoro nel settore pubblico
7. incentivi alle nuove attività d'impresa
8. sussidi di disoccupazione e sostegno del reddito ai disoccupati
9. schemi di pensionamento anticipato

Ad eccezione della prima tipologia che rappresenta una categoria a parte, i diversi interventi possono essere classificati in politiche attive (2-7) e politiche passive (8-9). La distinzione avviene sulla base della direzione verso la quale operano tali interventi: mentre i primi mobilitano l'offerta di lavoro per ridurre la disoccupazione, gli ultimi riducono gli incentivi ad una ricerca attiva del lavoro.

Non vi è sempre una separazione netta tra le due tipologie, alcune delle politiche esistenti sono piuttosto una combinazione di entrambi questi aspetti; ne è un esempio il modello del “benefit transfer” proposto da Snower (1994) che prevede che i benefici diretti ai lavoratori siano trasferiti alle imprese che li assumono.

Le politiche **passive** hanno come obiettivo la riduzione del disagio economico e sociale che deriva dalla disoccupazione; gli interventi con i quali si realizzano hanno una duplice natura, assicurativa perché assicurano un sostegno in termini di reddito e una natura distributiva nel senso che impediscono che il livello del reddito di alcuni individui scenda sotto un livello minimo essenziale nel loro intero ciclo di vita. I sussidi, garantendo un reddito ai lavoratori, permettono loro di effettuare una ricerca più accurata di un’occupazione, migliorando di conseguenza il *matching* tra domanda e offerta; d’altra parte hanno un impatto negativo sull’offerta di lavoro in quanto riducono gli incentivi ad una ricerca attiva dello stesso.

Le politiche **attive** sono interventi finalizzati al rientro nel mercato del lavoro dei soggetti che ne sono al margine fornendo loro maggiori possibilità occupazionali. Questi interventi non mirano esclusivamente ad aumentare la forza lavoro ma si preoccupano anche di fornire una formazione ai lavoratori disoccupati per migliorare la qualifica personale ed permettere loro di accedere a lavori più qualificati e meglio retribuiti. In una visione più generale, le politiche attive possono essere utilizzate come uno strumento di riduzione dei disincentivi conseguenti alle politiche passive. In quest’ottica si definiscono due principali strategie di protezione dei lavoratori.

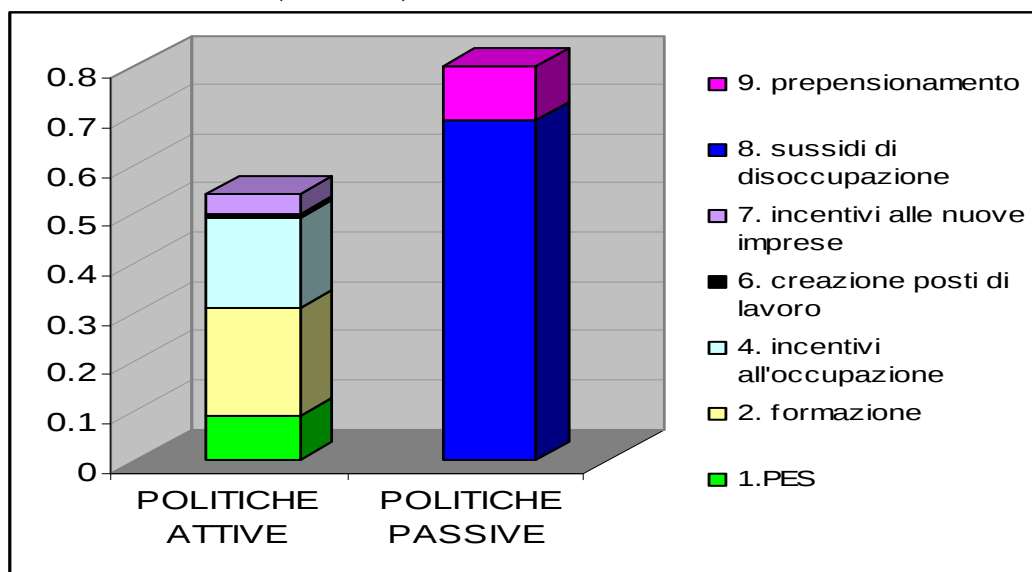
- Sistema di tutela ‘*on the job*’ : il lavoratore è protetto sul posto di lavoro attraverso forti vincoli al licenziamento (per il quale l’impresa deve pagare costi molto alti), mentre i sussidi per i lavoratori disoccupati sono modesti. Si tratta di un sistema tipico di un mercato molto rigido che ritroviamo in molti Paesi dell’Europa Occidentale.

- Sistema di tutela *'in the market'*: la legislazione per le assunzioni e i licenziamenti è molto liberale, nel periodo di disoccupazione d'altra parte sono previsti elevati sussidi ai lavoratori. E' un sistema che ritroviamo in molti paesi scandinavi caratterizzati da un mercato più flessibile dove la spesa pubblica destinata alle politiche sociali è elevata.

### 1.1 Le politiche del lavoro attive in Italia

Le politiche del lavoro in Italia sono maggiormente orientate alla protezione dei lavoratori quando questi risultano occupati (sistema di tutela *on the job*), mentre è scarsamente sviluppato il sistema di ammortizzatori sociali. Nell'ultimo decennio il governo ha cominciato ad aumentare la spesa destinata ad interventi di politica passiva che nel 2006 hanno raggiunto il 60% della spesa totale per le politiche del lavoro, rimanendo tuttavia ben al di sotto della spesa media europea.

Fig 1.1 Spesa per le politiche del lavoro attive e passive in Italia nel 2006 per tipologia di interventi.(% del PIL)



Fonte OECD

Lo strumento principale delle politiche passive è costituito dalla Cassa Integrazione Guadagni Ordinaria e Straordinaria rivolto a lavoratori che vengono sospesi temporaneamente dal lavoro per difficoltà produttive dell'impresa al fine di evitarne il licenziamento; la politica si realizza attraverso sussidi molto generosi con un importo pari all'80% dell'ultima retribuzione. Un sussidio di pari entità è previsto anche per i lavoratori inseriti nel programma Liste di Mobilità (con un tasso di rimpiazzo che scende al 64% negli anni successivi al primo); tale ammortizzatore sociale, che sarà oggetto di questo studio, prevede interventi volti sia ad agevolare il rientro del lavoratore nel mercato del lavoro, sia a sostenere il suo reddito tramite indennità di mobilità di cui possono usufruire solo i lavoratori licenziati da imprese medio-grandi<sup>7</sup>. Tutte queste misure si attuano solo in situazioni straordinarie; per tutti i lavoratori che non possiedono i requisiti necessari per usufruirne, il trattamento di disoccupazione ordinario prevede l'erogazione di un sussidio pari al 50% dell'ultima retribuzione per i primi 6 mesi (per poi diminuire progressivamente nei mesi successivi)<sup>8</sup>. Infine un ultimo strumento previsto dalla legislazione italiana è il pensionamento anticipato che, a causa degli alti costi di attuazione, è scarsamente utilizzato. Il sistema di assicurazione del reddito in Italia presenta numerose carenze. In primo luogo, dal punto di vista quantitativo, le entità dei sussidi ordinari sono molto modeste; inoltre il livello di copertura è limitato: tra la popolazione di disoccupati sono esclusi da ogni forma di assicurazione i giovani che devono ancora entrare nel mercato del lavoro o che non hanno maturato i requisiti necessari, per i quali non è previsto nessun tipo di sussidio.

Per quanto riguarda le politiche attive, in Italia si realizzano attraverso incentivi all'assunzione e alla stabilizzazione dell'occupazione, corsi di formazione personale, gestiti principalmente dalle Regioni, diversi programmi che prevedono l'inserimento di particolari categorie di lavoratori (donne, giovani, immigrati) e infine i contratti causa mista che affiancano un impegno di formazione alla prestazione lavorativa (apprendistato, contratto di formazione e lavoro). Tra

---

<sup>7</sup> Per una descrizione dettagliata delle Liste di Mobilità si rimanda al capitolo successivo

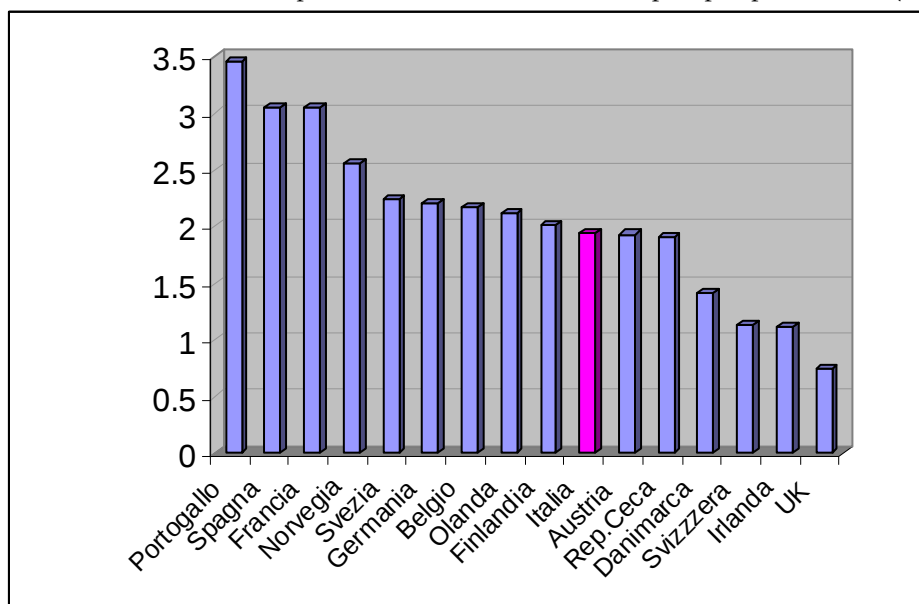
<sup>8</sup> Valori in uso dal 2005. Nel periodo considerato nello studio il tasso di rimpiazzo era solo del 30% per una durata massima di 6 mesi.

questi, gli interventi che più gravano sul bilancio dello Stato sono gli incentivi all'occupazione, che si realizzano in misura prevalente tramite sgravi contributivi, e la formazione professionale che nel 2006 hanno rappresentato rispettivamente il 33% e il 40% del totale delle spese per le politiche attive in Italia.

## 1.2 Un confronto con l'Europa

La legislazione italiana prevede un alto livello di protezione dei lavoratori quando questi risultano occupati. Un indicatore utile per un confronto a livello europeo è l'”Employment Protection Legislation” utilizzato dall'OECD. Dall'inizio degli anni '90 tuttavia, il livello dell'EPL per l'Italia è caratterizzato da un trend negativo, in calo dal 3.1 del 1991 al 2.7 del 1998 fino al 1.9 del 2003. Mentre nello scorso decennio quindi l'Italia si collocava al di sopra della maggior parte degli stati europei, attualmente si colloca attorno alla media dei paesi OECD, molto al di sotto di Francia, Spagna e Portogallo.

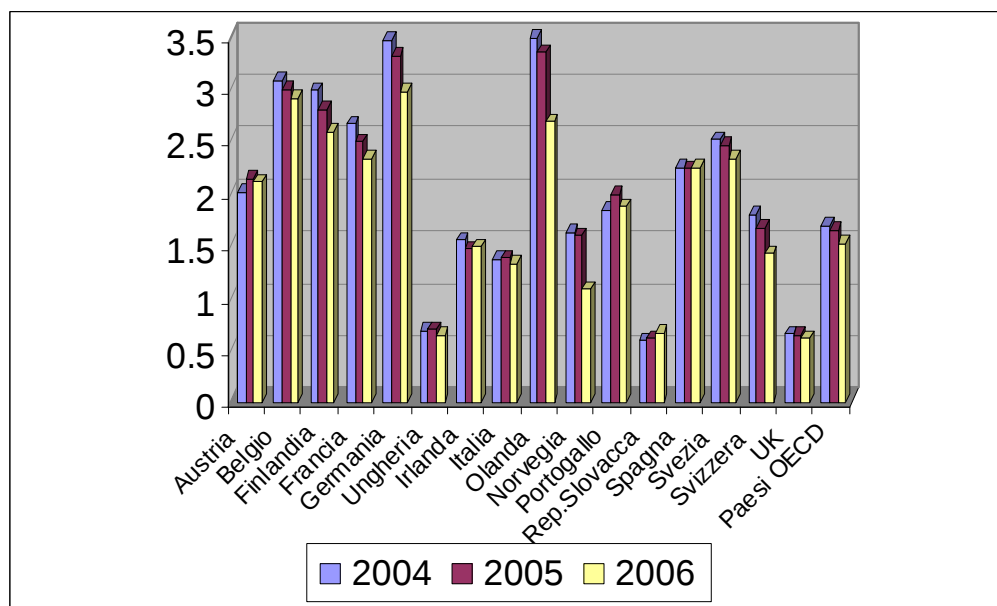
Fig 1.2 Misura del livello di protezione dei lavoratori occupati per paese, EPL (2003)



Fonte OECD

La spesa per le politiche del lavoro previste dalla legislazione italiana è inferiore rispetto alla media dei paesi OECD. L'Italia nel 2004 ha speso solo l'1.36% del proprio Pil per le politiche del lavoro e la percentuale si è mantenuta pressoché costante nei due anni successivi. Tutti i paesi dell'Europa Occidentale destinano una percentuale molto superiore alle LMP<sup>9</sup>, prime fra tutti Olanda e Francia in cui la spesa raggiunge il 4% del Pil. Solo Inghilterra, Repubblica Ceca e Repubblica Slovacca, tra i paesi presi in considerazione, hanno un valore inferiore.

Fig 1.3 Spesa destinata alle politiche del lavoro nei Paesi europei dal 2004 al 2006 (% del PIL)



Fonte OECD

<sup>9</sup>Labour market policy

## Capitolo 2

### Le liste di Mobilità

All'inizio degli anni '90, con il D.Lgs 223/1991, è stata introdotta in Italia una nuova politica del lavoro a sostegno dei disoccupati, conosciuta come *Liste di Mobilità*. La legge prevede che le imprese con più di 15 dipendenti (che verranno nel seguito indicate come *grandi imprese*) possano attuare una riduzione del personale per trasformazione, riduzione o cessazione di attività, ed iscrivere automaticamente i lavoratori licenziati nelle liste di mobilità regionali, create al fine di sostenere economicamente i lavoratori in difficoltà nel periodo della disoccupazione e agevolare il loro reinserimento lavorativo. La possibilità di usufruire di questi benefici è stata poi estesa tramite il D. Lgs. 236/1993 ai lavoratori licenziati, non solo collettivamente ma anche singolarmente, da piccole imprese (definite come tutte le imprese con meno di 15 dipendenti) per giustificato motivo oggettivo. I soggetti che possiedono questi requisiti non accedono automaticamente al programma ma possono iscriversi alle liste su base volontaria.

La mobilità si inserisce nel complesso degli ammortizzatori sociali, a fianco dell'indennità di disoccupazione standard, della Cassa Integrazione Guadagni Ordinaria e Straordinaria, già presenti nel panorama delle politiche a sostegno del reddito, regolate dalla legislazione italiana. Gli interventi previsti, a differenza ad esempio dell'indennità di disoccupazione, non sono costituiti esclusivamente dal semplice sostegno economico per sopperire alla mancanza del reddito da lavoro (indennità di mobilità); lo strumento principale di questo programma è

rappresentato al contrario da un sistema di incentivi studiati per sostenere e favorire il reinserimento dei lavoratori disoccupati nel mercato del lavoro.

Il programma è di competenza dell'Istituto Nazionale per la Previdenza Sociale (INPS) che finanzia la mobilità con il sostegno delle imprese che sono tenute a versare una quota per contribuire all'indennità di mobilità per ogni lavoratore licenziato.

### **2.1 Soggetti interessati e criteri di ammissibilità**

La mobilità è nata inizialmente per far fronte alle esigenze e ai bisogni di due attori del mercato del lavoro: da una parte le imprese in difficoltà a cui viene consentito il licenziamento di tutti o di solo una parte del personale in particolari momenti di difficoltà; dall'altra parte gli stessi lavoratori licenziati che devono essere sostenuti nel periodo di disoccupazione. Entrambi i soggetti coinvolti devono possedere determinati requisiti per poter essere ammessi al programma, stabiliti con precisione dalla normativa italiana.

#### **2.1.1 Le imprese**

La regolamentazione prevede che possano avviare la procedura di mobilità le imprese con più di 15 dipendenti che si trovino nella condizione necessaria di effettuare un licenziamento collettivo<sup>10</sup> per riduzione, trasformazione o cessazione dell'attività. Inoltre tutte le aziende che possiedano i requisiti necessari per far parte della Cassa Integrazione Guadagni Straordinaria (CIGS) ma che dimostrino l'impossibilità di garantire la riassunzione dei lavoratori sospesi, possono usufruire della mobilità. Di conseguenza il passaggio nelle liste può avvenire sia

---

<sup>10</sup> Per licenziamento collettivo si intende il licenziamento di almeno 5 dipendenti in 120 giorni in una o più unità produttive

in maniera diretta sia successivamente ad un periodo di transizione nella Cassa Integrazione Straordinaria.

Prima di poter effettuare il licenziamento le imprese sono tenute a contattare le associazioni sindacali e la Direzione regionale del lavoro per l'approvazione e la selezione del personale in esubero; in mancanza di accordi sindacali le imprese operano una scelta in base a precisi criteri quali carichi familiari, anzianità ed esigenze di organizzazione e produzione. Agli enti per il lavoro regionali spetta infine il compito di elaborare e gestire le liste di mobilità<sup>11</sup>.

### **2.1.2 I lavoratori**

Hanno diritto all'indennità i lavoratori che sono stati assunti con contratto a tempo indeterminato, che lavorano all'interno dell'azienda da almeno un anno (compresi periodi di lavoro con altre forme contrattuali prima dell'assunzione a tempo indeterminato) e che hanno almeno sei mesi di lavoro effettivo nell'impresa, comprese le sospensioni per ferie, festività o infortuni.

Inoltre i lavoratori che soddisfano questi requisiti e che sono stati licenziati senza giusta causa oggettiva da un'impresa con meno di 15 dipendenti possono decidere volontariamente di entrare nelle liste ed usufruire dei benefici previa richiesta all'Inps entro 60 giorni dalla data del licenziamento. A questa tipologia di lavoratori è tuttavia preclusa la possibilità di accedere alla parte passiva della politica, non hanno infatti diritto a ricevere l'indennità di mobilità. Evidenze empiriche dai dati dell'agenzia regionale per il lavoro veneta indicano che fondamentalmente tutti i lavoratori idonei licenziati da piccole imprese si registrano alle liste (Paggiaro, Trivellato 2002).

---

<sup>11</sup> Le liste gestite dalla Direzione regionale del lavoro sono le uniche previste dalla legge. Non esiste un database a livello Nazionale.

## **2.2 Le diverse componenti della politica**

La politica è costituita principalmente da due componenti: la parte attiva prevede una sostanziale riduzione dei contributi sociali per l'impresa che assume a tempo indeterminato un lavoratore iscritto alle liste, la parte passiva consiste nell'erogazione di un contributo finanziario diretto al lavoratore per sostenerlo nel periodo di disoccupazione. Le modalità di applicazione delle due componenti del programma e il periodo di permanenza nelle liste variano in relazione all'età del lavoratore e alla dimensione dell'azienda dalla quale è stato licenziato. Questi aspetti rendono questa politica complessa da analizzare in quanto pacchetti di benefici diversi sono destinati a gruppi di soggetti altrettanto diversi tra loro. Sarà necessario di conseguenza distinguere tra i reali effetti della politica e le differenze tra i gruppi considerati<sup>12</sup>.

### **2.2.1 La componente passiva (indennità di mobilità)**

A tutti i lavoratori licenziati da imprese con almeno 15 dipendenti viene erogato un contributo monetario al fine di mantenere il loro livello di reddito precedente all'episodio di disoccupazione. L'indennità di mobilità prevede un alto tasso di sostituzione, nel primo anno di iscrizione alle liste infatti, il contributo monetario si attesta attorno all'80% dell'ultimo salario percepito (il 100% dell'indennità prevista dalla CIGS) e si abbassa al 64% nel secondo o terzo anno (l'80% dell'indennità prevista dalla CIGS) per i lavoratori che risultano ancora iscritti alle liste.

Per i lavoratori che superano i 50 anni è prevista un'ulteriore variante che prende il nome di "mobilità lunga"; coloro che hanno maturato almeno 28 anni di contributi sociali possono decidere di allungare il loro periodo di permanenza e

---

<sup>12</sup> Per una valutazione più estesa delle problematiche connesse all'analisi di questa politica si rimanda al paragrafo 2.4

ricevere il sussidio fino al raggiungimento dell'età pensionabile utilizzando quindi questo programma come ponte verso la pensione<sup>13</sup>.

### **2.2.2 La componente attiva**

Questa parte prevede delle agevolazioni alle aziende che assumono lavoratori in mobilità.

Le agevolazioni consistono in una riduzione delle aliquote contributive (il cui livello nel 1998 si attestava attorno al 35%). Questi incentivi si ottengono indipendentemente dall'età del lavoratore, dalla dimensione dell'azienda dalla quale è stato licenziato e dalla durata di permanenza nelle liste; tuttavia sono applicati in maniera diversa in base al tipo e alla durata del contratto con il quale il lavoratore viene assunto:

- con contratto a tempo indeterminato: i contributi dovuti sono quelli previsti per gli apprendisti (circa il 3% delle aliquote contributive standard), per un periodo di 18 mesi.
- con contratto a termine per non più di 12 mesi: i contributi previdenziali a carico del datore di lavoro sono quelli previsti per gli apprendisti;
- con contratto a termine, trasformato a tempo indeterminato prima della scadenza: i contributi a carico del datore di lavoro sono quelli dovuti per gli apprendisti per ulteriori 12 mesi;

Inoltre le imprese che assumono con contratto a tempo indeterminato lavoratori che percepiscono l'indennità di mobilità, hanno il diritto di ricevere il 50% dei

---

<sup>13</sup> Si tratta di una forma di prepensionamento. Per dettagli si rimanda a Anastasia et al. (2004) pp.55-56

contributi monetari che sarebbero stati corrisposti al soggetto se fosse rimasto nelle liste. La parte di indennità che può essere trasferita alle imprese che assumono può raggiungere un valore massimo pari al 50% della somma totale dell'indennità che è possibile ricevere in un anno per lavoratori di età inferiori a 49 anni, in due anni per lavoratori che superano i 50 anni.

### **2.3 Modalità di distribuzione dei benefici**

Le diverse parti di cui si compone la politica non sono applicate universalmente alla popolazione eleggibile ma variano in base all'età dei lavoratori, alla zona geografica e alla dimensione dell'azienda che ha originato il licenziamento.

#### **2.3.1 Periodo di permanenza nelle liste**

La durata massima del periodo nel quale i soggetti risultano iscritti alle liste e quindi hanno il diritto di usufruire del programma, si calcola in base all'età del lavoratore al momento del licenziamento e risulta di

- un anno per lavoratori di età inferiore ai 40 anni
- due anni per lavoratori di età compresa tra 40 e 49 anni
- tre anni per lavoratori che superano i 50 anni

Per i lavoratori nelle zone del Mezzogiorno, considerate economicamente svantaggiate<sup>14</sup>, il periodo di idoneità si estende di 12 mesi per tutte le fasce.

---

<sup>14</sup> Definite dal D.P.R: 6 Marzo 1978, n.218

Il periodo di permanenza nelle liste può essere esteso se il lavoratore è assunto a tempo determinato. Nel caso in cui accetti un contratto di questo tipo, di durata non superiore ad un anno, il lavoratore può mantenere l'iscrizione alla lista. Inoltre il periodo di lavoro non viene valutato come periodo trascorso nelle liste quindi, nel caso risultasse idoneo, non riceve il sussidio monetario per tutta la durata del contratto. Il conteggio riprende nel momento in cui il lavoratore torna ad essere disoccupato, la permanenza nel programma verrà quindi prorogata per il periodo equivalente alla durata del rapporto di lavoro.

La politica, al fine di spingere i soggetti interessati verso una ricerca attiva del lavoro, prevede un controllo sul loro comportamento e precise regole a questo riguardo. Infatti, il diritto di godere del programma cessa di valere se il soggetto si rifiuta di partecipare a corsi di formazione organizzati dalla Regione, rifiuta offerte di lavoro appropriate rispetto alla posizione ricoperta con un salario non inferiore al 10% del salario percepito nel lavoro precedente o un impiego di pubblica utilità. Altre motivazioni possono essere la mancata comunicazione all'INPS entro 5 giorni dalla sottoscrizione ad un eventuale contratto a tempo determinato o la mancata risposta alle convocazioni dei Centri per l'Impiego. Se si realizza uno di questi casi, è prevista come sanzione l'abbandono del programma. Poiché in realtà non esiste un vero controllo sui lavoratori, non si ha evidenza empirica dell'applicazione di queste disposizioni.

Il lavoratore esce dal programma prima della scadenza del termine di permanenza nella lista in caso di assunzione a tempo indeterminato, riscossione dell'indennità in un'unica soluzione o ottenimento del diritto alla pensione di anzianità, vecchiaia o invalidità.

### 2.3.2 La dimensione dell'azienda: i due sotto-programmi

Oltre ad essere definite dall'età del lavoratore, le modalità di attuazione del programma dipendono dalla dimensione dell'azienda che licenzia: le aziende vengono distinte in piccole aziende se hanno meno di 15 dipendenti o grandi aziende se hanno almeno 15 dipendenti. A partire da questa distinzione la politica può essere considerata come l'unione di due sotto programmi separati indirizzati a popolazioni distinte tra loro.

**Sotto programma 1:** politica rivolta a lavoratori che sono stati licenziati collettivamente da **grandi imprese**. I soggetti destinatari beneficiano di tre provvedimenti:

- a. Sono automaticamente iscritti alle liste di mobilità dopo il licenziamento e ricevono un'indennità monetaria
- b. Le aziende che assumono un lavoratore iscritto alle liste ricevono il 50% dei benefici monetari residui del lavoratore.
- c. Le stesse aziende hanno inoltre diritto ad un forte sgravio degli oneri sociali per due anni

**Sotto programma 2:** politica rivolta a lavoratori licenziati collettivamente o anche individualmente da **piccole imprese**, che

- a. si possono iscrivere al programma su base volontaria nel caso in cui abbiano lavorato almeno un anno con un contratto a tempo indeterminato.
- b. Portano in beneficio sgravi contributivi per una durata di due anni alle imprese che li assumono (parte attiva)
- c. non beneficiano della parte passiva. Se idonei, ricevono un'indennità di disoccupazione ordinaria<sup>15</sup>.

---

<sup>15</sup> dal 1995 al 1998 (periodo di interesse del nostro studio) pari al 30% della retribuzione media degli ultimi 3 mesi per un massimo di 6 mesi. L'indennità è aumentata fino ad arrivare dal 1°

## 2.4 La valutazione del programma: obiettivi e problemi di identificazione

La valutazione della politica ha l'obiettivo di stabilire in che misura i diversi pacchetti di benefici alternativi, proposti ai lavoratori in relazione alla loro età o all'impresa di provenienza, hanno prodotto effetti sia sulla probabilità di essere reinseriti nel mondo del lavoro sia, nel caso in cui ciò avvenga, nella qualità del lavoro svolto nel periodo successivo alla mobilità.

Le variabili di interesse possono essere molteplici: il tempo di permanenza nelle liste prima del passaggio ad un'occupazione a tempo indeterminato, la probabilità di occupazione, la qualità del lavoro. L'effetto finale della politica su queste quantità non è prevedibile in quanto è il risultato della somma di due effetti che influiscono in direzioni opposte sul fenomeno di interesse.

Consideriamo ad esempio l'impatto sul tasso di rioccupazione degli iscritti. Da una parte la componente passiva della politica ha un effetto negativo sulla probabilità di essere occupati: la percezione dell'indennità di mobilità provoca un aumento dei salari di riserva e permette al soggetto di essere più selettivo nei confronti dei lavori che gli vengono proposti. Al contrario la possibilità di poter usufruire degli sgravi fiscali e del trasferimento del beneficio (nei casi in cui sia previsto), aumentano l'interesse delle imprese nei confronti dei lavoratori iscritti alle liste con un conseguente effetto positivo sulla loro occupazione. La "direzione" dell'impatto finale è incerta poiché non è possibile sapere a priori quale tra i due effetti prevalga sull'altro.

Il modo con il quale la politica è stata disegnata rende la sua analisi molto complessa e di particolare interesse allo stesso momento.

Il primo limite è costituito dalla sua applicazione universale. Poiché tutti i lavoratori iscritti alla lista usufruiscono dei benefici (benché con modalità diverse) manca un gruppo di confronto adeguato, ossia un insieme di lavoratori non idonei alla politica ma con le stesse caratteristiche rispetto al gruppo degli esposti. Se una

---

gennaio 2008 ad un livello pari al 50% dell'ultima retribuzione per i primi sei mesi; 40% per i successivi tre; 30% per gli ulteriori mesi. (DI n. 35, 14-03-2005).

valutazione della politica nel suo insieme non è realizzabile, si può concentrare l'analisi sugli effetti differenziali, ossia sul confronto tra i pacchetti di benefici alternativi rispetto

- al tempo di permanenza nella lista. La quantità di interesse è l'impatto differenziale che ha l'eleggibilità per un anno, due o tre anni sulla variabile risultato
- alla dimensione dell'impresa. In questo caso si parla anche di *effetto indennità*<sup>16</sup>.

In entrambi i casi, confrontare pacchetti alternativi significa confrontare gruppi potenzialmente diversi tra loro. Le differenze osservabili nella variabile risultato possono essere causate da una parte dal trattamento, dall'altra dalle differenze tra gli stessi soggetti.

In questa tesi ci concentreremo sull'effetto differenziale rispetto al periodo di idoneità. Considerando che la durata dell'eleggibilità è determinata esclusivamente dall'età del lavoratore al momento di iscrizione alla lista, per separare l'effetto della diversa durata di permanenza nelle liste dall'effetto dell'età, utilizzeremo il metodo del *regression discontinuity design*<sup>17</sup>. Questo metodo si adatta bene a disegni nei quali il processo di assegnazione è una funzione deterministica di una variabile osservata. Per tale variabile viene stabilito un valore soglia (in questo caso i valori soglia sono i valori dell'età pari a 40 anni o 50 anni), tutti i soggetti che si posizionano al di sotto della soglia sono assegnati al gruppo di controllo, mentre tutti i soggetti per i quali si misura un valore superiore sono assegnati al gruppo di esposti (o viceversa).

L'analisi che sarà proposta in questo studio sarà soggetta ad un'altra fonte di distorsione che in econometria prende il nome di selezione endogena.

---

<sup>16</sup> Cfr Anastasia et al. (2004) p.64

<sup>17</sup> Per un approfondimento sul metodo del Regression Discontinuity Design si rimanda al capitolo 3

Considereremo come variabile di interesse la qualità del lavoro, misurata dal livello dei salari a tre anni dall'entrata in lista. I salari palesemente sono osservabili solo per i soggetti che lavorano mentre la stessa variabile è non osservabile per i soggetti che al momento della misurazione risultano disoccupati. Verosimilmente il gruppo dei lavoratori occupati presenta caratteristiche diverse (che possono essere ad esempio abilità, formazione) rispetto alla popolazione totale di soggetti per ogni singola classe di età; i sottogruppi per i quali è osservabile il valore della variabile obiettivo non sono perciò confrontabili. Per risolvere questo problema di selezione endogena utilizzeremo un approccio denominato Principale Stratificati (Rubin, Frangakis 1999, 2002) che ci permetterà di valutare l'impatto differenziale per l'unico gruppo di soggetti confrontabili ossia per coloro che risulterebbero occupati indipendentemente dal tempo massimo di permanenza nelle liste.

## **2.5 Ricerche precedenti**

I primi studi si sono basati esclusivamente sui dati raccolti dalle Agenzie Regionali del Lavoro, un insieme molto povero di informazioni e di scarsa qualità. Le variabili rilevate erano caratteristiche individuali dei soggetti, dimensione e localizzazione delle imprese che avevano effettuato il licenziamento e una tra le seguenti tre condizioni del lavoratore: uscito della lista per scadenza del periodo massimo stabilito, uscito della lista per assunzione a tempo indeterminato, attualmente iscritto alla lista (senza specificare eventuali periodi di occupazione temporanea). La variabile oggetto del primo ciclo di studi è stata il tempo di permanenza nelle liste, intercorso tra il momento dell'iscrizione e l'eventuale assunzione con contratto permanente.

Tra gli studi effettuati possiamo citare Brunello e Miniaci (1997) per la Lombardia, Caruso (2001) per l'Umbria e Paggiaro e Trivellato (2002) per il

Veneto, che utilizzano un approccio molto simile tra loro<sup>18</sup>. In particolare Paggiaro e Trivellato si sono serviti di una stima non parametrica della funzione di sopravvivenza e hanno stimato un modello a rischi proporzionali. Le analisi sono state condotte in quattro sottogruppi, condizionatamente al sesso dei lavoratori e al diritto a ricevere l'indennità (dimensione dell'azienda responsabile del licenziamento); mentre per distinguere tra l'effetto differenziale del trattamento e l'effetto dell'età hanno incluso tra le variabili esplicative un polinomio nell'età.

Le conclusioni a cui pervengono sono le medesime: un anno aggiuntivo di benefici monetari ha un impatto negativo sulla probabilità di trovare lavoro a tempo indeterminato; tutti gli studi sostengono quindi la prevalenza dell'effetto della parte passiva rispetto all'effetto positivo degli incentivi alle imprese<sup>19</sup>.

L'abbinamento tra i dati amministrativi dei centri regionali e le informazioni contenute in Netlabor (vedi Paggiaro e Trivellato, 2001) hanno permesso di arricchire le informazioni disponibili, aggiungendo al set di variabili già presenti informazioni relative sia alla storia lavorativa precedente e successiva all'entrata in lista sia ai singoli eventi lavorativi occorsi durante la permanenza nelle liste. La variabile di interesse nel successivo ciclo di studi diventa lo status occupazionale nei 36 mesi successivi all'iscrizione mentre la valutazione si concentra sulla stima dell'effetto differenziale dovuto all'anno aggiuntivo di permanenza nel programma LM<sup>20</sup>. In questo ciclo di studi si inseriscono Paggiaro, Rettore e Trivellato (2005), Martini, Rettore e Trivellato (2006), Martini e Mo Costabella (2007). La stima dell'impatto è calcolata con il metodo del *matching* tramite *propensi score* calcolato sulla base delle caratteristiche osservate. L'utilizzo di un più ampio insieme informativo però non è sufficiente ad eliminare completamente l'effetto età; in Paggiaro, Rettore e Trivellato (2005) un test sull'assunzione di

---

<sup>18</sup> Per una descrizione dettagliata degli studi che appartengono al primo ciclo si veda Anastasia et al. (2004)

<sup>19</sup> Si tratta del confronto tra soggetti con un'età inferiore a 40 anni e i soggetti con età compresa tra 40 e 49 anni. Per i soggetti con età superiore a 50 anni le analisi necessitano di considerazioni maggiori dovute alla presenza della mobilità lunga.

<sup>20</sup> La letteratura sulla stima dell'effetto indennità è ancora piuttosto scarsa

dimostra che le stime ottenute continuano ad essere affette da selection bias. Le analisi hanno comunque confermato i risultati ottenuti nel primo ciclo: per i lavoratori che godono anche dei benefici monetari, la politica ha un impatto negativo sia sul tempo di attesa prima di un lavoro permanente sia sulla probabilità di reimpiego.

Infine Paggiaro, Rettore e Trivellato (2008) fanno uso del Regression Discontinuity Design<sup>21</sup> applicato alla nuova base di dati costruita tramite l'abbinamento di diverse fonti amministrative: agenzie regionali, Netlabor e Inps. I risultati a cui pervengono sono in contrasto con i risultati ottenuti nelle ricerche precedenti; non si registrano infatti effetti significativi dell'estensione del periodo di eleggibilità di un anno sulla probabilità di occupazione a tre anni dall'entrata in lista. Gli autori ampliano la valutazione considerando anche l'effetto sul livello dei salari. In questo modo si aggiunge un ulteriore problema di selezione endogena del campione<sup>22</sup> in quanto i salari sono osservabili solo per coloro che risultano occupati al momento della rilevazione che possono distinguersi per caratteristiche personali dall'intero gruppo di lavoratori dal quale provengono che comprende sia i lavoratori occupati sia i lavoratori disoccupati. Per verificare che siano effettivamente confrontabili, gli autori hanno confrontato i due gruppi (lavoratori occupati al momento della rilevazione con età inferiore a 40 anni o superiore a 40 anni e inferiore a 50 anni) sulla base di caratteristiche misurate prima del trattamento e quindi per definizione non influenzate dal trattamento stesso. L'analisi realizzata conclude che i due gruppi di lavoratori occupati a tre anni dall'entrata in lista sono confrontabili e l'effetto dell'anno aggiuntivo dei benefici sui salari è nullo.

---

<sup>21</sup> Per la notazione precisa si rimanda al capitolo 3.2

<sup>22</sup> Di questo problema tratteremo nel dettaglio nei capitoli 3 e 4



## Capitolo 3

### Breve panoramica sui metodi per la valutazione degli effetti di politiche

Una politica è un intervento o un insieme di attività finalizzate a modificare la condizione o il comportamento dei soggetti destinatari.

La valutazione d'impatto ha l'obiettivo di confrontare la condizione dei soggetti esposti al trattamento (la *variabile risultato*<sup>23</sup>) osservata in un momento successivo alla realizzazione del programma, con la condizione che si sarebbe osservata se gli stessi soggetti non fossero stati esposti al trattamento (Rettore, Trivellato, Martini 2002). Nell'ambito delle politiche del lavoro la condizione che si è interessati a modificare può essere la condizione di occupato o disoccupato, il livello delle retribuzioni o particolari comportamenti sociali come la discriminazione delle donne nel mercato del lavoro (il progetto "Maggiori Opportunità" della regione Veneto).

L'effetto che gli analisti sono interessati a stimare è un effetto netto, ripulito da tutti i fattori che possono contribuire ad una differenza nei risultati tra gruppo di trattati e gruppo di controllo. Vedremo nel capitolo come in letteratura è stata impostata l'analisi delle politiche, quali sono i problemi che possono emergere e infine quali sono i metodi sviluppati sinora per ottenere delle stime corrette nei diversi casi.

---

<sup>23</sup> In questa tesi utilizzeremo i termini *variabile risultato* e *variabile obiettivo* come sinonimi

### 3.1 Il problema fondamentale dell'inferenza causale

Definiamo come

- *evento fattuale* il valore osservato della variabile risultato a seguito dell'esposizione al trattamento per i soggetti esposti alla politica oppure il valore osservato in assenza del trattamento per i soggetti non esposti alla politica.
- *evento controfattuale* è invece definito come il valore della variabile risultato che si sarebbe osservata sui soggetti esposti se questi non avessero ricevuto il trattamento oppure sui soggetti non esposti se questi fossero stati esposti al trattamento.

Nell'approccio del modello causale elaborato da Rubin, i valori della variabile risultato che si sarebbero verificati sia in assenza sia in presenza della politica sono definiti *eventi potenziali*. Si definisce infine *effetto causale* del trattamento sulla variabile obiettivo il confronto tra i due risultati potenziali misurati sullo stesso gruppo di unità (Rubin, 1974). Non è possibile ovviamente osservare entrambi le situazioni potenziali in quanto non possiamo sapere quale sarebbe la condizione dei soggetti esposti nel caso in cui questi non fossero stati esposti. Questo problema è definito in letteratura come “*the fundamental problem of causal inference*” (Holland, 1986): la non osservabilità di uno dei due risultati potenziali implica, infatti, l'impossibilità di effettuare un confronto tra evento fattuale e controfattuale e la conseguente impossibilità di osservare l'impatto di una politica.

Una strategia utilizzata per superare questo problema consiste nella costruzione di un adeguato gruppo di controllo, ossia di un gruppo di soggetti non esposti alla politica che possano approssimare bene l'evento controfattuale. Questo procedimento conduce a risultati validi solo se i soggetti appartenenti al gruppo di confronto non presentano differenze sistematiche nei confronti dei soggetti esposti; è necessario che, se presenti, le differenze tra i due gruppi siano

ric conducibili esclusivamente all'intervento e non siano dovute invece al processo attraverso il quale i gruppi sono stati formati. Per costruire adeguati gruppi di confronto è possibile utilizzare metodi sperimentali (campionamento casuale) o non sperimentali. Tra questi ultimi ci concentreremo sul *regression discontinuity design* che si adatta bene al disegno della politica della mobilità, oggetto di questa tesi.

### 3.2 Notazione

Sia  $\{\omega\}$  la popolazione obiettivo della politica che può potenzialmente essere esposta al trattamento dalla quale si estrae un campione di numerosità  $N$ . Definiamo  $Y$  la caratteristica dei soggetti della popolazione nei confronti della quale viene condotta la valutazione d'impatto; ci riferiremo a questa quantità come la variabile obiettivo o variabile risultato.

Per ogni soggetto  $i$  ( $i = 1, \dots, N$ ) definiamo le seguenti quantità:

- $X$  è un insieme di caratteristiche osservabili misurate prima del trattamento
- $Y_i(1)$  è il valore che assumerebbe la variabile obiettivo se l' $i$ -esima unità fosse esposta al trattamento
- $Y_i(0)$  è il valore che assumerebbe la variabile obiettivo nel caso in cui l' $i$ -esima unità non fosse esposta al trattamento
- $Z_i$  è una variabile binaria che indica l'esposizione al trattamento

$$Z_i = \begin{cases} 1 & \text{se l'i - esima unità è esposta al trattamento} \\ 0 & \text{se l'i - esima unità non è esposta al trattamento} \end{cases}$$

Il valore che assume questa variabile determina quale dei due eventi potenziali risulta essere il valore fattuale,  $Y_i(1)$  se  $Z_i = 1$  o  $Y_i(0)$  se  $Z_i = 0$ , e quale il valore controfattuale,  $Y_i(0)$  se  $Z_i = 1$  o  $Y_i(1)$  se  $Z_i = 0$ .

L'impatto della politica per l' $i$ -esima unità è definito come

$$\beta_i = Y_i(1) - Y_i(0) \quad (1)$$

Poiché l'impatto è la differenza tra due eventi, uno dei quali non osservabile, l'impatto stesso diventa non osservabile per il singolo soggetto  $i$  (problema fondamentale dell'inferenza causale). La valutazione si concentra quindi sui momenti della distribuzione, primo fra tutti la media. L'effetto medio calcolato sull'intero campione è noto in letteratura come *Average Treatment Effect* (ATE); se l'analisi si concentra su sottocampioni della popolazione ci si riferisce invece all'effetto medio per i trattati, *Average Treatment Effect on Treated* (ATT), oppure per i non trattati, *Average Treatment Effect on Not Treated* (ATNT).

Ipotizziamo che l'obiettivo della valutazione sia l'impatto medio del trattamento sul sottogruppo costituito dai soggetti trattati; in base alla notazione fin qui definita l'ATT è pari alla quantità

$$E[\beta | Z = 1] = E[Y(1) - Y(0) | Z = 1] = E[Y(1) | Z = 1] - E[Y(0) | Z = 1] \quad (2)$$

dove l'elemento  $E[Y(0) | Z = 1]$  non è osservabile perché rappresenta il valore medio della variabile obiettivo che si sarebbe osservata per il gruppo dei soggetti esposti se questi non fossero stati trattati.

L'unico confronto possibile, basato sulle quantità osservabili, è pari alla differenza del valore medio che assume la variabile risultato da trattati per i soggetti sottoposti al trattamento e il valore della variabile risultato da non trattati per i soggetti che non hanno usufruito del trattamento

$$E[Y(1) | Z = 1] - E[Y(0) | Z = 0] \quad (3)$$

Non è possibile interpretare questa quantità in termini di impatto della politica; le differenze osservate infatti sono pari all'effetto del trattamento a cui si sommano le differenze tra i gruppi che si sarebbero osservate in assenza della politica. Questo risultato si dimostra tramite pochi passaggi; aggiungendo e sottraendo l'elemento non osservabile  $E[Y(0) | Z = 1]$  alla (2) otteniamo la relazione esistente tra la quantità osservabile e la quantità di interesse ATT:

$$E[\beta | Z = 1] = E[Y(1) | Z = 1] - E[Y(0) | Z = 0] - (E[Y(0) | Z = 1] - E[Y(0) | Z = 0]) \quad (4)$$

I primi due termini sono i risultati medi osservabili nei due gruppi, l'ultimo termine compreso tra parentesi indica la differenza tra le medie dei trattati e le medie dei non trattati che si sarebbe osservata se entrambi i gruppi non fossero stati esposti alla politica. Questo termine è indicato in letteratura come *selection bias* e rappresenta tutte le differenze presenti tra i gruppi oggetto dell'analisi non riconducibili all'intervento. Non è identificabile in quanto coinvolge una quantità non identificabile, il risultato potenziale da non esposti calcolato per il gruppo degli esposti.

La distorsione è causata dal processo di selezione dei soggetti coinvolti nella politica; i gruppi che devo confrontare differiscono per particolari caratteristiche che influiscono sulla variabile risultato e che, nello stesso momento, dipendono dal modo in cui sono stati selezionati i gruppi. In questo modo i soggetti esposti e i soggetti non esposti sono caratterizzati da diversi valori potenziali della variabile obiettivo Y anche in caso di assenza della politica.

Per poter ricondurre le differenze osservate all'impatto della politica è possibile procedere in due modi: nelle situazioni in cui sia possibile, si costruisce un gruppo di confronto che, in assenza del trattamento, non presenta differenze dal gruppo assegnato al trattamento in modo che il selection bias sia nullo. Alternativamente

si possono utilizzare informazioni aggiuntive per calcolare il termine di disturbo e sottrarlo alle differenze osservate per ottenere la stima corretta dell'impatto.

Nell'ambito del disegno sperimentale è possibile ottenere degli stimatori corretti dell'impatto medio del trattamento tramite la costruzione di campioni casuali; questo procedimento permette di interpretare la differenza osservata nei due gruppi come l'impatto della politica. La condizione che ci permette di affermare che in caso di randomizzazione il selection bias è nullo è la seguente:

$$Y(0), Y(1) \perp Z \quad (5)$$

Nell'ambito della ricerca sociale, la selezione casuale dei gruppi è spesso non applicabile per problemi che possono essere di origine etica o semplicemente perché non è possibile controllare il processo di selezione (i soggetti possono decidere di entrare nel gruppo al quale non sono stati assegnati o la politica può consentire l'autoselezione degli individui). In letteratura sono stati elaborati diversi metodi non sperimentali per la costruzione di stimatori consistenti dell'impatto; questi si dividono in due diverse tipologie a seconda che il processo di selezione sia stato generato da variabili osservabili o non osservabili. Nel nostro studio ci concentreremo sulla prima tipologia di stimatori.

### 3.3 Regression Discontinuity Design

*“The regression discontinuity (RD) design is a quasi-experimental design with the defining characteristic that the probability of receiving treatment changes discontinuously as a function of one or more underlying variables”* (Hahn, Todd, Var der Klaauw, 1991).

Questo metodo è applicabile quanto il processo di selezione è una funzione deterministica di una caratteristica osservabile  $X$  misurata prima del trattamento.

Il soggetto  $i$ -esimo viene assegnato al gruppo di trattamento o al gruppo di controllo qualora la variabile pretrattamento assuma valore  $X_i$  inferiore o superiore al valore soglia stabilito dalla politica (il valore soglia può essere il valore di un test ed hanno accesso alla trattamento solo coloro che raggiungono un punteggio superiore oppure può essere un particolare livello di reddito che non bisogna superare per poter usufruire di particolari benefici). Se identifichiamo con  $x^*$  il valore soglia, il trattamento può essere così definito

$$Z_i = \begin{cases} 0 & \text{se } X_i \leq x^* \\ 1 & \text{se } X_i > x^* \end{cases}$$

Il metodo si basa sull'assunzione che in prossimità della soglia la variabile pre-trattamento e di conseguenza il processo di selezione sia indipendente dai valori potenziali della variabile risultato.

$$Y(0), Y(1) \perp Z \mid X = x^* \quad (6)$$

La (6) significa che, il processo di selezione del gruppo di trattati e non trattati non è legato ai risultati potenziali dei soggetti in prossimità della soglia. Verosimilmente si può infatti pensare che, poiché l'assegnazione al trattamento avviene solo in base ad una particolare caratteristica  $X$ , i soggetti che stanno appena al di sopra o al di sotto della soglia non presentino differenze sistematiche. L'effetto stimato sarà dunque imputabile alla sola politica e non ad altri fattori.

Questo metodo ha una forte validità interna per le assunzioni molto semplici sulle quali si basa; al contrario presenta forti limiti di validità esterna. I risultati sono corretti per il sottogruppo della popolazione che consideriamo (i soggetti attorno alla soglia) ma non possono essere generalizzati a tutta la popolazione che presenta caratteristiche diverse dai sottogruppi utilizzati per la stima.



## Capitolo 4

### Principal Stratification

La Principal Stratification<sup>24</sup> è un nuovo metodo, elaborato nell'ambito dell'analisi sperimentale, per trattare problemi di selezione endogena non casuale che possono sorgere in un momento successivo all'assegnazione al trattamento (Frangakis, Rubin 1999, 2002). Questo approccio ha trovato fin da subito applicazione prevalentemente nell'ambito delle politiche sociali o in generale in tutti gli studi che coinvolgono esseri umani. In questi casi infatti i diversi comportamenti che possono assumere i soggetti beneficiari della politica possono modificare le modalità con le quali questa è applicata e complicare o addirittura invalidare le successive analisi: i soggetti coinvolti per esempio possono decidere di non entrare a far parte del gruppo al quale sono stati assegnati (fenomeno che prende il nome di *noncompliance*) oppure possono decidere di non rispondere, di abbandonare lo studio o in generale di assumere un comportamento che non permetta la misurazione della condizione o dell'evento oggetto dello studio. Questi processi non sono usualmente controllabili dall'analista e, nel momento in cui si verificano in maniera non casuale, danno origine a gruppi non equivalenti in assenza dell'intervento. Di conseguenza ci ritroviamo nella situazione, largamente

---

<sup>24</sup> Nel corso dello studio useremo come sinonimi i termini stratificazione principale o la sigla PS

descritta nel capitolo precedente, nella quale i gruppi di soggetti esposti o meno al trattamento non sono confrontabili e l'impatto della politica non è identificabile.

In letteratura, l'analisi di problemi di selezione endogena non casuale nell'ambito della valutazione di politiche aveva già portato alla realizzazione di diversi stimatori. Nei casi di non compliance ad esempio, l'analisi classica prende il nome di *as-treated analysis* in quanto è realizzata considerando il trattamento effettivamente ricevuto come se fosse il trattamento assegnato casualmente e non invece come il risultato di scelte individuali dei soggetti coinvolti. Un'evoluzione di questo approccio consiste nel confrontare i gruppi considerando esclusivamente l'assegnazione casuale e ignorando il reale trattamento ricevuto; questo metodo prende il nome di *intention to treat analysis*. Nel caso in cui la non compliance sia combinata con problemi di mancata risposta<sup>25</sup>, si definisce come *respondent based intention to treat analysis* un'analisi dell'effetto dell'assegnazione al trattamento basata solo sui rispondenti, che ignora sia la non aderenza al trattamento sia la presenza di dati mancanti. Nessuno di tali approcci riesce però a produrre una stima non distorta dell'impatto.

Lo sviluppo della Principal Stratification, definita dagli stessi autori come “...with respect to a posttreatment variable[...] a cross- classification of the units based on their joint potential values of that variable under each of the treatments being compared”, ha portato alla costruzione di una nuova procedura di stima dell'impatto che mantiene buone proprietà rispetto agli stimatori utilizzati in precedenza (Frangakis and Rubin 1999, 2002).

L'obiettivo iniziale degli autori era quello di sviluppare un nuovo metodo al fine di costruire un “surrogato” della variabile risultato, ossia un suo indicatore che fosse possibile misurare in anticipo rispetto al primo. La definizione della stratificazione principale e degli effetti principali hanno permesso loro di costruire un nuovo tipo di indicatore, chiamato appunto *principal surrogate* che è

---

<sup>25</sup> Per mancata risposta intendiamo la mancata osservazione del valore della variabile risultato a causa dell'abbandono del programma da parte dei soggetti oppure della loro volontà di non rendere noto tale valore.

caratterizzato da buone capacità predittive e di proprietà migliori rispetto agli indicatori già esistenti.

L'ambito di applicazione si è dimostrato in seguito molto vasto. Nel campo delle politiche sociali, la Stratificazione Principale ha permesso di trattare diversi problemi di selezione endogena non casuale tra i quali possiamo elencare

- non aderenza dei soggetti al trattamento assegnato.
- non osservabilità della variabile risultato, principalmente in studi che si estendono a lungo nel tempo con un alto rischio di abbandono dei soggetti.
- *Censoring by death*, termine con il quale si intende l'impossibilità che l'evento che siamo interessati a studiare si verifichi per particolari gruppi di soggetti<sup>26</sup>.
- Autoselezione delle unità assegnate al trattamento.

Recenti studi si sono inoltre focalizzati sulla relazione tra PS e l'analisi degli effetti diretti e indiretti (Rubin 2004,2005) e sull'applicazione del metodo delle variabili strumentali in questo ambito (Mealli, Pacini 2008).

In questo capitolo descriveremo inizialmente l'elaborazione teorica del metodo della Principal Stratification specificando la tipologia delle variabili coinvolte, la definizione di *strati principali* e di *effetti principali* e le assunzioni sulle quali si basa la validità di questo metodo. Nelle sezioni successive analizzeremo nel dettaglio il problema di selezione endogena non casuale che prende il nome di "censoring by death", le implicazioni che questo ha nella valutazione d'impatto e nella correttezza delle stime. Infine mostreremo l'applicazione della PS nel caso specifico sopra descritto, per eliminare la distorsione da selezione e pervenire alla costruzione di gruppi confrontabili.

---

<sup>26</sup> I gruppi di soggetti sono classificati in base al valore che assume una variabile intermedia. Per un approfondimento di rimanda al paragrafo 4.2

## 4.1 Effetti Causali e Variabili Intermedie

Riprendiamo la notazione utilizzata nel capitolo precedente. Sia  $\{\omega\}$  la popolazione di riferimento dalla quale viene estratto il campione di unità  $i = 1, \dots, N$ . La variabile  $Z$  indica il trattamento al quale ogni soggetto può essere potenzialmente esposto ed assume i valori  $Z=0$  nel caso in cui sia assegnato a non ricevere il trattamento,  $Z=1$  altrimenti. Siano infine  $X$  un insieme di variabili relative ad un periodo precedente il trattamento e  $Y$  la variabile obiettivo relativa ad un determinato momento o intervallo temporale successivo all'attuazione della politica.

Rispetto alla notazione definita nel capitolo precedente definiamo un'ulteriore covariata  $S$ , una variabile post-trattamento binaria<sup>27</sup>, che può essere una misura della compliance del soggetto, un indicatore della presenza o meno del valore dell'outcome, un surrogato della variabile obiettivo. Per semplicità assumiamo che si possano realizzare le seguenti due situazioni,  $S_i(Z) = 0$  oppure  $S_i(Z) = 1$ .

Per ogni soggetto, in sintesi, esistono le due coppie di valori potenziali:

- $Y_i(0), S_i(0)$  se l'individuo  $i$ -esimo è assegnato al gruppo di controllo e quindi non riceve il trattamento
- $Y_i(1), S_i(1)$  se l'individuo  $i$ -esimo è assegnato al gruppo di trattamento

Come già descritto nel precedente capitolo, l'analista è interessato a valutare l'effetto che ha il trattamento sulla variabile  $Y$  rispetto alla situazione che si verificherebbe se lo stesso trattamento non venisse applicato. La quantità sulla quale si basa la stima è il confronto tra i valori che assume la variabile risultato

---

<sup>27</sup> In questo studio ci limiteremo a considerare  $Z$  e  $S$  come variabili binarie ma è possibile generalizzare al caso dell'esistenza di più livelli di trattamento e di più modalità per la variabile intermedia.

per il gruppo di esposti al trattamento e i valori che assume la stessa variabile, misurata nello stesso momento, per il gruppo di controllo. Perché questo confronto possa essere interpretabile come effetto causale è necessario che i due gruppi siano identici.

In situazioni di questo tipo, l'approccio standard utilizzato in precedenza si serviva della variabile intermedia per correggere la stima dell'impatto. La quantità di interesse era l'effetto netto dell'assegnazione al trattamento aggiustato per la variabile post-trattamento e si basava sul confronto tra le distribuzioni

$$pr\{Y_i^{obs} | S_i^{obs} = s, Z_i = 0\} \text{ e } pr\{Y_i^{obs} | S_i^{obs} = s, Z_i = 1\} \quad (7)$$

dove  $Y_i^{obs} = Y_i(Z_i)$  e  $S_i^{obs} = S_i(Z_i)$ .

Questo approccio si basa principalmente su due assunzioni che assumeremo valide anche nel corso di questo capitolo.

Assunzione 1 :Stable Unit Treatment Value Assumption (SUTVA)

Le unità non interferiscono tra di loro, l'effetto della politica per il singolo soggetto non è influenzata dall'effetto che la stessa politica ha avuto per gli altri soggetti esposti al trattamento.

Assunzione 2: Assegnazione casuale

$$Z \perp Y(0), Y(1), S(0), S(1) | X \quad (8)$$

Si suppone che il trattamento sia assegnato in maniera completamente casuale entro le celle definite dalle variabili pre-trattamento  $X$ .

Il confronto di interesse ha luogo quindi tra le seguenti distribuzioni:

$$pr\{Y_i(0) | S_i(0) = s, X\} \text{ e } pr\{Y_i(1) | S_i(1) = s, X\} \quad (9)$$

Anche nel caso in cui le assunzioni siano entrambe valide, il confronto proposto non può essere considerato un effetto causale in tutti i casi in cui la variabile post trattamento, che ha un effetto su  $Y$ , è influenzata nel contempo dal trattamento stesso. Tutti i soggetti sono caratterizzati dallo stesso valore della variabile post trattamento  $S=s$ , però la realizzazione di questo particolare evento avviene in due situazioni diverse. Il primo gruppo infatti ottiene un valore per la variabile intermedia pari a  $s$  quando sottoposto al trattamento, mentre il secondo gruppo ottiene lo stesso risultato quando non è sottoposto allo stesso trattamento. I sottogruppi di esposti e non esposti alla politica,  $\{i : S_i(1) = s, X\}$  e  $\{i : S_i(0) = s, X\}$ , sebbene siano caratterizzati dallo stesso valore della variabile  $S$ , non sono identici e il loro confronto non può essere identificato con l'effetto causale della politica<sup>28</sup>. Questo fenomeno è noto in letteratura come distorsione dovuta a selezione post trattamento (Rosenbaum, 1984).

Un ulteriore approccio proposto in passato consiste nell'utilizzare delle quantità contro fattuali per  $Y$  definite a priori (Robins e Greenland, 1992). Questo metodo ha un ambito di applicazione molto limitato in quanto si può applicare esclusivamente nel caso in cui l'analista possa controllare direttamente la variabile intermedia ed esistano dei valori reali che sia sensato utilizzare come controfattuali.

Vedremo nel seguito del capitolo come la PS costituisca un approccio completamente innovativo che ha permesso di superare i limiti esistenti nei metodi di stima fin qui descritti.

---

<sup>28</sup> Così come definito nel capitolo 3.1 nell'ambito del modello causale di Rubin

## 4.2 Gli effetti principali

La stratificazione principale utilizza la variabile intermedia  $S$  in un modo del tutto nuovo rispetto ai metodi precedenti al fine di ottenere dei campioni confrontabili e pervenire quindi ad una stima consistente dell'effetto causale.

### *Definizione 1*

(a) la stratificazione principale di base  $P_0$  rispetto ad una variabile post trattamento  $S$  è la partizione delle unità  $i = 1, \dots, n$  in modo tale che entro ogni gruppo  $P_0$  le unità hanno in comune il valore del vettore  $(S_i(0), S_i(1))$ .

(b) La **stratificazione principale**  $P$  rispetto ad una variabile post trattamento  $S$  è una partizione delle unità i cui elementi sono unione dei gruppi della stratificazione di base  $P_0$ .

Le unità vengono classificate quindi in quattro sottogruppi<sup>29</sup>  $S^P$  rispetto alla coppia di valori potenziali assunti dalla variabile intermedia  $(S_i(0), S_i(1))$ .

Sebbene possiamo definire gli strati, non è possibile osservare a quale di questi ogni soggetto appartenga in quanto palesemente risulta osservabile solo uno dei due valori potenziale che definiscono ogni strato. Per esempio un soggetto non esposto al trattamento per il quale la variabile intermedia assume valore  $S=1$  potrebbe appartenere allo strato costituito da individui per i quali la variabile intermedia assume sempre valore pari ad 1 indipendentemente dal gruppo al quale sono assegnati, oppure allo strato costituito da individui per cui la variabile intermedia assume valore pari a 1 se non esposti al trattamento e valore nullo in caso contrario.

---

<sup>29</sup> Nel caso in cui sia il trattamento sia la variabile intermedia siano variabili binarie. Nel caso generale è possibile avere più sottogruppi, in relazione al numero di modalità che possono assumere le variabili.

*Definizione 2*

Sia  $P$  una stratificazione principale rispetto ad una variabile post trattamento  $S$  e sia  $S_i^P$  un indicatore dello strato di  $P$  a cui appartiene l'unità  $i$ . Un **effetto principale**, rispetto a questa stratificazione principale, è definito come la differenza tra i risultati potenziali in caso di mancata assegnazione al trattamento rispetto al caso di assegnazione al trattamento entro lo strato principale  $\varsigma$  in  $P$ , ossia un confronto tra i gruppi ordinati

$$\{Y_i(1) | S_i^P = \varsigma, X\} \text{ e } \{Y_i(0) | S_i^P = \varsigma, X\} \quad (10)$$

Tramite il condizionamento allo strato principale è possibile superare il problema della distorsione da selezione. La coppia di valori che assume la variabile post-trattamento è indipendente dal trattamento stesso in quanto per ogni soggetto definisco quali valori assumerebbe la variabile intermedia sia se il soggetto fosse esposto al trattamento sia in caso contrario. Se quindi il singolo valore osservato della variabile post-trattamento riflette sia caratteristiche individuali sia l'effetto del trattamento; la coppia di valori potenziali è indipendente dal processo di assegnazione, per ogni stratificazione principale  $P$ . Lo strato principale di appartenenza riflette esclusivamente le caratteristiche individuali dei soggetti, può quindi essere utilizzato come una qualsiasi variabile pretrattamento mentre, condizionatamente ai singoli strati, non esistono altre caratteristiche non osservabili che possano essere legate al processo di selezione. In questo modo la distorsione da selezione non ha più ragione di esistere: i gruppi di soggetti esposti o meno al trattamento sono identici e di conseguenza gli effetti principali possono essere considerati degli effetti causali.

### 4.3 Il problema del ‘Censoring by death’

La principal stratification è stata inizialmente studiata e poi applicata in problemi di selezione endogena non casuale dovuti a noncompliance e mancata risposta da parte dei soggetti.

Nel nostro problema useremo questo approccio per risolvere una diversa causa di distorsione da selezione che in letteratura prende il nome di ‘*Censoring due to death*’ (Zhang e Rubin 2003, Rubin 2006) e che si realizza quando la variabile obiettivo è osservabile solo per i soggetti per i quali si verifica un particolare evento. Il nome si deve alla larga applicazione in ambito medico dove in molti casi la variabile obiettivo (ad esempio la qualità della vita, QOL, a seguito dell’utilizzo di un nuovo trattamento) è osservabile solo per i soggetti che sono sopravvissuti fino al momento della rilevazione.

Come già descritto, negli studi medici sulla qualità della vita la variabile intermedia può indicare se il soggetto è ancora in vita o meno fino al momento della rilevazione, ma esempi di questa particolare tipologia di dati sono presenti anche in altri ambiti. Nelle politiche sociali la variabile intermedia può indicare se il lavoratore risulta occupato o disoccupato quando la variabile di interesse è il salario, oppure ancora può indicare se lo studente ha partecipato o meno ad un test in cui si valutano le sue abilità in una politica finalizzata ad un miglioramento del livello di istruzione. Il livello del salario si osserva solo per i soggetti occupati mentre il punteggio acquisito nel test può essere rilevato solo per i soggetti che hanno effettivamente svolto il test.

Nella notazione adottata in questo studio, il censoring by death si traduce nell’impossibilità di osservare il valore di  $Y_i(z)$  nel momento in cui, per il soggetto  $i$ , si realizza l’evento  $S_i(z) = 0$ .

Per affrontare problemi di questo genere si erano già sviluppati diversi approcci<sup>30</sup>. Una prima tipologia, basata semplicemente sulla differenza nelle medie tra il gruppo di trattati e il gruppo di controllo, comprende metodi quali l'imputazione dei valori della variabile obiettivo nei casi in cui è censurata oppure l'utilizzo di una nuova variabile che sia definita per tutti i soggetti e che includa sia l'informazione sulla variabile intermedia sia sul valore della variabile troncata (Rubin 2006), per esempio nel caso dei salari e dello status occupazionale posso considerare la variabile *reddito* definita per tutti invece della variabile *salari* definita solo per i soggetti occupati. Il primo procedimento è concettualmente errato in quanto non esiste un valore per la variabile obiettivo che non osserviamo da poter sostituire; nel secondo caso potrebbe non esistere una variabile che soddisfa tali proprietà.

Approcci più complessi si focalizzano sulla stima dell'impatto solo per particolari sottogruppi per i quali la variabile risultato è sempre definita, senza pretese di allargare i risultati all'intera popolazione. Valutazioni di questo tipo si rifanno da una parte all'ambito dei *Selection models* (Heckman, 1979), ampiamente approfonditi nella letteratura econometrica, e dall'altro alla Principal Stratification nell'ambito del quale si sviluppano metodi che Mealli e Pacini (2008) valutano più robusti e caratterizzati da una maggiore facilità nell'interpretazione e nella validazione delle assunzioni rispetto alla prima classe di metodi.

#### **4.4 L'approccio della Principal Stratification nei problemi di 'Censoring by Death'**

Assumiamo innanzitutto che valgano le ipotesi di non interferenza tra le unità (SUTVA) e che possiamo considerare l'assegnazione al trattamento come casuale

---

<sup>30</sup> Per una panoramica sui metodi esistenti nel caso di truncation by death si rimanda a McConnell et al. (2008)

condizionatamente ai valori delle variabili pretrattamento  $X$ . Nella parte successiva dello studio, senza perdita di generalità, ipotizziamo di condurre tutte le analisi entro le celle definite dalle variabili  $X$ .

L'approccio della principal stratification suddivide il campione di riferimento in quattro sottogruppi, gli strati principali, definiti dalla coppia di valori che assumerebbe la variabile intermedia nel caso in cui i soggetti siano esposti al trattamento oppure non siano esposti allo stesso.

Gli strati principali  $S^P$  definiti in base alla coppia di valori potenziali  $(S_i(0), S_i(1))$  sono i seguenti

$$\begin{aligned} 11 &= \{i : S_i(0) = S_i(1) = 1\}, & 00 &= \{i : S_i(0) = S_i(1) = 0\} \\ 10 &= \{i : S_i(0) = 1, S_i(1) = 0\}, & 01 &= \{i : S_i(0) = 0, S_i(1) = 1\} \end{aligned}$$

Definiamo inoltre le proporzioni di soggetti che costituiscono ogni strato come  $\pi_{11}, \pi_{10}, \pi_{01}, \pi_{00} = 1 - \pi_{11} - \pi_{10} - \pi_{01}$ .

Poiché, come descritto nella sezione precedente, nel contesto di problemi di 'censura per morte' la variabile risultato è definita solo nel caso in cui la variabile intermedia assuma valore pari a uno, tra tutti gli strati, solo il primo ha entrambi i valori potenziali della variabile risultato definiti sia nel caso in cui i soggetti siano assegnati al trattamento sia nel caso in cui gli stessi siano assegnati al gruppo di controllo,  $\{Y_i(0) | S_i^P = 11\}$  e  $\{Y_i(1) | S_i^P = 11\}$ . Per i soggetti che appartengono allo strato  $S_i^P = 00$  non ho mai valori definiti per la variabile risultato mentre per i rimanenti due strati sono definiti rispettivamente solo i valori  $\{Y_i(0) | S_i^P = 10\}$  e  $\{Y_i(1) | S_i^P = 01\}$ .

In questo contesto non è sensato calcolare gli effetti della politica per l'intera popolazione in quanto l'assenza di almeno uno dei due valori potenziali della variabile di interesse  $Y$  preclude un possibile confronto. L'obiettivo dell'analisi risulta essere quindi la valutazione dell'effetto della politica esclusivamente per il

sottogruppo di soggetti per i quali è definita in entrambi i casi la variabile risultato, in questo caso specifico per i soggetti che appartengono allo strato  $S^P = 11$ . Questa quantità è definita in letteratura come *Survival Average Causal Effect*<sup>31</sup> (Rubin 2000)

$$SACE = E[Y(1) - Y(0) | S^P = 11] \quad (11)$$

Il nome ancora una volta rimanda all'ambito medico dove (riprendendo l'esempio del paragrafo precedente) si può interpretare come l'effetto del trattamento sulla qualità della vita per i soggetti che sarebbero sopravvissuti sia nel caso fossero stati sottoposti al nuovo trattamento, sia in caso contrario.

L'appartenenza allo strato non è osservabile ma è considerata come una variabile latente. Per identificare a quale strato appartengano i singoli soggetti è utile studiare la corrispondenza che esiste tra gli strati latenti e i gruppi osservati, definiti dalle coppie di valori assunte dalle variabili Z ed S.

Tabella 4.1. Corrispondenza tra gruppi osservati e strati latenti

| <b>Gruppi osservati:</b><br><b>G(Z,S)</b> | <b>Variabile risultato</b><br><b>Y(Z)=Y<sup>obs</sup></b> <sup>32</sup> | <b>Strati principali</b> |
|-------------------------------------------|-------------------------------------------------------------------------|--------------------------|
| $G(1,1) = \{i : Z_i = 1, S_i = 1\}$       | Y(1)                                                                    | 11 o 01                  |
| $G(1,0) = \{i : Z_i = 1, S_i = 0\}$       | *                                                                       | 00 o 10                  |
| $G(0,1) = \{i : Z_i = 0, S_i = 1\}$       | Y(0)                                                                    | 11 o 10                  |
| $G(0,0) = \{i : Z_i = 0, S_i = 0\}$       | *                                                                       | 01 o 00                  |

<sup>31</sup> Nel seguito dello studio ci si riferirà a questo effetto con la sigla SACE

<sup>32</sup> Il simbolo \* indica che il valore della variabile Y non esiste per tutti i soggetti che appartengono al sottogruppo corrispondente.

Si può notare come ogni gruppo osservato sia costituito dalla mistura di due strati. Il primo gruppo, ad esempio, è costituito dai soggetti che sono stati assegnati al trattamento e per i quali la variabile intermedia assume valore pari ad uno. Allo stesso gruppo appartengono i soggetti che provengono sia dallo strato  $S^P = 11$  sia dallo strato  $S^P = 01$ , ossia rispettivamente i soggetti per i quali la variabile intermedia assume sempre valore uno e i soggetti per i quali la stessa variabile assume valore uno solo nel caso in cui siano sottoposti al trattamento. Con lo stesso procedimento possiamo attribuire a tutti i gruppi i diversi strati che li compongono.

I metodi sviluppati sono finalizzati ad isolare la componente di interesse dai due sottogruppi osservati  $G(11)$  e  $G(01)$  per i quali è definita ed è osservabile la variabile risultato. Dal primo gruppo è necessario isolare la quantità di interesse  $E[Y(1)|S^P = 11]$  mentre dal secondo si estrae  $E[Y(0)|S^P = 11]$ . Il confronto dei due valori attesi così calcolati costituisce l'effetto causale della politica per il particolare sottogruppo di soggetti appartenenti allo strato  $S^P = 11$ .

A partire dal confronto analizzato in Tabella 1, senza ulteriori assunzioni non è possibile calcolare il valore esatto delle proporzioni degli strati e del valore della variabile risultato entro gli strati, ma è possibile un'identificazione solo parziale (Mealli e Pacini, 2008). L'utilizzo di particolari assunzioni sugli strati oppure sulla distribuzione parametrica della variabile  $Y$  entro gli strati permettono, in alcuni casi, l'identificazione puntuale dell'effetto causale di interesse.

In particolare possiamo classificare i metodi basati sull'approccio della PS in due tipologie: ad una prima classe di metodi non parametrici utile per un'identificazione solo parziale dell'effetto si contrappone una classe di metodi parametrici per l'identificazione puntuale dello stesso. Entrambi i metodi si basano sulle assunzioni di assegnazione casuale del trattamento, condizionatamente all'insieme di variabili osservabili  $X$ , e di non interferenza tra i soggetti (SUTVA).

#### 4.4.1 Metodi non parametrici: intervalli di confidenza

Basandosi esclusivamente sulle quantità osservate è possibile costruire un intervallo di valori plausibili per l'effetto SACE. In Zhang e Rubin (2003) l'intervallo è costruito calcolando i valori minimi e massimi per il valore medio della variabile obiettivo  $Y$  per i soggetti appartenenti allo strato 11, rispettivamente nel gruppo di trattamento e nel gruppo di controllo,  $\bar{Y}_{11}(Z=1)$  e  $\bar{Y}_{11}(Z=0)$ .

L'ampiezza dell'intervallo può eventualmente essere ristretta sulla base di assunzioni aggiuntive che possono variare a seconda del contesto al quale questo approccio è applicato e per le quali è necessaria una verifica in relazione alla loro plausibilità. In genere le assunzioni mirano principalmente a diminuire il numero di strati che compongono particolari sottogruppi di soggetti, oppure a limitare i valori che la variabile risultato può assumere entro particolari strati, senza mai imporre alcun modello parametrico per la distribuzione della variabile  $Y$ .

L'assunzione più frequentemente utilizzata e plausibile nell'ambito degli studi sociali prende il nome di *monotonicity*. Questa assunzione prevede che il trattamento non possa essere dannoso e quindi non possa provocare un peggioramento nella variabile che definisce il processo di "censura"; in pratica quindi implica l'assenza dello strato 10. La validità di questa assunzione, seppur plausibile in molti studi, deve essere sostenuta e argomentata in ogni singolo caso.

In altri studi, l'utilizzo congiunto di diverse assunzioni, ad esempio l'ipotesi di dominanza stocastica combinato a *monotonicity*, consente una stima puntuale dell'effetto causale (Zhang e Rubin 2003, Grilli e Mealli 2008).

#### 4.4.2 Metodi parametrici: il modello di mistura

Successivamente è stato sviluppato un approccio più elaborato e flessibile che consiste nella modellazione dell'appartenenza agli strati attraverso modelli di mistura (Zhang, Rubin e Mealli 2006, Rubin 2006). Questo metodo parametrico

prevede la formulazione congiunta di due diversi modelli: un modello per l'appartenenza agli strati e un modello per la variabile risultato condizionatamente allo strato e al trattamento ricevuto. In particolare, condizionandoci ai valori delle variabili  $X$ , la distribuzione congiunta dei risultati potenziali è pari al prodotto della distribuzione degli stati principali condizionata alle variabili pretrattamento e la distribuzione della variabile risultato potenziale condizionata sia alle variabili pretrattamento sia agli strati. Se invece del campione completo considero separatamente i due gruppi di soggetti esposti o meno al trattamento per cui osservo la variabile risultato, riprendendo la notazione utilizzata precedentemente, ottengo i due modelli distinti

$$\begin{aligned} f(Y(1) | X) = & \Pr(S^P = 11 | X) * f(Y(1) | S^P = 11, X) + \\ & + \Pr(S^P = 01 | X) * f(Y(1) | S^P = 01, X) \end{aligned} \quad (12)$$

$$\begin{aligned} f(Y(0) | X) = & \Pr(S^P = 11 | X) * f(Y(0) | S^P = 11, X) + \\ & + \Pr(S^P = 10 | X) * f(Y(0) | S^P = 10, X) \end{aligned} \quad (13)$$

L'identificazione dei modelli si basa sulla teoria dei modelli di misture finite (Everitt, Hand 1981; McLachlan, Peel 2000) che permettono di tenere in considerazione la presenza di sottogruppi, gli strati, caratterizzati da differenti distribuzioni della variabile  $Y$  entro lo stesso gruppo osservabile di soggetti. La procedura di stima utilizzata più frequentemente, utilizzata anche in questo studio, è l'algoritmo Expectation-Maximization (Dempster et al. 1977). Altri metodi che ritroviamo in letteratura sono il metodo Markov Chain Monte Carlo (Gelman et al. 2004) oppure metodi bayesiani che permettono di ottenere stime più precise sfruttando le informazioni a priori (Dominici et al., 2006).



## Capitolo 5

### I dati

Questa tesi prende le mosse dall'analisi descritta e sviluppata in Paggiaro, Rettore e Trivellato (2008). Il loro studio si basa su un dataset nuovo rispetto a quello utilizzato nelle analisi svolte nei primi anni di implementazione della politica, caratterizzato da una qualità più alta e da una maggiore ricchezza di informazioni; la disponibilità dei dati sui livelli dei salari relativi sia al periodo precedente sia a quello successivo al momento dell'iscrizione alle Liste di ogni soggetto ha permesso agli autori di muoversi dalle analisi iniziali che avevano caratterizzato il decennio precedente e realizzare una valutazione d'impatto più affidabile e completa, dove alla valutazione sul tasso di occupazione si affianca l'analisi degli effetti della politica sulla qualità del lavoro. L'utilizzo di un insieme più ampio di informazioni è stato reso possibile da un molteplici abbinamento tra tre diverse fonti amministrative.

Il campione utilizzato è rappresentato dai lavoratori iscritti alle liste nella regione Veneto tra il 1995 e il 1998; la finestra temporale è stata scelta sulla base di precisi criteri di affidabilità e disponibilità delle informazioni necessarie. Nel capitolo daremo una breve sintesi delle fonti utilizzate nella costruzione del dataset; successivamente descriveremo la composizione del campione sul quale verrà poi realizzata la valutazione d'impatto.

## 5.1 Le Fonti Amministrative

I dati utilizzati provengono da differenti archivi amministrativi gestiti a carattere regionale o nazionale. Dalle prime analisi svolte, molti passi in avanti sono stati compiuti sulla qualità e affidabilità dei dati. A partire dall' utilizzo dei semplici dati raccolti a livello regionale, l'insieme informativo utilizzato si è allargato tramite un abbinamento prima con il sistema informatico di gestione dei Centri per l'Impiego (CPI), Netlabor, e successivamente con la banca dati dell'istituto per la previdenza sociale italiano (Inps). Per quanto sia possibile allargare l'insieme informativo, la natura amministrativa dei dati utilizzati per l'analisi di questo programma ha fin da subito costituito un limite per una loro manipolazione a fini statistici. Inoltre non esiste un archivio in Italia che raccolga in maniera standardizzata i dati relativi alle Liste di Mobilità; le informazioni sono gestite solo a livello regionale dalle singole agenzie per l'impiego che curano la raccolta dei dati secondo una propria metodologia organizzativa, impedendo di fatto un'analisi comparativa a carattere nazionale (Brunello, Miniaci 2007).

Vediamo nel dettaglio quali sono le diverse fonti amministrative dalle quali provengono i dati utilizzati in questo studio.

La legge 123/1991 affidava alle Commissioni Regionali il compito di gestire le liste attraverso gli **archivi amministrativi regionali**, basi di dati informatizzate che fungono esclusivamente da mero elenco dei soggetti destinatari del programma. Nati esclusivamente per scopi amministrativi, gli archivi contengono informazioni relative ad iscrizioni, avviamento e cancellazione dei lavoratori e vengono aggiornate in maniera continuativa. In particolare, per ogni singolo lavoratore, sono indicate le caratteristiche socio-anagrafiche (sesso, età, istruzione, provincia di residenza), settore dell'impresa di provenienza, posizione lavorativa nella stessa, data di iscrizione alla Lista, diritto o meno all'indennità di mobilità e status corrente tra cui si può distinguere tra le tre situazioni: assunto con contratto a tempo permanente, cancellato o iscritto alla lista. La natura amministrativa degli archivi li rende uno strumento carente dal punto di vista

statistico; le informazioni infatti sono spesso raccolte in maniera poco accurata e, in particolare negli archivi relativi ai primi anni di attivazione del programma, si registra una significativa presenza di dati mancanti. Inoltre, poiché la gestione avviene a livello regionale e non segue una procedura universale a livello nazionale, l'utilizzo dei soli dati estratti dagli archivi preclude la possibilità di un confronto tra la politica nelle diverse Regioni e una valutazione del programma nel complesso.

L'insieme informativo si è successivamente ampliato attraverso l'abbinamento degli archivi regionali con **Netlabor**, un archivio informatizzato, contenente dati individuali su lavoratori e aziende, gestito dai CPI (Trivellato, 2001). Questa fonte è costituita da informazioni aggiuntive, relative non solo ai singoli eventi occorsi nel periodo di permanenza in Lista ma anche alla storia lavorativa precedente e successiva allo stesso. Per ogni episodio lavorativo si rileva la data di assunzione, la tipologia del contratto e la qualifica professionale del lavoratore, il settore dell'impresa che assume. Inoltre sono presenti caratteristiche demografiche che completano quelle presenti nell'archivio delle liste quali ad esempio titolo di studio o data di nascita. Sebbene di qualità migliore, i dati continuano a manifestare tutti i limiti propri dei dati di natura amministrativa, non adeguati a fini di ricerca.

Per far fronte ad esigenze crescenti di monitoraggio e valutazione della politica, è stato realizzato negli ultimi anni un ulteriore abbinamento tra gli archivi amministrativi, Netlabor e la base di dati dell' **Inps** che racchiude le informazioni relative alla storia lavorativa e ai salari dal 1975 fino al 2001. In questo modo è stato possibile utilizzare le informazioni annue sul salario percepito, il numero di settimane retribuite, lo status di occupato o disoccupato relativo ad ogni mese al fine di ricostruire la carriera lavorativa dei singoli lavoratori e la dinamica retributiva per un intervallo temporale molto ampio.

L'utilizzo integrato di Inps e le precedenti fonti utilizzate consente di sviluppare un nuovo filone di analisi rispetto alle valutazioni d'impatto effettuate in precedenza. Per la prima volta infatti è possibile effettuare un'analisi a medio e lungo termine sull'impatto della politica perché l'orizzonte temporale dei dati

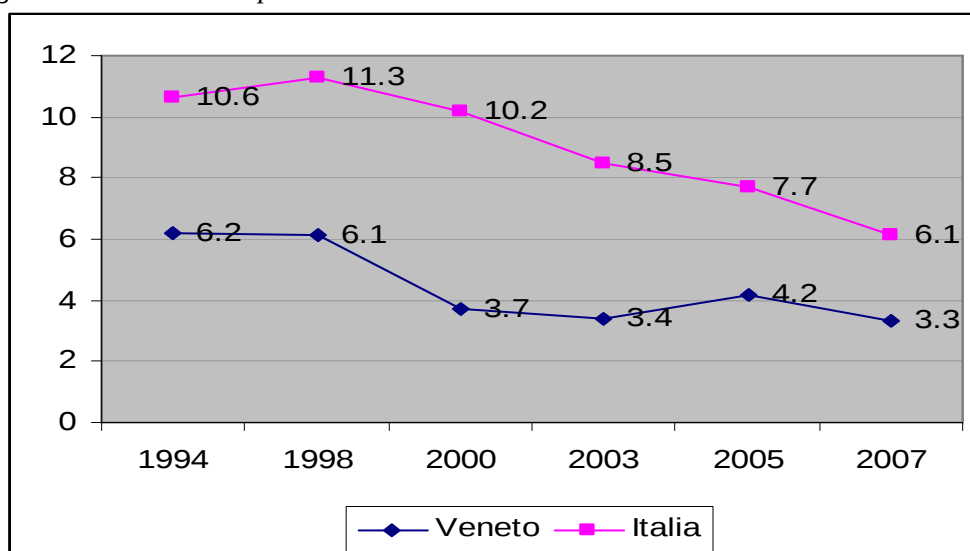
copre anche gli anni successivi al momento di uscita dalle Liste; inoltre è possibile allargare la valutazione dell'impatto della politica sui livelli dei salari eventualmente percepiti come indice della qualità del lavoro ottenuto a seguito della permanenza nelle Liste. Infine i dati a disposizione permettono di effettuare un'analisi di validazione degli assunti sui quali poggia la strategia di valutazione utile per saggiare la correttezza dei risultati stimati. Nel nostro studio, in particolare, considereremo il livello dei salari a tre anni dall'entrata nelle Liste come variabile di interesse ed utilizzeremo l'informazione relativa alla storia lavorativa passata per verificare che i gruppi di confronto utilizzati per l'analisi siano il più possibile simili al gruppo di trattamento al fine di ottenere una valutazione corretta dell'impatto.

## **5.2 Il campione di riferimento**

Anche se sono utilizzabili tutti i dati dal 1991, anno in cui la politica è stata attivata, restringiamo il campione di interesse ai lavoratori iscritti alle Liste di Mobilità della Regione Veneto dal 1995 fino al 1998. In questi anni è possibile trovare dati di qualità superiore grazie ad un miglioramento nel processo di raccolta dei dati. Inoltre nel 1994 la legislazione che regola la politica ha subito una modifica, attraverso la legge 451/1994 si sono proibite forme di frode che prendono il nome di "collective hires", le quali consistevano nella riassunzione di gruppo dei lavoratori licenziati nel medesimo giorno dalla stessa impresa allo scopo di poter usufruire dei benefici della mobilità (vedi Brunello e Miniaci 1997). L'orizzonte temporale considerato si ferma al 1998 perché in questo modo, per tutti i soggetti del campione, è possibile rilevare i valori del salario dopo tre anni (come già specificato infatti l'abbinamento è stato realizzato con la base di dati dell'Inps che copre le informazioni sui salari fino al 2001).

Dal punto di vista occupazionale il Veneto presenta un quadro largamente differente e positivo se rapportato al profilo medio italiano. Il tasso di occupazione nella popolazione tra i 15 e i 64 anni della regione è sempre superiore al tasso di occupazione a livello nazionale (nel 2007 il dato regionale è di 65.8% contro il dato nazionale di 58.7%). La disoccupazione per contro è un problema di entità inferiore rispetto alla media nazionale, nel 2007 il tasso di disoccupazione in Veneto è di poco superiore alla metà del tasso di disoccupazione italiano (3.3% contro 6.1%).

Fig 5.1 Tasso di disoccupazione<sup>33</sup> annuale in Veneto e in Italia



Le peculiarità che caratterizzano il mercato del lavoro della regione Veneto non ci permettono di estendere i risultati ottenuti in questo studio a tutta la popolazione italiana che usufruisce della mobilità.

Dalla popolazione iniziale di 29540 soggetti, l'abbinamento tra Netlabor e Inps porta ad una iniziale esclusione del solo 0.7% della popolazione totale. Tra i soggetti rimanenti consideriamo solo coloro per i quali l'entrata in lista coincide con un episodio di disoccupazione registrato nell'archivio Inps. Eliminati infine

<sup>33</sup> Percentuale sulla popolazione 15-64 anni

gli individui per i quali si verificano incongruenze relative alle storie lavorative o per problemi di dati mancanti, il campione finale conta 23.644 lavoratori, corrispondenti all'80% della popolazione iniziale di tutti i soggetti iscritti alle Liste nel periodo considerato.

Il campione è composto prevalentemente da donne che rappresentano il 59.5% del campione totale, mentre sono più numerosi i soggetti che provengono da piccole imprese, rappresentativi del 61.8% dei lavoratori totali, rispetto ai soggetti che possono godere dell'indennità di mobilità sia tra gli uomini (50.8%) sia tra le donne (69.3%). La prima fascia d'età (soggetti con età inferiore a 40 anni) è la più numerosa sia per genere sia per ammissibilità ai benefici monetari e comprende più di due terzi di tutto il campione, con percentuali entro i singoli gruppi che variano dal 45.2% per gli uomini con indennità fino a raggiungere l'81.1% tra le donne senza indennità. L'ultima fascia d'età al contrario ha una numerosità molto bassa.

*Tabella 5.1. Lavoratori iscritti alle liste del Veneto dal 1995 al 1998 per genere, ammissibilità all'indennità di mobilità e gruppo di età.*

| Età    | Uomini                    |      |                             |      | Donne                     |      |                             |      | Totale |      |
|--------|---------------------------|------|-----------------------------|------|---------------------------|------|-----------------------------|------|--------|------|
|        | Con indennità di mobilità |      | Senza indennità di mobilità |      | Con indennità di mobilità |      | Senza indennità di mobilità |      |        |      |
|        | N                         | %    | N                           | %    | N                         | %    | N                           | %    | N      | %    |
| 14-39  | 2127                      | 45.2 | 2929                        | 60.2 | 2808                      | 65.1 | 7921                        | 81.1 | 15785  | 66.7 |
| 40-49  | 1392                      | 29.6 | 1185                        | 24.4 | 1037                      | 24.1 | 1366                        | 14   | 4980   | 21.1 |
| +50    | 1183                      | 25.2 | 751                         | 15.4 | 465                       | 10.8 | 480                         | 4.9  | 2879   | 12.2 |
| Totale | 4702                      | 100  | 4865                        | 100  | 4310                      | 100  | 9767                        | 100  | 23644  | 100  |

In questo studio le analisi verranno condotte separatamente per genere e distintamente per i due sottoprogrammi. Saranno quindi effettuate quattro diverse analisi per i quattro sottogruppi così creati, mentre non sarà preso in considerazione alcun confronto tra uomini e donne oppure tra soggetti che beneficino di sostegno monetario e soggetti che non ne hanno diritto. L'unico

confronto preso in considerazione sarà quello relativo alle diverse durate di fruizione dei benefici che si differenziano per età dei soggetti beneficiari.

Le variabili contenute nel dataset, oltre al codice identificativo di ogni singola unità, sono

- Anno di ingresso in mobilità
- Genere
- Età
- Indicatore della percezione o meno dell'indennità di mobilità
- Il salario nei 72 mesi precedenti all'entrata in lista
- Il salario nei 36 mesi successivi all'entrata in lista

Il livello del salario relativo ad ogni mese assume valore pari a zero nel caso in cui il soggetto risulti disoccupato oppure abbia lavorato meno di una settimana durante il mese considerato. Nel caso in cui il soggetto risulti occupato, il valore mensile del salario è calcolato come il salario indicato negli archivi Inps (dove è indicato esclusivamente il valore del salario annuale) diviso per il numero di settimane di lavoro effettivamente svolto per ogni singolo evento lavorativo.



## **Capitolo 6**

### **La stima dell'impatto dell'anno aggiuntivo**

L'analisi che sviluppiamo in questo studio è una valutazione dell'effetto differenziale di un anno aggiuntivo di ammissibilità alle liste di mobilità sulle successive retribuzioni. Anche se questa politica è stata oggetto di numerosi studi, a causa dell'utilizzo solo recente dei dati sui salari, sono ancora limitate le analisi che si concentrano su questo aspetto. Le informazioni sui salari vengono utilizzate con due differenti finalità, non solo per valutare gli effetti della politica ma anche per verificare le assunzioni sulle quali si basa la correttezza della stima dell'impatto.

La valutazione viene effettuata in relazione al livello dei salari a tre anni dall'iscrizione alle liste, senza distinzioni per tipologia di contratto; l'effetto è identificato dalla differenza tra i salari dei lavoratori che hanno potuto usufruire del trattamento per due anni rispetto ai salari dei lavoratori che sono potuti rimanere in mobilità per un periodo massimo di un anno. I gruppi soggetti del confronto sono dunque i lavoratori di età compresa tra i 14 e i 39 anni che possono rimanere in lista al massimo un anno e i soggetti tra i 40 e i 49 anni con un limite massimo di permanenza in lista pari a due anni. Non saranno presi in considerazione i lavoratori che appartengono alla terza fascia di età per i quali la

presenza del programma *mobilità lunga*, utilizzato come ponte verso la pensione, ha degli effetti molto alti sulla probabilità di occupazione successiva all'entrata in lista.

Poiché palesemente i salari sono misurabili solo per coloro che hanno un'occupazione al momento della rilevazione, i soggetti sui quali verrà effettuata l'analisi saranno i soli lavoratori occupati a tre anni dall'iscrizione alle liste. Tra le prime due fasce d'età risultano occupati il 65.7% delle donne e il 77.9% degli uomini per un totale di 14581 lavoratori (70.2% del campione totale di lavoratori con età inferiore a 49 anni). La proporzione degli occupati risulta maggiore per gli uomini, sia in presenza sia in assenza di indennità di mobilità, rispetto alle donne entro tutte le classi d'età.

*Tabella 6.1. Uomini tra i 14 e i 49 anni iscritti alle liste di mobilità nella regione Veneto dal 1995 al 1998 secondo l'ammissibilità ai benefici monetari, la classe di età e la situazione occupazionale*

|             | Uomini con<br>indennità di<br>mobilità<br><40 anni |       | Uomini con<br>indennità di<br>mobilità<br>40-49 anni |       | Uomini senza<br>indennità di<br>mobilità<br><40 anni |       | Uomini senza<br>indennità di<br>mobilità<br>40-49 anni |       | Totale |       |
|-------------|----------------------------------------------------|-------|------------------------------------------------------|-------|------------------------------------------------------|-------|--------------------------------------------------------|-------|--------|-------|
| Occupati    | 1737                                               | 81.6% | 1024                                                 | 73.5% | 2251                                                 | 76.8% | 936                                                    | 78.9% | 5948   | 77.9% |
| Disoccupati | 390                                                | 18.4% | 368                                                  | 26.5% | 678                                                  | 23.2% | 249                                                    | 21.1% | 685    | 22.1% |
| Totale      | 2127                                               | 100%  | 1392                                                 | 100%  | 2929                                                 | 100%  | 1185                                                   | 100%  | 7633   | 100%  |

*Tabella 6.2. Donne tra i 14 e i 49 anni iscritte alle liste di mobilità nella regione Veneto dal 1995 al 1998 secondo l'ammissibilità ai benefici monetari, la classe di età e la situazione occupazionale*

|             | Donne con<br>indennità di<br>mobilità<br><40 anni |       | Donne con<br>indennità di<br>mobilità<br>40-49 anni |       | Donne senza<br>indennità di<br>mobilità<br><40 anni |       | Donne senza<br>indennità di<br>mobilità<br>40-49 anni |       | Totale |       |
|-------------|---------------------------------------------------|-------|-----------------------------------------------------|-------|-----------------------------------------------------|-------|-------------------------------------------------------|-------|--------|-------|
| Occupati    | 1878                                              | 66.8% | 594                                                 | 57.2% | 5323                                                | 67.2% | 838                                                   | 61.3% | 8633   | 65.7% |
| Disoccupati | 930                                               | 33.2% | 443                                                 | 42.8% | 2598                                                | 32.8% | 528                                                   | 38.7% | 4499   | 34.3% |
| Totale      | 2808                                              | 100%  | 1037                                                | 100%  | 7921                                                | 100%  | 1366                                                  | 100%  | 13132  | 100%  |

Il capitolo è organizzato come segue. Nel paragrafo 6.1 verrà presentato il disegno di valutazione nel quale si definiranno le variabili utilizzate, saranno descritte le possibili fonti di distorsione e il modo con il quale il *regression discontinuity design* e la stratificazione principale possono essere applicate per trarre delle conclusioni anche in presenza di problemi di selezione non casuale. Nel paragrafo 6.2 verranno presentate le stime dell'effetto di un anno aggiuntivo di permanenza nelle liste sul livello dei salari divise per i quattro sottogruppi della popolazione classificati per genere e ammissibilità ai benefici monetari.

## 6.1 Il disegno di valutazione

Consideriamo come trattamento l'anno aggiuntivo di permanenza in lista a cui hanno diritto solo i lavoratori con età superiore a 40 anni<sup>34</sup>. Sfruttando il disegno della politica, il trattamento è definito come segue

$$Z_i = \begin{cases} 0 & \text{se } 14 \leq \text{età}_i < 40 \\ 1 & \text{se } 40 \leq \text{età}_i < 50 \end{cases}$$

La variabile intermedia, che determina l'osservabilità o meno della variabile risultato, è un indicatore dello stato occupazionale del lavoratore

$$S_i = \begin{cases} 0 & \text{se il soggetto } i\text{-esimo è disoccupato a tre anni dall'iscrizione} \\ 1 & \text{se il soggetto } i\text{-esimo è occupato a tre anni dall'iscrizione} \end{cases}$$

La variabile oggetto dell'analisi è il salario a tre anni dall'entrata in lista del lavoratore. Consideriamo un intervallo temporale di tre anni in quanto è il valore più lontano nel tempo dal momento dell'iscrizione al programma per il quale

---

<sup>34</sup> Come già specificato ci limitiamo a studiare i lavoratori tra i 14 e i 49 anni, escludendo i lavoratori con età superiore a 50 anni che hanno diritto a tre anni di iscrizione alle liste.

abbiamo a disposizione l'informazione sui salari per l'intero campione. Se il lavoratore ha età superiore a 40 anni riceve il trattamento e risultano osservabili i suoi risultati potenziali  $Y_i^T, S_i^T$ , in caso contrario risultano osservabili i suoi risultati potenziali  $Y_i^{NT}, S_i^{NT}$ . Infine indichiamo con  $X$  il vettore di caratteristiche osservabili identificate da genere e dal diritto di percepire l'indennità di mobilità che, come ricordiamo, è prevista solo per i lavoratori licenziati da imprese con più di 15 addetti.

La valutazione d'impatto di questa politica è soggetta a diversi problemi di selezione non casuale, sia per il modo in cui è applicata, sia per la tipologia di variabile obiettivo che prendiamo in considerazione. La strategia di valutazione dovrà quindi tenere conto dei diversi problemi al fine di creare due gruppi, rispettivamente gruppo di esposti e gruppo di non esposti alla politica, confrontabili tra loro, per i quali calcolare l'impatto.

### **6.1.1 Le fonti di distorsione**

La prima fonte di distorsione è costituita dal processo di assegnazione al trattamento che risulta essere una funzione deterministica dell'età dei soggetti nel momento in cui vengono licenziati. L'età è una caratteristica rilevante per lo studio che stiamo affrontando in quanto ha degli effetti sia sulla probabilità di rioccupazione sia sul livello dei salari. E' plausibile infatti che lavoratori più giovani siano preferibili rispetto a lavoratori più anziani da parte delle imprese e che quindi i primi abbiano una probabilità maggiore di essere reinseriti nel mercato del lavoro. Il salario d'altra parte presenta spesso una correlazione positiva con l'età, quindi è probabile rilevare un salario medio superiore per i lavoratori che appartengono alla classe d'età più alta. Il semplice confronto osservato nell'intero campione rifletterà in parte l'effetto della politica ed in parte l'effetto dell'età. Il secondo problema di distorsione è rappresentato dall'osservabilità dei salari per i soli individui che risultano occupati al momento

della rilevazione; qualora i soggetti occupati abbiano caratteristiche rilevanti differenti rispetto al gruppo dal quale provengono, le stime risulteranno distorte poiché relative a campioni non confrontabili tra loro. Questo problema di selezione non casuale rientra nell'ambito del 'censoring by death' dove un particolare valore della variabile intermedia, in questo caso lo stato di disoccupazione, preclude la possibilità di osservare il valore della variabile obiettivo, il salario. Non sono invece presenti episodi di noncompliance in quanto l'assegnazione al trattamento è vincolata all'età e quindi nessun lavoratore può decidere di ricevere un trattamento diverso da quello a cui è assegnato dalla politica mentre il dataset utilizzato nell'analisi non presenta dati mancanti relativi a stato occupazionale e salari.

Nello studio sono condotte analisi separate per i quattro sottogruppi di individui classificati per genere e per tipologia di benefici ricevuti al fine di controllare per le altre fonti di eterogeneità tra i soggetti. Tutte le assunzioni e le analisi descritte nel corso del capitolo saranno effettuate entro ciascuno dei quattro sottogruppi; i metodi proposti hanno l'obiettivo di annullare le differenze sistematiche che persistono anche successivamente al condizionamento al vettore di caratteristiche osservabili  $X$ .

### **6.1.2 Il controllo dell'effetto età tramite *regression discontinuity design***

Al fine di controllare l'effetto età, come suggerito da precedenti analisi di valutazione della mobilità, utilizziamo il *regression discontinuity design*. Il metodo è applicabile in questo studio a causa del disegno particolare della politica per la quale la durata massima di permanenza in lista passa da uno a due anni se il lavoratore ha un'età superiore a 40 anni al momento dell'iscrizione. L'utilizzo dello *sharp rdd* si basa sull'assunzione di indipendenza condizionale secondo la quale l'assegnazione al trattamento non ha nessuna relazione con i risultati potenziali per i soggetti in prossimità del valore soglia di 40 anni.

$$Y^T, Y^{NT} \perp Z \mid età = 40 \quad (14)$$

Ammettendo che l'assunzione sia valida, è possibile annullare l'effetto età e ricondurre l'esercizio di valutazione al caso di assegnazione casuale considerando solo i soggetti attorno alla soglia.

La restrizione di identificazione poggia sull'assunto che, in assenza della politica, i soggetti con età di poco inferiore e di poco superiore a 40 anni avrebbero avuto lo stesso salario medio. I nostri dati potrebbero violare questa assunzione principalmente per due motivi. Possiamo anzitutto pensare che i licenziamenti non avvengano in maniera casuale ma che imprese e sindacati prendano degli accordi in base alle prospettive future e ai benefici di cui i lavoratori poi usufruiranno nel momento della disoccupazione. Non sarebbe improbabile quindi che i lavoratori appena al di sopra della soglia mostrino abilità o competenze minori rispetto ai lavoratori che si trovano appena al di sotto della soglia in quanto, in un momento successivo, la minore abilità verrebbe compensata dalla fruizione di benefici per un periodo più lungo e da maggiori incentivi ad una futura assunzione. Un'altra possibile causa di violazione della restrizione potrebbe essere il processo di abbinamento tra archivi e dati dell'Inps qualora fosse caratterizzato da un tasso di successo diverso per le due classi di età in maniera non casuale rispetto a caratteristiche rilevanti dei soggetti. Per testare la validità dell'assunzione Paggiaro, Rettore e Trivellato (2008) hanno confrontato i tassi di occupazione e i valori medi dei salari per i lavoratori tre anni prima dell'entrata in lista. Poiché la politica non può necessariamente avere un effetto causale sulla storia lavorativa passata dei soggetti, qualsiasi discontinuità sulla soglia può essere attribuita alla diversa composizione del campione per caratteristiche rilevanti in relazione allo status occupazionale e al salario. Il confronto è stato realizzato applicando un *local linear regression*<sup>35</sup> per stimare i

---

<sup>35</sup> Per stimare i polinomi locali sono stati usati pesi che seguono una funzione tricubica (Cleveland, 1979) mentre l'ampiezza dell'intervallo è stata scelta tramite una versione del criterio AIC (Hurvich e Simonoff, 1998)

valori attesi condizionati per la probabilità di occupazione e per il salario medio. Dall'analisi non si riscontrano discontinuità, i (futuri) gruppi di esposti e controllo sono identici per le caratteristiche prese in considerazione e qualsiasi confronto tra loro può essere ricondotto ad un effetto della politica attuata.

Dopo aver verificato la plausibilità delle assunzioni sulle quali poggia il *rdd*, procediamo restringendo il campione ai soli soggetti con età vicino alla soglia dei 40 anni; in particolare per ogni gruppo a destra e a sinistra della soglia considero due intervalli per l'età di diversa ampiezza, pari rispettivamente ad uno o due anni. I gruppi di controllo e trattamento che considereremo nel seguito sono rispettivamente

| Controllo                                 | Trattamento                               |
|-------------------------------------------|-------------------------------------------|
| $\{i : Z_i = 0, X_i, età_i = 39\}$        | $\{i : Z_i = 1, X_i, età_i = 40\}$        |
| $\{i : Z_i = 0, X_i, età_i = \{38,39\}\}$ | $\{i : Z_i = 1, X_i, età_i = \{40,41\}\}$ |

Non consideriamo intervalli di ampiezza maggiore in quanto all'aumentare delle differenze di età tra i soggetti inclusi nell'analisi, aumenta il rischio che l'assunzione di indipendenza condizionata non sia valida<sup>36</sup>. Verranno successivamente realizzati entrambi i confronti, in modo da poter verificare la robustezza dei risultati all'aumentare dell'ampiezza dell'intervallo per l'età considerato.

---

<sup>36</sup> Le analisi hanno verificato che, nel momento in cui vengono inclusi anche i soggetti con età pari a 37 e 42 anni il metodo della stratificazione principale non è più applicabile in quanto l'assunzione di assegnazione casuale del trattamento non risulta più valida quando si considera un intervallo troppo ampio per l'età

### 6.1.3 Identificazione degli strati principali e degli effetti causali

Anche dopo aver eliminato le differenze sistematiche legate all'età, il confronto osservabile sulla soglia tra i salari dei soggetti ammissibili o meno al secondo anno di permanenza in lista è ancora affetto da distorsione a causa della censura della variabile risultato<sup>37</sup>. Il livello del salario  $Y_i$  per il soggetto  $i$ -esimo è osservabile solo nel caso in cui si verifichi la condizione  $S_i^{obs} = 1$ ; in caso contrario il soggetto risulta disoccupato e  $Y_i$  non è osservabile. Nello studio proposto è possibile osservare i salari medi dei due gruppi di soggetti, controllo e trattamento, che risultano occupati a tre anni dall'ingresso nelle liste

$$\{i : Z_i = 0, S_i = 1\} \text{ e } \{i : Z_i = 1, S_i = 1\} \quad (15)$$

Il confronto tra il salario medio dei due gruppi di lavoratori così definiti<sup>38</sup> non identifica l'effetto causale in quanto la variabile intermedia che definisce lo status occupazionale verosimilmente è il risultato di un effetto sia di caratteristiche dei soggetti sia del trattamento. Un'azienda, ad esempio, può scegliere un lavoratore dalle liste da un lato per le sue competenze personali, dall'altro per i benefici dei quali può godere, in modo diverso per ogni soggetto. Per ognuno di questi è impossibile distinguere in che misura i due aspetti abbiano determinato il rientro nel mercato del lavoro. Di conseguenza è possibile che i due gruppi di occupati, ammissibili rispettivamente ad uno o due anni nelle liste, presentino delle differenze che si sarebbero osservate anche in assenza del trattamento. Una possibile critica potrebbe essere mossa a partire dai risultati di Paggiaro et al. (2008) che hanno dimostrato che l'anno aggiuntivo di permanenza in lista non ha alcun effetto sulla probabilità di occupazione a tre anni per i lavoratori sulla soglia

---

<sup>37</sup> Si tratta del problema di selezione endogena non casuale, *censoring by death*, ampiamente descritto nella sezione 4.3 di questa tesi

<sup>38</sup> Ricordiamo che a questo punto del disegno di valutazione tutte le considerazioni sono effettuate solo per i soggetti con età vicino alla soglia condizionatamente a genere e ammissibilità ai benefici monetari.

di 40 anni. Questo risultato tuttavia non implica che vi sia indipendenza tra trattamento e stato occupazionale dei singoli soggetti in quanto è possibile che si mantenga lo stesso tasso di occupazione attorno alla soglia, ma che i soggetti che costituiscono il gruppo di occupati tra il gruppo di controllo abbiano caratteristiche diverse dai soggetti trattati occupati.

La distorsione in questo caso è causata dal processo che porta gli individui che appartengono allo stesso gruppo ad essere occupati o disoccupati, il quale non si realizza in maniera casuale ma è correlato con caratteristiche che possono avere un'influenza sulla qualità del lavoro. E' possibile pensare ad esempio che tra i lavoratori che percepiscono l'indennità di mobilità, a parità di genere ed età, le imprese decidano di assumere tra i soggetti che possono usufruire per un anno della politica e dai quali quindi possono trarre minori benefici, solo i lavoratori che presentano abilità o competenze maggiori, ai quali verrà offerto un lavoro ben retribuito. Al contrario l'assunzione di soggetti che hanno diritto anche al secondo anno di trattamento comporta un trasferimento maggiore di benefici alle imprese che quindi saranno incentivate ad assumere tra questi lavoratori anche coloro meno competenti o specializzati ai quali sarà offerto un lavoro con una retribuzione minore. In questo modo è verosimile ipotizzare la presenza di differenze sistematiche tra il gruppo di soggetti trattati e occupati a tre anni dall'entrata in lista e l'intero gruppo di soggetti trattati del quale fanno parte; lo stesso si può pensare per il gruppo di non trattati.

$$E[Y_i^T | S_i^T = 1] \neq E[Y_i^T] \quad (16)$$

$$E[Y_i^{NT} | S_i^{NT} = 1] \neq E[Y_i^{NT}] \quad (17)$$

Anche in caso di assegnazione casuale del trattamento, i soggetti occupati dopo aver avuto il diritto di usufruire di due anni della politica saranno sistematicamente diversi dai soggetti occupati ammessi alle liste per un periodo

massimo di un anno in quanto la formazione dei gruppi avviene in maniera non casuale da gruppi casuali.

Per ottenere una stima corretta dell'impatto in presenza di un problema di selezione endogena utilizziamo la stratificazione principale. Prima di procedere all'applicazione del metodo, verifichiamo se nello studio di valutazione che stiamo effettuando possono essere valide le assunzioni sulle quali poggia tale approccio.

**Assunzione 1. SUTVA:** questa ipotesi può essere considerata valida in quanto solo una piccola percentuale dei disoccupati italiani entra ogni anno nelle liste di mobilità; in particolare in Veneto i lavoratori iscritti rappresentano meno del 10% del livello medio di disoccupati della regione (Paggiaro et al. 2008). Possiamo quindi ipotizzare che non ci sia interazione tra i soggetti che partecipano e che l'implementazione della politica non comporti conseguenze a livello macroeconomico.

**Assunzione 2. Assegnazione casuale:** l'assegnazione al trattamento per i lavoratori sulla soglia (così come definiti nel paragrafo precedente), per la validità dell'indipendenza condizionale alla base del *rdd*, può essere considerata alla pari di un assegnazione casuale.

La validità delle assunzioni ci garantisce che la stratificazione rispetto all'indicatore della situazione occupazionale ci conduce ad una stima corretta dell'impatto sui salari per lo strato definito dai lavoratori che risulterebbero sempre occupati a prescindere dalla durata di ammissibilità nelle liste. Definiamo in primo luogo i quattro strati definiti dalla coppia di valori potenziali  $S_i^P = (S_i^{NT}, S_i^T)$  rappresentata dallo status occupazionale che ciascun soggetto sperimenterebbe nel caso in cui sia possa rimanere iscritto rispettivamente fino ad un massimo di uno o due anni nelle liste. Definiamo i quattro strati nel seguente modo; seguendo la terminologia utilizzata da Zhang e Rubin (2003), per ognuno

indichiamo quali valori potenziali della variabile risultato sono definiti in  $\mathfrak{R}^+$  e quali non lo sono.

- $EE = \{i : S_i(NT) = S_i(T) = 1\}$ , soggetti che risulterebbero occupati sia se esposti al trattamento sia se non esposti<sup>39</sup>, per i quali i valori  $Y_{i,EE}^{NT}, Y_{i,EE}^T$  sono definiti in  $\mathfrak{R}^+$ .
- $UU = \{i : S_i(NT) = S_i(T) = 0\}$ , soggetti che non sarebbero mai occupati a prescindere dal trattamento al quale sono assegnati, per i quali il valore del salario non è mai definito in  $\mathfrak{R}^+$ .
- $EU = \{i : S_i(NT) = 1, S_i(T) = 0\}$ , soggetti che risulterebbero occupati solo se non potessero estendere l'iscrizione in lista al secondo anno per i quali si osserva solo il valore  $Y_{i,EU}^{NT}$  mentre  $Y_{i,EU}^T$  non è definito in  $\mathfrak{R}^+$ .
- $UE = \{i : S_i(NT) = 0, S_i(T) = 1\}$ , soggetti che risulterebbero occupati solo se ammissibili al secondo anno di permanenza in lista per i quali si osserva solo il valore  $Y_{i,UE}^T$  mentre  $Y_{i,UE}^{NT}$  non è definito in  $\mathfrak{R}^+$ .

definiamo infine le proporzioni degli strati nell'intera popolazione come  $\pi_{EE}$ ,  $\pi_{UE}$ ,  $\pi_{EU}$ ,  $\pi_{UU} = 1 - \pi_{EE} - \pi_{UE} - \pi_{EU}$ .

Poiché il salario è definito solo per i soggetti occupati, l'unico strato per il quale sono definiti entrambi i valori potenziali è lo strato  $S_i^P = EE$  per il quale è possibile calcolare l'effetto causale dell'anno aggiuntivo di permanenza in lista. Utilizzando la notazione

$$E[Y_i(T) | S_i^P = EE] = \bar{Y}_{EE}^T$$

$$E[Y_i(NT) | S_i^P = EE] = \bar{Y}_{EE}^{NT}$$

<sup>39</sup> Nel seguito ci riferiremo a questo gruppo di soggetti anche con il termine 'sempre occupati'

l'effetto causale medio per i soggetti 'sempre occupati' si calcola come la differenza tra i valori medi tra i soggetti trattati e non trattati che appartengono al medesimo strato

$$SACE = \bar{Y}_{EE}^T - \bar{Y}_{EE}^{NT} \quad (18)$$

L'effetto non è osservabile direttamente, la variabile che indica l'appartenenza di ogni soggetto agli strati è una variabile latente. Ciò che possiamo osservare sono i gruppi definiti dall'ammissibilità o meno all'anno aggiuntivo in lista e dalla situazione occupazionale. L'idea da cui parte il metodo utilizzato è che, in base al modo con il quale sono definiti, è possibile capire quale coppia di strati principali compone ogni gruppo osservato.

*Tabella 6.3 Composizione dei gruppi osservati per strati principali*

| <b>Gruppi osservati: G(Z,S)</b>      | <b>Variabile risultato</b><br>$Y^{obs}$ | <b>Strati principali</b> |
|--------------------------------------|-----------------------------------------|--------------------------|
| $G(T,E) = \{i : Z_i = 1, S_i = 1\}$  | $Y^T$                                   | EE o UE                  |
| $G(T,U) = \{i : Z_i = 1, S_i = 0\}$  | *                                       | UU o EU                  |
| $G(NT,E) = \{i : Z_i = 0, S_i = 1\}$ | $Y^{NT}$                                | EE o EU                  |
| $G(NT,U) = \{i : Z_i = 0, S_i = 0\}$ | *                                       | UE o UU                  |

Per ogni gruppo per il quale la variabile obiettivo risulta definita,  $G(T,E)$  e  $G(NT,E)$ , si stima un modello di mistura che identifica le due componenti dalle quali è formato, ognuna delle quali identifica uno strato. Poiché la componente che si riferisce allo strato EE è presente in entrambi i gruppi, il passaggio successivo consiste nell'associare le componenti stimate al rispettivo strato principale al fine di individuare quali sono i parametri che identificano la distribuzione dei salari nello strato di interesse da utilizzare per il calcolo dell'effetto causale.

L'analisi si sviluppa a partire della teoria dei modelli di misture finite nell'ambito del quale stimiamo sia la probabilità di appartenere agli strati sia la distribuzione del salario entro ciascuno di questi tramite il metodo della massima verosimiglianza. Il metodo prevede prima di tutto di assegnare una distribuzione parametrica alla variabile risultato. Sulla base della letteratura econometrica possiamo ipotizzare che i logaritmi dei salari, che nel seguito denoteremo  $W$ , seguano una distribuzione normale. Definiamo le distribuzioni entro gli strati come segue

- $W_{i,EE}^T = \log(Y_{i,EE}^T) \sim N(\mu_{EE}^T; \sigma_{EE}^T)$
- $W_{i,UE}^T = \log(Y_{i,UE}^T) \sim N(\mu_{UE}; \sigma_{UE})$
- $W_{i,EE}^{NT} = \log(Y_{i,EE}^{NT}) \sim N(\mu_{EE}^{NT}; \sigma_{EE}^{NT})$
- $W_{i,EU}^{NT} = \log(Y_{i,EU}^{NT}) \sim N(\mu_{EU}; \sigma_{EU})$

Per lo strato EE i parametri variano in base al trattamento per consentire che i lavoratori che vi appartengono possano avere un salario diverso in relazione alla durata del periodo di ammissibilità (ossia che sia presente un effetto significativo della politica). Per i gruppi UE e EU i parametri sono definiti solo nel caso in cui i soggetti siano ammissibili o meno all'anno aggiuntivo di permanenza.

Prendiamo in considerazione singolarmente i gruppi per i quali osserviamo i salari. Il gruppo di soggetti che hanno diritto a un periodo massimo di permanenza nelle liste pari a due anni e che risultano occupati a tre anni dall'entrata in lista,  $G(T,E)$ , è costituito dai lavoratori che risulterebbero occupati a prescindere dalla durata massima di permanenza nelle liste e dai soggetti che risulterebbero occupati solo nel caso in cui potessero usufruire della politica per due anni mentre risulterebbero disoccupati altrimenti. Per ogni soggetto che appartiene a questo gruppo posso scrivere la densità del logaritmo del salario come la mistura di due normali univariate eteroschedastiche. Se definiamo il vettore dei parametri

$$\vartheta^T = (\beta^T, \mu_{EE}^T, \sigma_{EE}^T, \mu_{UE}, \sigma_{UE}) \quad (19)$$

$$\text{con } \beta^T = \frac{\Pr(i \in S_i^P = \{EE\})}{\Pr(i \in G(T, E))} = \frac{\pi_{EE}}{\pi_{EE} + \pi_{UE}} \quad (20)$$

la distribuzione della mistura dei salari e la verosimiglianza si possono scrivere come

$$f(W^T | Z, S, X; \vartheta^T) = \beta^T \cdot \phi(w_i^T; \mu_{EE}^T, \sigma_{EE}^T) + (1 - \beta^T) \cdot \phi(w_i^T; \mu_{UE}, \sigma_{UE}) \quad (21)$$

$$L(\vartheta^T | W, Z, S, X) = \prod_{i \in G(T, E)} \beta_i^T \cdot \phi(w_i; \mu_{EE}^T, \sigma_{EE}^T) + (1 - \beta_i^T) \cdot \phi(w_i; \mu_{UE}, \sigma_{UE}) \quad (22)$$

dove  $\phi(w_i; \mu, \sigma)$  rappresenta la densità di una normale univariata con media  $\mu$  e deviazione standard  $\sigma$

Successivamente consideriamo il gruppo  $G(NT, E)$ . I soggetti che hanno diritto ad una durata massima del periodo di permanenza nelle liste pari ad un anno e che risultano occupati possono essere ancora una volta dai lavoratori che risulterebbero occupati in entrambi i casi oppure lavoratori che risulterebbero occupati solo se potessero rimanere nelle liste per un periodo non superiore ad un anno mentre risulterebbero disoccupati se venisse concesso loro di rimanere un anno aggiuntivo. Il vettore dei parametri, la distribuzione dei salari e la verosimiglianza sono definiti come segue

$$\vartheta^{NT} = (\beta^{NT}, \mu_{EE}^{NT}, \sigma_{EE}^{NT}, \mu_{EU}, \sigma_{EU}) \quad (23)$$

$$\text{con } \beta^{NT} = \frac{\Pr(i \in S_i^P = \{EE\})}{\Pr(i \in G(NT, E))} = \frac{\pi_{EE}}{\pi_{EE} + \pi_{EU}} \quad (24)$$

$$f(W^{NT} | Z, S, X; \vartheta^{NT}) = \beta^{NT} \cdot \phi(w_i^{NT}; \mu_{EE}^{NT}, \sigma_{EE}^{NT}) + (1 - \beta^{NT}) \cdot \phi(w_i^{NT}; \mu_{EU}, \sigma_{EU}) \quad (25)$$

$$L(\vartheta^{NT} | W, Z, S, X) = \prod_{i \in G(NT, E)} \beta_i^{NT} \cdot \phi(w_i; \mu_{EE}^{NT}, \sigma_{EE}^{NT}) + (1 - \beta_i^{NT}) \cdot \phi(w_i; \mu_{EU}, \sigma_{EU}) \quad (26)$$

Come già descritto in precedenza, non è possibile osservare a quale componente appartenga ogni singola unità. Le stime di massima verosimiglianza, non possono essere calcolate in forma esplicita ma è necessario utilizzare procedure iterative; le più semplici sono l'algoritmo Expectation-Maximization (Dempster et al., 1977) e il metodo Newton-Raphson. Nello studio utilizziamo entrambe le procedure di stima, i risultati sono sempre robusti alle alternative.

Questi algoritmi in genere presentano problemi di stima nel caso in cui le componenti siano molto simili tra loro. Nelle analisi effettuate, anche se le distribuzioni in alcuni sottogruppi saranno caratterizzate da valori molto simili delle medie, non si verificano problemi di convergenza e si raggiunge sempre una stima dei parametri. I valori iniziali per le procedure iterative sono fissati in base ad un'analisi grafica dell'istogramma dei logaritmi dei salari per ogni gruppo.

Successivamente alla stima dei parametri è necessario associare ciascuna componente allo strato che rappresenta in modo da identificare i parametri dello strato di interesse EE sia in caso di assegnazione al trattamento sia in caso contrario. Questo studio di valutazione presenta un forte limite nell'applicazione del metodo della stratificazione principale in quanto non è mai possibile identificare i singoli strati. Consideriamo infatti il tasso di occupazione a tre anni dall'iscrizione nelle liste per i soggetti che hanno età superiore ed inferiore a 40 anni, pari rispettivamente a  $P_{T,E} = \pi_{EE} + \pi_{UE}$  e  $P_{NT,E} = \pi_{EE} + \pi_{EU}$ , identificabili per i soggetti sulla soglia mediante *rdd*. Nel nostro caso sappiamo dai risultati ottenuti in Paggiaro et al.(2008) che non ci sono effetti sui tassi di occupazione per nessuno dei sottogruppi considerati, di conseguenza si avrà necessariamente che la proporzione di lavoratori occupati solo se assegnati al trattamento è uguale alla proporzione di lavoratori occupati solo se non assegnati allo

stesso ( $\pi_{UE} = \pi_{EU}$ ). Senza ulteriori assunzioni non sarà mai possibile identificare quale sia la componente della mistura che si riferisce al gruppo EE e quale rispettivamente al secondo strato che costituisce ogni gruppo. In alternativa si potrebbero imporre dei vincoli sui parametri del modello, ad esempio si potrebbe ipotizzare che il valore medio del salario nel gruppo EE sia superiore al salario medio degli altri strati. Nel nostro studio non utilizziamo assunzioni aggiuntive; il procedimento è il seguente. Prima di tutto si stimano, separatamente per gruppo di controllo e trattamento, i parametri delle quattro componenti della mistura, associate ai tre strati EE, UE, EU. Poiché vale l'ipotesi di assegnazione casuale al trattamento, dal momento che il gruppo EE è presente in entrambi i gruppi di trattamento e controllo e in questo particolare caso gli altri due strati che formano la mistura comprendono la stessa proporzione di soggetti, il passo successivo consiste nell'abbinare ogni componente stimata a destra della soglia con una componente stimata a sinistra della soglia caratterizzata dalla stessa proporzione. In questo modo si formano due coppie di componenti, queste rappresentano le coppie  $\{EE, EE\}$ ,  $\{EU, UE\}$  che tuttavia continuiamo a non poter identificare. Infine si confrontano le medie delle distribuzioni degli elementi all'interno di ogni coppia, si avranno di conseguenza due valori stimati, tra i quali solo uno è l'effetto causale sullo strato EE.

## 6.2 Analisi e commento dei risultati

Dopo aver definito le caratteristiche fondamentali della politica, il set di dati e i metodi da utilizzare, passiamo ora all'analisi empirica dei dati. La valutazione viene realizzata separatamente per i quattro sottogruppi divisi per genere e diritto all'indennità di mobilità. Per ognuno di questi vengono considerati i soli soggetti occupati a tre anni dall'iscrizione in lista con età vicina alla soglia di 40 anni, in particolare i lavoratori con età inferiore e superiore alla soglia di uno o due anni.

Sui sottocampioni così costruiti si stima la distribuzione dei log-salari come mistura di due distribuzioni normali univariate eteroschedastiche e si confrontano le stime dei parametri delle componenti che presentano la stessa proporzione stimata entro i gruppi osservati.

Preliminarmente al confronto tra le stime, per ogni distribuzione stimata, viene svolta un'analisi della bontà di adattamento alla distribuzione empirica dei dati basata su procedure grafiche<sup>40</sup>. I valori empirici vengono confrontati con i quantili della distribuzione stimata, tramite il Q-Q plot. Una seconda procedura per il controllo empirico del modello è basata sul confronto tra stima della funzione di ripartizione empirica  $\hat{F}_n(w)$  e la funzione di ripartizione stimata  $F(w, \hat{\theta})$ . Poiché quest'ultima non è definita a partire da parametri noti ma dai valori stimati dal campione, è possibile individuare delle bande di confidenza non parametriche, sulla base della teoria legata al test di Kolmogorov-Smirnov<sup>41</sup>. L'intervallo di confidenza per la funzione di ripartizione teorica esatta  $F(w)$  è pari a

$$[F(w, \hat{\theta}) - d_{\alpha, n}, F(w, \hat{\theta}) + d_{\alpha, n}] \quad (27)$$

dove  $d_{\alpha, n} = \frac{1.358}{\sqrt{n}}$  è il valore critico approssimato per  $n > 50$ , ad un livello di significatività del 5%, derivato dalle tavole di Kolmogorov.

### 6.2.1 Uomini che percepiscono l'indennità di mobilità

Il primo gruppo sul quale vengono svolte le analisi è costituito dagli uomini che percepiscono l'indennità di mobilità per il periodo di permanenza nelle liste,

<sup>40</sup> Riportiamo tutti i grafici per l'analisi della bontà delle stime in Appendice

<sup>41</sup> Si tratta di un test *distribution free*, per la verifica dell'adattamento che confronta le funzioni di ripartizione empirica e teorica sotto l'ipotesi di uguaglianza e si basa sulla quantità  $D_n = \sup |\hat{F}_n(w) - F(w, \hat{\theta})|$ .

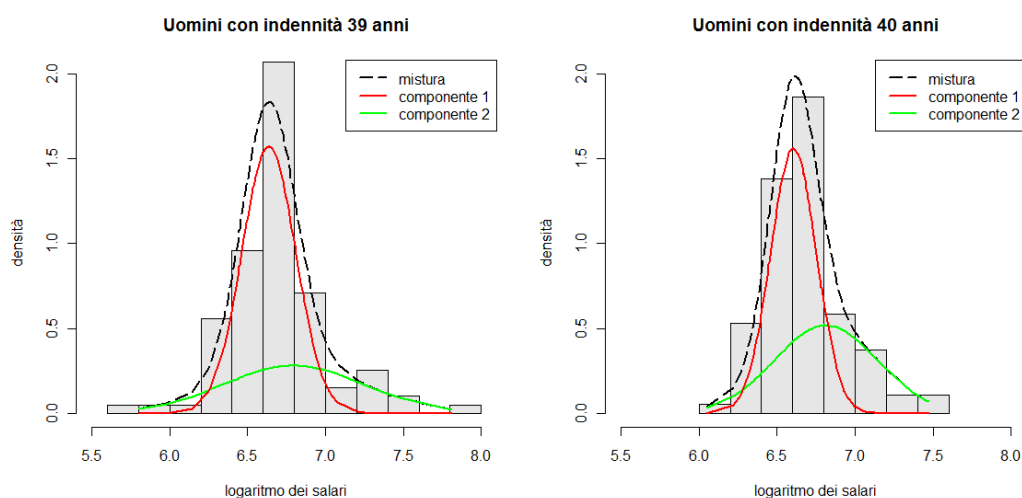
occupati a tre anni dall'ingresso in mobilità con età al momento dell'iscrizione che varia dai 38 ai 41 anni.

*Tabella 6.4 Uomini che percepiscono l'indennità di mobilità per età e situazione occupazionale tre anni dopo l'ingresso in lista.*

|             | 39 anni |       | 40 anni |       | 38-39 anni |       | 40-41 anni |       |
|-------------|---------|-------|---------|-------|------------|-------|------------|-------|
|             | N       | %     | N       | %     | N          | %     | N          | %     |
| Occupati    | 101     | 78.9% | 94      | 81.7% | 209        | 81.6% | 199        | 78.6% |
| Disoccupati | 27      | 21.1% | 21      | 18.3% | 47         | 18.4% | 253        | 21.4% |
| Totale      | 128     | 100 % | 115     | 100 % | 256        | 100 % | 1185       | 100 % |

Riportiamo in tabella le numerosità dei gruppi di controllo e trattamento che consideriamo nell'analisi. I soggetti occupati sui quali si svolge lo studio rappresentano circa l'80% di ogni singolo sottogruppo. Dal campione eliminiamo quattro osservazioni, due delle quali relative a lavoratori con età pari a 39 anni e due a soggetti con età pari a 41 anni per i quali è rilevato un salario inferiore a 300 euro mensili che si discosta ampiamente dai valori dei salari dei rispettivi gruppi. Sulle unità rimanenti stimiamo le componenti della mistura

*Figura 6.1. Distribuzione stimata dei salari a tre anni dall'iscrizione alle liste per gli uomini che percepiscono l'indennità di mobilità per età (39 e 40 anni). Stima della mistura e delle componenti.*



Inizialmente stimiamo la mistura considerando solo i soggetti con età pari a 39 anni per il gruppo di controllo ed età pari a 40 anni per il gruppo di trattamento. In tabella riportiamo le stime di massima verosimiglianza con i relativi standard error indicati tra parentesi. Il parametro  $\hat{\beta}$  rappresenta la probabilità che le unità del campione appartengano alla prima componente, definita dai parametri  $\hat{\mu}_1$  e  $\hat{\sigma}_1$ . La seconda componente è definita invece dai parametri  $\hat{\mu}_2$  e  $\hat{\sigma}_2$  ed ha una probabilità associata pari a  $1 - \hat{\beta}$ .

Tabella 6.5 Stime ML dei parametri che identificano le componenti della mistura per età (39,40 anni)

|                  | Età=39<br>$\vartheta^{NT}$ | Età=40<br>$\vartheta^T$ |
|------------------|----------------------------|-------------------------|
| $\hat{\beta}$    | 0.681<br>(0.128)           | 0.574<br>(0.179)        |
| $\hat{\mu}_1$    | 6.636<br>(0.027)           | 6.604<br>(0.032)        |
| $\hat{\mu}_2$    | 6.794<br>(0.104)           | 6.812<br>(0.092)        |
| $\hat{\sigma}_1$ | 0.173<br>(0.025)           | 0.146<br>(0.034)        |
| $\hat{\sigma}_2$ | 0.454<br>(0.078)           | 0.328<br>(0.046)        |

La distribuzione del logaritmo dei salari nel gruppo di controllo risulta composta per il 68% da una componente con distribuzione  $N(6.6, 0.17^2)$  e per il rimanente 32% da una  $N(6.8, 0.45^2)$ . Per quanto riguarda il gruppo di controllo, i soggetti che risulterebbero sempre occupati e i soggetti che troverebbero lavoro solo se ammissibili ad un anno nelle liste presentano lo stesso salario medio e differiscono tra loro solo per la variabilità dei dati. La distribuzione nel gruppo di trattamento si compone per un 57% di una  $N(6.6, 0.15^2)$  e per il 43% di una  $N(6.8, 0.33^2)$ ; i soggetti appartenenti ai due strati EE e UE hanno distribuzioni diverse sia in media sia in varianza. Prima di tutto è necessario abbinare le componenti stimate nel gruppo di

controllo (soggetti con 39 anni) alle componenti stimate nel gruppo di trattamento (soggetti con 40 anni) per i quali la proporzione è uguale. L'ipotesi di uguaglianza tra le stime delle proporzioni, come le successive ipotesi di uguaglianza tra medie e deviazioni standard, viene verificata tramite un test basato sulla distribuzione approssimata normale degli stimatori ML e sull'indipendenza tra gruppi di controllo e trattamento. In questo particolare caso tutte le componenti stimate risultano avere la medesima proporzione; in questo modo non solo non è possibile identificare gli strati ma non è possibile nemmeno effettuare un unico abbinamento tra le coppie che sappiamo avere la stessa proporzione. Ci limitiamo dunque a confrontare la media di ciascuna componente del gruppo di controllo con la media di ciascuna componente del gruppo di trattati. Per qualsiasi combinazione il test accetta sempre l'ipotesi nulla di uguaglianza dei parametri. Possiamo concludere che per questo sottocampione sembra non ci siano effetti sui salari medi dei lavoratori che appartengono allo strato EE.

Replichiamo il processo di stima includendo nei gruppi di controllo e trattamento utilizzati nella prima analisi i lavoratori con età pari rispettivamente a 38 e 41 anni.

*Figura 6.2. Distribuzione stimata dei salari a tre anni dall'iscrizione alle liste per gli uomini che percepiscono l'indennità di mobilità per età (38-39 e 40-41 anni). Stima della mistura e delle componenti.*

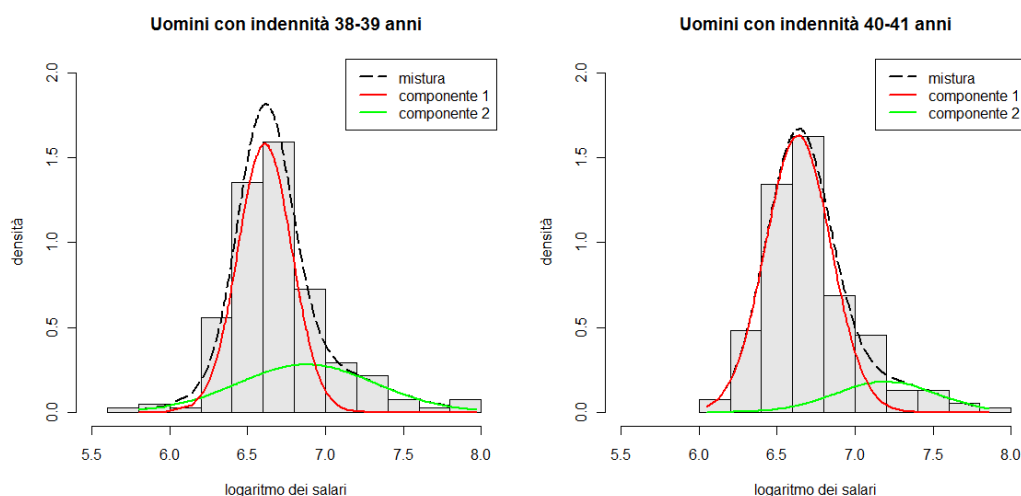


Tabella 6.6 Stime ML dei parametri che identificano le componenti della mistura per età (38-39,40-41 anni)

|                  | Età=38-39<br>$\vartheta^{NT}$ | Età=40-41<br>$\vartheta^T$ |
|------------------|-------------------------------|----------------------------|
| $\hat{\beta}$    | 0.687<br>(0.082)              | 0.853<br>(0.155)           |
| $\hat{\mu}_1$    | 6.611<br>(0.020)              | 6.636<br>(0.037)           |
| $\hat{\mu}_2$    | 6.879<br>(0.078)              | 7.198<br>(0.403)           |
| $\hat{\sigma}_1$ | 0.173<br>(0.017)              | 0.208<br>(0.024)           |
| $\hat{\sigma}_2$ | 0.444<br>(0.049)              | 0.323<br>(0.165)           |

A partire dal campione più numeroso ottengo delle stime differenti. Innanzitutto le medie degli strati che appartengono al gruppo di controllo sono diverse tra loro; a differenza del caso precedente, nel caso di non assegnazione al trattamento, i lavoratori che sarebbero sempre occupati hanno un livello medio dei salari diverso dai soggetti che sono occupati solo perché non sono ammessi a restare nelle liste un anno aggiuntivo. Lo stessa differenza non si riscontra, nel caso in cui siano ammessi nelle liste anche il secondo anno, tra i lavoratori che sarebbero sempre occupati e i lavoratori occupati solo se trattati; tale risultato è dovuto alla scarsa precisione nella stima della media della seconda componente per la bassa numerosità di soggetti che ne fanno parte. Anche i valori delle proporzioni cambiano in modo tale da rendere possibile un unico abbinamento delle componenti tra le quali poi effettuare il confronto. Poiché dal test di uguaglianza risulta che  $\hat{\beta}^T = \hat{\beta}^{NT}$ , dove  $\beta^T$  e  $\beta^{NT}$  sono le proporzioni della componente 1 sia nel gruppo di controllo sia nel gruppo di trattamento, si ha che

- se lo strato EE corrisponde alla componente 1 l'effetto causale è identificato dalla differenza  $\hat{\mu}_1^T - \hat{\mu}_1^{NT}$

- se lo strato EE è associato alla componente 2 l'effetto causale è identificato dalla differenza  $\hat{\mu}_2^T - \hat{\mu}_2^{NT}$

In entrambi i casi la differenza non è significativa. Le analisi effettuate considerando un intervallo più largo per l'età attorno alla soglia confermano che un anno aggiuntivo di permanenza in lista non ha nessun effetto sul livello medio dei salari degli uomini che hanno il diritto di percepire benefici monetari e che risulterebbero sempre occupati a prescindere dal trattamento al quale sono assegnati. E' interessante rilevare come anche le varianze risultano uguali per entrambe le coppie; questo significa che i salari dei soggetti che appartengono agli strati EU ed UE sono caratterizzati dalla stessa distribuzione.

### 6.2.2 Uomini senza indennità di mobilità

*Tabella 6.7 Uomini senza indennità di mobilità per età e situazione occupazionale tre anni dopo l'ingresso in lista.*

|             | 39 anni |       | 40 anni |       | 38-39 anni |       | 40-41 anni |       |
|-------------|---------|-------|---------|-------|------------|-------|------------|-------|
| Occupati    | 85      | 73.9% | 108     | 78.2% | 171        | 72.7% | 198        | 79.2% |
| Disoccupati | 30      | 26.1% | 30      | 21.8% | 64         | 27.3% | 52         | 20.8% |
| Totale      | 115     | 100%  | 138     | 100%  | 235        | 100%  | 250        | 100%  |

Tutti i sottogruppi definiti dall'età sono caratterizzati da un tasso di occupazione superiore al 70%. Eliminiamo dal campione due valori per il salario particolarmente elevati rispetto al campione di riferimento, associati a due lavoratori di 40 e 41 anni.

Come per il primo gruppo stimiamo i parametri della mistura considerando inizialmente un intervallo per l'età pari ad un anno dalla soglia.

Figura 6.3. Distribuzione stimata dei salari a tre anni dall'iscrizione alle liste per gli uomini che non percepiscono l'indennità di mobilità per età (39 e 40 anni). Stima della mistura e delle componenti.

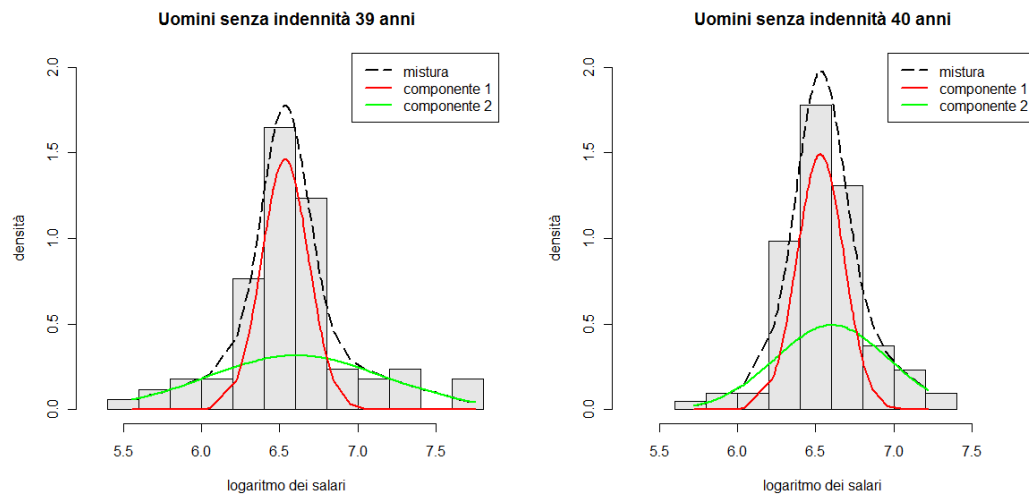


Tabella 6.8 Stime ML dei parametri che identificano le componenti della mistura per età (39,40anni)

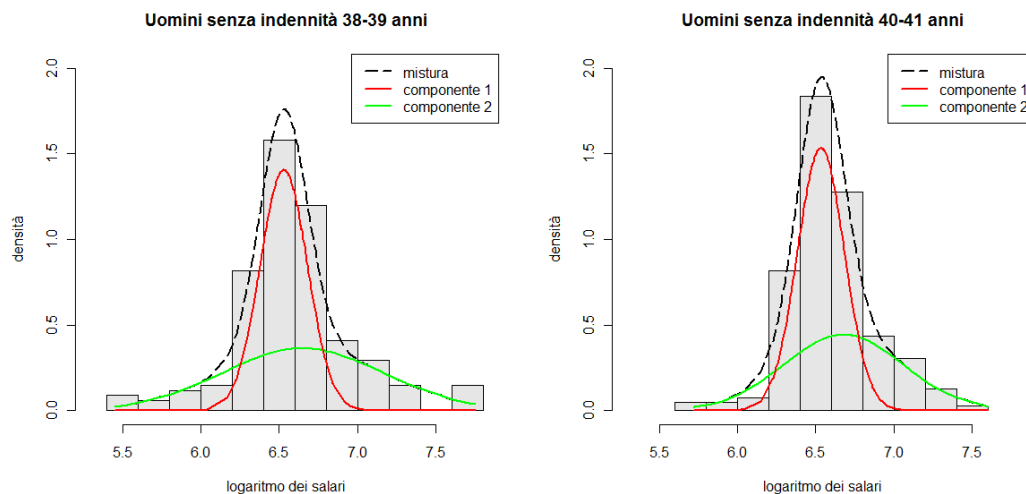
|                  | Età=39<br>$\vartheta^{NT}$ | Età=40<br>$\vartheta^T$ |
|------------------|----------------------------|-------------------------|
| $\hat{\beta}$    | 0.550<br>(0.104)           | 0.555<br>(0.155)        |
| $\hat{\mu}_1$    | 6.534<br>(0.029)           | 6.53<br>(0.030)         |
| $\hat{\mu}_2$    | 6.603<br>(0.095)           | 6.602<br>(0.059)        |
| $\hat{\sigma}_1$ | 0.150<br>(0.023)           | 0.148<br>(0.023)        |
| $\hat{\sigma}_2$ | 0.569<br>(0.081)           | 0.359<br>(0.056)        |

Per questo campione, sia in caso di assegnazione al trattamento sia in caso contrario, l'algoritmo stima una mistura di normali nella quale le componenti stimate per lo stesso gruppo hanno una distribuzione dei salari caratterizzata dalla stessa media e da varianze diverse. Le proporzioni di tutte le componenti sono uguali tra loro quindi non è possibile abbinare le coppie di parametri da confrontare; dal confronto di tutte le possibili combinazioni di coppie risulta che non vi sono differenze tra nessuna delle medie stimate mentre gli strati ovviamente differiscono solo per i valori delle varianze.

Anche senza la possibilità di identificare lo strato EE per il quale calcolare l'effetto casuale possiamo concludere che non esiste un effetto dell'anno aggiuntivo di permanenza in lista sui salari medi se ci condizioniamo ai soli lavoratori di 39 e 40 anni.

Verifichiamo che i risultati siano confermati anche dalle analisi sul campione allargato ai lavoratori con età compresa tra i 38 e i 41 anni.

*Figura 6.4. Distribuzione stimata dei salari a tre anni dall'iscrizione alle liste per gli uomini che non percepiscono l'indennità di mobilità per età (38-39 e 40-41 anni). Stima della mistura e delle componenti.*



*Tabella 6.9 Stime ML dei parametri che identificano le componenti della mistura per età (38-39,40-41 anni)*

|                  | Età=38-39        | Età=40-41        |
|------------------|------------------|------------------|
| $\hat{\beta}$    | 0.530<br>(0.083) | 0.576<br>(0.102) |
| $\hat{\mu}_1$    | 6.529<br>(0.022) | 6.533<br>(0.021) |
| $\hat{\mu}_2$    | 6.648<br>(0.061) | 6.677<br>(0.052) |
| $\hat{\sigma}_1$ | 0.150<br>(0.019) | 0.150<br>(0.017) |
| $\hat{\sigma}_2$ | 0.516<br>(0.052) | 0.381<br>(0.041) |

Come effettuato nel campione ristretto verifico l'uguaglianza delle proporzioni. Ancora una volta non è possibile individuare un unico abbinamento tra le componenti ed è necessario confrontare le media di ogni componente stimata per i salari del gruppo di controllo con le medie di ogni componente che contribuisce alla distribuzione per i salari dei soggetti trattati. Tra tutti i possibili confronti, solo la differenza tra  $\hat{\mu}_1^{NT}$  e  $\hat{\mu}_2^T$  è significativa. Non essendo possibile sapere a quale strato si riferisca ogni componente, la stratificazione della variabile intermedia non ci permette di trarre conclusioni sull'impatto della politica per questo particolare gruppo di soggetti.

### 6.2.3 Donne con indennità di mobilità

Rispetto al gruppo di uomini, le donne presentano delle differenze significative sia nei tassi di occupazione sia nella distribuzione dei salari. Nei diversi sottogruppi solo circa il 64% risulta occupato a tre anni dall'iscrizione in lista con un valore leggermente maggiore per le sole donne di 39 anni per le quali si osserva un tasso di occupazione di poco inferiore al 70%.

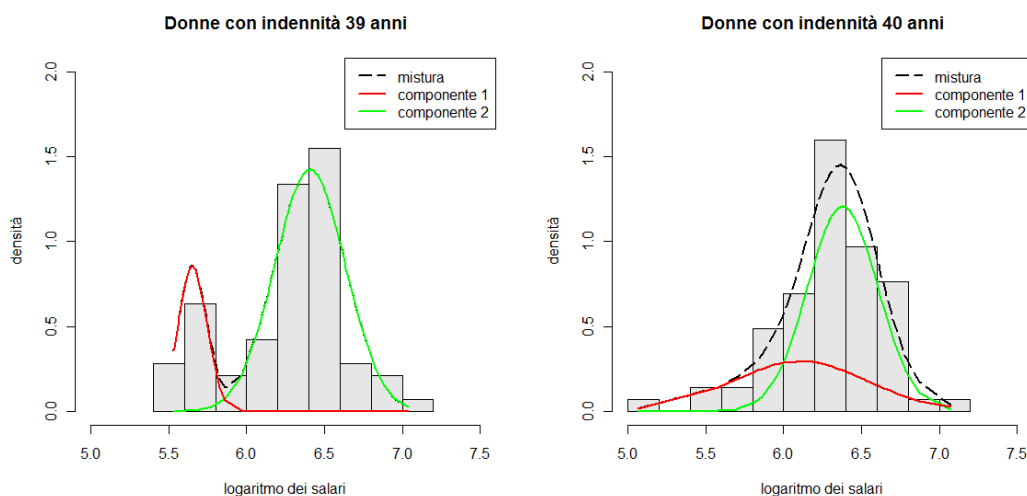
*Tabella 6.10 Uomini senza indennità di mobilità per età e situazione occupazionale tre anni dopo l'ingresso in lista.*

|             | 39 anni  | 40 anni  | 38-39 anni | 40-41 anni |
|-------------|----------|----------|------------|------------|
| Occupati    | 72 69.2% | 72 64.3% | 141 64.4%  | 143 63%    |
| Disoccupati | 32 30.8% | 40 35.7% | 78 35.6%   | 84 37%     |
| Totale      | 104 100% | 112 100% | 219 100%   | 227 100%   |

Nel gruppo di controllo i salari sono caratterizzati da una distribuzione bimodale per entrambi gli intervalli per l'età considerati. Le donne che risulterebbero sempre occupate quindi hanno un valore medio dei salari che

differisce significativamente dal valore medio dei salari per le donne che risulterebbero occupate solo nel caso in cui potessero usufruire della mobilità per un solo anno. Per il gruppo di trattamento le medie degli strati EE e UE differiscono tra loro in maniera meno evidente. Per quanto riguarda quest'ultimo gruppo, le stime calcolate sono fortemente sensibili a piccole variazioni del campione; l'eliminazione di pochi valori anomali provoca un cambiamento consistente nelle stime. Nelle considerazioni successive decidiamo di considerare il campione completo, sarà necessario però tener conto del problema di instabilità delle stime.

*Figura 6.5. Distribuzione stimata dei salari a tre anni dall'iscrizione alle liste per le donne che percepiscono l'indennità di mobilità per età (39 e 40 anni). Stima della mistura e delle componenti.*



Se limitiamo l'analisi alle sole lavoratrici di 39 e 40 anni, ancora una volta il metodo della stratificazione principale non ci permette di trarre conclusioni certe sull'impatto della politica sui salari per lo strato EE.

Tabella 6.11 Stime ML dei parametri che identificano le componenti della mistura per età (39,40 anni)

|                  | Età=39           | Età=40           |
|------------------|------------------|------------------|
| $\hat{\beta}$    | 0.197<br>(0.050) | 0.318<br>(0.345) |
| $\hat{\mu}_1$    | 5.653<br>(0.029) | 6.106<br>(0.25)  |
| $\hat{\mu}_2$    | 6.406<br>(0.032) | 6.378<br>(0.06)  |
| $\hat{\sigma}_1$ | 0.091<br>(0.022) | 0.431<br>(0.096) |
| $\hat{\sigma}_2$ | 0.224<br>(0.026) | 0.225<br>(0.075) |

Nessuna differenza tra le proporzioni è statisticamente significativa mentre i test di confronto tra le medie sono tutti a favore dell'uguaglianza tranne per una coppia  $(\hat{\mu}_1^{NT}, \hat{\mu}_2^T)$  per la quale la differenza è non nulla. In questo caso l'impossibilità di identificare l'effetto del trattamento è dovuto alla non applicabilità del metodo in questo campione da una parte e alla forte sensibilità delle stime ai valori estremi dall'altra.

Figura 6.6. Distribuzione stimata dei salari a tre anni dall'iscrizione alle liste per le donne che percepiscono l'indennità di mobilità per età (38-39 e 40-41 anni). Stima della mistura e delle componenti.

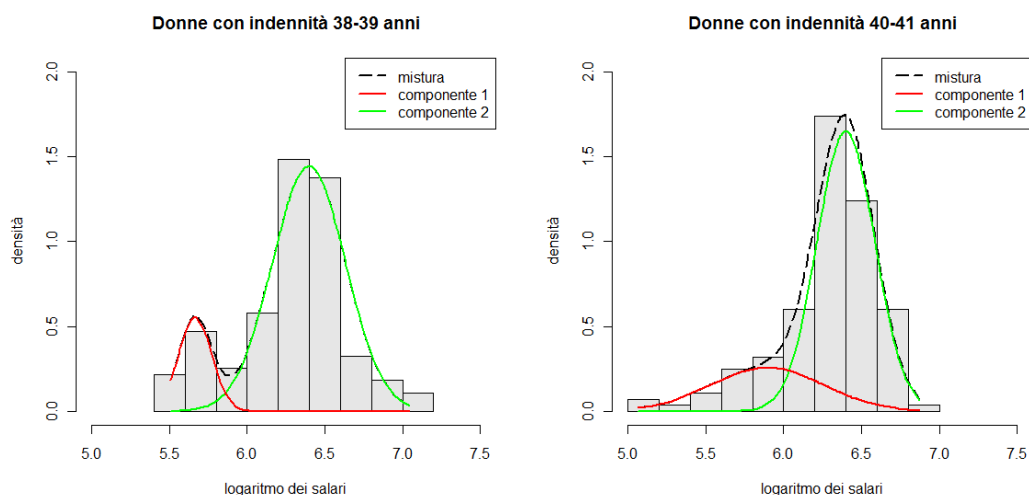


Tabella 6.12 Stime ML dei parametri che identificano le componenti della mistura per età (38-39, 40- 41 anni)

|                  | Età=38-39        | Età=40-41        |
|------------------|------------------|------------------|
| $\hat{\beta}$    | 0.148<br>(0.033) | 0.231<br>(0.110) |
| $\hat{\mu}_1$    | 5.666<br>(0.028) | 5.894<br>(0.199) |
| $\hat{\mu}_2$    | 6.397<br>(0.023) | 6.399<br>(0.024) |
| $\hat{\sigma}_1$ | 0.106<br>(0.021) | 0.359<br>(0.087) |
| $\hat{\sigma}_2$ | 0.235<br>(0.018) | 0.185<br>(0.019) |

Le stime sono meno sensibili se consideriamo il campione più ampio.. La distribuzione dei salari per il gruppo di controllo è composta per il 14.8% da una  $N(5.7, 0.11^2)$  e per il restante 85.2% da una  $N(6.4, 0.23^2)$ ; per il gruppo di trattamento invece le componenti stimate sono una  $N(5.9, 0.36^2)$  e una  $N(6.4, 0.18^2)$  con probabilità associate pari rispettivamente al 23% e al 77%. La differenza delle proporzioni stimate tra le componenti all'interno del gruppo di controllo e del gruppo di trattamento rispettivamente, consentono un abbinamento tra le componenti a destra e a sinistra della soglia. Entrambe le coppie, abbinate per valori della proporzione uguali, risultano avere lo stesso valore medio del logaritmo dei salari. Nel campione più numeroso per le donne che usufruiscono dei benefici monetari, il metodo è applicabile e permette di concludere che l'impatto medio sui salari di un anno aggiuntivo di ammissibilità alle liste per le lavoratrici 'sempre occupate' è nullo.

### 6.2.4 Donne senza indennità di mobilità

Le donne che non percepiscono l'indennità di mobilità costituiscono il sottogruppo con la numerosità più alta, caratterizzato un tasso di occupazione osservato che varia circa tra il 60% e il 70% all'interno di ogni gruppo d'età.

*Tabella 6.13 Donne senza indennità di mobilità per età all'ingresso in lista e situazione occupazionale tre anni dopo.*

|             | 39 anni |       | 40 anni |       | 38-39 anni |       | 40-41 anni |       |
|-------------|---------|-------|---------|-------|------------|-------|------------|-------|
| Occupati    | 139     | 68.8% | 124     | 61.4% | 289        | 66.6% | 235        | 61.5% |
| Disoccupati | 63      | 0.32% | 78      | 38.6% | 145        | 33.4% | 147        | 38.5% |
| Totale      | 202     | 100%  | 202     | 100%  | 434        | 100%  | 382        | 100%  |

Anche in questo caso le stime relative alle donne con età pari a 40 anni sono fortemente sensibili ai dati estremi. Decidiamo di considerare il campione completo senza eliminare alcun valore.

*Figura 6.7. Distribuzione stimata dei salari a tre anni dall'iscrizione alle liste per le donne che non percepiscono l'indennità di mobilità per età (39 e 40 anni). Stima della mistura e delle componenti.*

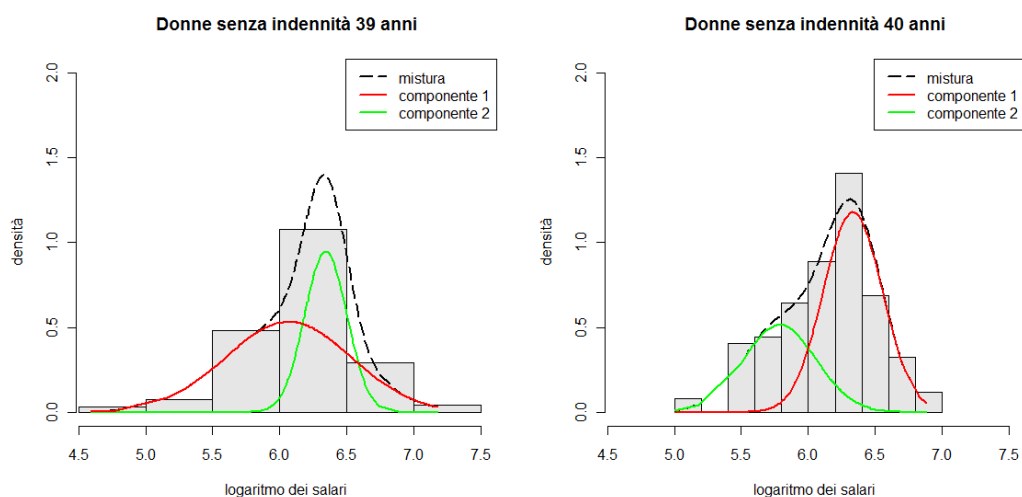


Tabella 6.14 Stime ML dei parametri che identificano le componenti della mistura per età (39,40 anni)

|                    | Età=39           | Età=40           |
|--------------------|------------------|------------------|
| $\hat{\beta}$      | 0.626<br>(0.069) | 0.648<br>(0.162) |
| $\hat{\mu}_1$      | 6.064<br>(0.046) | 6.332<br>(0.055) |
| $\hat{\mu}_2$      | 6.342<br>(0.024) | 5.787<br>(0.163) |
| $\hat{\sigma}_1^2$ | 0.469<br>(0.029) | 0.219<br>(0.030) |
| $\hat{\sigma}_2^2$ | 0.157<br>(0.020) | 0.272<br>(0.086) |

Ancora una volta emerge il limite di applicazione del metodo della stratificazione che non permette di identificare nessuno degli abbinamenti possibili a causa dell'uguaglianza delle proporzioni stimate per tutti gli strati. I test di uguaglianza tra le medie concludono a favore dell'assenza di un effetto solo per alcune delle possibili combinazioni tra le componenti, impedendo di trarre alcuna conclusione.

Figura 6.8. Distribuzione stimata dei salari a tre anni dall'iscrizione alle liste per le donne che non percepiscono l'indennità di mobilità per età (38-39 e 40-41 anni). Stima della mistura e delle componenti.

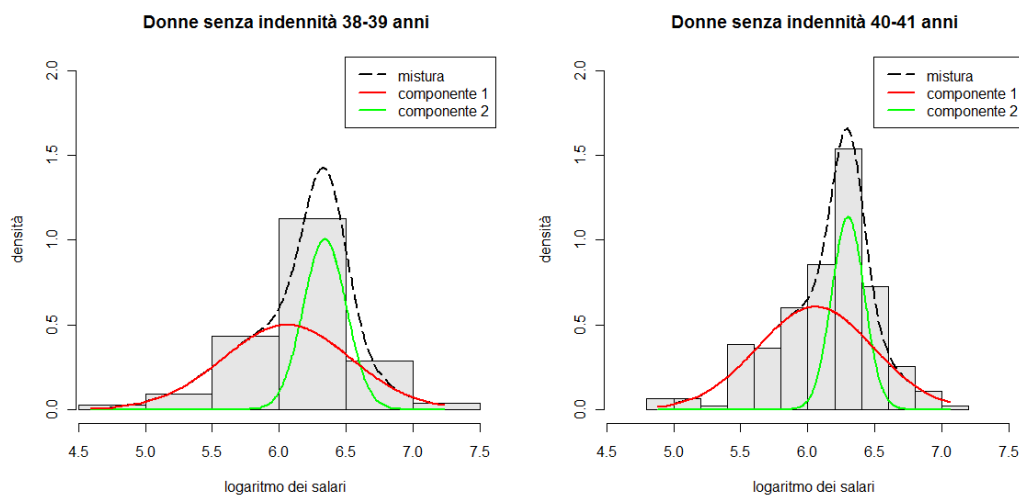


Tabella 6.15 Stime ML dei parametri che identificano le componenti della mistura per età (38-39,40-41 anni)

|                    | Età=38-39         | Età=40-41        |
|--------------------|-------------------|------------------|
| $\hat{\beta}$      | 0.595<br>(0.069)  | 0.655<br>(0.07)  |
| $\hat{\mu}_1$      | 6.056<br>(0.046)  | 6.055<br>(0.043) |
| $\hat{\mu}_2$      | 6.340<br>(0.024)  | 6.300<br>(0.022) |
| $\hat{\sigma}_1^2$ | 0.476<br>(0.0299) | 0.431<br>(0.027) |
| $\hat{\sigma}_2^2$ | 0.160<br>(0.020)  | 0.120<br>(0.022) |

Anche in questo caso le stime calcolate sulle donne con età compresa tra 38 e 41 anni sono meno sensibili. Inoltre tramite il confronto tra le proporzioni è possibile abbinare le coppie di componenti. La distribuzione del logaritmo dei salari per le lavoratrici nel gruppo di controllo è composto per il 59% da una  $N(6.1, 0.48^2)$  e per il rimanente 41% da una  $N(6.3, 0.16^2)$ . Nel gruppo di trattamento le componenti identificate sono una  $N(6, 0.43^2)$  per il 65% dei valori e una  $N(6.3, 0.12^2)$  per il rimanente 35%. Per entrambi i confronti accetto l'ipotesi di uguaglianza delle medie. L'effetto causale dell'anno aggiuntivo di ammissibilità alle liste per le donne che non percepiscono l'indennità di mobilità, che sarebbero sempre occupate a prescindere dal trattamento al quale sono assegnate, è nullo.



## Conclusioni e possibili sviluppi

L'interesse di questo studio si concentra sull'applicazione di un metodo di stima sviluppato nell'ambito della stratificazione principale, al fine di valutare l'impatto dell'estensione del periodo di ammissibilità al programma "Liste di Mobilità" sul livello dei salari. La politica è composta principalmente da due sottoprogrammi destinati a gruppi distinti di soggetti. I lavoratori licenziati da imprese con più di 15 addetti possono beneficiare di una riduzione degli oneri sociali previsti per le imprese che impiegano lavoratori dalle LM, inoltre ricevono un sussidio che può essere trasferito all'impresa che li assume. Per i lavoratori licenziati da imprese con meno di 15 addetti è prevista solo la riduzione degli oneri sociali per le imprese che assumono mentre i lavoratori disoccupati non hanno diritto a ricevere alcun beneficio monetario. In entrambi i gruppi, i lavoratori usufruiscono della politica per un periodo di tempo variabile in relazione alla loro età al momento dell'iscrizione alle liste. Il disegno della politica prevede che i lavoratori con età inferiore a 39 anni siano ammissibili al programma per un massimo di un anno, il periodo di ammissibilità si estende rispettivamente fino a due anni per i lavoratori dai 40 ai 49 anni e fino a tre anni per i lavoratori con età superiore a 50 anni. Lo studio si concentra solo sulle prime due fasce d'età e stima l'impatto del passaggio da uno a due anni di ammissibilità al programma sul salario dei lavoratori a tre anni dall'ingresso in lista. I gruppi di confronto sono stati costruiti, sfruttando il disegno della politica, tramite *rdd* con un valore soglia per l'età pari a 40 anni. In ogni gruppo si è poi applicata una stratificazione dei soggetti in relazione allo stato occupazionale; in questo modo i lavoratori sono stati suddivisi in quattro diversi gruppi (denominati appunto strati

principali) definiti dallo stato occupazionale che si sarebbe osservato per ogni soggetto se questo fosse stato sottoposto o meno al trattamento. L'applicazione del metodo consiste nell'individuare il valore medio dei salari per i soggetti che sarebbero stati comunque occupati a prescindere dalla lunghezza del periodo di ammissibilità al programma, sia in caso di assegnazione al gruppo di trattamento sia in caso di assegnazione al gruppo di controllo. La differenza tra i due valori medi è interpretabile come l'effetto causale sui salari per il gruppo dei 'sempre occupati'. L'idea sulla quale si basa il processo di identificazione dello strato per il quale calcolare la stima d'impatto è che il gruppo di confronto è formato dai 'sempre occupati' e dai soggetti che risulterebbero occupati solo se ammissibili alle liste per un solo anno mentre il gruppo di trattamento è formato dai 'sempre occupati' e dai lavoratori occupati solo nel caso possano rimanere nelle liste per due anni. Lo strato di interesse sia in caso di assegnazione al trattamento sia in caso contrario sarà necessariamente identificato, nel nostro particolare caso, dai due sottogruppi che presentano la stessa proporzione a destra e a sinistra della soglia.

Nel set di dati utilizzato in questo studio, la particolare composizione dei campioni che consideriamo comporta forti limitazioni nell'applicazione del metodo proposto. A causa dell'assenza di un effetto del trattamento sui tassi di occupazione a tre anni dall'ingresso in lista per i soggetti sulla soglia, sia per genere sia per benefici monetari percepiti, l'identificazione del sottogruppo di lavoratori 'sempre occupati' non è mai realizzabile. L'applicazione del metodo, così come esposto nell'analisi di Rubin (2006), consisteva infatti nel calcolare le proporzioni delle singole componenti e in seguito nell'estrarre le due componenti, la prima dal gruppo di controllo e la seconda dal gruppo di trattamento, per le quali erano stimate le stesse proporzioni. Queste componenti necessariamente identificavano lo strato EE presente in entrambi i gruppi; mentre le altre due componenti, che comprendevano una proporzione diversa di soggetti, identificavamo rispettivamente gli strati EU e UE. Nel nostro caso invece, anche tali strati sono caratterizzati dalla stessa proporzione. La strategia proposta consiste dunque nell'abbinare ogni componente stimata nel gruppo di controllo

con ogni componente stimata nel gruppo di trattamento. Tra tutte le possibili combinazioni consideriamo solo le coppie formate da componenti che presentano la stessa proporzione; infine, per ognuna di queste, calcoliamo un test di significatività sulla differenza tra le medie. Senza poter conoscere quale tra queste coppie identifichi il gruppo EE e di conseguenza quale differenza sia interpretabile come l'effetto causale SACE, è possibile trarre delle conclusioni sulla presenza o meno dell'impatto solo se i test per tutte le differenze considerate sono concordi a favore della loro nullità o significatività.

Per tutti i sottocampioni, sia uomini sia donne con o senza indennità di mobilità, per i quali il gruppo di controllo è stato costruito a partire dai soli lavoratori con 39 anni e il gruppo di trattamento dai soggetti con 40 anni, tutti gli strati sono caratterizzati dalla medesima proporzione. Per entrambi i gruppi degli uomini, i test effettuati su tutte le possibili combinazioni concludono a favore dell'ipotesi di uguaglianza. Per i lavoratori 'sempre occupati' tra i 39 e i 40 anni, il passaggio da uno a due anni di ammissibilità al trattamento non ha nessun effetto sui salari. Non è possibile al contrario trarre delle conclusioni per entrambi i gruppi delle donne; tra tutte le possibili combinazioni, infatti, i test di significatività portano a risultati discordi tra le diverse coppie di medie. La stratificazione principale dunque non è utilizzabile al fine di valutare se la politica ha avuto un impatto sui salari a tre anni dall'iscrizione in lista per le lavoratrici 'sempre occupate' tra i 39 e i 40 anni.

Laddove i gruppi di controllo e trattamento sono costruiti prendendo un campione più numeroso, nel quale sono inclusi anche i lavoratori con età pari rispettivamente a 38 e 41 anni, le combinazioni si riducono a un solo abbinamento possibile, i confronti sono limitati rispetto al caso precedente a solo due coppie di valori. Poiché entrambi i test sulle medie danno risultati concordanti, è possibile in questi casi trarre delle conclusioni sull'impatto della politica. Per gli uomini con diritto all'indennità di mobilità e per entrambi i gruppi delle donne, indistintamente dal diritto o meno di percepire benefici monetari, l'estensione di un anno del periodo di ammissibilità alle liste non ha nessun effetto sul livello dei salari. Questo significa che per i lavoratori 'sempre occupati' con età vicina alla

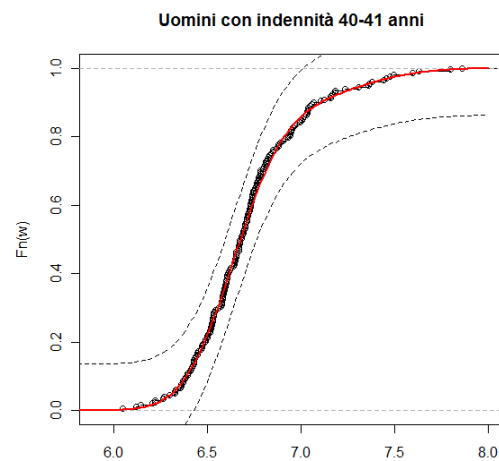
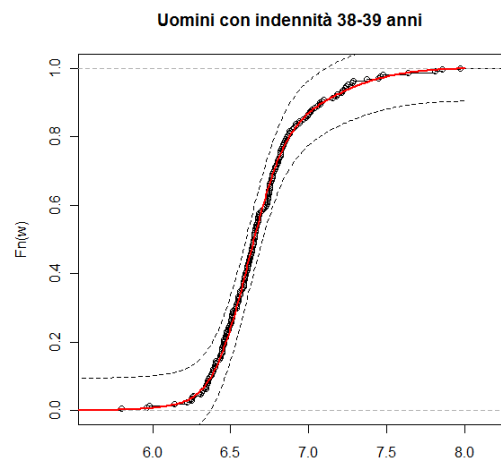
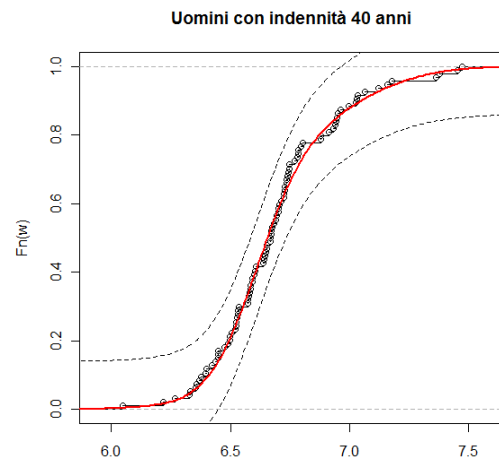
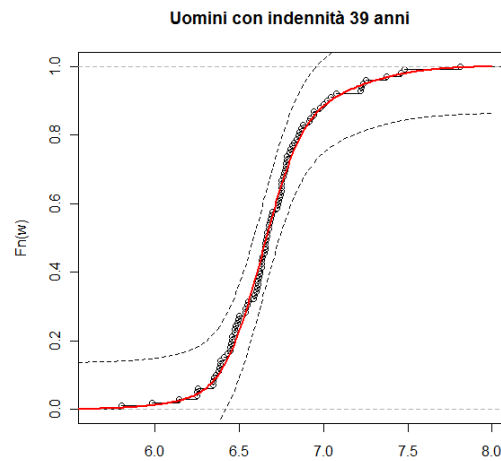
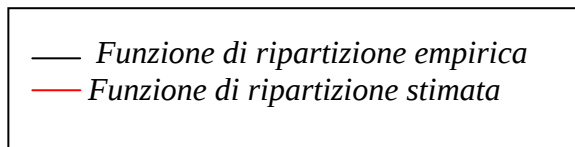
soglia, la permanenza prolungata nelle liste non ha effetti né sui tassi di occupazione né lavoro sulle retribuzioni eventualmente percepite tre anni dopo l'iscrizione alle liste. La valutazione basata sulla stratificazione, limitatamente alle donne e agli uomini che percepiscono l'indennità di mobilità, conferma i risultati ottenuti nello studio di Paggiaro et al. (2008) dove l'analisi dei salari è stata condotta tramite *rd* sullo stesso campione di soggetti. Il gruppo di uomini che non percepiscono indennità di mobilità tra i 38 e i 41 anni presenta invece problemi di non identificabilità del SACE; la discordanza tra i test sulle differenze preclude la possibilità di realizzare una valutazione dell'effetto causale.

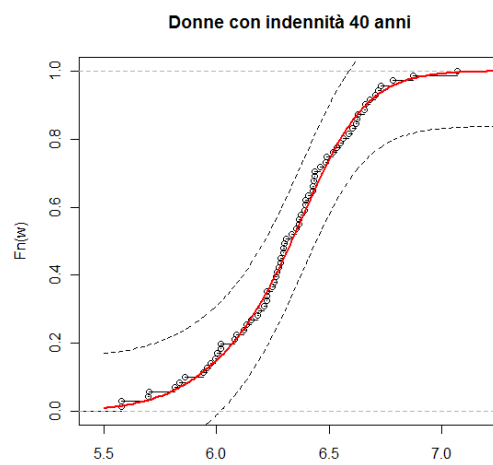
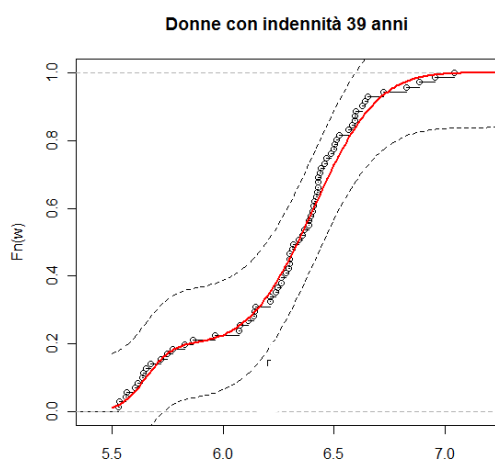
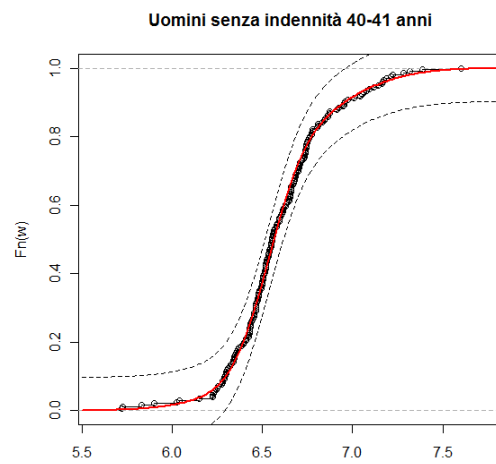
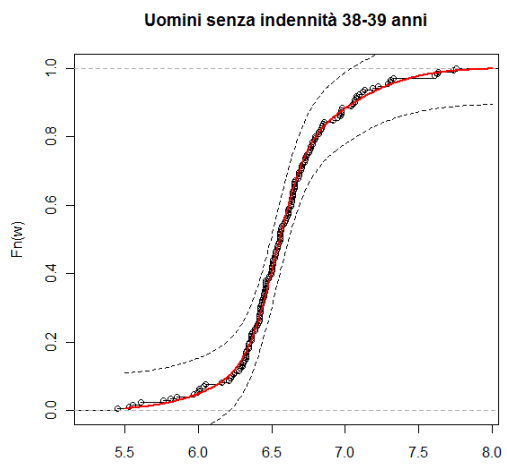
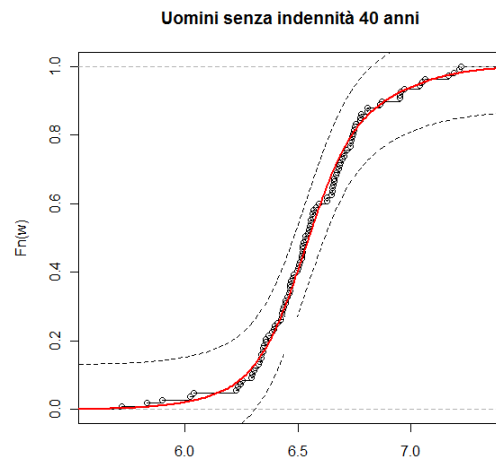
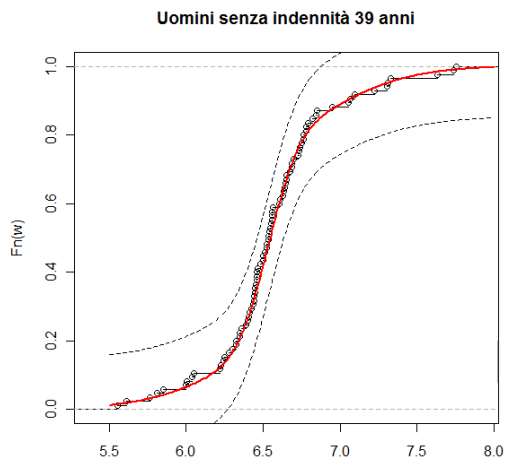
Le difficoltà di applicazione del metodo della stratificazione principale sono dovute alla particolare composizione degli strati nel campione utilizzato. Poiché la discriminazione tra gli strati avviene in base alla loro proporzione nella popolazione, l'uguaglianza delle stime delle proporzioni non permette di identificare i singoli strati. Il problema potrebbe essere superato imponendo ulteriori assunzioni sulla distribuzione dei dati o sulle proporzioni degli strati. L'assunzione di monotonicity, secondo la quale l'anno aggiuntivo di permanenza in lista non può portare un soggetto a passare dallo stato di occupato allo stato di disoccupato, non è verosimile in questo caso. In alternativa si potrebbe ipotizzare che il livello medio dei salari per il gruppo dei lavoratori 'sempre occupati' sia sempre maggiore rispetto al livello medio dei salari per i soggetti che appartengono ai rispettivi strati che compongono gruppo di controllo e gruppo di trattamento. Ogni assunzione ovviamente dovrebbe essere argomentata per sostenerne la plausibilità in ogni singolo caso. Infine per aumentare la scarsa precisione delle stime sarebbe interessante applicare ai salari un *local linear regression* con pesi non uniformi ma decrescenti attorno alla soglia al fine di utilizzare tutta l'informazione campionaria a nostra disposizione e non solo l'informazione sui salari dei soggetti con età più vicina alla soglia.

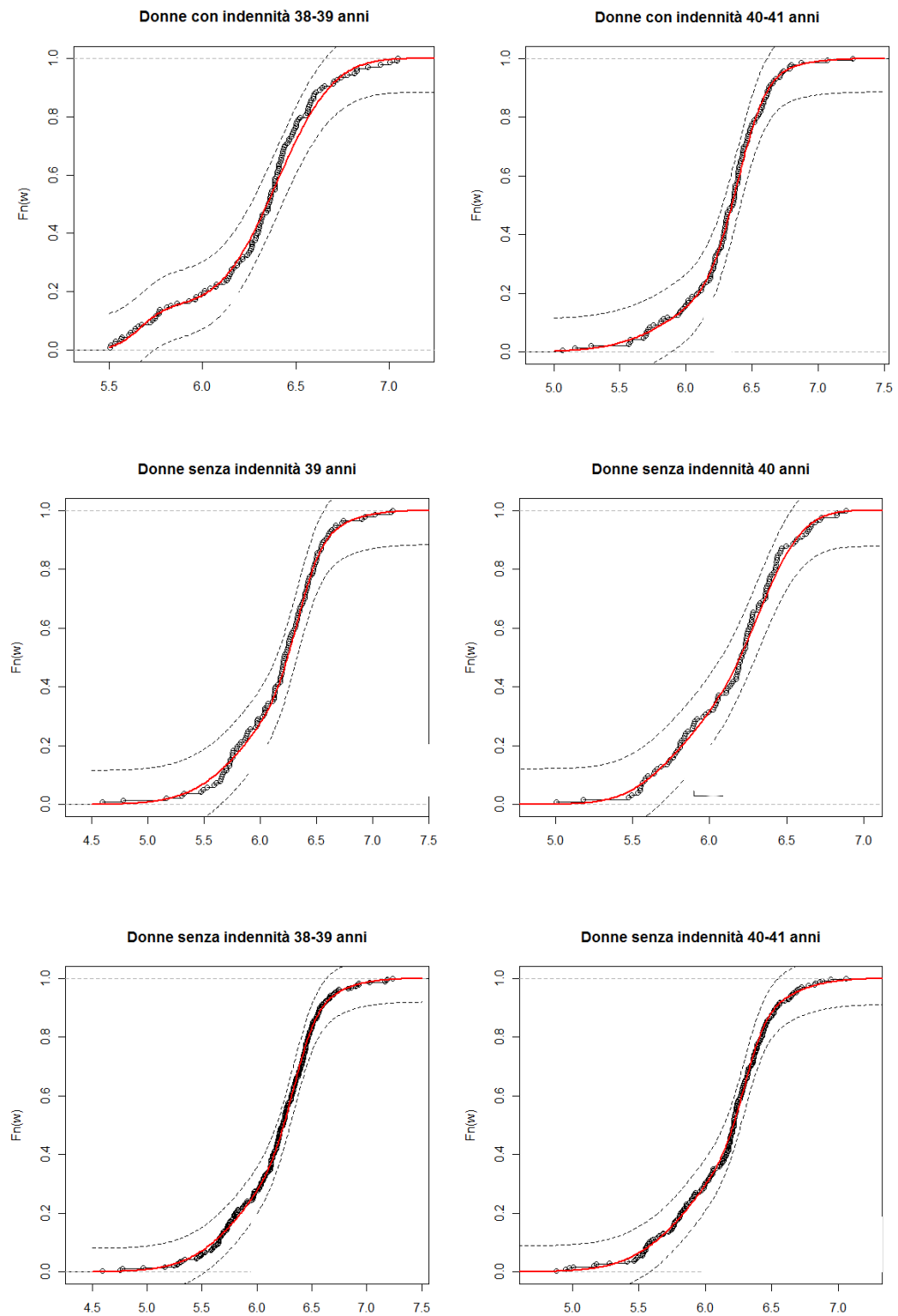
# APPENDICE

## ANALISI GRAFICA DELLA BONTÀ DI ADATTAMENTO DELLA DENSITÀ STIMATA

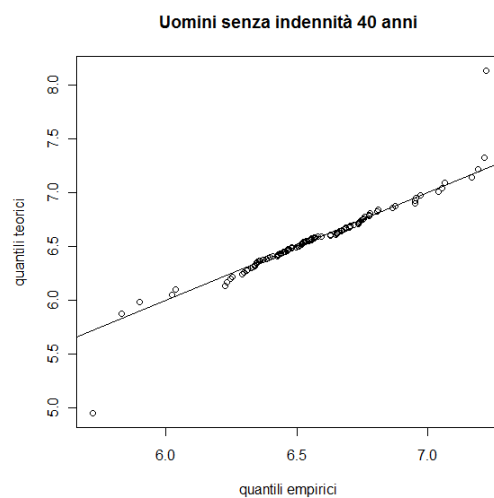
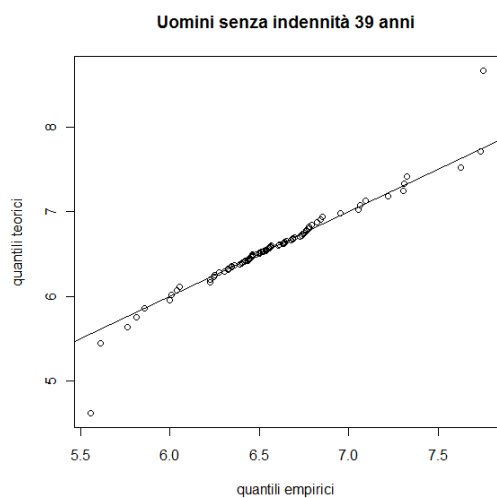
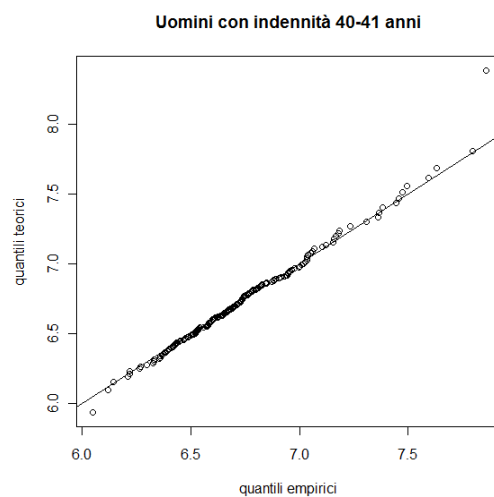
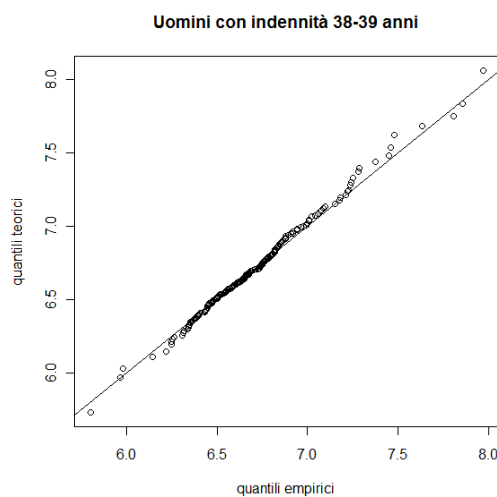
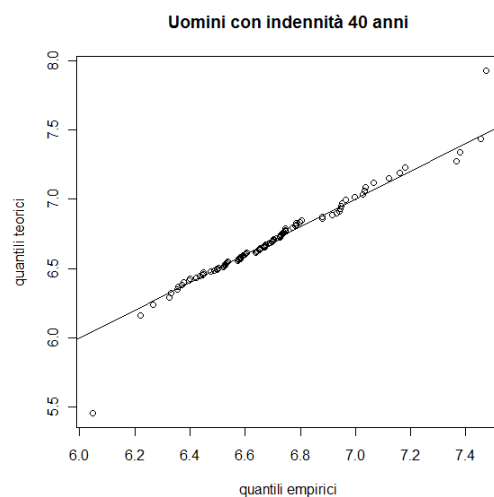
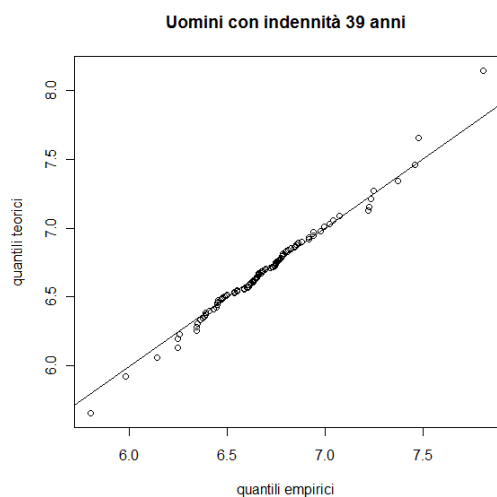
*Confronto tra la funzione di ripartizione empirica e la funzione di ripartizione stimata secondo l'ammissibilità ai benefici monetari, il genere e la classe di età.*



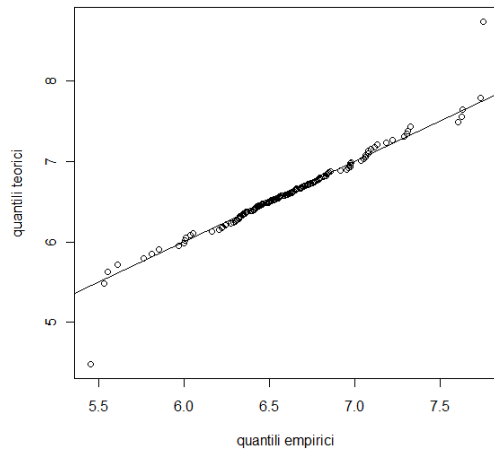




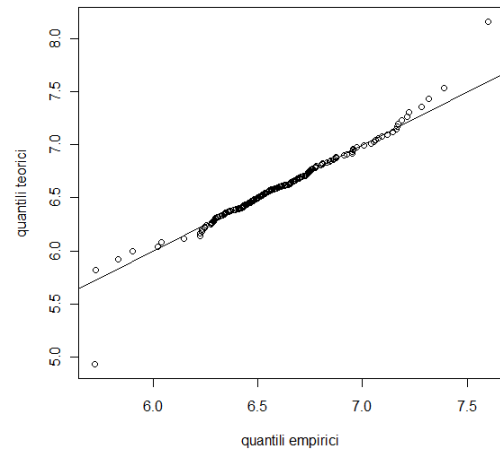
*Grafici quantile-quantile secondo l'ammissibilità ai benefici monetari, il genere e la classe di età.*



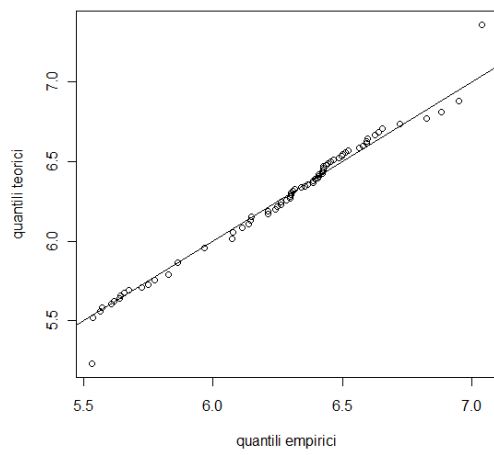
Uomini senza indennità 38-39 anni



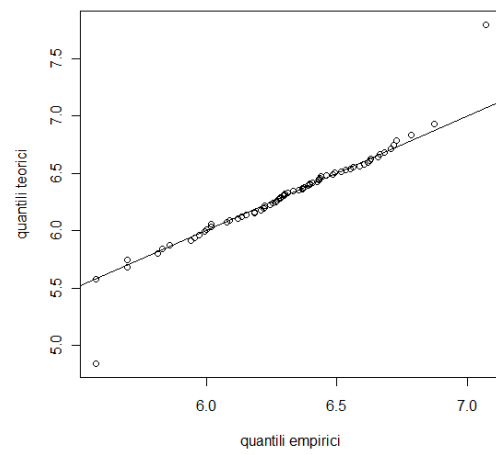
Uomini senza indennità 40-41 anni



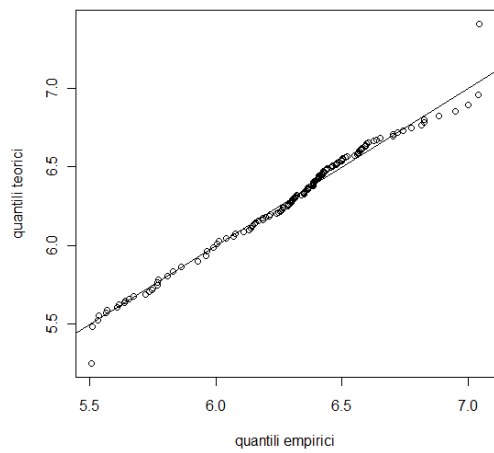
Donne indennità 39 anni



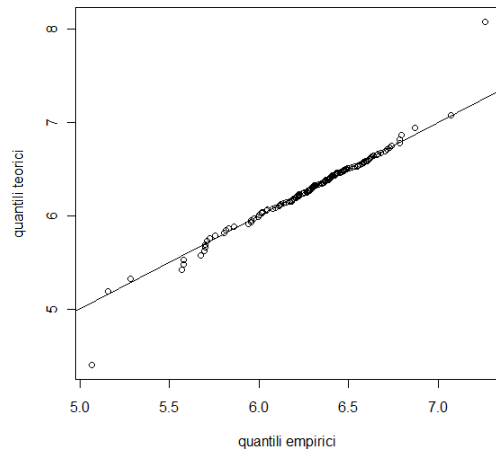
Donne con indennità 40 anni



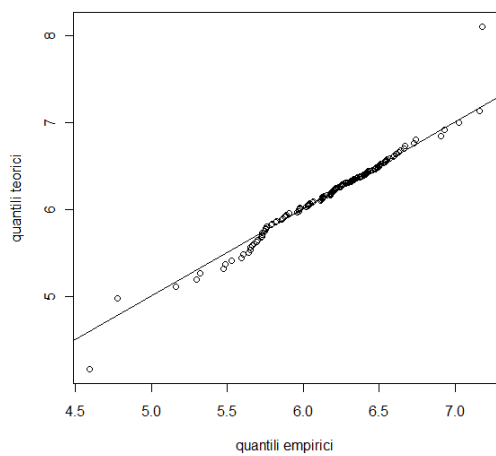
Donne indennità 38-39 anni



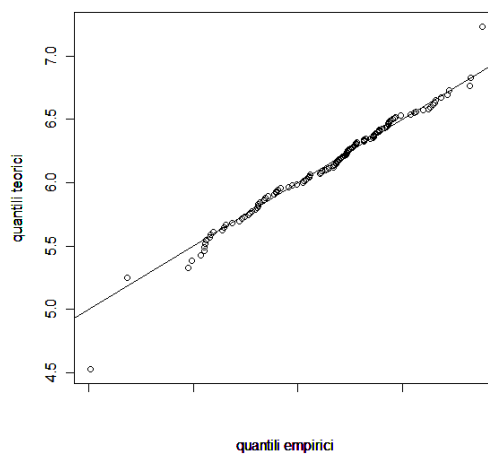
Donne con indennità 40-41 anni



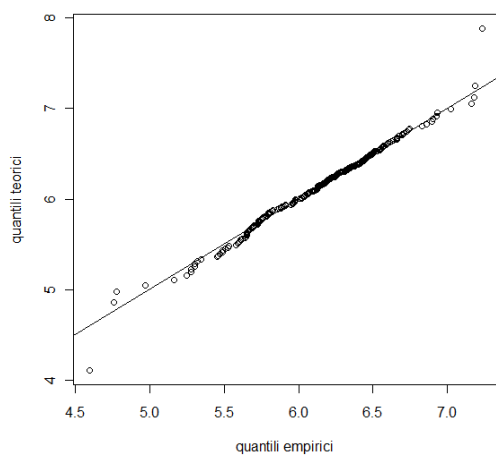
**Donne senza indennità 39 anni**



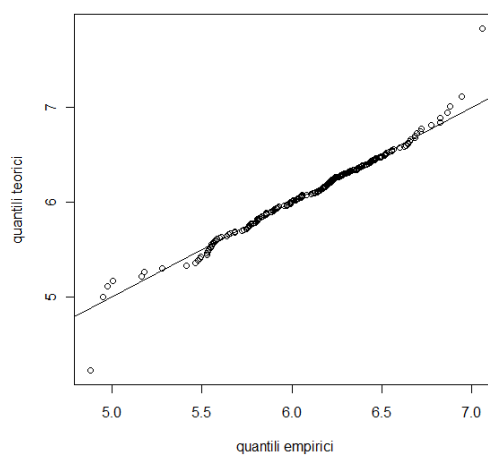
**Donne senza indennità 40 anni**



**Donne senza indennità 38-39 anni**



**Donne senza indennità 40-41 anni**



## Bibliografia

- Anastasia B., D. Cavapozzi, M. Gambuzza, D. Maurizio, A. Paggiaro, M. Rasera, E. Rettore e U. Trivellato (2004), “*Interazione fra sussidi passivi e incentivi al reimpiego: provenienze ed esiti di lavoratori iscritti nelle liste di mobilità. Rapporto finale*”, Venezia: Agenzia Veneto Lavoro (mimeo)
- Borzaga C. e Brunello G. (a cura di) (1997), *L’impatto delle politiche attive del lavoro in Italia*, Firenze: Edizioni Scientifiche Italiane.
- Brucchi Luchino B. (2001), *Manuale di economia del lavoro*, Bologna, Il mulino.
- Brunello G. e R. Miniaci (1997), “Benefit transfers in Italy: an empirical study of mobility lists in the Milan area”, *Oxford Bulletin of Economics and Statistics*, 59: 329-347.
- Caruso E. (2001), “Durata della disoccupazione e probabilità di impiego in presenza di incentivi all’assunzione: l’evidenza della lista di mobilità della regione Umbria”, *Politica Economica*, 17 (1): 73-95.
- Everitt and Hand (1981), *Finite mixture distributions*, London, New York: Chapman and Hall.
- Frangakis C.E. and D.B. Rubin (1999), “Addressing Complications of Intention-to-Treat Analysis in the Combined Presence of All-or-None Treatment-Noncompliance and Subsequent Missing Outcomes”, *Biometrika*, 86: 365–379.
- Frangakis, C.E. and D.B. Rubin (2002), “The defining role of ‘principal stratification and effects’ for comparing treatments adjusted for posttreatment variables: from treatment non-compliance to surrogate endpoints”, *Biometrics*, 58: 191–199.
- Frumento P., F. Mealli and B. Pacini (2008), “Evaluating the effect of training on wages in the presence of non-compliance”, Preliminary draft

- Grilli, L. and F. Mealli (2007), “Nonparametric Bounds on the Causal Effect of University Studies on Job Opportunities Using Principal Stratification”, *Journal of Educational and Behavioral Statistics*, 33 (1): 111-130
- Hahn J., P. Todd and W. Van der Klaauw (2001), “Identification and estimation of treatment effects with a regression-discontinuity design”, *Econometrica*, 69: 201-209.
- Holland P. W. (1986), “Statistics and causal inference”, *Journal of the American Statistical Association*, 81(396):945-960.
- Imbens G.W. and D.B. Rubin (1997), “Bayesian inference for causal effects in randomized experiments with non-compliance”, *Annals of Statistics*, 25: 305–327.
- Martini A. and L. Mo Costabella, (2007), “Una valutazione degli effetti indesiderati dell’istituto della mobilità su imprese e lavoratori”, *Politica economica*, 23 (3): 259-288.
- Martini A., E. Rettore and U. Trivellato (2006), “The effect of combining unemployment compensation and hiring subsidies on the employment of displaced workers: Evidence from Italy”, paper presented at the 28th Annual APPAM Research Conference “Designing, Implementing. Managing and Evaluating Effective Distributional Policies”, Madison, WI, November 2-4, 2006 (mimeo.).
- Martini A., E. Rettore ed U. Trivellato (2003), “La valutazione delle politiche del lavoro in presenza di selezione: migliorare la teoria, i metodi o i dati?”, *Politica Economica*, 3: 301-342.
- Mattei A., and F. Mealli (2007), “Application of the Principal Stratification Approach to the Faenza Randomized Experiment on Breast Self-Examination”, *Biometrics*, 63, 437-446.
- Mealli F. e B: Pacini (2008a), “Exploiting instrumental variables in causal inference with nonignorable outcome response using principal stratification”, Preliminary draft.
- Mealli F. e B: Pacini (2008b), “Comparing principal stratification and selection models in parametric causal inference with nonignorable missingness”, *Computational Statistics and Data Analysis*, doi:10.1016/j.csda.2008.09.005.
- McLachlan, G and D. Peel (2000), *Finite Mixture Models*, Wiley series in probability and statistics.

- McConnell S., E.A. Stuart and B. Devaney (2008), "The Truncation-by-Death Problem. What to do in an experimental evaluation when the outcome is not always defined", *Evaluation Review*, 32 (2), 157-186.
- Paggiaro A. e U. Trivellato (2001), "Il monitoraggio e la valutazione delle 'Liste di Mobilità': prime evidenze dalla integrazione con gli archivi Netlabor", in Trivellato U. (a cura di), *Servizi per l'impiego e ricerche sul lavoro. L'esperienza del Veneto*, *Quaderni di economia del lavoro*, 72, Milano, Franco Angeli.
- Paggiaro A., E. Rettore, U. Trivellato (2008), "*The effect of Extending the Duration of Eligibility in an Italian Labour Market Programme for Dismissed Workers*", CESIFO Working Paper No.2340
- Paggiaro A., E. Rettore, U. Trivellato (2005), "*The impact of the Italian "Mobility Lists" on employment chances: New evidence from linked administrative archive*", . Progetto MIUR 'Metodi e studi di valutazione degli effetti di politiche del lavoro, di aiuto alle imprese e di welfare', Working Paper No.65, Padova: Dipartimento di Scienze Statistiche, Università di Padova.
- Paggiaro A. e U. Trivellato (2002), "Assessing the effects of the 'Mobility Lists' programme by flexible duration models", *Labour*, 16 (2), 235-266.
- Rettore E., Trivellato U. e Martini A. (2003), "La valutazione delle politiche del lavoro in presenza di selezione: migliorare la teoria, i metodi o i dati?", , *Politica Economica*, 3, 301-342.
- Rubin D. B. (2006), "Causal Inference Through Potential Outcomes and Principal Stratification: Application to Studies with "Censoring" Due to Death", *Statistical Science*, 21, No.3, 299-309.
- Rubin D. B. (2005), "Causal Inference Using Potential Outcomes: Design, Modeling, Decisions", *Journal of the American Statistical Association*, 100, No. 469, 322-331.
- Rubin, D.B. (2004), "Direct and indirect causal effects via potential outcomes (with discussion)", *Scandinavian Journal of Statistics*, 31, 161-170.
- Snower D. (1994), "Converting unemployment benefits into employment subsidies", *American Economic Review*, 84 (2), 65-70.
- Veneto Lavoro (a cura di), *Il mercato del lavoro nel Veneto. Tendenze e politiche*. Rapporto 2005.

Zhang J., D.B. Rubin, and F. Mealli (2006), “Evaluating The Effects of Job Training Programs on Wages through Principal Stratification”, in D. Millimet, J. Smith, and E. Vytlacil (eds.), *Modeling and Evaluating Treatment Effects in Econometrics*, Advances in Econometrics, volume 21, Elsevier Science Ltd, UK.

Zhang J.L and D.B. Rubin (2003), “Estimation of causal effects via principal stratification when some outcomes are truncated by ‘death’”, *Journal of Educational and Behavioural Statistics*, 28, 353-368.

Ulteriori informazioni sono state tratte dai seguenti siti

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

[www.lavoro.gov.it](http://www.lavoro.gov.it)

[www.oecd.org](http://www.oecd.org)