

## **Analisi di sopravvivenza e modelli di durata della disoccupazione: l'evidenza empirica di un caso\***

Federica Balugani  
Barbara Pistoresi  
Università di Modena

*Sommario. In questo lavoro si utilizza l'analisi di sopravvivenza per stimare la durata della disoccupazione attraverso un modello parametrico di azzardo proporzionale con ipotesi di distribuzione Weibull. Le variabili considerate nell'analisi riguardano le caratteristiche personali e della storia lavorativa passata degli individui iscritti alle liste di disoccupazione. I dati sono relativi agli iscritti all'Ufficio di Collocamento della provincia di Modena nell'Aprile 1991. I risultati non mostrano l'esistenza di una significativa dipendenza dalla durata, qualora si tenga conto della eterogeneità tra i lavoratori. Ciò significa che non vi è dipendenza, nè positiva nè negativa, della probabilità di uscita dalla disoccupazione dal tempo trascorso nella stato di disoccupato.*

### **Introduzione**

Il tasso di disoccupazione è uno degli indicatori economici più importanti del funzionamento di una economia. Tuttavia, come ogni altro indicatore aggregato, esso non è sufficiente a cogliere importanti aspetti che caratterizzano il fenomeno della disoccupazione. Anche una analisi disaggregata del tasso di disoccupazione per derivare classificazioni di categorie sociali ed economiche non sarebbe in grado di cogliere l'aspetto dinamico della disoccupazione. Ai fini di interventi di politica economica e industriale è molto importante non solo sapere qual è il numero dei disoccupati, ma anche sapere se le persone disoccupate sono sempre le stesse oppure cambiano nel tempo. È importante sapere, cioè, come si muovono i lavoratori tra i vari stati del mercato e di quali dinamiche sociali e produttive un particolare fenomeno di disoccupazione è il prodotto.

---

\*Ringraziamo Michele Bruni e Giovanni Solinas che ci hanno aiutato a comprendere meglio la realtà del mercato del lavoro locale. Inoltre, siamo grate a Michele Lalla e Nicola Rossi per i preziosi suggerimenti sulla parte econometrica. Infine, un particolare ringraziamento va anche a Luigi Brighi e Mario Forni che hanno letto la prima stesura di questo lavoro, fornendoci utili indicazioni. Per gli errori ancora presenti nel testo rimaniamo le uniche responsabili. Anche se l'articolo è il risultato di un lavoro comune, Federica Balugani ha curato i paragrafi 2 e 3, mentre Barbara Pistoresi ha curato i paragrafi 1 e 4.

Un primo tipo di indagine sulla disoccupazione in questa direzione è l'analisi della durata media della disoccupazione, sia in aggregato che in riferimento a particolari categorie sociali e demografiche. A questo punto sorge un problema di corretta misurazione della durata media. Disponendo di dati longitudinali, ad esempio su un anno, la durata media è calcolata prendendo la media di tutti gli intervalli di disoccupazione che sono stati completati durante l'anno. In presenza, invece, di dati di tipo cross-section al momento della rilevazione si registrano anche le durate di intervalli interrotti, che verranno completati solo in futuro. Se si calcola la durata media in modo tradizionale in presenza di intervalli interrotti si incorre in una distorsione campionaria che porta a sovrastimare (length bias) o sottostimare (interruption bias) la durata media della disoccupazione. La sovrastima è dovuta al fatto che al momento della rilevazione è più probabile siano in corso gli intervalli più lunghi e questo porta ad una loro sovra-rappresentazione, mentre gli intervalli brevi tendono a non essere rappresentati; questo implica che la durata media degli intervalli interrotti porta a sovrastimare la durata media degli intervalli completati. Un'altra forma di distorsione che porta invece a sottostimare la durata, consiste nel fatto che i periodi interrotti hanno in genere una lunghezza inferiore a quelli completati, quindi la durata media degli intervalli interrotti porta ad una sottostima della durata degli intervalli completati.

In generale, l'effetto che i due tipi di distorsioni esercitano sulla media può essere ricondotto ad una caratteristica delle funzioni di densità di probabilità della lunghezza degli intervalli completati. Ad esempio, se un disoccupato ha sempre la stessa probabilità di rientrare nell'occupazione, indipendentemente dal periodo di tempo trascorso come disoccupato, allora non è difficile dimostrare che la distribuzione dei periodi completati è esponenziale. Nel caso di una distribuzione di probabilità esponenziale le due distorsioni tendono a compensarsi e le due medie, quella degli interrotti e quella dei completati, coincidono. Questo è il caso che si ritrova come risultato della analisi presentata nel paragrafo 4 di questo lavoro, e che ci porta ad asserire che non si è commesso alcun errore trattando gli intervalli interrotti al momento della nostra indagine come se fossero completati. Più in generale la teoria ci suggerisce che se la probabilità di uscita dalla disoccupazione aumenta all'aumentare della durata trascorsa nello stato i dati di tipo cross section sottostimano la durata degli intervalli completati, mentre se tale probabilità diminuisce col trascorrere della durata i dati tendono a sovrastimare la durata degli intervalli completati (Salant 1977).

Un tipo di analisi appropriato per studiare la durata della disoccupazione che tiene conto di questi problemi è l'analisi di sopravvivenza. Tale metodo consente di stimare la distribuzione dei completi in presenza di intervalli non ancora completati o di stimare la distribuzione dei completi avendo solo intervalli interrotti. Questo strumento, già utilizzato in altre discipline come la demografia, l'ingegneria e la medicina, negli anni '50 viene utilizzato da un gruppo di studiosi del Tavistock Institute e da Silcock (1954) per l'analisi del turnover aziendale<sup>(1)</sup>. Negli anni '70 e '80 questo approccio metodologico viene utilizzato nell'ambito della teoria della Job Search per convalidare l'idea che la disoccupazione sia un fenomeno di breve periodo e di natura volontaria, quindi si sposta l'attenzione dal problema della stabilità occupazionale, valutata attraverso l'analisi del turnover, al problema della disoccupazione.

In Italia prima degli anni '80 la letteratura si era principalmente concentrata sullo studio delle dinamiche del mercato del lavoro in relazione al ciclo economico e alle problematiche del lavoratore scoraggiato ed addizionale. Il dibattito, cioè, era legato all'idea che la probabilità di uscita dalla disoccupazione fosse essenzialmente un problema di carenza di domanda aggregata e di posti disponibili. Nello stesso periodo la letteratura straniera si era già indirizzata verso l'analisi delle dinamiche dell'offerta di lavoro mediante modelli di durata della disoccupazione. In Italia si comincia a porre attenzione ai modelli di durata solo agli inizi degli anni '80; i primi modelli formulati attraverso l'analisi di sopravvivenza ricalcano i modelli presentati dalla letteratura straniera e ne condividono sia i vantaggi sia i limiti.

In questo lavoro viene presentata una applicazione empirica dell'analisi di sopravvivenza allo studio della durata della disoccupazione nella realtà di Modena. L'analisi è relativa alla popolazione degli iscritti alle liste di collocamento dell'Ufficio provinciale di Modena ed è stata condotta su un campione di oltre trecento individui rilevato nell'Aprile del 1991. Il modello econometrico impiegato è quello di azzardo proporzionale e le stime sono state ottenute col metodo della Massima Verosimiglianza.

La nostra applicazione econometrica sulla durata si inserisce all'interno di una letteratura italiana di analisi dinamica del mercato del lavoro. A riguardo si ricordano i lavori di Lalla 1990 e Solinas 1990 sull'anzianità aziendale in Emilia Romagna, il lavoro di Flinn 1986 sempre sulla anzianità aziendale in riferimento a dati nazionali, il lavoro di Gay 1989 e di Dell'Aringa e Samek Lodovici 1988 sulla realtà disoccupazionale lombarda, il lavoro di De Stavola e Poli 1985 sulla occupazione giovanile e di Torelli e Trivellato 1988 sulla disoccupazione

giovanile in Lombardia. Il campione utilizzato nel nostro lavoro è simile a quello di Gay, anche se i dati sono relativi alla provincia di Modena; il modello di azzardo proporzionale con cui stimiamo la durata di disoccupazione è quello utilizzato anche da Dall'Aringa e Samek Lodovici, Flinn, Lalla e Solinas, Torelli e Trivellato.

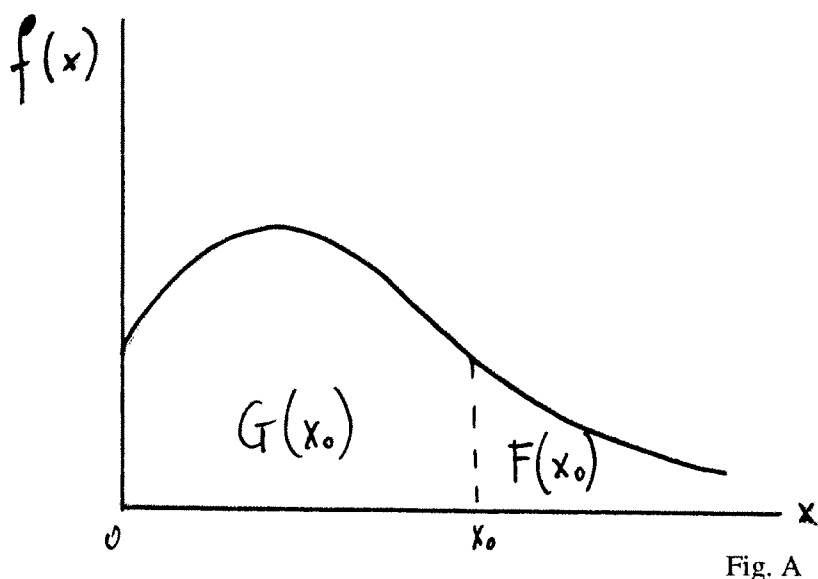
Nel primo paragrafo dell'articolo si illustrano brevemente alcuni concetti importanti per l'analisi di sopravvivenza e alcuni problemi connessi alla specificazione dell'eterogeneità fra gli individui per la stima corretta della durata; nel secondo paragrafo viene presentato il modello di azzardo proporzionale; il terzo paragrafo contiene una descrizione del campione e del metodo di rilevazione; nel quarto paragrafo si presentano i risultati dell'analisi empirica.

### **1. La dipendenza dalla durata e l'effetto eterogeneità nell'analisi di sopravvivenza**

L'analisi di sopravvivenza poggia su due importanti concetti: la funzione di sopravvivenza e la probabilità di uscita. La funzione di sopravvivenza  $F(x)$  indica la probabilità di rimanere nello stato (disoccupazione) almeno fino al tempo  $x$ . Se si indica  $G(x)$  la funzione di distribuzione che indica la probabilità di uscire dallo stato prima del tempo  $x$ ,  $F(x)$  risulta il complemento ad uno di  $G(x)$ . Se  $G(x)$  è differenziabile la sua funzione di densità di probabilità  $f(x)$  rappresenta la probabilità di uscire dalla disoccupazione "al margine", cioè nell'unità di tempo successiva ad  $x$ . Si può scrivere quindi

$$f(x) = G'(x) = -F'(x).$$

La relazione fra la funzione di sopravvivenza, di densità e di distribuzione è rappresentata nella seguente figura



La probabilità di uscita, o tasso di uscita dallo stato, indica la probabilità di uscire nell'unità di tempo successiva ad  $x$  condizionale al fatto di essere sopravvissuti nello stato fino ad  $x$ . Il tasso di uscita  $r(x)$  è definito nel modo seguente

$$r(x) = -\frac{F'(x)}{F(x)} = \frac{f(x)}{F(x)}$$

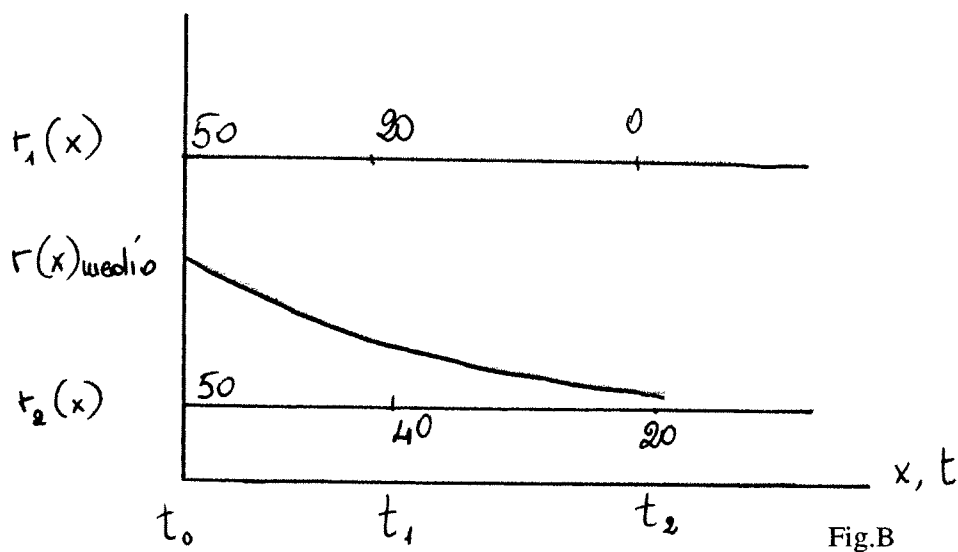
La funzione di sopravvivenza ha un andamento sempre decrescente, mentre l'andamento del tasso di uscita può essere crescente, costante o decrescente all'aumentare della durata di permanenza nello stato. I diversi orientamenti della letteratura sulla disoccupazione predicono andamenti diversi per il tasso di uscita. Ad esempio, la teoria della disoccupazione Job Search implica un andamento crescente o costante di  $r(x)$ . Questo significa che all'aumentare della durata aumenta o rimane costante la probabilità di uscita dalla disoccupazione<sup>(2)</sup>.

La maggior parte dei risultati empirici<sup>(3)</sup> mette in evidenza un andamento decrescente di  $r(x)$  rispetto alla durata della disoccupazione. Tuttavia, in generale le stime fanno riferimento ad un tasso di uscita *medio* relativo ad un gruppo di individui. L'andamento decrescente del tasso di uscita medio può avere due diverse spiegazioni:

A) esiste effettivamente una *dipendenza dalla durata* trascorsa nello stato. L'andamento di  $r(x)$  medio, in questo caso, è il risultato dell'aggregazione di  $r(x)$  individuali uguali per tutti gli individui e decrescenti, per i quali cioè l'andamento varia esclusivamente con la durata della disoccupazione già trascorsa<sup>(4)</sup>.

B) c'è un *effetto di eterogeneità* nella popolazione, cioè gli individui hanno tassi di uscita individuali costanti, ma diversi tra loro. In un gruppo di individui che rimangono disoccupati nello stesso momento vi sono persone con tassi di uscita relativamente più alti che lasciano la disoccupazione molto velocemente. Il tasso di uscita medio diminuisce perché col trascorrere del tempo restano nel gruppo solo gli individui che hanno modeste probabilità di trovare lavoro.

Ad esempio, supponiamo che la popolazione sia composta da due gruppi omogenei con due diversi tassi di uscita  $r_1$  ed  $r_2$  e di uguale numerosità (cinquanta individui al tempo  $t_0$ ) come è rappresentato nella Figura B. Supponiamo inoltre che il gruppo 1 sia caratterizzato da un tasso di uscita più alto. Al tempo  $t_1$  gli individui rimasti sono solo venti, mentre per il gruppo 2 gli individui sono ancora quaranta. Al tempo  $t_2$  gli individui del gruppo 1 sono tutti usciti, mentre nel gruppo 2 sono "sopravvissuti" ancora venti individui. Il tasso di uscita medio è decrescente, poiché al crescere della durata il gruppo con probabilità di uscita più bassa pesa sempre più. In  $t_0$ , il tasso medio è  $(r_1 + r_2)/2$ ; in  $t_1$  è  $(r_1 + 2r_2)/3$ ; in  $t_2$  è uguale a  $r_2$



Questo è un esempio eloquente di come l'aggregazione, in presenza di eterogeneità possa fornire dei risultati fuorvianti circa le effettive

dinamiche individuali. Se i comportamenti degli agenti non sono omogenei e la loro eterogeneità non è riconosciuta nel modello, si possono trarre conclusioni sbagliate stimando una dipendenza dalla durata più forte di quella che effettivamente esiste<sup>(5)</sup>.

Queste osservazioni sottolineano l'importanza di modellare esplicitamente l'eterogeneità dei comportamenti individuali nei modelli di durata. In generale nella letteratura l'eterogeneità è specificata da un insieme di variabili osservabili che descrivono le caratteristiche demografiche di ogni individuo. Tuttavia, la differenza nei comportamenti non è totalmente riducibile a questi fattori; l'eterogeneità tra disoccupati può essere determinata anche dalla storia lavorativa passata. Nel nostro modello la storia lavorativa entra come variabile esplicativa e, come si vedrà, ha un peso rilevante nella determinazione della durata della disoccupazione.

## 2. Il modello

Il modello econometrico impiegato nella analisi empirica consente di catturare separatamente la dipendenza dalla durata del tasso di uscita e gli effetti dovuti alla presenza di eterogeneità. Il tasso di uscita  $r(x, E)$  è funzione della durata  $x$  e di un vettore di variabili esplicative  $E$  che rappresentano le caratteristiche osservabili degli individui, come le caratteristiche demografiche e la storia lavorativa passata. In particolare si è utilizzato un modello di "azzardo proporzionale" (*Proportional Hazard Model*<sup>(6)</sup>) in cui il tasso di uscita (hazard rate<sup>(7)</sup>) è specificato nel modo che segue

$$r(x, E) = r_0(x)\phi(E, \beta),$$

dove  $r_0$  coglie la dipendenza dalla durata, mentre  $\phi$  cattura l'influenza dell'eterogeneità. In questa specificazione il tasso di uscita è uguale per tutti gli agenti a meno di un fattore moltiplicativo. Questa semplificazione è molto restrittiva, ma consente di rendere più agevole la stima ed è quella più utilizzata nella letteratura. La forma funzionale scelta per  $\phi$  è quella esponenziale:  $\phi(E, \beta) = \exp(E'\beta)$ , dove  $\beta$  è un vettore di parametri la cui componente  $i$ -esima  $\beta_i$  indica l'elasticità del tasso di uscita rispetto alla variabile esplicativa  $E_i$ . Nel caso di una variabile dummy, ad esempio, un valore positivo (negativo) del parametro  $\beta_i$  indica che il tasso di uscita per il gruppo della popolazione con caratteristiche  $E_i = 0$  è traslato verso l'alto (basso) rispetto al gruppo della popolazione di riferimento, come si può osservare dalla Figura C.

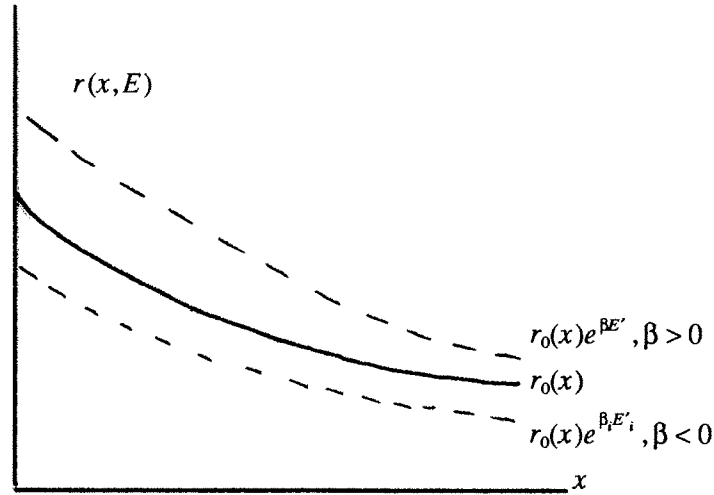


Fig.C

Per specificare il modello econometrico occorre introdurre una ipotesi sulla distribuzione di probabilità della durata della disoccupazione. Vi sono diversi tipi di distribuzione di probabilità che possono essere impiegati per stimare la durata della disoccupazione. La distribuzione esponenziale viene frequentemente adottata nella letteratura, tuttavia è fortemente restrittiva perché implica un tasso di uscita costante, cioè indipendente dalla durata. Al contrario, in questo lavoro si fa l'ipotesi che la funzione  $r_0(x)$  sia specificata da una distribuzione Weibull. Tale distribuzione è una generalizzazione della distribuzione esponenziale ed è compatibile con qualunque ipotesi sull'andamento del tasso di uscita. Sotto l'ipotesi di distribuzione Weibull il tasso di uscita assume la forma seguente

$$r_0(x) = \lambda \alpha (\alpha x)^{\lambda-1} = \lambda \alpha^\lambda x^{\lambda-1} \quad \alpha, \lambda > 0$$

Il parametro  $\lambda$  è una misura della dipendenza del tasso di uscita dalla durata (parametro di forma) e il parametro  $\alpha$  è un parametro di scala. Quando  $0 < \lambda < 1$  l'andamento del tasso di uscita è decrescente, se  $\lambda = 1$  si avrà un andamento costante e si riproduce la distribuzione esponenziale; per  $\lambda > 1$  l'andamento di  $r_0$  è crescente.

Quindi, date le ipotesi su  $r_0$  e su  $\phi$  il modello di azzardo proporzionale per la distribuzione Weibull è il seguente



$$r(x, E) = \lambda \alpha (\alpha x)^{\lambda-1} e^{-E^\beta}$$

Il metodo utilizzato per la stima è quello di Massima Verosimiglianza. La variabile dipendente è costituita dalla durata degli intervalli di disoccupazione interrotti, cioè ancora in corso durante la rilevazione<sup>(8)</sup>. In questo caso se si mantiene l'ipotesi che gli intervalli di disoccupazione una volta completati si distribuiscano secondo una Weibull e seguano un modello di azzardo proporzionale allora si può mostrare<sup>(9)</sup> che la funzione di densità degli intervalli interrotti assume la seguente forma

$$g(x, E) = \frac{\alpha \lambda}{\Gamma\left(\frac{1}{\lambda}\right)} e^{\frac{E^\beta}{\lambda}} e^{-(\alpha x)^\lambda e^{E^\beta}}$$

Il logaritmo della funzione di Verosimiglianza nel caso di campionamento casuale, per il modello con eterogeneità, è il seguente

$$1) \quad L(x, E, \beta) = \sum_i^n \log \left[ \frac{\alpha \lambda}{\Gamma\left(\frac{1}{\lambda}\right)} e^{\frac{E_i^\beta}{\lambda}} e^{-(\alpha x_i)^\lambda e^{E_i^\beta}} \right]$$

Invece, il logaritmo della funzione di Verosimiglianza nel caso del modello senza eterogeneità, in cui si stimano solo il parametro di scala e di forma, è il seguente

$$2) \quad L(x, \alpha, \lambda) = \log(\alpha) + \log(\lambda) - \log \Gamma\left(\frac{1}{\lambda}\right) - \log \sum_i^n (\alpha x_i)^\lambda$$

Per calcolare la durata media della disoccupazione, nel caso del modello con eterogeneità, una volta stimati i parametri, si utilizza la seguente espressione

$$3) \quad \mu = \frac{e^{-x^\beta \lambda} \Gamma\left(\frac{2}{\lambda}\right)}{\alpha \Gamma\left(\frac{1}{\lambda}\right)},$$

mentre si segue l'espressione 4) per il calcolo della durata nel caso del modello senza eterogeneità

$$4) \quad \mu = \frac{1}{\alpha} \frac{\Gamma\left(\frac{2}{\lambda}\right)}{\Gamma\left(\frac{1}{\lambda}\right)}.$$

### 3. I dati e le caratteristiche del campione

I dati utilizzati in questo lavoro sono dati dell'Ufficio di Collocamento e sono stati scelti perché hanno le caratteristiche dei dati longitudinali, cioè permettono di ricostruire le *storie lavorative* dei singoli

individui iscritti a partire dal loro ingresso nel mercato del lavoro<sup>(10)</sup>. Per ogni soggetto iscritto al Collocamento si può calcolare la durata di tutti i singoli intervalli di occupazione, disoccupazione e di non iscrizione, quindi si possono ricostruire le variabili di maggiore interesse per una analisi di flusso del mercato del lavoro in relazione alle caratteristiche personali dei soggetti (età, sesso, grado di istruzione, qualifica professionale, stato civile, numero dei componenti della famiglia).

Questa ricerca è volta ad individuare le caratteristiche qualitative della disoccupazione nel comune di Modena. Si è effettuata una indagine disaggregata sugli iscritti alle liste di collocamento attraverso una rilevazione delle informazioni presso l'Ufficio. Infatti, i dati statistici raccolti a fini amministrativi e forniti dall'Ufficio non sono sufficienti a spiegare i diversi comportamenti dei soggetti. Per cogliere il ruolo dei comportamenti individuali è stata necessaria una indagine su un campione di iscritti alle liste di Collocamento, dei quali si è appreso, tramite i dati contenuti nelle schede C2, la storia lavorativa passata. In questo modo può essere possibile individuare l'incidenza delle variabili dell'offerta sul percorso lavorativo individuale.

Il modello principale di raccolta dei dati dell'Ufficio di Collocamento è il modello C2<sup>(11)</sup>; questa scheda sintetizza le informazioni personali degli individui iscritti al collocamento o avviati al lavoro senza preiscrizione<sup>(12)</sup>. La popolazione presa in esame è costituita da tutti gli iscritti alle liste di disoccupazione disponibili (disoccupati, part-time, tempo determinato); la sua numerosità al momento della rilevazione era di 3836 soggetti.

La variabile endogena del modello è costituita dagli intervalli di disoccupazione interrotti; poiché nel campione sono presenti individui occupati part-time e a tempo determinato, e l'ufficio di Collocamento continua a considerare tali individui disponibili per una occupazione anche in questi periodi, la lunghezza degli intervalli interrotti di questi soggetti è data dalla somma dei loro periodi di occupazione in corso più la durata dell'ultimo intervallo di disoccupazione se precede immediatamente quello di occupazione in corso.

In questo modello si cerca di stimare la durata media della disoccupazione e di chiarire se l'andamento della distribuzione della durata dipenda dalla durata dell'intervallo in corso oppure dall'eterogeneità dei soggetti. Questa eterogeneità è espressa attraverso un insieme di variabili esplicative che riguardano sia le caratteristiche demografiche dei soggetti del campione, sia la loro storia lavorativa; l'ipotesi che si vuole verificare è se la storia lavorativa dei soggetti, insieme alla durata dell'intervallo

in corso, hanno un peso molto rilevante, e comunque preponderante rispetto alle caratteristiche demografiche, nel determinare l'andamento della distribuzione della durata.

Le variabili rilevate direttamente dalle schede C2, e introdotte nell'analisi come variabili esplicative, sono le variabili demografiche: il sesso, l'età, il comune di nascita, lo stato civile, il numero dei componenti del nucleo familiare, il titolo di studio, la qualifica.

Alcune variabili sopra indicate entrano nell'analisi come dummy. Il sesso è indicato dal valore uno per le femmine e dal valore zero per i maschi; analogamente, per lo stato civile si è attribuito il valore uno ai coniugati e il valore zero agli altri, non coniugati, separati, vedovi; dal comune di nascita si è diviso il campione in tre parti, cioè tra individui provenienti dal sud dell'Italia, dal nord-centro-estero, e individui nati a Modena, ognuna di queste tre variabili viene introdotta come dummy; la qualifica indica la manovalanza, gli impiegati amministrativi, gli impiegati tecnici, gli impiegati comuni e la manodopera generica, anche queste quattro categorie entrano come dummy nell'analisi.

Dalle informazioni registrate dalle C2 e in particolare dalle date di iscrizione, avviamento, licenziamento e cancellazione, è possibile ricostruire tutti gli spostamenti nel mercato del lavoro dei soggetti campionati. Le variabili utilizzate sono le seguenti:

- l'età di ingresso sul mercato, intesa come età anagrafica al momento della prima iscrizione o del primo avviamento;
- l'anzianità di presenza sul mercato calcolata come differenza tra data di rilevazione e data di ingresso sul mercato;
- i periodi di lavoro;
- i periodi di iscrizione, registrati come intervalli di disoccupazione;
- i periodi di non iscrizione a partire da una cancellazione che terminano in una nuova iscrizione o in un avviamento, registrati come periodi di cancellazione;
- i periodi documentati di non iscrizione a partire da un licenziamento che terminano con una nuova iscrizione o in un avviamento, registrati come intervalli di non disponibilità.

Sulla base di queste variabili disaggregate sono stati calcolati il numero e la durata dei singoli periodi di occupazione, disoccupazione e non iscrizione documentata e di tutte le altre variabili utilizzate nell'analisi. In particolare sono state calcolate, e introdotte nel modello come variabili esplicative, per ogni soggetto, le seguenti classi di variabili:

- indicatori di frequenza che permettono di valutare il numero dei rapporti di lavoro per ogni individuo, dei periodi di iscrizione e di non iscrizione documentata;

- indicatori di durata media dei periodi di permanenza nell'occupazione, disoccupazione e non iscrizione, calcolata dividendo i periodi complessivi per il numero dei singoli periodi di occupazione, disoccupazione e non iscrizione;
- indicatori di incidenza, permettono di valutare il peso della durata totale dei periodi lavorativi, di iscrizione e di non iscrizione documentata sull'anzianità di presenza sul mercato.

Di seguito vengono presentate le caratteristiche del campione dei disoccupati iscritti al Collocamento.

La numerosità campionaria è di 352 individui dei quali si hanno sia le caratteristiche di tipo demografico che la storia passata. Il campione è costituito prevalentemente da giovani, 55.4%; da donne, 73%; non coniugati, 65% e da individui che appartengono a nuclei familiari con tre o quattro componenti. Quasi la metà degli iscritti alle liste di disoccupazione non sono nati a Modena, 46%, tra questi il 26.7% proviene dall'Italia meridionale, in particolare il 13.6% dalla regione Campania. In genere, il livello di istruzione è di medio livello, infatti, il 71.9% ha conseguito almeno il diploma di scuola media inferiore ed il 38.1% ha un diploma di scuola superiore, tuttavia, si registra un'elevata presenza di individui con solo la licenza elementare, 28.1%. La maggioranza degli iscritti sono impiegati, 58.6%, gli altri si dividono tra manovalanza (operai), 25% e manodopera generica 16.5%. La descrizione del settore occupazionale è rilevante solo per gli operai, infatti rispetto alla manodopera generica hanno una specializzazione produttiva. I settori occupazionali principali in cui si colloca la manovalanza sono: i servizi, 28.4%, il settore metalmeccanico, 28.4%, il tessile abbigliamento 26.1%; i restanti occupano in modo esiguo gli altri settori.

L'età di ingresso nel mercato del lavoro si concentra tra i dodici e i trenta anni, 84.1%; l'anzianità di presenza sul mercato del lavoro è in generale non molto alta, infatti il 42.3% ha una anzianità inferiore ai tre anni; tuttavia, i disoccupati con una anzianità di presenza sul mercato superiore ai venti anni, non sono una percentuale trascurabile: l'8.2%.

Una parte dell'analisi strutturale del campione essendo riferita alla durata della disoccupazione dell'intervallo interrotto al momento della rilevazione e ad alcune variabili demografiche e della storia passata, verrà presentata insieme al commento dei parametri. Infatti, i risultati della stima spesso sono confermati anche dall'analisi strutturale. Passiamo quindi al commento dei risultati della stima.

#### 4. I risultati della stima

Nella Tabella 1 viene presentata la stima del parametro di forma  $\lambda$  e del parametro di scala  $\alpha$  senza il vettore di variabili esplicative. Tale stima è ottenuta dalla massimizzazione della Verosimiglianza data dalla equazione 2. Nella Tabella 2 oltre a  $\alpha$  e  $\lambda$  vengono aggiunte le variabili demografiche (sesso, titolo di studio, provenienza geografica degli iscritti, stato civile, numero di componenti della famiglia, età) per verificarne la loro significatività indipendentemente dalla storia lavorativa. Nella Tabella 3 vengono introdotte anche le variabili che spiegano la storia lavorativa dei disoccupati (l'anzianità di presenza sul mercato del lavoro, il numero dei periodi di occupazione, di disoccupazione, di cancellazione e di non disponibilità al lavoro, la durata media dell'occupazione, della disoccupazione, della cancellazione e della non disponibilità, l'incidenza della disoccupazione in riferimento all'intera vita lavorativa). Le stime relative alle Tabelle 2 e 3 sono date dalla massimizzazione della Verosimiglianza data dalla equazione 1.

Una prima osservazione riguarda il parametro di forma  $\lambda$ . Esso aumenta man mano che si introducono variabili esplicative, passando da un valore di 0.70 nel caso di stima dei soli due parametri, a 0.95 nel caso si aggiungano le sole variabili demografiche, a 1.13 quando vengono introdotte anche le variabili relative alla storia lavorativa. Ciò significa che se non si tiene conto dell'eterogeneità la probabilità di uscita dalla disoccupazione, il cui andamento è descritto da  $\lambda$ , è decrescente rispetto alla durata. Introducendo le variabili che colgono le differenze individuali la probabilità di uscita mostra un andamento crescente. Tuttavia, il valore di 1.13, che indica una dipendenza positiva dalla durata, pur essendo significativamente diverso da zero (t-statistic = 8.966), non è significativamente diverso da uno, come mostrato dal suo errore standard (E.S.= 0.125). Con un livello di confidenza del 95% il valore di  $\lambda$  è compreso tra 0.875 e 1.375 e quindi la crescita del parametro da 0.7 a 1.13 è sicuramente significativa. In sostanza, si può dire che, introducendo tutte le variabili esplicative, il tasso di uscita non dipende dalla durata infatti, è costante o debolmente crescente.

Un altro modo per verificare la significatività della crescita del parametro dovuta alla introduzione delle variabili esplicative è il test del rapporto di Verosimiglianza relativo a due specificazioni del modello: quella con i soli due parametri  $\lambda$  e  $\alpha$  e quella con tutte le variabili esplicative. Il logaritmo della Verosimiglianza per il primo modello è di - 1258.93 mentre per il secondo modello è di - 1206.52. Il doppio di tale differenza si distribuisce come un  $\chi^2_{22}$ , con tanti gradi di libertà quante sono le variabili esplicative. L'ipotesi nulla da confutare è che i parametri siano uguali a zero, cioè che le variabili esplicative non siano significative. Per rifiutare l'ipotesi nulla con una confidenza del 99% il valore

della differenza dei logaritmi moltiplicata per due deve essere maggiore di 40. Poiché, il doppio della differenza dei logaritmi è 104.82 si può rifiutare l'ipotesi nulla. Questo risultato conferma l'importanza delle variabili esplicative nella determinazione del parametro  $\lambda$ .

È molto ragionevole pensare ad un andamento non decrescente per la probabilità di uscita dalla disoccupazione in ragione della natura dei dati. Trattandosi di dati dell'Ufficio di Collocamento il funzionamento del processo di ricerca di lavoro è diverso da quello presente sul mercato. Chi è in lista d'attesa da più tempo matura un punteggio maggiore per l'assegnazione di un posto di lavoro ed ha, quindi, una maggiore probabilità di uscire dalla disoccupazione. Tuttavia, l'operare di questo meccanismo è frenato dal fatto che le imprese hanno diritto ad usufruire della chiamata nominativa per il cinquanta per cento dei lavoratori da assumere. Probabilmente la combinazione di questi due effetti dà luogo ad una debole dipendenza positiva dalla durata.

Si passi ora al commento dettagliato dei parametri stimati per il modello comprensivo di tutte le variabili: demografiche e della storia passata (Tabella 3).

Per quanto riguarda il sesso sembra che le donne siano più svantaggiate degli uomini in quanto il segno del parametro è negativo, tuttavia l'effetto della variabile nello spiegare la durata della disoccupazione non è statisticamente significativo.

Il parametro della variabile *titolo di studio* è positivo. Questo sta ad indicare che maggiore è il grado di istruzione maggiore è la probabilità di uscita dalla disoccupazione. Tuttavia, anche in questo caso il parametro non è significativo. La bassa significatività di questa variabile inizialmente si è pensato potesse essere attribuita ad un andamento non lineare. Per verificare questa ipotesi si è provato a introdurre delle dummies che cogliessero le differenze fra i vari livelli di istruzione, ma i risultati oltre ad essere statisticamente poco significativi, non sembrano confermare un andamento non lineare.

In sostanza, il sesso e il titolo di studio non sembrano incidere in modo consistente sulla probabilità di uscire dalla disoccupazione; se questo appare poco ragionevole per un mercato del lavoro regolare, può esserlo per l'Ufficio di Collocamento che segue regole non solo di mercato. Un'ulteriore conferma del risultato viene dall'analisi strutturale dei dati, da cui risulta che le donne sono leggermente svantaggiate rispetto agli uomini sulle durate lunghe di disoccupazione (superiore a sei mesi) e per il titolo di studio non si riscontra una relazione chiara che permetta di dire se sia meglio avere un grado di istruzione superiore per uscire prima dalla disoccupazione.

I valori stimati per le variabili relative alla zona di provenienza (nord-centro-estero e sud) indicano, in generale, che il non essere nati a Modena peggiora le probabilità di trovare una occupazione: entrambi i parametri stimati hanno, infatti, un segno negativo. Questo risultato è abbastanza ragionevole se si pensa che per chi proviene da altre città e non conosce il mercato modenese l'Ufficio di Collocamento costituisce spesso l'unico mezzo per trovare una occupazione. Per i nati a Modena la ricerca di un lavoro avviene spesso contemporaneamente sia attraverso le liste dell'ufficio che direttamente sul mercato. Infatti, i nati a Modena, che conoscono in genere abbastanza bene il mercato locale e hanno maggiori contatti diretti con le imprese, rispetto agli immigrati, riescono ad abbreviare i tempi di disoccupazione attraverso la ricerca autonoma. Inoltre, anche dall'analisi strutturale coloro che vengono da fuori Modena si collocano su periodi di disoccupazione superiori a sei mesi con maggiore frequenza. La relazione è confermata per la variabile nord-centro-estero: oltre il 70% di coloro che appartengono a tale categoria hanno una disoccupazione superiore a sei mesi. Ciò concorda con il fatto che nella stima l'influenza di questa variabile risulta significativa. Coloro che provengono dal sud sono anch'essi leggermente più svantaggiati rispetto ai modenesi, dato che il 59.6% di essi hanno disoccupazione superiore ai sei mesi. Tuttavia, lo scostamento dai modenesi è modesto, concordemente con il fatto che il parametro della variabile sud è meno significativo.

I parametri che specificano la diversa *qualifica* risultano tutti poco significativi. Eccezione viene fatta per gli impiegati comuni senza titolo di studio superiore, che risultano fra le categorie più svantaggiate. Anche questi risultati sono confermati dall'analisi strutturale, infatti, non risulta esserci una sostanziale differenza fra categorie professionali rispetto alle durate lunghe di disoccupazione, con l'eccezione degli impiegati comuni che risultano più svantaggiati degli altri (il 65% di questi ha durate superiori a sei mesi).

*Lo stato civile* influenza significativamente il tasso di uscita; gli individui non coniugati hanno maggiori probabilità di uscita di quelli coniugati e il parametro è significativo. Due sono le ragioni che spiegano questo risultato: la prima consiste nel fatto che, in generale, gli imprenditori sono più propensi ad assumere individui senza obblighi familiari, in modo particolare se si tratta di assumere donne, e all'Ufficio di Collocamento si iscrivono principalmente donne. La seconda ragione è che è ragionevole aspettarsi che nei nuclei familiari in cui vi sono donne sposate almeno un altro componente percepisca un reddito in modo continuativo; questo elemento, che costituisce un criterio di ordinamento delle liste, sfavorisce i coniugati. Inoltre gli individui sposati, il cui coniuge percepisce già un reddito, spesso non sono disposti ad

accettare qualunque tipo di lavoro che l'ufficio propone in quanto possono "permettersi" di prolungare l'attesa per un lavoro desiderato contando sul reddito del coniuge. Dall'analisi strutturale risulta che quasi il 70% dei coniugati ha durate di disoccupazione superiori ai sei mesi, risultato che conferma quanto trovato con la stima.

Anche la variabile *numero di componenti della famiglia* risulta essere importante per la determinazione del tasso di uscita, infatti individui che appartengono a famiglie numerose hanno maggiori probabilità di uscita. Questo risultato è in linea con la logica delle graduatorie del Collocamento. Infatti, una delle variabili di cui si tiene conto quando si devono dare dei punteggi agli iscritti è proprio il numero di componenti della famiglia. Anche dall'analisi strutturale il risultato è confermato: maggiore è il numero di componenti minore è la durata di disoccupazione. Tra coloro che hanno tre o meno componenti i due terzi (108 su 167) hanno una disoccupazione superiore a sei mesi. Invece, gli individui che vivono in nuclei familiari di tre o più componenti solo la metà (93 su 185) ha disoccupazione di lungo periodo. Questa variabile costituisce una proxy dell'effetto di altre due variabili di non facile rilevazione: il numero di componenti a carico e i redditi familiari. Anche per queste variabili valgono i ragionamenti fatti per lo stato civile.

Per l'*età*, il cui parametro è significativamente diverso da zero, ci si aspettava l'andamento classico a campana, cioè che per i giovani e gli anziani la probabilità di uscita dalla disoccupazione fosse più alta che per gli individui delle fasce centrali di età. Per questo motivo si era costruita una dummy che fosse in grado di cogliere questo aspetto; tuttavia, il risultato sembra suggerire un andamento lineare della variabile. Infatti, il parametro dei giovani risulta negativo, mentre quello degli anziani positivo, anche se poco significativo. Quindi anche se qualche dubbio sulla non linearità rimane, la variabile non dimostra di essere fortemente non lineare tanto da far risultare negativo il coefficiente riferito agli anziani.

Si passa ora all'analisi delle variabili che descrivono la storia lavorativa passata. La *anzianità di presenza sul mercato del lavoro* (anzianità di iscrizione al Collocamento, costruita tenendo presente la prima data di iscrizione alle liste di disoccupazione) risulta significativa. Questa variabile è la anzianità al momento dell'ingresso nello stato di disoccupato, cioè depurata dalla durata dell'ultimo intervallo così da non creare problemi di collinearità. Il segno del parametro è negativo, questo significa che maggiore è l'anzianità di presenza nelle liste di Collocamento più bassa è la probabilità di uscita dalla disoccupazione. Molto



ragionevolmente, se un individuo è sempre fuori e dentro da una occupazione e quindi ha una anzianità alta di presenza nelle liste (che non significa necessariamente che sia un anziano di età) è probabilmente uno svantaggiato.

Un'altra variabile molto significativa è l'*incidenza della disoccupazione* sofferta sulla anzianità di presenza sul mercato. Anche questo parametro ha il segno negativo, cioè indica che maggiore è l'incidenza della disoccupazione, minore è la probabilità di uscita. Se un individuo è stato molto tempo nella disoccupazione rispetto alla sua intera storia lavorativa è sicuramente più svantaggiato. Dall'analisi strutturale risulta che coloro che hanno una incidenza superiore al 60% si collocano principalmente su durate di disoccupazione lunghe, mentre coloro che hanno incidenza inferiore al 60% si ripartiscono equamente fra disoccupazione di lungo e di breve.

Sono state introdotte altre variabili della storia passata che ci si riteneva avessero un effetto sulla durata: *il numero dei periodi di disoccupazione, occupazione, cancellazione e non disponibilità e le rispettive durate medie di permanenza nei diversi stati*. Tuttavia, queste variabili dalla stima non risultano avere un peso sulla determinazione della attesa nelle liste. Per questo motivo non verranno trattati in questo contesto.

Come si è detto precedentemente si è stimato il modello degli intervalli interrotti come fossero completi (Tabella 4) per avere una misura dell'errore che si commette quando si stima la durata della disoccupazione utilizzando intervalli ancora in corso. Tuttavia, poiché si trova una dipendenza quasi nulla dalla durata, cioè un  $\lambda$  quasi uguale ad uno ed una distribuzione di probabilità esponenziale, si è quasi certi di non commettere gravi errori nella stima. La distribuzione degli intervalli interrotti e quella degli intervalli completi sono uguali.

Per calcolare la durata media della disoccupazione sulla base degli intervalli interrotti con le variabili esplicative l'equazione è la numero 3) del paragrafo 2. Assumendo che  $\lambda$  sia uguale ad uno, come ad esempio per una distribuzione di probabilità esponenziale, la durata media diventa il reciproco del parametro di scala moltiplicata per il vettore di variabili esplicative, cioè

$$5) \quad \mu = \frac{1}{\alpha} e^{-x\beta}.$$

Attraverso questa espressione si può calcolare la durata media di particolari categorie di individui, ricordando di sostituire al parametro  $\alpha$  il valore trovato dal modello con tutte le variabili esplicative (Tabella

3). Si ricorda, inoltre, che la media in questo particolare caso di distribuzione esponenziale ( $\lambda = 1$ ) è uguale sia per i completi che per gli interrotti. Invece, se si vuole trovare la durata media generale degli iscritti al Collocamento senza particolari precisazioni sulle caratteristiche demografiche o della storia lavorativa, si segue l'equazione 5) del paragrafo 3. I valori dei parametri da utilizzare in questo caso sono quelli del modello senza variabili esplicative (Tabella 1).

La durata media, calcolata con l'equazione 4, per un individuo generico è di poco più di 13 mesi.

Sulla base dell'equazione 3 si sono calcolate le durate medie di disoccupazione per alcune tipologie di iscritti al Collocamento.

L'iscritto "tipo" presenta le seguenti caratteristiche: è una donna, al di sotto dei trenta anni, non coniugata, di Modena, con una istruzione di grado superiore, che vive in un nucleo familiare di quattro componenti. La durata media di questo individuo è di circa 9 mesi. Un uomo con le stesse caratteristiche ha una durata media di 8 mesi. Questo risultato concorda con le indicazioni date dall'analisi strutturale: il sesso non è un elemento che influisce significativamente sul tasso di uscita dalla disoccupazione.

Un individuo con caratteristiche analoghe a quelle dell'iscritto tipo, ma in età matura ha tempi di attesa inferiori; infatti, la durata media è di circa 6 mesi. Conformemente ai risultati dell'analisi strutturale la provenienza, lo stato civile e la dimensione del nucleo familiare incidono molto sulla durata. Infatti, un individuo non di Modena, oppure che vive da solo, ha una durata di circa 13 mesi; per i coniugati la durata sale a 18 mesi e mezzo.

Nel campione vi sono anche tipologie di individui con durate particolarmente lunghe: ad esempio, una donna, giovane, coniugata, senza figli, non di Modena e con un grado di istruzione medio, attende un lavoro per circa 40 mesi.

## Tabelle dei risultati

Stime di Massima Verosimiglianza del tasso di uscita dalla disoccupazione per un campione di intervalli interrotti (modello Weibull; stima di  $\alpha$  e  $\lambda$ ).

---

Parametri	Stime	Errore Standard	t-statistic
$\alpha$	0.14803	0.03192	4.6366
$\lambda$	0.70203	0.04899	14.328

Logaritmo della Verosimiglianza = -1258.93

---

Tab. 1

Stime di Massima Verosimiglianza del tasso di uscita dalla disoccupazione per un campione di intervalli interrotti (modello Weibull; stime dei parametri  $\alpha$  e  $\lambda$ , variabili demografiche).

---

Variabili	Param.	Errore Standard	T-Statistic
femmine	-0.10319	0.12968	-0.79568
titolo studio	0.13287	0.07369	1.80310
nord-cent-est	-0.40033	0.14712	-2.72110
sud	-0.12012	0.12767	-0.94089
n.comp.famigl.	0.11575	0.04153	2.78690
stato civile	-0.72618	0.15131	-4.79930
giovani	-0.31282	0.14006	-2.23350
anziani	0.17655	0.12547	0.70706
$\alpha$	0.08178	0.02833	2.88650
$\lambda$	0.95286	0.09674	9.85000

Logaritmo della Verosimiglianza = -1227.29

---

Tab. 2

Stime di Massima Verosimiglianza del tasso di uscita dalla disoccupazione per un campione di intervalli interrotti (modello Weibull; stime dei parametri  $\alpha$  e  $\lambda$ , variabili demografiche e della storia lavorativa.)

Variabili	Param.	Errore Standard	T-Statistic
femmine	-0.04543	0.18735	-0.24249
titolo studio	0.03821	0.17286	0.22108
nord-cent-est	-0.43825	0.20691	-2.11800
sud	-0.18645	0.17323	-1.07630
man. generica	0.02043	0.21557	0.09480
impiegati amm.	0.18000	0.36065	0.49908
impiegati tec.	0.43831	0.41648	1.05240
impiegati com.	-0.06300	0.24458	-0.25752
stato civile	-0.69664	0.19436	-3.58440
n.comp.famigl.	0.15641	0.05527	2.82950
anz.sul m.d.l	-0.00368	0.00143	-2.16740
incidenza dis	-1.23070	0.34588	-3.55830
n.period.occ.	-0.10346	0.10376	-0.99707
n.period.dis.	-0.28175	0.07658	-0.36791
n.period.n.d.	0.16257	0.10109	1.60820
n.period.canc.	0.01279	0.09374	0.13649
media occup.	-0.00012	0.00011	-1.04630
media disocc.	0.00026	0.00026	0.98035
media non disp	0.00005	0.00015	0.34920
media cancell.	-0.00026	0.00022	-1.17120
giovani	-0.41129	0.18213	-2.25830
anziani	0.35552	0.31603	1.12500
$\alpha$	0.16134	0.08253	1.95470
$\lambda$	1.12500	0.12547	8.96620

Logaritmo della Verosimiglianza = -1206.52

Tab. 3

Stime di Massima Verosimiglianza del tasso di uscita dalla disoccupazione per un campione di intervalli interrotti trattati come completi (modello Weibull; stime dei parametri  $\alpha$  e  $\lambda$ , variabili demografiche e della storia lavorativa.)

Variabili	Param.	Errore Standard	T-Statistic
femmine	-0.04621	0.16563	-0.27897
titolo studio	0.04286	0.15000	0.28576
nord-cent-est	-0.44395	0.17908	-2.47910
sud	-0.18705	0.15138	-1.23560
man. generica	0.02405	0.19075	0.12609
impiegati ann.	0.17898	0.31867	0.56163
impiegati tec.	0.44008	0.36792	1.19610
impiegati com.	-0.06310	0.21490	-0.29363
stato civile	-0.70513	0.15332	-4.59910
n.comp.famigl.	0.15756	0.04683	3.36450
anz.sul m.d.l	-0.00368	0.00126	-2.92040
incidenza dis	-1.24390	0.28552	-4.35660
n.period.occ.	-0.10425	0.09097	-1.14590
n.period.dis.	-0.02827	0.06753	-0.41863
n.period.n.d.	0.16354	0.08745	1.87000
n.period.canc.	0.01237	0.08290	0.14923
media occup.	-0.00012	0.00009	-1.22060
media disocc.	0.00026	0.00023	1.12970
media non disp	0.00005	0.00013	0.38973
media cancell.	-0.00026	0.00019	-1.35440
giovani	-0.41698	0.15657	-2.66320
anziani	0.35964	0.27689	1.29890
$\alpha$	0.17791	0.07660	2.33490
$\lambda$	1.13270	0.05104	22.19500

Logaritmo della Verosimiglianza = -1203.16

Tab. 4

## Conclusioni

Per gli iscritti alle liste di Collocamento della provincia di Modena la probabilità di uscita dalla disoccupazione, qualora si tenga conto dell'eterogeneità, non dipende negativamente dalla durata. Ciò concorda con i risultati di Dell'Aringa Samek Lodovici (1988) per un campione di disoccupati iscritti all'Ufficio di Collocamento della provincia milanese. Il fatto che la probabilità di uscita non dipenda negativamente dalla durata è probabilmente dovuto all'operare del meccanismo delle liste di attesa del Collocamento: chi è in coda da più tempo matura un punteggio più alto<sup>(13)</sup>. Questa ipotesi è confortata dal lavoro di Torelli e Trivellato (1988), per la regione Lombardia, e di Lalla (1988), per la regione Emilia Romagna, in cui si utilizzano dati Istat. Questi lavori danno un'idea di cosa accade al tasso di uscita se si considera il mercato del lavoro esterno al Collocamento. La probabilità di uscita da loro stimata risulta essere decrescente al crescere della durata di disoccupazione sofferta dagli individui.

In Emilia Romagna i tempi di attesa dei disoccupati sembrano essere relativamente brevi. Nel lavoro di Lalla, sui disoccupati "in senso stretto", risulta che la durata media di attesa è di circa 7 mesi. Nel nostro lavoro la durata media della disoccupazione è di circa tredici mesi. La differenza dipende certamente dal fatto che tra gli iscritti da noi rilevati una quota consistente è costituita da giovani in cerca di prima occupazione che hanno normalmente tempi di attesa più lunghi. Nonostante la durata media sia breve, vi sono persone iscritte all'ufficio che presentano tempi di attesa notevolmente più lunghi della media: donne giovani, coniugate, senza figli, di media istruzione, provenienti da fuori Modena aspettano circa quaranta mesi.

Le variabili che incidono molto sulla durata della disoccupazione sono la provenienza, lo stato civile e il numero di componenti della famiglia. La provenienza discrimina fra individui che conoscono il mercato locale e possono avere un rapporto diretto con i datori di lavoro e chi non conoscendo il mercato non è in grado di operare una ricerca di lavoro personalmente. I coniugati hanno tempi di attesa maggiori forse perchè non essendo gli unici percettori di reddito hanno maggiore possibilità di attendere il lavoro più consono alle loro capacità. Il numero di componenti della famiglia è rilevante perchè se si hanno familiari a carico si sale nelle graduatorie di assegnazione dei posti di lavoro.

Le variabili demografiche come il sesso, il titolo di studio e la qualifica non sembrano giocare un ruolo significativo nella determinazione della durata di disoccupazione. Ciò contrasta con i risultati di Dell'Aringa e Samek Lodovici che trovano, per l'area milanese, differenze molto considerevoli tra uomini e donne. La storia lavorativa passata, costituita da anzianità di presenza sul mercato e incidenza della disoccupazione sull'anzianità, ha un peso notevole nel determinare la durata come trovato da Gay (1988) per la Lombardia.

Il modello da noi utilizzato è di tipo parametrico e all'interno di questi modelli è il più maneggevole dal punto di vista operativo come conferma il suo vasto utilizzo nella letteratura. Tuttavia, si basa sull'ipotesi semplificatrice che l'andamento dei tassi di uscita per le diverse tipologie di lavoratori sia proporzionale; per questo motivo i risultati sopra presentati devono essere considerati con una certa cautela. Un modello alternativo e più generale di stima della durata è quello di azzardo accelerato, che tuttavia non trova grande applicazione nella letteratura empirica perlomeno in Italia. Inoltre, assumere che la funzione di densità di probabilità degli intervalli completati di disoccupazione sia una Weibull, come nel nostro caso, o esponenziale o log-logistica come in altri esempi di trattazione dei modelli di durata, è anch'essa una forte semplificazione. Nella letteratura sull'argomento si trovano proposte di analisi di tipo non parametrico, cioè di ricerca della densità di probabilità partendo dai dati, ma anche in questo caso l'applicazione empirica di questi metodi non parametrici oltre ad essere molto complessa non è molto ricca di esempi operativi. Un altro limite del modello di azzardo proporzionale sta nell'assumere intervalli singoli per ogni individuo. Il procedimento corretto sarebbe quello di utilizzare un modello ad intervalli multipli in cui per ogni individuo si segue in modo esatto e completo il percorso lavorativo e quindi ad ogni individuo corrisponde più di un intervallo.

### Note

(1) Nei modelli ingegneristici l'analisi di sopravvivenza è usata per studiare il tasso di usura di certi materiali; nelle discipline biomediche, invece, è usata per valutare la probabilità di sopravvivenza di certi pazienti con determinate caratteristiche patologiche. Sull'analisi di sopravvivenza negli anni '50 del Tavistock Institute of Human Relation si vedano in particolare i lavori di Rice, Hill e Trist (1950), Hill (1951), Rice (1951, 1952), Rice e Trist (1952). Inoltre, una rassegna di studi sul turnover aziendale è contenuta in Silcock (1954) ed Hedberg (1961). Per un approfondimento dello strumento statistico si veda in



- particolare Cox (1967), Cox e Oakes (1984). Alcuni importanti riferimenti bibliografici sull'analisi della durata della disoccupazione attraverso l'analisi di sopravvivenza per gli anni '70 e '80 sono Kaitz (1970), Cripps e Tarling (1974), Salant (1977), Frank (1978), Clark e Summers (1979), Akerlof e Main (1980, 1981, 1983), Main (1981, 1982), Summers (1981, 1986), Carlson e Horrigan (1983), Bowers e Horvath (1984), Butler e McDonald (1986).
- (2) In generale, nel modello Job Search per i lavoratori in cerca di una occupazione maggiore è la durata della disoccupazione più è probabile che gli individui incontrino un'impresa che offre un salario uguale al loro salario atteso, perchè le informazioni ottenute con la ricerca attiva accelerano il processo di aggiustamento delle aspettative dei lavoratori alla distribuzione effettiva dei salari offerti. Anzi solitamente il loro salario atteso diminuisce col tempo e questo implica che il processo di ricerca si accorci. Sulla letteratura che riguarda la Job Search si veda Mortensen (1970, 1977, 1990), Lynch (1983), Narendranathan e Nickell (1984), Newmann e Westergard-Nielsen (1984), Pissarides (1985).
- (3) In particolare, si veda Kaitz (1970), Perry (1972), Salant (1977).
- (4) Una spiegazione di questo fenomeno può essere individuata nell'atteggiamento degli imprenditori nei confronti di disoccupati di medio e lungo periodo. Gli imprenditori possono pensare infatti, che con l'aumentare della durata della disoccupazione i lavoratori subiscano una perdita di qualifica, una perdita di motivazione a lavorare e una generale perdita di produttività. In più nell'imprenditore può sorgere il dubbio che il periodo di disoccupazione si sia protratto così a lungo perchè altri imprenditori hanno rifiutato di assumere quell'individuo dopo averne appurato le scarse qualità lavorative.
- (5) Queste osservazioni già avanzate da Silcock (1954) vengono riprese da Salant (1977), Lancaster (1979, 1985), Nickell (1979), Heckman e Borjas (1980), Lancaster e Nickell (1980), Heckman (1981), Heckman e Singer (1985), Kiefer (1988), Kiefer e Neumann (1989).
- (6) In riferimento al modello di azzardo proporzionale si veda Lancaster (1979), Nickell (1979), Lancaster e Nickell (1979), Elbers e Ridder (1982), Flinn (1986), Dell'Aringa e Samek-Lodovici (1988), Lalla (1990), Solinas (1990).
- (7) Il tasso di uscita nella letteratura recente è comunemente chiamato hazard rate, mentre negli anni '70 è chiamato più frequentemente escape rate. Nella letteratura degli anni '50 e '60 lo si trova spesso come leaving rate, ad esempio in Lane ed Andrew (1955) e Bartholomew (1959) o addirittura come rate of wastage, in Silcock (1954). Quest'ultima denominazione lascia intravedere l'origine ingegneristica del termine, infatti come già sottolineato precedentemente gli ingegneri usavano questo "tasso di spreco" per valutare i guasti dei materiali edili.
- (8) In questo articolo presentiamo le stime del modello con distribuzione Weibull riferito ai soli intervalli di disoccupazione interrotti al momento della rilevazione e non ancora terminati. Tuttavia, il campione consente anche la stima del modello su intervalli completati e interrotti, nel quale si modella la distribuzione degli interrotti sulla base della distribuzione di durata di quelli completati. I risultati di questa ulteriore indagine sono presentati nella nostra tesi di laurea: "Mercato del lavoro ed analisi di flusso" (1991). La scelta di presentare

solo il modello su intervalli interrotti dipende dal fatto che si è osservato che la distribuzione degli intervalli completati è molto diversa da quella che si otterrebbe dagli intervalli interrotti una volta completati; l'uguaglianza tra queste due distribuzioni è fondamentale per poter utilizzare il modello su intervalli completati e interrotti. Una possibile spiegazione di tale differenza va ricercata nel fatto che tra i periodi in cui si sono realizzati gli intervalli di disoccupazione completati rispetto a quelli in corso, sono intervenute modificazioni notevoli del mercato del lavoro. Inoltre, le stime sono state fatte usando un'altra distribuzione: la Log-Logistica. I risultati non si discostano molto da quelli ottenuti con la Weibull.

(9) La funzione di densità per gli intervalli interrotti è diversa da quella per gli intervalli completati:

per la distribuzione Weibull si ha

$$g(x) = \frac{\alpha\lambda}{\Gamma\left(\frac{1}{\lambda}\right)} e^{-(\alpha x)^\lambda}$$

ciò è stato dimostrato da Salant (1977); si veda anche Heckman e Singer (1985), Flinn (1986) e Lalla (1990).

(10) Parlando di storia lavorativa è necessario premettere che questa è determinata secondo le definizioni adottate dall'Ufficio di Collocamento e non dalle definizioni adottate in sede di indagine delle forze di lavoro dell'ISTAT. Infatti, le due definizioni non coincidono completamente. Gli iscritti al Collocamento non corrispondono a tutte le persone in cerca di occupazione secondo la definizione dell'ISTAT, viceversa non tutti gli iscritti sono ufficialmente in cerca di occupazione. Il concetto di disoccupato per il Collocamento non coincide con il disoccupato delle rilevazioni trimestrali, poiché essere disponibile, per l'Ufficio, significa solamente ricercare un'occupazione nel settore privato. Gli altri settori (lavoro autonomo, pubblica amministrazione ed altre categorie) non vengono considerati dall'Ufficio. D'altra parte gli individui non occupati attraverso il Collocamento non possono essere assimilati alle non forze di lavoro, poiché una quota di questi svolge sicuramente una attività lavorativa non colta dal Collocamento ed una altra quota può essere alla ricerca attiva di occupazione (disoccupato secondo l'ISTAT) senza essere iscritto. Tenendo presente questa distinzione i soggetti iscritti all'Ufficio di Collocamento possono essere distinti in

- occupato, soggetto avviato al lavoro e non ancora licenziato,
- disoccupato, soggetto iscritto come disponibile al lavoro,
- non iscritto documentato, individuo che mostra, per l'Ufficio di Collocamento, una apparente autoesclusione dal mercato.

(11) L'Ufficio utilizza un'altra scheda riassuntiva dei dati dei singoli soggetti: la CNSCRIZIONE. Tuttavia, tale modello non è disponibile per tutti i soggetti registrati, in quanto è stata inserita con la riforma degli Uffici di Collocamento del 1987. Essendo una scheda da compilarsi al momento della prima iscrizione, esiste solo per quei soggetti iscritti per la prima volta dopo l'anno 1987. Poiché lo scopo della raccolta dei dati è quello di avere più informazioni possibili sulla intera vita lavorativa dei soggetti, il modello CNSCRIZIONE non è stato utilizzato, poiché non è disponibile per tutti gli iscritti e di conseguenza avrebbe reso disomogeneo il numero delle variabili raccolte per ogni soggetto.

(12) Un individuo può essere avviato al lavoro senza preiscrizione nella disoccupazione; se egli trova autonomamente un'occupazione lo comunica all'Ufficio e gli viene segnata come prima data di iscrizione la data di inizio dell'occupazione.

(13) Va detto che il meccanismo delle liste di attesa è mitigato in parte dalla chiamata nominativa, concessa alle imprese, che impedisce vi sia una significativa dipendenza positiva dalla durata.

(14) Per approfondire l'analisi non parametrica si veda Liu e Ryzin (1985).

### **Bibliografia**

AKERLOF G.A. e MAIN G.M. [1980], "Unemployment Spells and Unemployment Experience", *The American Economic Review*, vol. 70, n. 5, pp. 885-93.

— [1981a], "An Experience-Weighted Measure of Employment and Unemployment Duration", *The American Economic Review*, vol. 71, n. 5, pp. 1003-10.

— [1983], "Measures of Unemployment Duration as Guides to Research and Policy: Reply", *The American Economic Review*, pp. 1151-1152.

BARTHOLOMEW D.J. [1959], "Note on the Measurement and Prediction of Labour Turnover", *Journal of Royal Statistical Society*, vol. 122, serie A, pp. 232-39.

BOWERS e HORVATH F.W. [1984], "Keeping Time: An Analysis of Errors in the Measurement of Unemployment Duration", *Journal of Business and Economic Statistics*, vol. 2, n. 2, aprile, pp. 140-149.

BUTLER R. J. e McDONALD J. B. [1986], "Trends in unemployment duration data", *The Review of Economics and Statistics*, Nov. n. 4, pp. 345-557.

CARLSON J.A. e HERRIGAN M.W. [1983], "Measures of Unemployment Duration as Guides to Research and Policy: Comment", *The American Economic Review*, Dec., pp. 1143-1150.

CLARK K.B. e SUMMERS L.H. [1979], "Labor Market Dynamics and Unemployment: A Reconsideration", *Brooking Papers on Economic Activity*, 1, pp. 13-72.

COX D.R. [1967], *Renewal Theory*, ed. by Methuen and Co., Ltd., London.

— e OCKES [1984], *Analysis of survival data*, London: Chapman and Hall.

CRIPPS T.F. e TARLING R.J. [1974], "An Analysis of the Duration of Male Unemployment in Great Britain 1932-73", *Economic Journal*, vol. 84, June, pp. 289-316.

DELL'ARINGA C. e SAMEK LODOVICI M. [1988], "Determinants of Unemployment Duration for Displaced Workers", *Labor* 2 (3), pp.91-111.

DE STAVOLA E POLI [1985], "Un'analisi longitudinale dell'occupazione giovanile", *Statistica*, n. 3.

ELBERS C. e RIDDER G. [1982], "True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model", *Review of Economic Studies*, vol. 49, pp. 403-10.