

Lavoro e disoccupazione: questioni di misura e di analisi

Progetto di ricerca cofinanziato dal Ministero per l'Università
e la Ricerca Scientifica e Tecnologica - Assegnazione: 1998
Coordinatore: Ugo Trivellato

**Alcuni modelli per l'analisi delle durate
complete e incomplete della
disoccupazione: il caso Emilia Romagna**

Michele Lalla*, Francesco Pattarin**

* *Dip. di Economia Politica, Univ. di Modena*

** *Dip. Di Economia Pubblica, Univ. Di Roma*

Working Paper n. 11

maggio 1999

Unità locali del progetto:

Dip. di Economia Politica, Univ. Di Modena

Dip. di Economia "S. Cognetti De Martiis", Univ. di Torino

Dip. Di Statistica, Univ "Ca' Foscari" di Venezia

Dip. di Metodi Quantitativi, Univ. di Siena

Dip. di Scienze Statistiche, Univ. di Padova

(coord. Michele Lalla)

(coord. Bruno Contini)

(coord. Tommaso Di Fonzo)

(coord. Achille Lemmi)

(coord. Ugo Trivellato)

Dip. di Scienze Statistiche
via S. Francesco 33, 35121 Padova

1. Introduzione¹

La rilevazione delle forze di lavoro, condotta dall'Istituto Centrale di Statistica (ISTAT), si attua con un'indagine trasversale rotata, ripetuta ogni trimestre, le cui unità campionarie sono le famiglie e lo schema di rotazione è 2-2-2. Per ogni componente della famiglia in età lavorativa si rilevano: la condizione professionale, le notizie sull'attività lavorativa, la durata della disoccupazione in corso, e le caratteristiche individuali che potrebbero influenzarla (Falorsi, Falorsi, 1994). La durata osservata risulta così *incompleta* perché la data di fine della disoccupazione è ignota al momento della rilevazione (*troncamento a destra*). Lo schema di rotazione delle famiglie nel campione permette di ricostruire un insieme di durate *complete* perché considerando più indagini successive si possono individuare quei periodi di disoccupazione di cui si conoscono le date di inizio e fine. I dati così rilevati presentano comunque alcune peculiarità: scarsa presenza dei periodi brevi e numerosità elevata dei periodi lunghi (distorsione campionaria delle durate); troncamento a destra; errori di memoria, quando non vengono riportati eventi accaduti; e effetti telescopici, quando gli eventi sono riferiti a date sbagliate sicché la distribuzione di frequenza delle durate mostra picchi anomali in certi valori (Torelli, Trivellato, 1993). La Figura 1 illustra un'indagine trasversale condotta al tempo τ e alcune durate della disoccupazione che presentano caratteristiche tipiche dei periodi osservabili: t_1 indica un'entrata diretta nel mercato del lavoro; t_2 rappresenta un periodo completo di disoccupazione che non può essere rilevato con un'indagine trasversale, ma con una indagine *panel* o retrospettiva; $\{t_3, t_7, t_9\}$ sono durate "catturate" da un'indagine trasversale e ancora in corso di svolgimento (t_7 rappresenta il caso raro in cui il periodo termina esattamente nella data di rilevazione τ); $\{t_4, t_5, t_6\}$ sono durate di cui non si conoscono le date di inizio e fine, di inizio, e di fine, rispettivamente; t_8 denota il caso, anch'esso raro, in cui il periodo di disoccupazione inizia esattamente in τ .

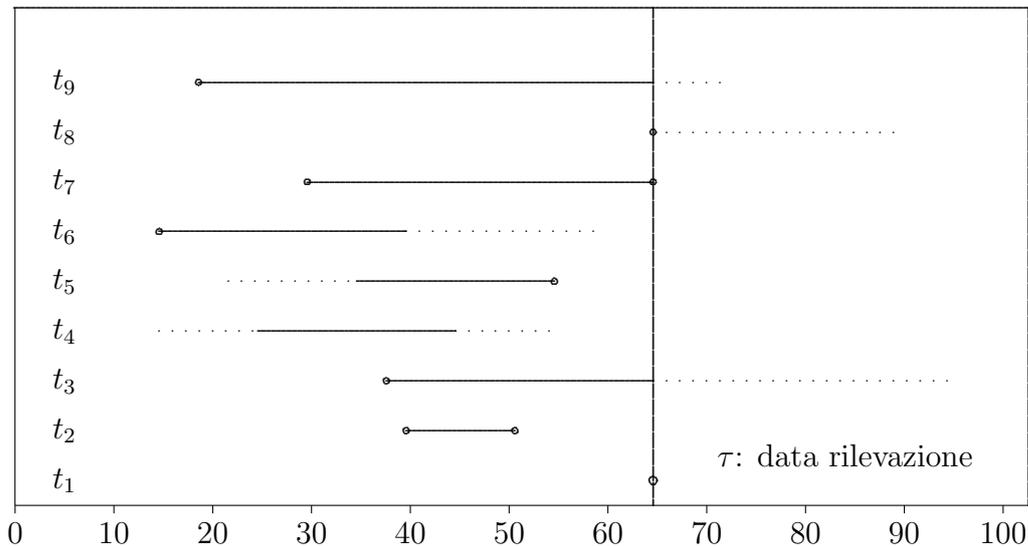
I metodi di analisi che consentono di modellare agevolmente i dati affetti da tali fenomeni utilizzano la funzione di azzardo e sono oramai consolidati nella letteratura (Lancaster, 1990). La costruzione dei modelli e la stima dei relativi parametri presentano, tuttavia, ancora aspetti problematici quando i periodi di disoccupazione sono tutti incompleti, tipici di un'indagine trasversale qual è la rilevazione trimestrale delle forze di lavoro.

In questo elaborato si descrive un procedimento per modellare proprio un insieme di durate incomplete in presenza di variabili esplicative, dove si applica in un contesto più esteso l'approccio già utilizzato, tra gli altri, da Salant (1977) e Flinn (1986) per analizzare la loro distribuzione in gruppi omogenei senza includere quindi le caratteristiche individuali. Si presentano, poi, anche i risultati dell'analisi dei dati delle indagini delle forze di lavoro raccolti dall'ISTAT nella re-

¹Il lavoro è il risultato della collaborazione dei due autori. Si precisa, tuttavia, che i paragrafi 2.2, 2.3, 2.4, 2.6, 3.2, 3.3, 3.3.1 sono stati redatti da Michele Lalla e i paragrafi 1, 2.1, 2.5, 3.1, 3.3.2, 4 sono stati redatti da Francesco Pattarin. Si ringrazia, inoltre, Giovanni Solinas per i preziosi suggerimenti che hanno contribuito a rendere il lavoro più chiaro e preciso in alcuni punti. Vale il solito *caveat*: gli errori e le omissioni si devono attribuire esclusivamente agli autori.

gione Emilia-Romagna, dal primo trimestre del 1993 al primo trimestre del 1995. In particolare, si stimano i parametri dei modelli di durata per ogni trimestre (in cui si hanno solo periodi incompleti di disoccupazione) e per l'insieme dei periodi completi (ricostruiti sfruttando lo schema di rotazione dell'indagine) con l'obiettivo di appurare le modifiche che intervengono nei parametri della distribuzione e delle possibili determinanti. Si possono così verificare la stabilità dei coefficienti nel tempo, le differenze che intercorrono tra i modelli relativi alle durate incomplete e complete, e anche gli eventuali effetti della stagionalità che possono manifestarsi quando si considerano solo gli insiemi di periodi incompleti. In linea di principio, i periodi incompleti dei diversi trimestri potrebbero essere cumulati in un unico campione, ma il cumulo non si può eseguire perché i campioni non sono indipendenti e per lo stesso motivo non si possono cumulare i periodi completi e incompleti; si possono cumulare i periodi incompleti dell'ultimo trimestre con i periodi completi, ma si tralascia per brevità.

Figura 1. Durate della disoccupazione rilevabili con un'indagine trasversale



In questo ambito si assume che la durata della disoccupazione sia distribuita secondo una Weibull perché l'assunto è supportato dai dati. Nell'introduzione dell'eterogeneità non osservata, poi, occorre individuare una distribuzione per la variabile casuale che la rappresenta: se si assume che sia ancora una Weibull, si ottengono funzioni di densità per le durate complete e incomplete che sembrano interessanti e somigliano, tra l'altro, a una distribuzione logistica. Si stimano, perciò, anche i parametri dei modelli ottenuti da queste due distribuzioni per accertare come cambia la struttura delle relazioni tra le determinanti delle durate.

L'esposizione continua secondo lo schema seguente. Nel paragrafo 2 si descrivono i modelli utilizzati nell'analisi delle durate sia complete (§2.1), sia incomplete (§2.2), che assumono rischi proporzionali senza l'eterogeneità non osservata; poi, si considerano l'introduzione dell'eterogeneità non osservata assumendo una distribuzione di Weibull per il parametro di scala della distribuzione della

durata (§2.3), un modello a rischi non proporzionali derivato dalla distribuzione log-logistica (§2.4), l'introduzione dell'effetto telescopico (§2.5), e la funzione di verosimiglianza 'generalizzata' per i vari modelli (§2.6). Nel paragrafo 3 si illustrano i risultati dell'analisi dei dati empirici relativi ai diversi insiemi di durate e ai diversi modelli considerati. Nel paragrafo 4 seguono le conclusioni.

2. L'analisi delle durate complete e incomplete

2.1. Per rappresentare a livello individuale il fenomeno della disoccupazione come un processo che avviene nel tempo, si assume che a un certo istante t_0 , che senza perdita di generalità si può porre uguale a zero, un soggetto i entra nella condizione di disoccupato (*stato*). Le caratteristiche personali (quali sesso, età, titolo di studio, stato civile), famigliari (quali reddito, numero dei figli, numero di occupati, numero di componenti), e lo stato del mercato del lavoro locale (quali tasso di disoccupazione, tasso di partecipazione) sono indicate con un vettore riga $\mathbf{x}(t)$ e denominate regressori o covariate. Per ogni $t > t_0$ si assume, inoltre, che il soggetto i compia azioni di ricerca di lavoro che generano una probabilità, $\omega[\mathbf{x}(t)]$, di ricevere un'offerta di lavoro nell'intervallo di tempo $(t, t + dt)$. Sia $P[\mathbf{x}(t)]$ la probabilità che tale offerta di lavoro venga accettata; allora la probabilità di transizione dallo stato di disoccupato allo stato di occupato nel periodo $(t, t + dt)$ è data da $\omega[\mathbf{x}(t)] P[\mathbf{x}(t)] = h[\mathbf{x}(t)] dt$; dove $h[\mathbf{x}(t)]$ è la funzione di rischio o di azzardo (*hazard function*), nel caso in cui vi sia un solo stato di destinazione, che indica la probabilità di uscita dallo stato di disoccupato nell'unità di tempo successivo a t (Cox, Oakes, 1984).

Il tempo intercorso tra l'ingresso e l'uscita dallo stato (*durata*) è una variabile casuale non negativa, T . Sia $f(t)$ la sua funzione di densità. La variabile osservata, invece, è $T_0 = \min(T, U)$, dove U indica la durata trascorsa nello stato per i soggetti ancora disoccupati al momento dell'indagine: per le durate non completate si ha $T > U$. Se i regressori sono invarianti nel tempo o vengono misurati al momento dell'ingresso nello stato, allora $\mathbf{x}(t) = \mathbf{x}$ e la funzione di azzardo, $h(t; \mathbf{x})$, per le durate complete è uguale a

$$h(t; \mathbf{x}) = \frac{f(t; \mathbf{x})}{1 - F(t; \mathbf{x})} = h_0(t; \boldsymbol{\theta}) \phi(\mathbf{x}; \boldsymbol{\beta}) \quad (1)$$

dove $f(t; \mathbf{x})$ e $F(t; \mathbf{x})$ indicano la funzione di densità e di ripartizione di T , rispettivamente; $h_0(t; \boldsymbol{\theta})$ denota la funzione di rischio di base e $\boldsymbol{\theta}$ rappresenta il vettore di parametri coinvolti nella funzione di densità di T quando $\mathbf{x} = 0$; $\phi(\mathbf{x}; \boldsymbol{\beta})$ esprime l'impatto delle covariate \mathbf{x} su $h_0(t; \boldsymbol{\theta})$ e $\boldsymbol{\beta}$ è il vettore di parametri utilizzati nell'espressione della $\phi(\cdot)$.

Le diverse specificazioni di $h_0(t; \boldsymbol{\theta})$ e $\phi(\mathbf{x}; \boldsymbol{\beta})$ consentono di adeguare i modelli alle situazioni concrete oggetto di analisi. In seguito, per la durata T si utilizza la distribuzione di Weibull, $T \sim W(t; \lambda, \alpha)$, che implica una funzione di rischio di base $h_0(t; \boldsymbol{\theta}) = \alpha \lambda (\lambda t)^{\alpha-1}$ e rappresenta un andamento monotono (crescente, costante, e decrescente); mentre, si è scelta la forma log-lineare del modello a rischi

proporzionali che comporta $\phi(\mathbf{x}; \boldsymbol{\beta}) = \exp(\mathbf{x}\boldsymbol{\beta})$ e appare una delle piú utilizzate (Cox, 1972). La scelta della distribuzione di Weibull per la durata T e il modello a rischi proporzionali implicano assunti sul processo generatore dei dati che richiedono di essere verificati perché potrebbero essere anche molto restrittive per la complessità del fenomeno che debbono rappresentare; ma se sono supportate dai dati, la funzione di densità da utilizzare per stimare i parametri del modello è

$$f(t, \mathbf{x}; \lambda, \alpha, \boldsymbol{\beta}) = \alpha \lambda (\lambda t)^{\alpha-1} \exp[\mathbf{x}\boldsymbol{\beta} - (\lambda t)^\alpha \exp(\mathbf{x}\boldsymbol{\beta})]. \quad (2)$$

Il valore atteso e la varianza di T sono uguali rispettivamente a

$$\begin{aligned} \mu_T &= \frac{1}{\lambda \alpha} \Gamma\left(\frac{1}{\alpha}\right) \exp\left(-\frac{\mathbf{x}\boldsymbol{\beta}}{\alpha}\right) \\ \sigma_T^2 &= \frac{1}{(\lambda \alpha)^2} \left[2\alpha \Gamma\left(\frac{2}{\alpha}\right) - \Gamma^2\left(\frac{1}{\alpha}\right)\right] \exp\left(-2\frac{\mathbf{x}\boldsymbol{\beta}}{\alpha}\right). \end{aligned}$$

Le indagini sulle forze di lavoro forniscono, in genere, dati trasversali e le durate osservate sono soltanto i periodi trascorsi dall'ingresso nello stato di disoccupato alla data di rilevazione; allora, il modello precedente non può essere piú applicato perché non è possibile valutare la funzione di rischio che si basa sull'uso delle durate complete. Un approccio possibile all'analisi di questi dati consiste nella formulazione di opportune ipotesi sul legame esistente tra i periodi trascorsi osservati e i corrispondenti periodi completi non osservati. In condizioni di stazionarietà e omogeneità è possibile ottenere la distribuzione delle durate trascorse dall'ingresso nello stato di disoccupato alla data di campionamento. Gli assunti sono, però, problematici perché l'eterogeneità non è catturata completamente dalle variabili esplicative e la stazionarietà è sensibile ai mutamenti frequenti cui è soggetto il mercato del lavoro; tuttavia, è possibile così formulare modelli abbastanza coerenti con la realtà e la teoria, che consentono di esplorare la struttura delle relazioni tra le variabili esplicative e la durata.

2.2. Per descrivere l'evento uscita dallo stato di disoccupato si può considerare un processo di Poisson o un processo di rinnovamento (Allison, 1985; Lancaster, 1990). Nel seguito si utilizza la relazione esistente tra la distribuzione della variabile casuale U , che indica i periodi trascorsi osservati, e la distribuzione della variabile casuale T , che indica la durata completa non osservata.

Sia τ la data della rilevazione. Sia uguale a 0 il minimo tra le diverse date di ingresso (osservate e rappresentate con la variabile casuale D) nello stato di disoccupato $\{d_1, \dots, d_n\}$ sicché $\tau - d_i$ indichi la durata trascorsa osservata, u_i . La variabile casuale $U = \tau - D$ può assumere, perciò, valori al piú uguali a τ e quando risulta $U < \tau$, allora $U \in [u, u + du]$ se e sole se si verifica un evento, ossia si trova un posto di lavoro, nell'intervallo di tempo $[(\tau - u - du), (\tau - u)]$, che ha probabilità $h(\tau - u; \mathbf{x}) du$; mentre il nuovo periodo che inizia e dura u ha probabilità $[1 - F(u)]$. Se si assume che l'inizio di ogni periodo di disoccupazione

abbia la stessa probabilità di accadere in qualunque istante di tempo e che i periodi siano estratti dalla stessa distribuzione indipendentemente dalla data del loro inizio, allora la (densità di) probabilità che un periodo di disoccupazione “catturato” da un’indagine sia incompleto di lunghezza u (Lancaster, 1990) è data da

$$f(u, \mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\beta}) = \frac{1}{E(T)} \int_u^\infty f(t, \mathbf{x}; \boldsymbol{\theta}, \boldsymbol{\beta}) dt \quad (3)$$

dove $E(T)$ indica il valore atteso dei periodi completi non osservati.

Si possono stimare così i parametri della funzione di rischio e il valore atteso delle durate complete non osservate, T ; ma individuare la distribuzione dei periodi completi, che non vengono osservati, non è facile. Inoltre, la funzione di densità, $f(u; \mathbf{x})$, risulta più complessa se l’integrale non è risolubile e bisogna ricorrere, pertanto, all’integrazione numerica. L’espressione di $f(u; \mathbf{x})$ diventa, invece, semplice nel caso in cui si assuma che la distribuzione dei periodi completi non osservati sia una Weibull, $T \sim W(t; \lambda, \alpha)$. Infatti, si ha

$$f(u, \mathbf{x}; \lambda, \alpha, \boldsymbol{\beta}) = \frac{\lambda}{\Gamma\left(1 + \frac{1}{\alpha}\right)} \exp\left[\frac{\mathbf{x}\boldsymbol{\beta}}{\alpha} - (\lambda u)^\alpha \exp(\mathbf{x}\boldsymbol{\beta})\right] \quad (4)$$

per cui il valore atteso e la varianza di U sono uguali rispettivamente a

$$\begin{aligned} \mu_U &= \frac{1}{\lambda \Gamma(1/\alpha)} \Gamma\left(\frac{2}{\alpha}\right) \exp\left(-\frac{\mathbf{x}\boldsymbol{\beta}}{\alpha}\right) \\ \sigma_U^2 &= \frac{1}{[\lambda \Gamma(1/\alpha)]^2} \left[\Gamma\left(\frac{3}{\alpha}\right) \Gamma\left(\frac{1}{\alpha}\right) - \Gamma^2\left(\frac{2}{\alpha}\right) \right] \exp\left(-2\frac{\mathbf{x}\boldsymbol{\beta}}{\alpha}\right). \end{aligned}$$

Per valutare il valore atteso dei periodi completi non osservati, T , quando si dispone soltanto dei periodi trascorsi si possono stimare i parametri di $f(u, \mathbf{x}; \lambda, \alpha, \boldsymbol{\beta})$ e sostituirli nell’espressione di μ_T , oppure utilizzare l’espressione

$$\mu_U = \mu_T \frac{\alpha \Gamma(2/\alpha)}{\Gamma^2(1/\alpha)} \quad (5)$$

che denota la relazione tra la media, μ_U , dei periodi incompleti osservati e la media, μ_T , dei periodi completi non osservati: $\alpha > 1$ implica $\mu_U < \mu_T$ e $\mu_U/\mu_T \rightarrow 1/2$ per $\alpha \rightarrow \infty$; $\alpha < 1$ implica $\mu_U > \mu_T$ e $\mu_U/\mu_T \rightarrow \infty$ per $\alpha \rightarrow 0^+$; $\alpha = 1$ implica $\mu_U = \mu_T$. Si sottolinea ancora che questi risultati dipendono dalle condizioni di stazionarietà nel processo di mobilità della forza lavoro tra gli stati e tale assunto è abbastanza plausibile nel contesto dell’analisi a breve termine come può essere il fenomeno della disoccupazione.

2.3. L’accertamento del ruolo delle qualità individuali nei risultati che si conseguono successivamente a una data di riferimento concerne due aspetti: la loro

persistenza temporale per le differenze tra i soggetti (eterogeneità); e la loro intensità nel tempo per l'accentuazione o l'attenuazione in base all'esperienza (dipendenza dallo stato). Per esempio, chi ha sperimentato un periodo di disoccupazione può avere una maggiore probabilità di sperimentare in futuro un altro periodo di disoccupazione (a causa della perdita di esperienza lavorativa).

L'eterogeneità dei soggetti è catturata dalle caratteristiche individuali, \mathbf{x} , osservate. L'eterogeneità rimanente, detta non osservata, si introduce nel modello assumendo che si manifesti nel parametro di scala, λ , considerato come una variabile casuale continua, Λ (Heckman, Singer, 1984a); ma non esistono criteri teorici per scegliere la sua distribuzione, che se è inadeguata comporta una distorsione nelle stime dei parametri del modello (Heckman, Singer, 1984b) per cui assunti diversi possono condurre a conclusioni (almeno) qualitativamente diverse in termini di proposizioni teoriche e modelli statistici. Senza criteri per individuare la distribuzione delle variabili non osservate, che sia statisticamente migliore (Hannan, Carroll, 1942, cap. 4), il vaglio delle diverse leggi è spesso condizionato dalla difficoltà di calcolo. Così, una distribuzione di probabilità di Λ , plausibile e conveniente, è ancora la Weibull, $\Lambda \sim W(\lambda; \vartheta, \alpha)$ con parametro di scala ϑ e lo stesso parametro di forma, α , della distribuzione delle durate (complete o trascorse). Allora, la funzione di densità delle durate sarà data da $\psi(\cdot, \mathbf{x}) = \int_0^\infty f(\cdot, \mathbf{x}; \lambda, \alpha, \boldsymbol{\beta}) \phi(\lambda; \vartheta, \alpha) d\lambda$, dove $\phi(\lambda; \vartheta, \alpha)$ è la funzione di densità di Λ , mentre $f(\cdot, \mathbf{x}; \lambda, \alpha, \boldsymbol{\beta})$ è la funzione di densità dei periodi completi o incompleti. Se si sostituiscono le espressioni corrispondenti e si integra, i risultati sono semplici perché non includono più l'operatore di integrazione:

$$\psi(t, \mathbf{x}; \vartheta, \alpha, \boldsymbol{\beta}) = \frac{\alpha t^{\alpha-1} \vartheta^\alpha \exp(-\mathbf{x}\boldsymbol{\beta})}{[t^\alpha + \vartheta^\alpha \exp(-\mathbf{x}\boldsymbol{\beta})]^2} \quad (6)$$

$$\psi(u, \mathbf{x}; \vartheta, \alpha, \boldsymbol{\beta}) = \frac{\vartheta^\alpha \exp(-\mathbf{x}\boldsymbol{\beta})}{[u^\alpha + \vartheta^\alpha \exp(-\mathbf{x}\boldsymbol{\beta})]^{1+\frac{1}{\alpha}}}. \quad (7)$$

La funzione di rischio che si ottiene da queste funzioni di densità è

$$h_e(t, \mathbf{x}; \vartheta, \alpha, \boldsymbol{\beta}) = \frac{\alpha t^{\alpha-1}}{t^\alpha + \vartheta^\alpha \exp(-\mathbf{x}\boldsymbol{\beta})}$$

che presenta una forma flessibile, riportata in Figura 2 per alcuni valori di ϑ e α ; infatti, ha un andamento non monotono: vale 0 in $t=0$, cresce fino a un massimo che si verifica in $t_{\max} = \sqrt[\alpha]{\alpha - 1} \vartheta \exp(-\mathbf{x}\boldsymbol{\beta}/\alpha)$, quindi decresce e $h_e(t, \mathbf{x}) \rightarrow 0$ per $t \rightarrow \infty$. I risultati sembrano così soddisfacenti; tuttavia, non solo la scelta $\Lambda \sim W(\lambda; \vartheta, \alpha)$ è arbitraria, ma media e varianza riportate di seguito e calcolate con le funzioni di densità delle durate $\psi(t, \mathbf{x})$ e $\psi(u, \mathbf{x})$ non sono definite in alcuni punti:

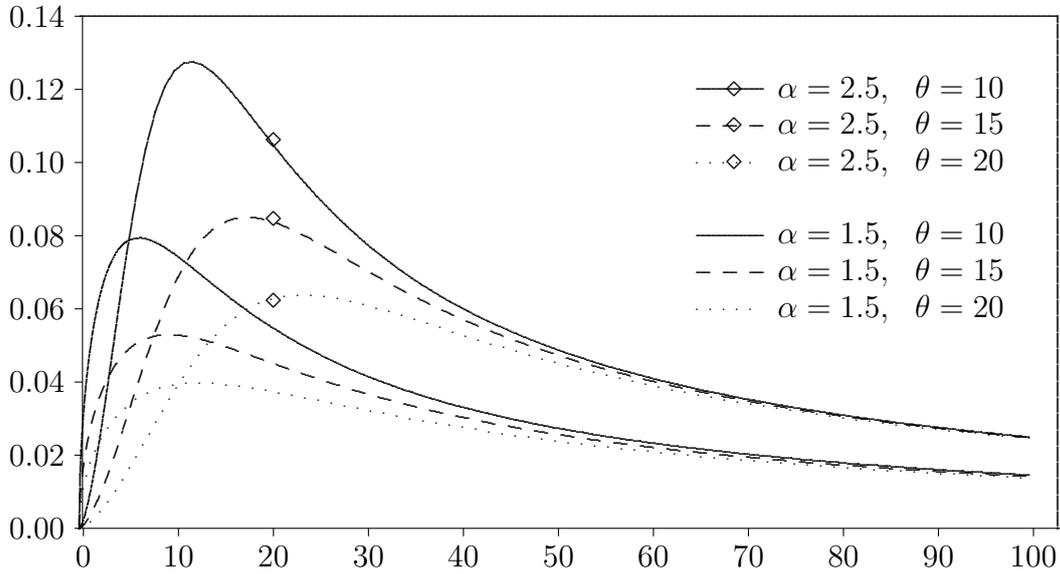
$$\begin{aligned} \mu_{T,e} &= \frac{\vartheta}{\alpha} \exp\left(-\frac{\mathbf{x}\boldsymbol{\beta}}{\alpha}\right) \pi \operatorname{cosec}\left(\frac{\pi}{\alpha}\right) \\ \sigma_{T,e}^2 &= \frac{\vartheta^2}{\alpha^2} \exp\left(-2\frac{\mathbf{x}\boldsymbol{\beta}}{\alpha}\right) \left[2\pi\alpha \operatorname{cosec}\left(2\frac{\pi}{\alpha}\right) - \pi^2 \operatorname{cosec}^2\left(\frac{\pi}{\alpha}\right)\right] \end{aligned}$$

$$\mu_{U,e} = \frac{\vartheta \exp(-\mathbf{x}\boldsymbol{\beta}/\alpha) \left[\Gamma(1 - \frac{1}{\alpha}) \Gamma(\frac{2}{\alpha}) \right]}{\alpha \Gamma(1 + \frac{1}{\alpha})}$$

$$\sigma_{U,e}^2 = \frac{\vartheta^2 \exp(-2\mathbf{x}\boldsymbol{\beta}/\alpha) \left[\alpha \Gamma(1 + \frac{1}{\alpha}) \Gamma(1 - \frac{2}{\alpha}) \Gamma(\frac{3}{\alpha}) - \Gamma^2(1 - \frac{1}{\alpha}) \Gamma^2(\frac{2}{\alpha}) \right]}{\alpha^2 \Gamma^2(1 + \frac{1}{\alpha})}.$$

In particolare, la media dei periodi completi, $\mu_{T,e}$, non è definita nei punti $\alpha = 1/k$ con $k \in \mathbb{N}$ (dove \mathbb{N} indica l'insieme dei numeri naturali); mentre la media dei periodi incompleti, $\mu_{U,e}$, non è definita, per esempio, nei punti $\alpha = 1$, $\alpha = 1/2$, e altri minori di $1/2$; pertanto, l'uso delle funzioni di densità $\psi(t, \mathbf{x})$ o $\psi(u, \mathbf{x})$ per modellare le durate di disoccupazione non sembra del tutto soddisfacente.

Figura 2. Funzione di rischio relativa alla densità $\psi(t; \theta, \alpha)$ per $\mathbf{x} = 0$



2.4. La media e la varianza di T e U non esistono quando i parametri assumono alcuni valori non soltanto per le funzioni di densità $\psi(t, \mathbf{x})$ e $\psi(u, \mathbf{x})$, ma anche per altre funzioni di azzardo di base molto usate nell'analisi delle durate della disoccupazione, come la log-logistica, $LL(\cdot)$. Infatti, quando $T \sim LL(t; \lambda, \alpha)$, la funzione di ripartizione

$$F(t; \lambda, \alpha) = 1 - \frac{1}{1 + (\lambda t)^\alpha}, \quad (8)$$

appartiene alla famiglia delle distribuzioni tipo XII di Burr (1942). La funzione di azzardo

$$h_{LL}(t; \lambda, \alpha) = \frac{\lambda \alpha (\lambda t)^{\alpha-1}}{1 + (\lambda t)^\alpha} \quad (9)$$

risulta molto flessibile perché presenta andamenti differenziati (anche non monotoni) al variare dei parametri λ e α , come mostrano le curve della Figura 3 tra le

quali ve ne sono alcune simili a quelle riportate nella Figura 2. Se si adotta la log-logistica come funzione di azzardo di base nel modello di Cox a rischi proporzionali, allora la funzione di ripartizione assume la forma

$$F(t, \mathbf{x}; \lambda, \alpha, \boldsymbol{\beta}) = 1 - \frac{1}{[1 + (\lambda t)^\alpha]^{\phi(\mathbf{x}; \boldsymbol{\beta})}}. \quad (10)$$

L'espressione diventa piú semplice e generale se l'eterogeneità osservata viene introdotta sul parametro di scala λ della 8 ponendo $\lambda^\alpha = \omega^\alpha \phi(\mathbf{x}; \boldsymbol{\beta})$ per cui il modello diventa a rischi non proporzionali e la funzione di ripartizione,

$$F_{LL}(t, \mathbf{x}; \omega, \alpha, \boldsymbol{\beta}) = 1 - \frac{1}{1 + \phi(\mathbf{x}; \boldsymbol{\beta}) (\omega t)^\alpha}, \quad (11)$$

appartiene ancora alla famiglia delle distribuzioni tipo XII di Burr, che per $\phi(\mathbf{x}; \boldsymbol{\beta}) = \exp(\mathbf{x} \boldsymbol{\beta})$ fornisce una media e una varianza rispettivamente pari a

$$\begin{aligned} \mu_{T-LL} &= \frac{\exp(-\mathbf{x} \boldsymbol{\beta} / \alpha)}{\omega \alpha} \pi \operatorname{cosec} \left(\frac{\pi}{\alpha} \right) \\ \sigma_{T-LL}^2 &= \frac{\exp(-2 \mathbf{x} \boldsymbol{\beta} / \alpha)}{\omega^2 \alpha^2} \left[2 \pi \alpha \operatorname{cosec} \left(2 \frac{\pi}{\alpha} \right) - \pi^2 \operatorname{cosec}^2 \left(\frac{\pi}{\alpha} \right) \right]. \end{aligned}$$

Anche per la distribuzione log-logistica, la media dei periodi completi, μ_{T-LL} , e la varianza, σ_{T-LL}^2 , non sono definite nei punti $\alpha = 1/k$ con $k \in \mathbb{N}$, per cui potrebbe risultare insoddisfacente.² Si noti che $F(t, \mathbf{x}; \lambda, \alpha, \boldsymbol{\beta})$ risulta da un miscuglio di distribuzioni di Weibull secondo una Gamma (Lancaster, 1990, pp. 65–70) e, quindi, per includere nei modelli l'eterogeneità non osservata si sono utilizzate le funzioni di densità corrispondenti che rappresentano un modello a rischi non proporzionali

$$f_{LL}(t, \mathbf{x}; \omega, \alpha, \boldsymbol{\beta}) = \frac{\alpha \omega \exp(\mathbf{x} \boldsymbol{\beta}) (\omega t)^{\alpha-1}}{[1 + \exp(\mathbf{x} \boldsymbol{\beta}) (\omega t)^\alpha]^2} \quad (12)$$

$$f_{LL}(u, \mathbf{x}; \omega, \alpha, \boldsymbol{\beta}) = \frac{\omega \exp(\mathbf{x} \boldsymbol{\beta} / \alpha)}{\Gamma(1 - \frac{1}{\alpha}) \Gamma(1 + \frac{1}{\alpha})} \frac{1}{1 + \exp(\mathbf{x} \boldsymbol{\beta}) (\omega u)^\alpha} \quad (13)$$

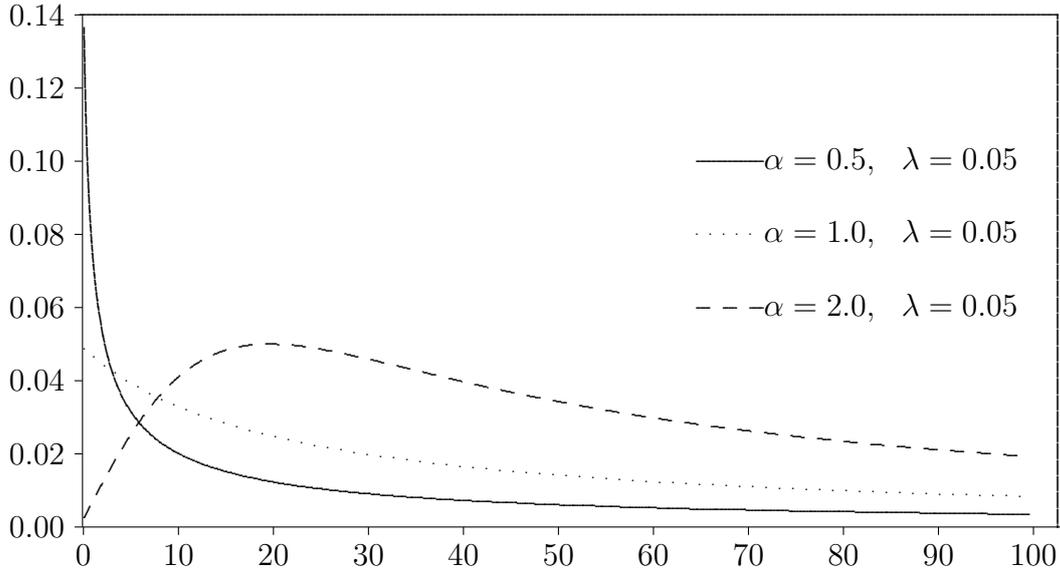
dove $f_{LL}(t, \mathbf{x}; \omega, \alpha, \boldsymbol{\beta})$ coincide ora con $\psi(t, \mathbf{x}; \vartheta, \alpha, \boldsymbol{\beta})$, a meno di una diversa denominazione dei parametri.

La dipendenza dallo stato è espressa dalla dipendenza dal tempo della funzione di rischio: cosí, nel caso in cui $T \sim W(t; \lambda, \alpha)$ e non si introduce l'eterogeneità non osservata, un valore di α statisticamente minore (maggiore) di 1 implica una diminuzione (un aumento) della probabilità di trovare lavoro all'aumentare della durata di permanenza nello stato di disoccupato; un valore di α uguale all'unità implica l'assenza di dipendenza dal tempo per cui il processo non ha memoria (del passato) e un soggetto appena entrato nel mercato del lavoro ha la stessa probabilità

²Le espressioni di μ_{T-LL} e σ_{T-LL}^2 si possono scrivere in una forma diversa per le relazioni esistenti tra le funzioni beta, gamma, e cosecante: $B(1 - \eta, 1 + \eta) = \Gamma(1 - \eta) \Gamma(1 + \eta) = \pi \eta \operatorname{cosec}(\pi \eta)$ (Johnson, Kotz, Balakrishnan, 1995, pp. 147–154).

di uscire dallo stato di uno che aspetta da molto tempo. L'introduzione dell'eterogeneità non osservata con una variabile casuale $\Lambda \sim W(t; \vartheta, \alpha)$ sul parametro di scala di $W(t; \lambda, \alpha)$ conduce a una dipendenza dal tempo espressa da $h_e(t; \vartheta, \alpha, \beta)$ che risulta simile a quella espressa da $h_{LL}(t; \omega, \alpha, \beta)$; i loro andamenti sono riportati in Figura 2 e Figura 3, rispettivamente.

Figura 3. Funzione di rischio relativa alla densità $f_{LL}(t; \lambda, \alpha)$



2.5. La lunghezza del periodo di disoccupazione è soggetta all'effetto telescopico che può comportare una concentrazione anomala delle frequenze in certi valori delle durate (*heaping effect*), rispetto alla loro distribuzione che si presume regolare. Infatti, la durata vera (completa o incompleta) non osservata, t_i , può essere "arrotondata" dall' i -esimo rispondente a un altro valore, t_{Hij} (Sudman, Bradburn, 1973; Sikkel, 1985). Un semplice modello del processo stocastico di accumulazione delle durate (Torelli, Trivellato, 1993) è $t_{Hij} = t_i + \delta_{ij} y_i$ con $\delta_{ij} = h_j - t_i$ tale che $|h_j - t_i|$ sia minimo per $j = 1, \dots, J$, dove $h_1 < h_2 < \dots < h_J$, sono i valori delle durate in cui si è osservata un'accumulazione (*heaped values*). La variabile casuale y_i è dicotoma: $y_i = 1$ con probabilità $H(t_i; \gamma)$ se la durata osservata coincide con un punto di accumulazione, altrimenti $y_i = 0$ con probabilità $[1 - H(t_i; \gamma)]$; pertanto, $H(t_i; \gamma)$, detta anche "funzione di accumulazione" (*heaping function*), è la funzione di ripartizione che esprime la probabilità di osservare una durata di lunghezza t_i contaminata dall'effetto telescopico e dipende dal vettore di parametri γ . Per la funzione accumulazione si può scegliere una distribuzione di probabilità semplice e adeguata, quale la distribuzione esponenziale: $H(t; \gamma) = 1 - \exp(-\gamma t)$. Tale distribuzione potrebbe dipendere anche dalle caratteristiche individuali, \mathbf{x} , ma per evitare una sovra-parametrizzazione del modello questa possibilità è stata ignorata.

Per ogni soggetto i le grandezze rilevate tramite l'indagine sono $\{t_i, c_i, \mathbf{x}_i\}$,

dove c_i indica una variabile dicotoma che vale 1 se il periodo osservato t_i è completo e 0 altrimenti. I punti di accumulazione si possono determinare soltanto, dopo avere raccolto i dati, esaminando la distribuzione delle frequenze relative (o assolute) delle durate. Si tratta perciò di una questione prevalentemente empirica. Le distribuzioni degli insiemi di durate analizzate in seguito sono riportate: nella Figura 4 e nella Figura 5 quelle relative alle durate incomplete o trascorse, nella Figura 6 e nella Figura 7 quelle relative alle durate complete, rispettivamente per i maschi e per le femmine. Dalle distribuzioni delle durate osservate, e esposte nei grafici menzionati, si individuano i punti di accumulazione h_1, \dots, h_J .

La durata dell' i -esimo soggetto, t_i , può essere uguale a un $h_j \in \{h_1, \dots, h_J\}$ oppure no. Si crea perciò una variabile y_i che vale 1 se t_i coincide con un punto di accumulazione e 0 altrimenti. Se si ha coincidenza con un certo h_j , si possono determinare anche t_{HLij} e t_{HUij} che formano un intervallo contenente t_i , che sarà indicato pertanto con t_{Hij} . L'estremo inferiore dell'intervallo, t_{HLij} , corrisponde alla media dei valori dei punti di accumulazione h_{j-1} e h_j , ossia $(h_{j-1} + h_j)/2$. L'estremo superiore dell'intervallo, t_{HUij} , corrisponde alla media dei valori dei punti di accumulazione h_j e h_{j+1} , ossia $(h_j + h_{j+1})/2$. Allora, per ogni soggetto i le grandezze disponibili diventano $\{t_i, c_i, y_i, t_{HLij}, t_{HUij}, \mathbf{x}_i\}$, dove t_{HLij} e t_{HUij} assumono i valori che gli competono quando y_i vale 1 e assumono il valore 0 altrimenti, ossia quando y_i vale 0.

2.6. I parametri dei modelli relativi alla durata della disoccupazione si possono stimare con il metodo di massima verosimiglianza in quanto sono note le funzioni di densità che devono includere ora anche l'effetto telescopico. L'espressione della funzione di verosimiglianza per il modello di durata, in generale, diventa

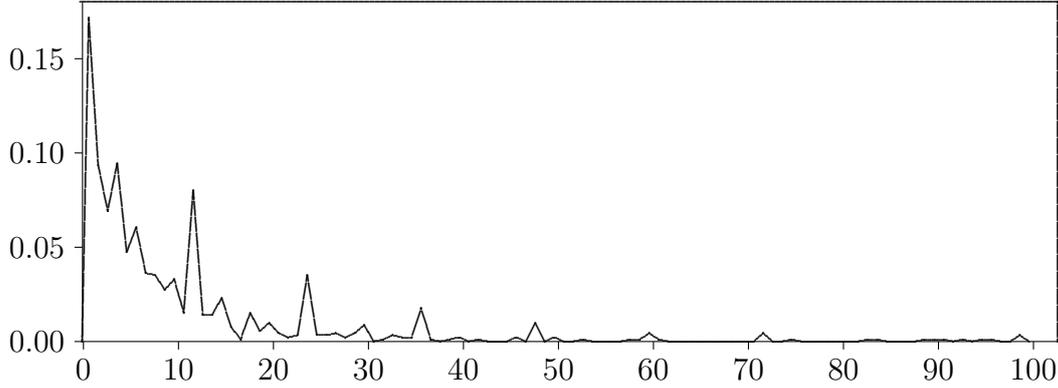
$$\begin{aligned}
L(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) &= \prod_{i=1}^n \left\{ [f(t_i, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})]^{c_i} [S(t_i, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})]^{1-c_i} [1 - H(t_i; \boldsymbol{\gamma})] \right\}^{1-y_i} \\
&\cdot \left\{ [f(t_{Hij}, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})]^{c_i} [S(t_{Hij}, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})]^{1-c_i} [1 - H(t_{Hij}; \boldsymbol{\gamma})] \right\}^{y_i} \quad (14) \\
&+ \int_{t_{HLij}}^{t_{HUij}} [f(v, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})]^{c_i} [S(v, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})]^{1-c_i} [H(v; \boldsymbol{\gamma})] dv \Big\}^{y_i}
\end{aligned}$$

dove n è il numero totale di osservazioni; come in precedenza, $\boldsymbol{\theta}$ è il vettore di parametri coinvolti nella funzione di rischio di base; mentre $f(\cdot, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})$ è la funzione di densità di T e $S(\cdot, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})$ è la funzione di sopravvivenza nelle quali i valori osservati delle durate sono stati indicati con t_i quando non coincidono con i punti di accumulazione (termini inclusi nella prima parentesi graffa).³ La forma della funzione di verosimiglianza si modifica di poco quando i periodi sono tutti

³Per illustrare la derivazione delle componenti della funzione di verosimiglianza si considera la probabilità che la variabile casuale T assuma valori nell'intervallo infinitesimo $(t_i, t_i + dt)$, dato che $T \geq t_i$. Sia D l'evento $\{T \in (t_i, t_i + dt) | T \geq t_i\}$. La probabilità $P(D)$ si può valutare con il prodotto della funzione di densità per l'infinitesimo dt ; ma per semplicità il ragionamento si imposta soltanto in termini probabilistici. Si considerino gli altri eventi che possono "influenzare" D : A che denota $t_i \notin \{h_1, \dots, h_J\}$; B che denota t_i non troncato; e C che denota l'avvenuta

incompleti: si assume sempre $c_i = 1$ e si sostituisce $f(t_i, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})$ con $f(u_i, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})$ e $f(t_{Hij}, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})$ con $f(u_{Hij}, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})$. Non si riportano, per brevità, le espressioni estese delle funzioni di verosimiglianza per le durate incomplete e per le durate complete.

Figura 4. Distribuzione di frequenza relativa della durata
Periodi incompleti – Maschi



alterazione del valore vero di t_i . L'inclusione del fenomeno di accumulazione deve considerare: sia gli eventi \bar{A} e A che indicano rispettivamente se il valore di t_i è un picco della distribuzione di frequenza o non lo è; sia gli eventi \bar{C} e C che indicano se non c'è stata alterazione del valore di t_i o c'è stata, rispettivamente. La probabilità si può esprimere come in una distribuzione di Bernoulli, in base alla quale si ha

$$\begin{aligned} P(D) &= \{P[D|A(C \cup \bar{C})]\}^{1-y_i} \{P[D|\bar{A}(C \cup \bar{C})]\}^{y_i} \\ &= \{P[D|A\bar{C}]\}^{1-y_i} \{P[D|\bar{A}\bar{C}] + P[D|\bar{A}C]\}^{y_i} \end{aligned}$$

dove il termine $P[D|AC] = 0$ perché se vale A , allora si esclude che possa esserci stata alterazione del valore osservato t_i in quanto si assume che ciò possa accadere solo quando $t_i \in \{h_1, \dots, h_J\}$, ossia vale \bar{A} . Nell'espressione precedente si può introdurre il possibile troncamento della durata osservata, sempre in analogia con la distribuzione di Bernoulli, sicché si ottiene

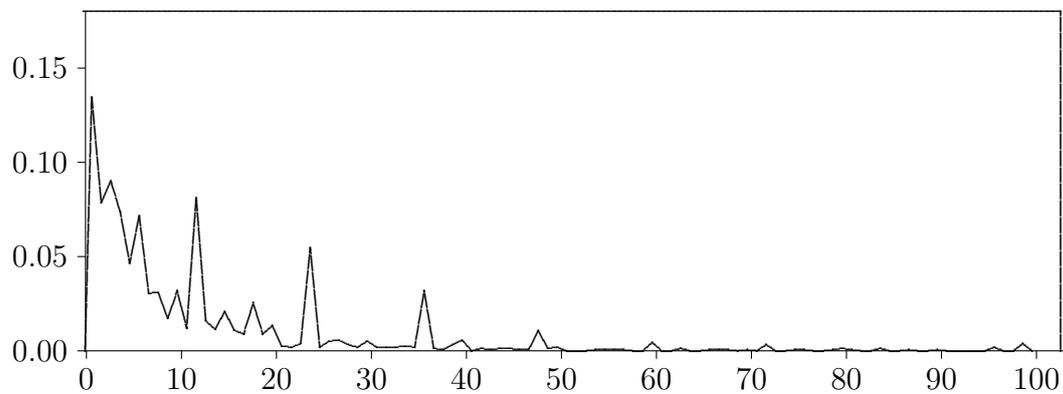
$$\begin{aligned} P(D) &= \left\{ [P(D|AB\bar{C})]^{c_i} [P(D|A\bar{B}\bar{C})]^{1-c_i} \right\}^{1-y_i} \\ &\times \left\{ [P(D|\bar{A}B\bar{C})]^{c_i} [P(D|\bar{A}\bar{B}\bar{C})]^{1-c_i} + [P(D|\bar{A}BC)]^{c_i} [P(D|\bar{A}\bar{B}C)]^{1-c_i} \right\}^{y_i}. \end{aligned}$$

Si può utilizzare ora la "corrispondenza" esistente tra le probabilità e le funzioni di densità relative agli eventi considerati; si ottiene, così, la funzione di densità per l' i -esimo soggetto. Per esempio, la probabilità $P(D|AB\bar{C})$ coinvolge, a livello di intorno di t_i , la funzione di densità quando non vi sia né troncamento e né accumulazione per cui corrisponde a $f(t_i, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta}) [1 - H(t_i; \boldsymbol{\gamma})]$. Analogamente, la probabilità $P(D|A\bar{B}\bar{C})$ si riferisce alla funzione di densità quando vi sia troncamento senza accumulazione, che corrisponde a $S(t_i, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta}) [1 - H(t_i; \boldsymbol{\gamma})]$. Il prodotto di questi due termini elevati ai rispettivi esponenti dà il primo termine della funzione di verosimiglianza:

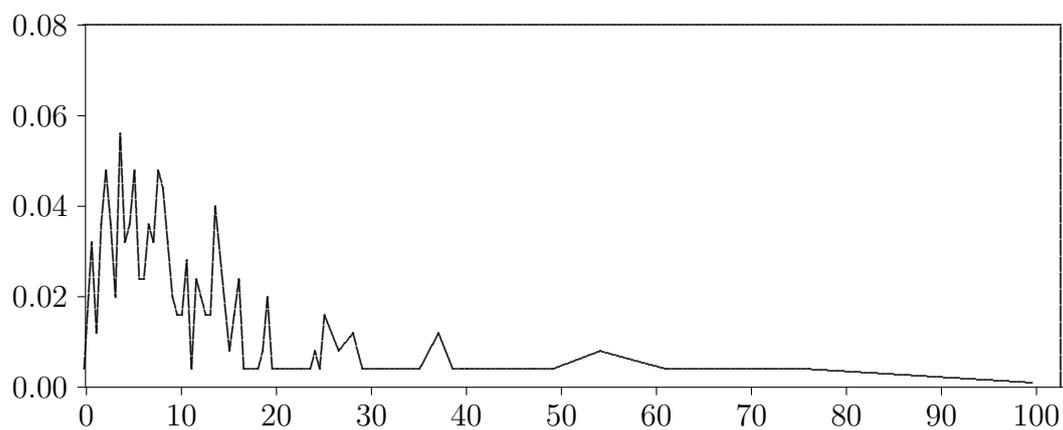
$$\left\{ [f(t_i, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})]^{c_i} [S(t_i, \mathbf{x}_i; \boldsymbol{\theta}, \boldsymbol{\beta})]^{1-c_i} [1 - H(t_i; \boldsymbol{\gamma})] \right\}^{1-y_i}.$$

Per ottenere gli altri termini si procede allo stesso modo.

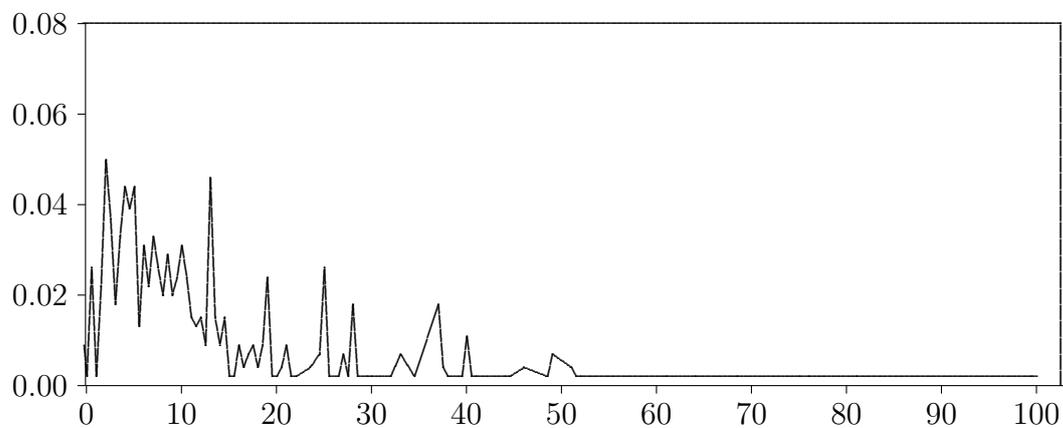
*Figura 5. Distribuzione di frequenza relativa della durata
Periodi incompleti – Femmine*



*Figura 6. Distribuzione di frequenza relativa della durata
Periodi completi – Maschi*



*Figura 7. Distribuzione di frequenza relativa della durata
Periodi completi – Femmine*



3. I risultati delle stime dei parametri

3.1. L'insieme delle persone in cerca di occupazione, determinato in base alla definizione adottata dall'ISTAT da ottobre del 1992, comprende coloro che: dichiarano di non essere occupati e di essere in cerca di occupazione; hanno effettuato almeno un'azione di ricerca di lavoro entro i 30 giorni precedenti l'intervista; sono pronti (entro 2 settimane) a accettare un'offerta di lavoro. Le persone in cerca di occupazione sono classificate in tre sottogruppi: i disoccupati, sono le persone con età maggiore o uguale a 15 anni che hanno perso una occupazione precedente per licenziamento, dimissioni, scadenza di un contratto a termine; i nuovi ingressi (nel mercato del lavoro) sono le persone in cerca di prima occupazione con età maggiore o uguale a 15 anni che non hanno mai esercitato un'attività lavorativa; i rientri, sono le persone che hanno smesso un'attività lavorativa e per un certo periodo di tempo non hanno più cercato un impiego diventando non forze di lavoro, quindi, hanno ricominciato a cercare lavoro.

Nella Tabella 1 sono riportate alcune statistiche descrittive per le covariate, utilizzate nei modelli esplicativi delle durate della disoccupazione, che si riferiscono sia all'insieme dei periodi completi, sia all'insieme dei periodi incompleti ottenuto dall'unione di tutti i campioni trimestrali per semplificare l'esposizione. Il numero di periodi completi, ricostruiti dal primo trimestre 1993 al primo trimestre 1995, si riferisce a 655 soggetti: lo 0,3% ha sperimentato tre periodi di disoccupazione, il 7,5% ne ha sperimentati due e il resto (604 soggetti) uno.⁴ I valori relativi ai caratteri osservati nell'insieme dei periodi completi sono assai simili a quelli osservati nell'insieme dei periodi incompleti e possono considerarsi come valori "medi" rispetto al periodo esaminato. Ciò nonostante, nella descrizione si citeranno spesso entrambi perché il confronto con i dati dell'Italia si può eseguire quasi sempre solo per quelli incompleti; mentre i dati relativi ai completi mostrano come incide grosso modo il funzionamento del mercato del lavoro nel tempo sull'insieme delle persone in cerca di lavoro.

La durata media della ricerca di lavoro è di poco superiore a 13 mesi e risulta circa uguale alla durata media nazionale (Favro-Paris, Gennari, Oneto, 1996). Il valore in sé non denota una situazione drammatica, com'è in Italia, perché in Emilia-Romagna si è vicino al pieno impiego; infatti, nell'intervallo di tempo esaminato l'entità del tasso di disoccupazione è circa il 6% contro il 10-12% dell'Italia, e l'incidenza della disoccupazione di lungo periodo è il 25-30% contro il 60-65% in Italia.⁵ La durata media mostra che, anche in condizione di prossimità al pieno

⁴Il numero totale di periodi completi è 708. Le elaborazioni si basano solo su 655 soggetti e per quelli con periodi multipli si è usato un solo periodo di disoccupazione per avere osservazioni tra loro indipendenti; per convenzione, si è preso il primo. Le stesse elaborazioni si sono eseguite anche sui 708 periodi, ma i risultati restavano essenzialmente invariati.

⁵Il tasso di disoccupazione in Italia è stato pari al 10,4% nel 1993 e all'11,3% nel 1994; mentre in Emilia-Romagna si rilevavano valori del 6,0% e del 6,1%, rispettivamente. Per quanto concerne la disoccupazione di lungo periodo, definita come un'attesa con durata superiore a un anno, il numero di persone in cerca di lavoro che aspettano più di un anno è stato pari al 24,7% del totale dei disoccupati nel 1993 e al 29,8% nel 1994; mentre in Italia è stato pari al 61,2% nel 1993 e al 64,9% nel 1994; nel Nord, invece, si è avuto un valore intorno al 50% (ISTAT, 1994, 1995a).

impiego, vi possono essere componenti dell'offerta con durate lunghe per varie ragioni: un alto reddito familiare che consente di cercare piú a lungo per trovare un'occupazione adeguata, la presenza di gruppi con caratteristiche diverse da quelle richieste dalla domanda, e cosí via. La media dei periodi incompleti è circa uguale a 12 mesi: un mese inferiore a quella dei periodi completi che potrebbe indicare anche una certa indipendenza dallo stato e un valore del parametro α circa uguale a 1 quando si assume che la durata sia distribuita secondo una Weibull, $T \sim W(t; \lambda, \alpha)$. La media dei periodi completi uguaglia, infatti, la media dei periodi incompleti quando la funzione di rischio è costante (Salant, 1977), che implica un valore di α uguale a 1.

L'età media dei soggetti che hanno completato il periodo di disoccupazione (quindi, non piú disoccupati) è 29,6 anni, che è circa uguale a quella dei soggetti che non hanno ancora completato il periodo (quindi, in cerca di occupazione). L'età media dei disoccupati in Italia è presumibilmente un po' piú alta.⁶

Le femmine in cerca di occupazione sono piú numerose dei maschi, con una proporzione del 64,7% (si veda la Tabella 1) che è piú elevata di quella relativa a tutta l'Italia pari al 52,7% nel 1993 e al 51,2% nel 1994. La tendenza è comune a tutte le regioni del Centro e del Nord, mentre nel Sud la composizione si ribalta e si riscontra una prevalenza dei maschi, specie in Campania, Puglia, e Sicilia. Presumibilmente, ciò dipende dal piú elevato tasso di partecipazione femminile, ma anche dal piú basso tasso di disoccupazione.

Il titolo di studio piú diffuso è la scuola dell'obbligo, conseguita dal 51,8% del totale dei disoccupati che hanno trovato lavoro; il 42,1% ha conseguito un diploma di maturità e il 6,1% la laurea. La distribuzione dei titoli di studio tra le persone in cerca di lavoro è simile; ma rispetto al dato nazionale, che presenta una distribuzione analoga a quella relativa al Nord, si ha una maggiore presenza di laureati (uno o due punti percentuali) e una rilevante presenza di diplomati (oltre il 10% in piú).

La posizione di capofamiglia riguarda il 14,1% del totale; quella di coniuge il 24,1%; e quella di figlio o altra condizione il 61,8%. Per indicare lo stato complessivo delle singole situazioni familiari, anche dal punto di vista della condizione economica, si è considerata la percentuale di occupati in famiglia; i suoi valori medi, riportati in Tabella 1, sono intorno al 30% sia per coloro che hanno trovato lavoro, sia per coloro che lo stavano ancora cercando.

La provenienza dei soggetti che hanno trovato lavoro presenta la modalità "disoccupati" con la massima frequenza, quasi la metà (43,7%) delle persone in cerca di lavoro; mentre i nuovi ingressi sono il 26,1% e i rientri sono il 30,2%. Anche tra coloro che cercano un lavoro si ha sempre una preponderanza dei disoccupati

L'incidenza della disoccupazione di lungo periodo tra coloro che cercavano e hanno trovato lavoro è stata pari al 33,7% nel 1993 e al 33,2% nel 1994 (per $n = 708$); ovviamente, sono un po' piú alti dei corrispondenti valori calcolati per i disoccupati che non hanno completato la ricerca.

⁶Tutti i dati riportati nel seguito, relativi all'Italia o alle sue tradizionali ripartizioni geografiche, sono stati ottenuti dalle pubblicazioni dell'ISTAT (1995b, 1995c). L'età è suddivisa in classi di cui una è aperta: "maggiore o uguale a 50 anni". L'età media delle persone in cerca di lavoro, calcolata dalla distribuzione di frequenze assolute e escludendo la classe aperta, è uguale a 28 anni circa; pertanto, con l'aggiunta della classe esclusa l'età media diventa molto piú alta.

(47,3%) che è piú alta di quella nazionale: 36,2% nel 1993 e 38,4% nel 1994. Le persone in cerca di prima occupazione sono il 25,8% e rappresenta un valore molto piú basso del corrispondente dato nazionale: 43,1% nel 1993 e 40,9% nel 1994. Le persone che rientrano nel mercato del lavoro e cercano un'occupazione corrispondono a un ammontare del 26,9%, che è maggiore di quello dell'Italia che registra un valore del 20% circa. Le differenze osservate sembrano indicare un mercato del lavoro regionale piú "dinamico" di quello nazionale: facilita l'inserimento dei giovani in cerca di prima occupazione che possono cosí accumulare un'esperienza e incoraggia a rientrare perché c'è speranza di trovare un impiego; ciò potrebbe spiegare una presenza maggiore dei disoccupati tra le persone in cerca di lavoro.⁷

Il tipo di orario di lavoro preferito è l'impiego a tempo pieno con un'incidenza del 62%; il lavoro a tempo parziale è stato scelto dal 23,2%. Le preferenze per l'orario di lavoro possono denotare la percezione relativa al mercato esistente dei soggetti coinvolti nella ricerca perché con l'aumentare della disoccupazione dovrebbe aumentare la propensione a accettare qualsiasi orario. Infatti, a livello nazionale le percentuali sono rovesciate: il 37% cerca un lavoro a tempo pieno, il 5% cerca un lavoro a tempo parziale e il 58% cerca un lavoro a qualsiasi orario nel 1993; le percentuali restano quasi invariate nel 1994. Se si considera, tuttavia, la distribuzione degli occupati si ha che soltanto il 5,5% lavora a tempo parziale, con una notevole differenza tra maschi e femmine: nel 1994 si ha il 2,7% dei maschi contro l'11,7% delle femmine. Anche in Emilia-Romagna l'incidenza percentuale delle assunzioni a tempo parziale sul totale (escluso il settore dell'agricoltura), riferito al primo semestre dell'anno, risulta simile: il 5,7% nel 1993 e il 6,2% nel 1994, ma negli ultimi anni mostra una tendenza alla crescita (Osservatorio del mercato del lavoro, 1997, pp. 67-72).

La distribuzione della preferenza per il luogo di lavoro mostra che la noncuranza per la sede è espressa solo dall'11,1% tra coloro che hanno trovato lavoro e dal 9,7% tra coloro che cercano un'occupazione. Si tratta di un valore che è comune a tutto il Nord; mentre, al Centro sale al 16,2% nel 1993 e al 19,6% nel 1994; valori anche piú alti del 23% si riscontrano nel Sud piú le isole. Questo potrebbe rappresentare ancora una indicazione di un mercato che offre buone opportunità sicché il lavoratore può porre piú vincoli di mobilità. La preferenza per il comune di residenza è dichiarata dal 38,9% di coloro che hanno trovato lavoro e dal 37,3% di coloro che cercano un'occupazione; tali valori sono un po' inferiori a quelli del Nord e molto piú bassi di quelli nazionali. Coloro che desiderano lavorare in un comune raggiungibile giornalmente (anche all'estero) sono quindi oltre il 50%; tale valore è piú elevato, sia rispetto al Nord nel suo insieme, sia rispetto all'Italia.⁸

⁷La letteratura sull'economia e sul mercato del lavoro della regione Emilia-Romagna è vasta. Per gli anni di interesse ci si limita a citare solo uno studio di Capecchi (1995) sulla domanda e sulla qualità delle assunzioni che è la ripetizione di indagini condotte e pubblicate anche negli anni 1993 e 1992; inoltre, si cita una pubblicazione periodica dell'Osservatorio del mercato del lavoro (1995) che presenta dati commentati e saggi sull'economia della regione Emilia-Romagna.

⁸Tali differenze potrebbero rivelare una disponibilità alla mobilità "locale" della forza lavoro, ma non contraddicono il dato precedente. Infatti, le possibilità di scelta tra piú opportunità potrebbero aumentare gli spostamenti della forza lavoro perché i soggetti tendono a andare dove vi sono vantaggi maggiori. Inoltre, si deve rilevare che si tratta di un'area molto integrata, in

Tabella 1 — *Media, deviazione standard (D.S.), minimo (min), massimo (max), e numero di casi (n) relativi alle variabili impiegate nell'analisi nei due campioni: uno di periodi completi della disoccupazione e l'altro di periodi incompleti*

Variabili	Periodi completi					Periodi incompleti ^(a)				
	Media	D.S.	Min	Max	n ^(b)	Media	D.S.	Min	Max	n ^(b)
Durata	13,36	13,63	0,0	100,5	655	11,85	14,95	0,5	99,0	2482
Sesso										
Maschio	0,353	0,478	0	1	231	0,366	0,482	0	1	909
Femmina	0,647	0,478	0	1	424	0,634	0,482	0	1	1573
Età	29,61	11,25	15	67	655	29,77	10,75	15	74	2482
Istruzione										
Laurea	0,061	0,240	0	1	40	0,069	0,253	0	1	171
Diploma	0,421	0,494	0	1	276	0,407	0,491	0	1	1009
Obbligo	0,518	0,500	0	1	339	0,525	0,499	0	1	1302
Stato-Famiglia										
C.F.	0,141	0,348	0	1	92	0,151	0,357	0	1	373
Coniuge	0,241	0,428	0	1	158	0,225	0,418	0	1	559
Figlio/a	0,618	0,486	0	1	405	0,624	0,484	0	1	1550
P.O.	0,310	0,225	0	0,8	655	0,304	0,224	0	1,0	2482
Provenienza										
N.I.	0,261	0,440	0	1	171	0,258	0,438	0	1	641
Rientri	0,302	0,460	0	1	198	0,269	0,444	0	1	668
Disoccupati	0,437	0,496	0	1	286	0,473	0,499	0	1	1173
Orario										
T.P.o	0,620	0,486	0	1	406	0,647	0,478	0	1	1606
T.P.e	0,232	0,422	0	1	152	0,191	0,393	0	1	475
Qualsiasi	0,148	0,355	0	1	97	0,162	0,368	0	1	401
Sede										
S.C.Re.	0,389	0,488	0	1	255	0,373	0,484	0	1	927
S.C.Ra.	0,499	0,500	0	1	327	0,529	0,499	0	1	1314
Ovunque	0,111	0,315	0	1	73	0,097	0,296	0	1	241
Contratto										
C.T.I.	0,646	0,479	0	1	423	0,689	0,463	0	1	1709
C.T.e	0,076	0,266	0	1	50	0,054	0,226	0	1	134
Qualsiasi	0,278	0,448	0	1	182	0,257	0,437	0	1	639
Azioni	2,373	1,497	1	9	655	2,494	1,557	1	11	2482
I.U.C.	0,730	0,444	0	1	177	0,769	0,422	0	1	574
Sussidio	0,085	0,280	0	1	56	0,085	0,279	0	1	211

^(a) I dati si riferiscono all'unione dei campioni trimestrali relativi al periodo 1993:1 – 1995:1

^(b) Per le variabili dicotome è riportato il numero di casi in cui assumono valore pari a 1

Leggenda: C.F. (Capo Famiglia); P.O. (Percentuale di Occupati in famiglia); N.I. (Nuovi Ingressi); C.T.I. (Contratto a Tempo Indeterminato); C.T.e (Contratto a Termine); T.P.o (Tempo Pieno); T.P.e (Tempo Parziale); S.C.Re. (Sede preferita per il lavoro è il Comune di Residenza); S.C.Ra. (Sede preferita per il lavoro è un Comune Raggiungibile in giornata); I.U.C. (Iscritto all'Ufficio di Collocamento).

posizione centrale, con un sistema di trasporti relativamente diffuso che consente agevolmente il pendolarismo per cui diventa piú facile per i disoccupati accettare di spostarsi giornalmente per lavorare. La mobilità della forza lavoro meriterebbe un'attenzione maggiore di quanto le si possa dedicare qui, ma la letteratura è assai vasta (tra gli altri: Attanasio e Padoa-Schioppa, 1991; Sestito, 1995; Campiglio, 1996; Faini, Galli, Rossi, 1996).

La maggioranza cerca un impiego con contratto a tempo indeterminato: il 64,4% di coloro che hanno trovato lavoro e il 68,9% di coloro che cercano un'occupazione; si tratta di un valore un po' piú alto di quello del Nord che è pari al 65% circa e di quello dell'Italia che è 64,4% nel 1993 e 61,8% nel 1994. Coloro che accetterebbero qualunque forma contrattuale ammontano al 25,7% in Emilia-Romagna, mentre in Italia superano il 30% e i valori sono ancora piú elevati al Sud. Coloro che accetterebbero un contratto a termine sono il 5,4% e la percentuale è piú elevata di quella dell'Italia che è inferiore al 4%. Le differenze non sono notevoli, ma coerenti con la buona condizione economica della regione.⁹

L'iscrizione all'ufficio di collocamento riguarda il 73%. L'indennità o il sussidio di disoccupazione è stato percepito, invece, soltanto dall'8,5%: sia tra coloro che hanno trovato lavoro, sia tra coloro che lo cercano ancora. Il numero medio di azioni di ricerca effettuate da coloro che hanno poi trovato lavoro è 2,4 che corrisponde, in media, a circa 2 azioni all'anno. Un valore analogo si ottiene per coloro che stanno ancora cercando un'occupazione; presumibilmente, il valore medio per l'Italia non è molto diverso.¹⁰

Si sono esaminate le funzioni di rischio relative ai maschi, alle femmine, e al totale per verificare se il loro andamento risultava compatibile con la scelta della distribuzione di Weibull per la durata della disoccupazione. La stima di $h(t)$ dai dati osservati si può ottenere con il metodo delle tavole di sopravvivenza (Cox, Oakes, 1984) in base al quale il supporto di T è suddiviso in K intervalli dai punti $\{t_0 < t_1 < \dots < t_K\}$, dove $t_0 = 0$ e $t_K = \infty$, e la stima è data da

$$\hat{h}(t_{mi}) = \frac{2(1 - \hat{p}_i)}{(t_i - t_{i-1})(1 + \hat{p}_i)}$$

dove t_{mi} è il punto di mezzo dell'intervallo (t_{i-1}, t_i) , mentre \hat{p}_i è la probabilità stimata di trovare lavoro durante il periodo di tempo (t_{i-1}, t_i) per $i = 1, \dots, K - 1$ (si è trascurato così l'ultimo intervallo). Siano n_i la dimensione effettiva del campione, ossia il numero di disoccupati che entrano nell'intervallo (t_{i-1}, t_i) , e d_i il numero di disoccupati che trovano lavoro in quel periodo, allora $\hat{p}_i = d_i/n_i$. I valori di $\hat{h}(t_{mi})$ sono sensibili alla dimensione campionaria effettiva, all'ampiezza degli intervalli $(t_i - t_{i-1})$, e sono molto instabili alle durate piccole e grandi, perciò nella Figura 8 si sono riportati i valori delle funzioni di rischio cumulate¹¹ che sommando i valori di $\hat{h}(t_{mi})$ compensa queste variabilità o instabilità "strutturali". Le funzioni stimate

⁹Si noti, tra l'altro, che sia la preferenza per un orario ridotto di lavoro e sia il numero di occupati a tempo parziale presentano un'incidenza superiore tra le femmine (Di Nicola, 1995; McCall, 1996). Inoltre, la differenza tra la percentuale di occupati a tempo parziale e la percentuale di preferenza per un lavoro a tempo parziale mostra che il desiderio di trovare un lavoro che consenta una qualche forma di flessibilità oraria sembra insoddisfatto.

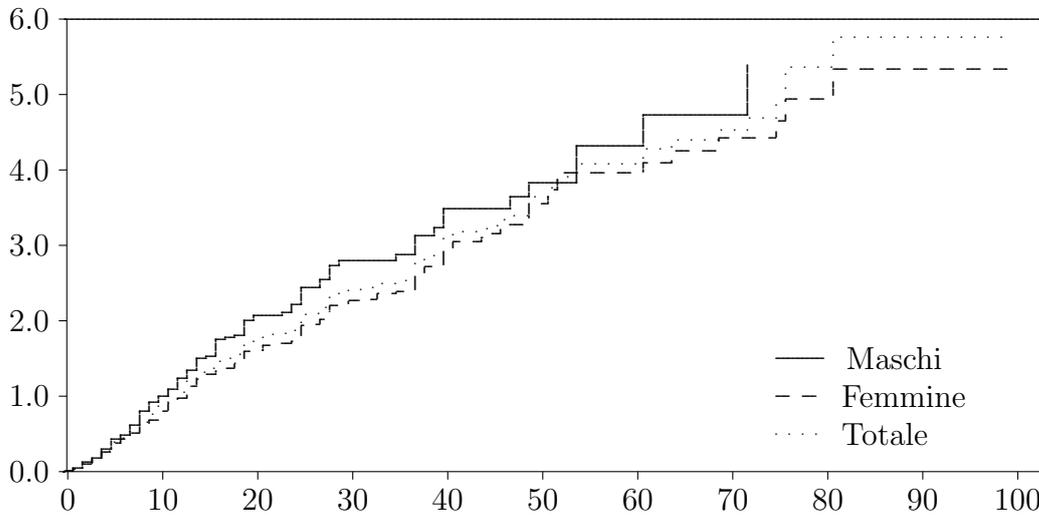
¹⁰Nelle pubblicazioni dell'ISTAT il numero di azioni di ricerca presenta una classe aperta: "maggiore o uguale a 5 azioni". La media del numero di azioni, calcolata dalla distribuzione di frequenze assolute e escludendo la classe aperta, è uguale a 1,9 nel 1993; solo considerando una media di 10 azioni per la classe aperta, si ottiene una media totale simile a quella della regione Emilia-Romagna. I valori restano quasi invariati nel 1994.

¹¹I valori di $\hat{h}(t_{mi})$ sono stati ottenuti con PROC LIFETEST del SAS (1992) e da essi si sono derivati, poi, quelli della funzione cumulata, $\hat{H}(t_{mi}) = \sum_{j=1}^i \hat{h}(t_{mj})$.

mostrano un andamento abbastanza regolare, crescente, e approssimativamente lineare o con una lieve concavità rivolta verso il basso; allora, la funzione di rischio (che è la derivata di $H(t)$ rispetto al tempo) sarà a sua volta quasi costante o tendenzialmente decrescente sicché se ne deduce che la distribuzione di Weibull sembra adeguata per la rappresentazione delle durate osservate della disoccupazione.

L'andamento della funzione di rischio $\hat{h}(t_{mi})$, così calcolata, si riferisce alle durate senza considerare le variabili esplicative, ma potrebbe accadere che per sottogruppi omogenei i comportamenti siano differenziati per cui si considerano, in seguito, sia i modelli con eterogeneità non osservata, sia i modelli a funzioni di rischio non proporzionali (v. *infra*).

Figura 8. Funzione di rischio cumulata per maschi, femmine, e totale



3.2. Nella Tabella 2 sono riportate le stime, $\hat{\lambda}$ e $\hat{\alpha}$, dei parametri, λ e α , della funzione di densità dei periodi trascorsi dal primo trimestre del 1993 al primo trimestre del 1995. I valori di α mostrano un andamento oscillante e a volte sono più piccoli dell'unità implicando una dipendenza di stato negativa. Per i maschi, $\hat{\alpha}$ è più piccolo di quello delle femmine e risulta statisticamente diverso per il secondo trimestre del 1993 e 1994; ma per le femmine, la durata media della disoccupazione è sempre più elevata di, ma non sempre statisticamente diversa da, quella dei maschi. La relazione esistente tra μ_U e μ_T , data dalla (5), indica che per i trimestri in cui si ha $\hat{\alpha} < 1$ la media delle durate complete non osservate risulta minore della media delle durate incomplete osservate. A questo stadio dell'analisi, ciò potrebbe derivare dal mancato controllo dell'eterogeneità; infatti, la media delle durate complete osservate è poco più alta, ma si tratta di un valore medio nell'intervallo di tempo considerato e poiché almeno in un trimestre di ogni anno si è avuto un $\hat{\alpha}$ maggiore di 1, i risultati mostrati sinora sono coerenti tra loro.

L'introduzione dell'eterogeneità nel modello (dati non riportati) comporta una stima del parametro α maggiore dell'unità implicando che la probabilità di

trovare lavoro cresce al crescere del tempo di attesa nei primi mesi e poi decresce quando il tempo supera un certo valore; le medie delle durate sono piú alte e, in particolare, le stime della media delle durate complete (non osservate) sono piú alte delle durate incomplete (osservate) di un ammontare che varia dal 30% al 60% circa di queste ultime; mentre la variabilità delle stime risulta piú elevata e talvolta non si può neanche calcolare. Le stime del parametro α per le durate complete (ricostruite) della disoccupazione sono un po' piú elevate dell'unità: sia per i maschi, sia per le femmine, sia per il totale; sebbene i valori di $\hat{\alpha}$ siano molto vicino a 1, risultano comunque statisticamente diverse dall'unità.¹² Il risultato è degno di nota, ma ha un valore soltanto indicativo perché non si sono ancora eseguiti i controlli per l'eterogeneità, né osservata e né non osservata.

Tabella 2 — *Stime dei parametri della distribuzione Weibull per i periodi incompleti di disoccupazione (in mesi) dal primo trimestre 1993 al primo trimestre 1995, e per i periodi completi ricostruiti (CR) nello stesso intervallo di tempo, relativi a maschi, femmine, e totale*

	λ_M	α_M	L	λ_F	α_F	L	λ	α	L
	(e.s.)	(e.s.)	n	(e.s.)	(e.s.)	n	(e.s.)	(e.s.)	n
93:1	0,253 (0,040)	0,621 (0,033)	-274,8 82	0,190 (0,019)	0,688 (0,027)	-599,2 178	0,207 (0,018)	0,665 (0,021)	-874,1 260
93:2	0,155 (0,019)	0,740 (0,039)	-380,8 112	0,081 (0,006)	1,034 (0,056)	-616,4 178	0,103 (0,007)	0,894 (0,034)	-999,1 290
93:3	0,132 (0,017)	0,830 (0,053)	-309,5 93	0,103 (0,010)	0,856 (0,043)	-546,3 155	0,115 (0,009)	0,837 (0,033)	-856,7 248
93:4	0,278 (0,038)	0,654 (0,032)	-326,7 105	0,299 (0,032)	0,590 (0,020)	-619,8 186	0,298 (0,025)	0,606 (0,017)	-947,6 291
94:1	0,163 (0,020)	0,730 (0,038)	-378,6 112	0,157 (0,015)	0,704 (0,027)	-671,5 192	0,160 (0,012)	0,712 (0,022)	-1050,5 304
94:2	0,136 (0,017)	0,728 (0,039)	-381,5 107	0,077 (0,007)	0,903 (0,045)	-662,1 178	0,095 (0,007)	0,824 (0,030)	-1044,9 285
94:3	0,136 (0,018)	0,836 (0,057)	-272,5 83	0,076 (0,007)	1,038 (0,064)	-489,7 139	0,097 (0,007)	0,933 (0,042)	-764,0 222
94:4	0,336 (0,048)	0,566 (0,024)	-364,6 109	0,221 (0,024)	0,620 (0,023)	-589,8 169	0,262 (0,023)	0,596 (0,017)	-955,0 278
95:1	0,169 (0,022)	0,720 (0,038)	-357,9 106	0,134 (0,012)	0,751 (0,030)	-695,7 198	0,146 (0,011)	0,738 (0,024)	-1054,1 304
CR	0,081 (0,005)	1,150 (0,060)	-795,7 231	0,067 (0,003)	1,115 (0,044)	-1546,2 424	0,071 (0,003)	1,121 (0,035)	-2345,6 655

¹²Le stime dei parametri sono state ottenute con la procedura *constr*(·) di MATLAB, nei programmi di ottimizzazione (Optimization Toolbox), che utilizza il metodo sequenziale di programmazione quadratica per la ricerca del massimo (Grace, 1994: 2/22-32, 3/9-12).

Si introduce quindi nel modello l'effetto telescopico per accertare come si modificano le stime di λ e α e qual è la funzione di accumulazione più adeguata. Nelle distribuzioni dei periodi incompleti, separate per maschi e femmine dal primo trimestre del 1993 al primo trimestre del 1995, i punti di accumulazione derivanti dalla presenza dell'effetto telescopico non sono sempre molto evidenti e distinti, ma sembrano abbastanza "stabili" in ogni trimestre e simili a quelli delle distribuzioni ottenute dall'unione dei campioni trimestrali riportate nella Figura 4 e nella Figura 5. I punti di accumulazione utilizzati per determinare la funzione di verosimiglianza sono $\{h_j: 1, 4, 6, 12, 15, 18, 24, 36, 48, 60, 72, 84, 96, 99\}$. Per le durate uguali a 30 e 40 si osserva solo una lieve tendenza all'accumulazione, come anche per le durate uguali a 10 e 20 solo per le femmine; in entrambi i casi, però, si rilevano intensità trascurabili nei singoli trimestri e, pertanto, tali punti sono stati ignorati. I primi punti della lista sembrano tipici arrotondamenti, ma non si può escludere con certezza che non indichino ingressi periodici nel mercato del lavoro. Nella distribuzione dei periodi completi, invece, i punti di accumulazione non si manifestano sempre a valori così ben distinti e vi è qualche differenza tra maschi e femmine (Figura 6 e Figura 7). I punti di accumulazione $\{h_j: 1; 2,5; 4; 5,5; 19,5; 25,5; 28,5; 37,5\}$ sono comuni ai maschi e alle femmine; aggiungendo quattro valori, $\{8; 8,5; 14; 16,5\}$, si ottiene l'insieme relativo ai maschi; aggiungendo sette valori, $\{4,5; 7,5; 9; 10,5; 13,5; 40,5; 49,5\}$, si ottiene l'insieme relativo alle femmine. La scelta della funzione di accumulazione si è basata sul confronto di una distribuzione esponenziale, $H(t; \lambda) = 1 - \exp(-\gamma t)$, e di una distribuzione di Weibull, $H(t; \lambda, \alpha) = 1 - \exp[-(\gamma t)^\delta]$.

Le stime dei parametri della funzione di accumulazione esponenziale e di Weibull sono riportate nella Tabella 3 e nella Tabella 4, rispettivamente, dove si può osservare che i valori di $\hat{\lambda}$ e $\hat{\alpha}$ non subiscono cambiamenti rilevanti, mentre aumentano considerevolmente gli errori standard delle stime per l'incertezza introdotta dall'effetto telescopico. Le stime del parametro δ sono risultate sempre minori dell'unità; infatti, come si supponeva, il fenomeno risulta "dipendente" dal tempo: è più intenso nei valori piccoli delle durate e diminuisce di intensità nei valori grandi.¹³ Tuttavia, si deve rilevare che in alcuni trimestri i dati sono coerenti con l'ipotesi nulla $\delta = 1$; quando $\hat{\delta}$ è statisticamente inferiore a 1, poi, i valori di $\hat{\gamma}$ sono tanto piccoli da rendere plausibile una distribuzione degenera con $\gamma = 0$. La distribuzione esponenziale è più parsimoniosa e non presenta alcuna difficoltà.

La stima del parametro γ della funzione esponenziale di accumulazione per i periodi completi non è risultata, invece, significativa per i maschi; mentre per le femmine è risultata statisticamente diverso da 0, ma la sua entità è pressoché trascurabile in quanto la ricostruzione delle durate complete ha disperso nell'arco di un anno o più il marcato effetto telescopico che si può osservare nelle durate incomplete, come si nota anche dall'elenco dei punti sopra riportato in cui ve ne sono alcuni molto vicini.

¹³Si noti che, se si considerano i disoccupati in base alla definizione precedente al 1992, le stime del parametro δ risultano molto più vicino a 1 e quasi sempre statisticamente uguali a 1.

Tabella 3 — *Stime dei parametri della distribuzione esponenziale per l'effetto telescopico e della distribuzione Weibull per i periodi incompleti di disoccupazione (in mesi) dal primo trimestre 1993 al primo trimestre 1995, e per i periodi completi ricostruiti (CR) nello stesso intervallo di tempo, relativi a totale (T), maschi (M), e femmine (F)*

		λ	(e.s.)	α	(e.s.)	γ	(e.s.)	L
1993:1	T	0,203	(0,050)	0,669	(0,062)	0,030	(0,005)	-825,2
	M	0,248	(0,119)	0,624	(0,102)	0,026	(0,009)	-261,0
	F	0,186	(0,053)	0,693	(0,078)	0,033	(0,007)	-563,9
1993:2	T	0,103	(0,016)	0,895	(0,078)	0,025	(0,004)	-951,8
	M	0,154	(0,048)	0,743	(0,100)	0,017	(0,006)	-370,5
	F	0,081	(0,014)	1,035	(0,118)	0,031	(0,006)	-578,2
1993:3	T	0,118	(0,025)	0,820	(0,089)	0,029	(0,005)	-808,4
	M	0,131	(0,043)	0,832	(0,139)	0,027	(0,008)	-295,8
	F	0,108	(0,030)	0,826	(0,117)	0,030	(0,006)	-511,6
1993:4	T	0,294	(0,080)	0,608	(0,055)	0,032	(0,005)	-887,6
	M	0,273	(0,111)	0,659	(0,098)	0,037	(0,011)	-306,6
	F	0,296	(0,106)	0,592	(0,067)	0,030	(0,006)	-579,7
1994:1	T	0,162	(0,035)	0,708	(0,061)	0,029	(0,005)	-987,0
	M	0,172	(0,059)	0,709	(0,099)	0,029	(0,008)	-358,3
	F	0,154	(0,042)	0,710	(0,079)	0,030	(0,006)	-628,5
1994:2	T	0,101	(0,020)	0,795	(0,074)	0,020	(0,003)	-994,0
	M	0,154	(0,057)	0,685	(0,100)	0,018	(0,006)	-365,4
	F	0,079	(0,017)	0,888	(0,108)	0,021	(0,004)	-627,1
1994:3	T	0,099	(0,020)	0,913	(0,107)	0,018	(0,004)	-741,0
	M	0,136	(0,046)	0,836	(0,145)	0,019	(0,008)	-267,0
	F	0,079	(0,019)	1,000	(0,157)	0,018	(0,005)	-472,2
1994:4	T	0,278	(0,082)	0,583	(0,054)	0,020	(0,004)	-910,5
	M	0,334	(0,159)	0,568	(0,081)	0,015	(0,005)	-354,2
	F	0,244	(0,090)	0,596	(0,072)	0,023	(0,005)	-555,3
1995:1	T	0,149	(0,032)	0,729	(0,066)	0,016	(0,003)	-1023,6
	M	0,168	(0,063)	0,721	(0,114)	0,015	(0,006)	-349,0
	F	0,139	(0,036)	0,737	(0,081)	0,016	(0,004)	-674,2
CR	T	0,071	(0,003)	1,114	(0,036)	0,007	(0,002)	-2328,6
	M	0,080	(0,006)	1,142	(0,062)	0,005	(0,003)	-792,6
	F	0,067	(0,003)	1,110	(0,044)	0,007	(0,002)	-1532,3

Le medie delle durate della disoccupazione, riportate nella Tabella 5, si riferiscono ai valori campionari e alle stime ottenute dai modelli, sia senza e sia con l'inclusione dell'eterogeneità. Le medie campionarie sono lievemente sottostimate dai modelli che non includono l'eterogeneità e le medie dei completi non osservati, ma stimati, risultano molto più basse di quelle campionarie. Questo risultato sarebbe plausibile in un mercato del lavoro che offre molte possibilità di occupazione

e, quindi, solo coloro che pongono vincoli molto stringenti non trovano un impiego. Allora, se la distribuzione di Weibull fosse adeguata, dovrebbe risultare $\hat{\alpha} < 1$ e la media dei periodi completi non osservati minore di quella degli incompleti osservati.

Tabella 4 — *Stime dei parametri della distribuzione di Weibull per l'effetto telescopico e della distribuzione Weibull per i periodi incompleti di disoccupazione (in mesi) dal primo trimestre 1993 al primo trimestre 1995, e per i periodi completi ricostruiti (CR) nello stesso intervallo di tempo, relativi a totale (T), maschi (M), e femmine (F)*

		λ	(e.s.)	α	(e.s.)	γ	(e.s.)	δ	(e.s.)	L
1993:1	T	0,203	(0,067)	0,670	(0,078)	0,008	(0,011)	0,268	(0,133)	-817,6
	M	0,246	(0,160)	0,626	(0,132)	0,015	(0,016)	0,448	(0,274)	-259,8
	F	0,186	(0,071)	0,693	(0,098)	0,005	(0,012)	0,200	(0,153)	-557,3
1993:2	T	0,103	(0,022)	0,895	(0,096)	0,022	(0,006)	0,804	(0,190)	-951,3
	M	0,154	(0,067)	0,743	(0,125)	0,010	(0,010)	0,612	(0,328)	-369,9
	F	0,081	(0,019)	1,035	(0,147)	0,031	(0,007)	0,957	(0,239)	-578,2
1993:3	T	0,117	(0,025)	0,826	(0,086)	0,015	(0,010)	0,457	(0,184)	-806,2
	M	0,130	(0,050)	0,836	(0,159)	0,015	(0,016)	0,537	(0,339)	-295,2
	F	0,107	(0,028)	0,833	(0,104)	0,012	(0,014)	0,364	(0,210)	-509,8
1993:4	T	0,294	(0,080)	0,608	(0,055)	0,032	(0,005)	0,032	(0,005)	-887,6
	M	0,272	(0,154)	0,660	(0,136)	0,035	(0,014)	0,887	(0,404)	-306,6
	F	0,296	(0,106)	0,592	(0,067)	0,030	(0,006)	0,030	(0,006)	-579,7
1994:1	T	0,161	(0,043)	0,709	(0,073)	0,018	(0,009)	0,444	(0,137)	-981,4
	M	0,173	(0,071)	0,709	(0,105)	0,009	(0,016)	0,305	(0,217)	-355,3
	F	0,153	(0,055)	0,712	(0,101)	0,021	(0,010)	0,512	(0,176)	-625,6
1994:2	T	0,101	(0,023)	0,797	(0,079)	0,011	(0,007)	0,489	(0,168)	-989,9
	M	0,153	(0,067)	0,686	(0,104)	0,013	(0,010)	0,636	(0,310)	-364,8
	F	0,078	(0,020)	0,890	(0,121)	0,009	(0,009)	0,416	(0,200)	-623,3
1994:3	T	0,097	(0,018)	0,924	(0,085)	0,000	—	0,108	(0,186)	-735,2
	M	0,135	(0,049)	0,840	(0,137)	0,000	—	0,100	(0,330)	-262,6
	F	0,078	(0,016)	1,017	(0,116)	0,000	—	0,113	(0,254)	-470,8
1994:4	T	0,276	(0,101)	0,584	(0,064)	0,016	(0,006)	0,704	(0,206)	-909,7
	M	0,339	(0,227)	0,565	(0,109)	0,000	—	0,115	(0,230)	-352,3
	F	0,243	(0,107)	0,597	(0,081)	0,021	(0,006)	0,876	(0,274)	-555,1
1995:1	T	0,148	(0,040)	0,731	(0,081)	0,002	(0,004)	0,327	(0,177)	-1019,0
	M	0,168	(0,093)	0,721	(0,174)	0,000	—	0,146	(0,246)	-347,2
	F	0,138	(0,044)	0,738	(0,097)	0,004	(0,006)	0,380	(0,221)	-671,1
CR	T	0,071	(0,003)	1,114	(0,038)	0,007	(0,704)	1,055	(0,604)	-2328,6
	M	0,080	(0,006)	1,141	(0,065)	0,006	(1,994)	1,110	(1,187)	-792,6
	F	0,067	(0,003)	1,110	(0,047)	0,008	(0,746)	1,029	(0,493)	-1532,3

I valori delle medie campionarie sono sovrastimate di due tre mesi, e in alcuni casi anche di cinque sei mesi, dai modelli che includono l'eterogeneità. Le

stime dei periodi completi non osservati risultano maggiori del 70% circa, in media, delle medie campionarie. Il risultato è enfatizzato dalla sovrastima iniziale, ma pare coerente con il senso comune che si aspetta una media dei completi piú alta degli incompleti. D'altronde, la media dei periodi completi ricostruiti è circa uguale a quella degli incompleti e, pertanto, si può evincere che occorrono ulteriori analisi per valutare l'impatto della distorsione campionaria delle durate.

Tabella 5 — *Medie campionarie stimate senza e con eterogeneità non osservata (con effetto telescopico) per i periodi incompleti di disoccupazione (in mesi) dal primo trimestre 1993 al primo trimestre 1995, e per i periodi completi ricostruiti nello stesso intervallo di tempo, relativi a maschi (M), femmine (F), e totale (T)*

	μ	Valori campionari			Stime senza eterogeneità			Stime con eterogeneità		
		M	F	T	M	F	T	M	F	T
1993:1	$\widehat{\mu}_U$	11,1	11,0	11,1	11,0	11,0	11,0	13,1	12,6	12,7
	$\widehat{\mu}_T$				5,8	6,9	6,5	18,6	18,3	18,4
1993:2	$\widehat{\mu}_U$	11,3	11,7	11,6	11,2	11,8	11,6	11,1	11,8	11,5
	$\widehat{\mu}_T$				7,8	12,2	10,3	17,4	19,7	18,8
1993:3	$\widehat{\mu}_U$	10,3	12,6	11,7	10,4	12,7	11,8	11,2	14,0	13,1
	$\widehat{\mu}_T$				8,4	10,2	9,4	17,7	22,4	20,7
1993:4	$\widehat{\mu}_U$	8,6	10,9	10,1	8,5	10,9	10,0	11,0	17,1	14,6
	$\widehat{\mu}_T$				4,9	5,2	5,0	15,3	22,0	19,3
1994:1	$\widehat{\mu}_U$	11,1	12,5	12,0	11,2	12,4	12,0	12,3	14,6	13,7
	$\widehat{\mu}_T$				7,2	8,1	7,7	18,3	21,3	20,2
1994:2	$\widehat{\mu}_U$	13,3	15,2	14,5	13,6	15,3	14,7	14,1	16,6	15,7
	$\widehat{\mu}_T$				8,4	13,5	11,3	21,7	26,2	24,5
1994:3	$\widehat{\mu}_U$	9,9	12,5	11,5	9,9	12,6	11,6	10,2	13,2	12,3
	$\widehat{\mu}_T$				8,1	12,6	10,5	16,9	23,1	20,7
1994:4	$\widehat{\mu}_U$	11,3	12,7	12,1	11,1	12,9	12,2	15,9	18,5	17,6
	$\widehat{\mu}_T$				4,9	6,2	5,6	20,9	24,6	23,2
1995:1	$\widehat{\mu}_U$	11,0	12,6	12,0	11,4	12,6	12,1	14,4	14,8	14,7
	$\widehat{\mu}_T$				7,3	8,7	8,2	20,1	21,8	21,2
Completi	$\widehat{\mu}_T$	11,7	14,3	13,4	11,8	14,5	13,4	13,1	16,6	15,3

3.3. Le durate incomplete della disoccupazione sono state analizzate con un insieme identico di covariate per ognuna delle nove indagini trimestrali tramite: (a) la $f(u, \mathbf{x}; \lambda, \alpha, \boldsymbol{\beta})$ e le stime dei parametri sono riportate in Tabella 6; (b) la $\psi(u, \mathbf{x}; \vartheta, \alpha, \boldsymbol{\beta})$ e le stime dei parametri sono riportate in Tabella 7; (c) la $f_{LL}(u, \mathbf{x}; \omega, \alpha, \boldsymbol{\beta})$ e le stime dei parametri sono riportate in Tabella 8. Le durate complete (ricostruite)

della disoccupazione sono state analizzate con le funzioni corrispondenti, $f(t, \mathbf{x}; \cdot)$ o $\psi(t, \mathbf{x}; \cdot)$ o $f_{LL}(t, \mathbf{x}; \cdot)$, sia per il campione totale e sia separatamente per maschi e femmine; le stime dei parametri sono riportate nella Tabella 9. Tutti i modelli nelle Tabelle 6–9 includono l’effetto telescopico.¹⁴

Le variabili che possono influenzare il processo di ricerca del lavoro sono numerose e quelle considerate sono state elencate nella Tabella 1. La loro scelta è stata eseguita in base a alcuni criteri, in parte arbitrari, elaborando l’insieme delle durate complete ricostruite: si sono tralasciate le variabili di destinazione occupazionale nell’analisi dei periodi completi per avere un modello “confrontabile” con quello relativo ai periodi incompleti; si sono trascurate altresí le variabili di provenienza per i disoccupati e i rientri al fine di avere caratteri comuni a quelli concernenti i nuovi ingressi. Tra le rimanenti covariate si è proceduto quindi a una selezione automatica, con il criterio delle stime di massima verosimiglianza parziale e il metodo *backward* nel modello proporzionale di Cox, per individuare quelle che risultavano determinanti nell’insieme delle durate complete. Un numero piccolo di covariate ha mostrato un impatto significativo sulla funzione di azzardo di base, $h_0(t; \boldsymbol{\theta})$. A queste si sono aggiunte le altre, ritenute comunque importanti nel processo di ricerca del lavoro.¹⁵ Si può ritenere, quindi, che l’insieme ottenuto sia incompleto per cui il controllo sull’eterogeneità sia inefficace perché non solo si sono trascurate alcune covariate, ma l’informazione disponibile sulla storia lavorativa pregressa è largamente insufficiente.

Le stime del parametro di forma, $\hat{\alpha}$, riportate nella Tabella 6 e relative alle durate incomplete, variano ancora da un trimestre all’altro e risultano quasi sempre inferiori a 1, ma statisticamente uguali tra loro e all’unità, eccetto nel quarto trimestre del 1993 e del 1994, dove il valore di $\hat{\alpha}$, è significativamente minore di 1. Quando il valore di $\hat{\alpha}$ è minore di 1 si ha una dipendenza dallo stato negativa prevista da alcune teorie (Salant, 1977) e dalla presenza di eterogeneità non osservata, ma tuttora non si sono conseguite evidenze empiriche sufficientemente stabili: in alcuni casi emerge una dipendenza dallo stato negativa (Ham, Rea, 1987; Torelli, Trivellato, 1988; Lynch, 1989); in altri casi si osserva una dipendenza dallo stato positiva (Hujer, Schneider, 1989); e in altri ancora si verifica una dipendenza dallo stato non significativa (Flinn, 1986; Groot, 1990). Si potrebbe ipotizzare che i valori stimati di α minori di 1 derivino proprio dallo scarso controllo dell’eterogeneità perché gli insiemi di dati analizzati contengono sottoinsiemi molto caratterizzati: come i nuovi ingressi e i rientri, giovani e meno giovani. Tuttavia, nei periodi completi si ha la stessa composizione di gruppo e la stima del parametro di forma della Weibull, riportata nella terza colonna della Tabella 9, risulta ancora maggiore di 1 implicando una dipendenza dallo stato positiva, che sembra coerente con l’andamento della funzione di rischio cumulata ottenuta dai dati (Figura 8). Si noti, tra

¹⁴I coefficienti significativamente diversi da 0 sono in grassetto e quelli con una probabilità di significatività, p , compresa tra 0,05 e 0,10 sono contrassegnati dall’asterisco.

¹⁵Per semplicità, non si sono considerate alcune covariate che non erano mai significative come il coniuge, il tipo di contratto a tempo parziale, la sede preferita nel comune di residenza, e l’iscrizione all’ufficio di collocamento; mentre sono state inserite in un modello riportato in un altro lavoro (Lalla e Pattarin, 1998).

l'altro, che i valori di $\hat{\alpha}$ ottenuti per i periodi completi implicano una durata media dei periodi incompleti osservati un po' piú bassa della durata media dei periodi completi, che corrisponde a quanto riportato nella Tabella 5. I valori delle stime del parametro di scala, λ , mostrano variazioni che dipendono sia dai valori di $\hat{\alpha}$, sia dalla durata media della disoccupazione: valori piccoli di $\hat{\lambda}$ denotano una durata media piú elevata rispetto ai valori grandi. Le stime del parametro della funzione esponenziale di accumulazione, $\hat{\gamma}$, risultano ancora significativamente piú piccole per le durate complete implicando un impatto trascurabile dell'effetto telescopico.

Le stime dei parametri di forma $\hat{\vartheta}$ e $\hat{\omega}$, riportate nelle Tabelle 7–8 e relative alle durate incomplete modellate con $\psi(u, \mathbf{x}; \cdot)$ e con $f_{LL}(u, \mathbf{x}; \cdot)$, variano anch'esse da un trimestre all'altro e mostrano valori che indicano un andamento della funzione di rischio crescente per i primi mesi, fino a raggiungere il massimo in un punto, che in base alle stime si trova tra 3 e 15 mesi; dopo comincia a decrescere (Figure 9–11). Nel quarto trimestre di ogni anno il valore delle stime è piú piccolo e risulta anche significativamente diverso dalle stime che presentano il valore piú elevato. Le stime di $\hat{\alpha}$ rimangono, invece, pressoché immutate nei due trimestri corrispondenti. Gli andamenti delle funzioni di rischio di base per $\mathbf{x} = 0$ e il modello $\psi(u, \mathbf{x}; \cdot)$ mostrano, poi, una intensità piú bassa (quasi la metà) di quelle relative ai modelli $f_{LL}(u, \mathbf{x}; \cdot)$ nei momenti iniziali della durata, come si vede nelle Figure 9 e 10.

3.3.1. Il modello relativo ai periodi incompleti senza eterogeneità, presenta covariate significative che spesso variano da un trimestre all'altro. In prevalenza, coefficienti negativi sono risultati per i nuovi ingressi, e i rientri.¹⁶ Ne consegue che chi entra per la prima volta nel mercato del lavoro, o rientra dopo un certo periodo di assenza, incontra piú difficoltà di chi ha un'esperienza lavorativa e rimane sempre sul mercato del lavoro. L'ipotesi di uguaglianza dei parametri relativi ai vari trimestri è coerente con i dati osservati. Coefficienti negativi e significativi solo in qualche trimestre si sono avuti per le donne, il contratto a tempo indeterminato, il numero di azioni di ricerca (di un posto di lavoro), e l'essere capofamiglia. Anche se tali evidenze possono solo indicare delle tendenze non necessariamente attendibili, il segno negativo per le donne era in un certo senso atteso perché è noto che le donne siano piú svantaggiate nella ricerca di lavoro e aspettano in media piú degli uomini; lo stesso vale per la ricerca di un contratto a tempo indeterminato. Il numero di azioni di ricerca e la condizione di capofamiglia avrebbero dovuto mostrare un segno positivo. Ora, sebbene il numero di volte in cui la stima è risultata significativamente diversa da zero non sia elevato, il risultato potrebbe essere di interesse perché indica che piú si cerca e piú diminuisce la probabilità di trovare lavoro (Si utilizzano canali inefficaci? Si hanno vincoli stringenti? Non si hanno professionalità adeguate?); così pure, l'essere capofamiglia comporta un

¹⁶Nei modelli a rischi proporzionali relativi a $f(\cdot)$ e $\psi(\cdot)$, le covariate con coefficienti negativi comportano una diminuzione di $h_0(t; \boldsymbol{\theta})$ e, quindi, della probabilità di trovare lavoro implicando un aumento dell'attesa: “frenano” l'uscita dallo stato. Viceversa, le covariate con coefficienti positivi comportano un aumento di $h_0(t; \boldsymbol{\theta})$ e, quindi, della probabilità di trovare lavoro implicando una diminuzione dell'attesa: “accelerano” l'uscita dallo stato. La relazione è piú complessa nei modelli a rischi non proporzionali espressi con $f_{LL}(\cdot)$.

ritardo nella ricerca di lavoro (Ci sono vantaggi per loro?).

Coefficienti positivi si sono osservati quasi stabilmente per quelli che percepiscono un sussidio. Anche in questo caso il segno non è proprio quello atteso perché ci si aspettava che il disoccupato con sussidio tendesse a ritardare il suo ingresso sul lavoro; tuttavia, numerosi studi hanno trovato risultati incoerenti: Nickell (1979a,b) stima un coefficiente negativo per il sussidio, significativo per i periodi brevi (inferiori a 20 settimane) e positivo per i periodi lunghi (superiori a 20 settimane); Narendranathan, Nickell, e Stern (1985) presentano ancora un coefficiente negativo per i giovani sotto i venti anni e positivo per gli altri anche se non è sempre diverso da zero in modo significativo;¹⁷ Ham e Rea (1987) ottengono un segno positivo che ritengono errato e lo attribuiscono alla variazione insufficiente della variabile; Groot (1990) stima un effetto negativo significativo; Narendranathan e Stewart (1993) trovano di nuovo un coefficiente negativo che dipende anche dalla durata della disoccupazione. I disturbi che possono influenzare gli effetti del sussidio sulla durata sono diversi e di natura varia. Il sussidio potrebbe incentivare la non ricerca del lavoro fino a quando la scadenza del diritto a percepirlo è lontana; quando si avvicina, invece, la ricerca potrebbe diventare più frenetica e l'accettazione dei posti trovati meno soggetta a vincoli; che è, poi, quanto previsto dalla teoria della ricerca del lavoro (*job-search*).¹⁸ Coefficienti positivi, solo saltuariamente significativi, sono risultati per i laureati, i diplomati, la ricerca di un lavoro a tempo pieno, la preferenza per una sede in un comune raggiungibile. Ne consegue che i laureati, i diplomati, e coloro che sono disposti a lavorare anche in un comune raggiungibile sono favoriti nella ricerca di un lavoro, come pure coloro che cercano un impiego a tempo pieno. In particolare, per chi si mostra disponibile a spostarsi per lavorare si può presumere che ciò sia un indicatore di determinazione nella ricerca di lavoro che consente di accelerare l'uscita dalla disoccupazione. L'impatto positivo sulla funzione di rischio della ricerca di un lavoro a tempo pieno deriva presumibilmente dalla scarsa disponibilità di posti a tempo parziale per cui trovano "più facilmente" coloro che cercano un lavoro a tempo pieno; infatti, si è visto che c'è uno scarto non trascurabile tra la proporzione di disoccupati che preferiscono un orario a tempo parziale e quella degli occupati a tempo parziale.¹⁹

¹⁷Il segno negativo è attribuito al basso livello dell'offerta ricevuta dal gruppo dei giovani. Ne può conseguire un abbassamento del salario di riserva all'aumentare del tempo di ricerca, che porterebbe a accettare la prima offerta a prescindere dal sussidio percepito.

¹⁸Il segno positivo ottenuto potrebbe derivare dalla presenza di lavoratori stagionali (del settore agricolo e turistico) che percepiscono il sussidio nei periodi di assenza dal lavoro, ma aspettano in genere poco tra due occupazioni. Tra l'altro, la variabile è dicotoma; pertanto, è come se individuasse una "categoria" e i percettori sono meno del 10%. Questa possibilità potrebbe essere corroborata eseguendo la stima dei modelli sul campione ridotto che esclude i settori "agricoltura, caccia, silvicoltura e pesca" e "Alberghi e ristoranti". Una diminuzione di casi nei campioni trimestrali renderebbe molto instabile le stime sicché la questione, per ora, è tralasciata o ignorata.

¹⁹Il panorama legislativo forse non offre ampi margini di flessibilità alle imprese o non c'è convenienza all'uso del tempo parziale nell'organizzazione del lavoro. Per esempio, solo recentemente la Legge n. 196 del 24 giugno 1997 (c.d. "Pacchetto Treu", S.O. n. 136L alla G.U. n. 154 del 4 luglio) ha esteso e potenziato la gamma dei contratti "atipici" (Biagi, Tiraboschi, 1998).

I modelli stimati non mostrano sempre un buon adattamento con la distribuzione ipotizzata; il valore osservato della statistica di Kolmogorov-Smirnov conduce talvolta a sostenere l'incoerenza dei dati con l'ipotesi: per il modello $f(u, \mathbf{x}; \cdot)$ ciò accade nel quarto trimestre del 1993 e 1994 e nel primo del 1995 (Tabella 6); per i modelli $\psi(u, \mathbf{x}; \cdot)$ e $f_{LL}(u, \mathbf{x}; \cdot)$ ciò accade nei trimestri terzo e quarto del 1993 e nel terzo del 1994 (Tabelle 7-8).

L'età al momento dell'ingresso nello stato di disoccupato (divisa per 10) è stata inserita con un polinomio di secondo grado: il termine lineare presenta un coefficiente negativo e il termine quadratico presenta un coefficiente positivo. La probabilità di trovare lavoro rispetto all'età ha, quindi, un andamento non lineare: cresce al crescere dell'età fino a raggiungere un massimo a una data età, oltre la quale comincia a diminuire. L'età è risultata significativa solo nel primo trimestre del 1995; tuttavia, il segno dei coefficienti (anche se non significativi) è risultato quasi sempre coerente con quello atteso.

3.3.2. Nel modello relativo ai periodi completi senza eterogeneità, coefficienti negativi sono risultati per le donne ($p = 0,086$), i nuovi ingressi, e i rientri. Ne consegue che le donne tendono a aspettare più degli uomini e sono quindi svantaggiate nella ricerca di lavoro; infatti, presentano una durata media significativamente più alta degli uomini e una dipendenza più complessa dalle covariate. I soggetti che entrano per la prima volta nel mercato del lavoro, o rientrano dopo un certo periodo di assenza, incontrano più difficoltà di chi ha un'esperienza lavorativa e rimane sempre sul mercato del lavoro. Coefficienti positivi sono risultati per i laureati, per la ricerca di un lavoro a tempo pieno, per chi preferisce una sede in un comune raggiungibile, per chi percepisce un sussidio. Come già detto, i laureati, chi è disposto a lavorare anche in un comune raggiungibile, e chi cerca un lavoro a tempo pieno risultano ancora favoriti nella ricerca di un lavoro. Anche il sussidio mostra ancora un impatto positivo sulla funzione di rischio. I modelli relativi alle donne e al totale non mostrano, però, un buon adattamento con la distribuzione ipotizzata: il valore osservato della statistica di Kolmogorov-Smirnov comporta l'incoerenza dei dati con l'ipotesi. Infine, nella Tabella 9 si può notare che il modello per i maschi è molto più semplice di quello per le femmine: sono significative soltanto le covariate "preferenza per la sede in un comune raggiungibile" (al 5%), "nuovi ingressi" e laureati (al 10%).

I modelli relativi ai periodi completi che includono l'eterogeneità non osservata o che utilizzano la log-logistica risultano coerenti con la distribuzione adottata. I parametri delle funzioni di rischio di base sono, sí, diversi, ma gli andamenti risultanti coincidono, come si può vedere in Figura 11: le linee tratteggiate e punteggiate non si distinguono. I coefficienti delle covariate sono essenzialmente identici per i modelli $\psi(t, \mathbf{x}; \cdot)$ e $f_{LL}(t, \mathbf{x}; \cdot)$ e risultano statisticamente uguali a quelli del modello $f(t, \mathbf{x}; \cdot)$; mentre gli errori dei modelli che includono l'eterogeneità sono più elevati. Questo vale anche per le stime dei parametri delle covariate relativi ai periodi incompleti. L'uguaglianza dei coefficienti delle covariate nei vari modelli sembra indicare che quando dall'esame dei periodi completi si individua una distribuzione "tipica" delle durate, allora si possono elaborare soltanto le durate incomplete con

risultati soddisfacenti. Inoltre, può dare anche “robustezza” alle stime ottenute e alle ipotesi adottate per la distribuzione delle durate; tuttavia, bisogna sempre tenere conto dei numerosi effetti concomitanti (non sempre considerati) che possono alterare i risultati in modo imprevedibile (Heckman, Singer, 1984b), e l’eterogeneità non osservata può non essere correttamente rappresentata dal modello utilizzato.

Figura 9. Funzioni di rischio stimate nei modelli con eterogeneità

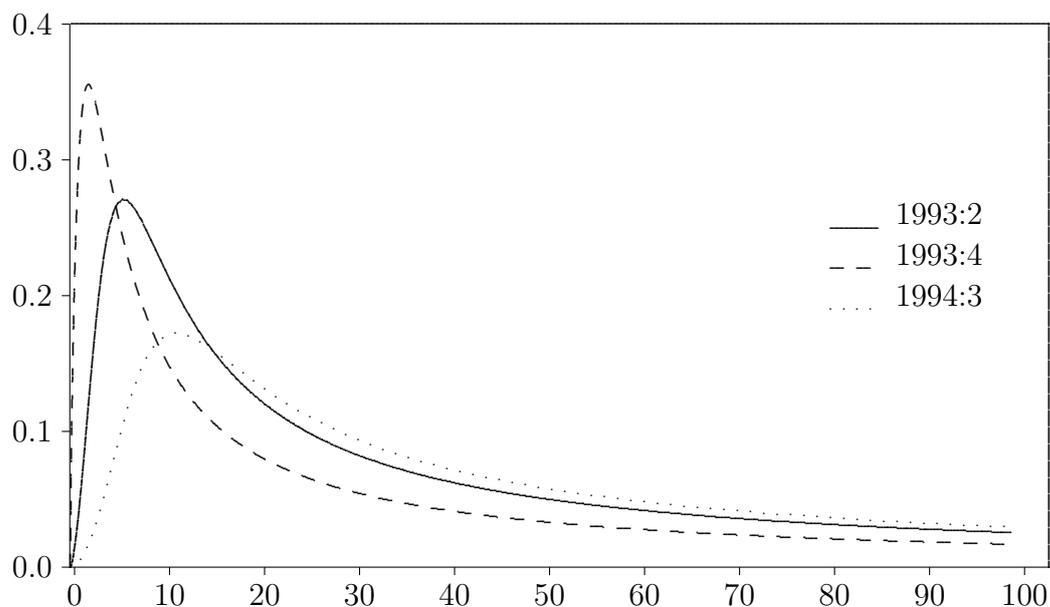


Figura 10. Funzioni di rischio stimate nei modelli con la log-logistica

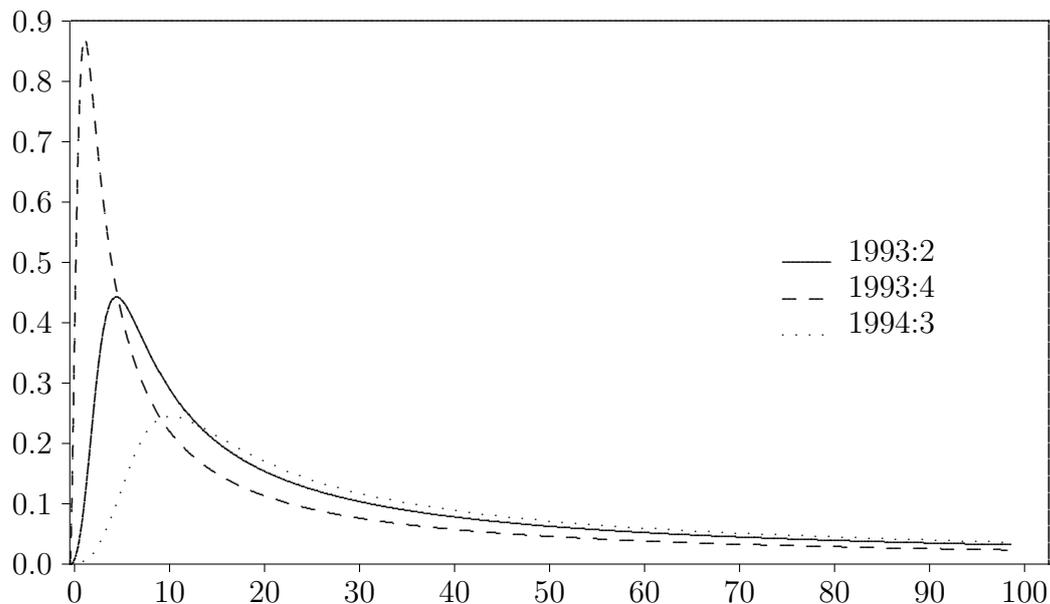
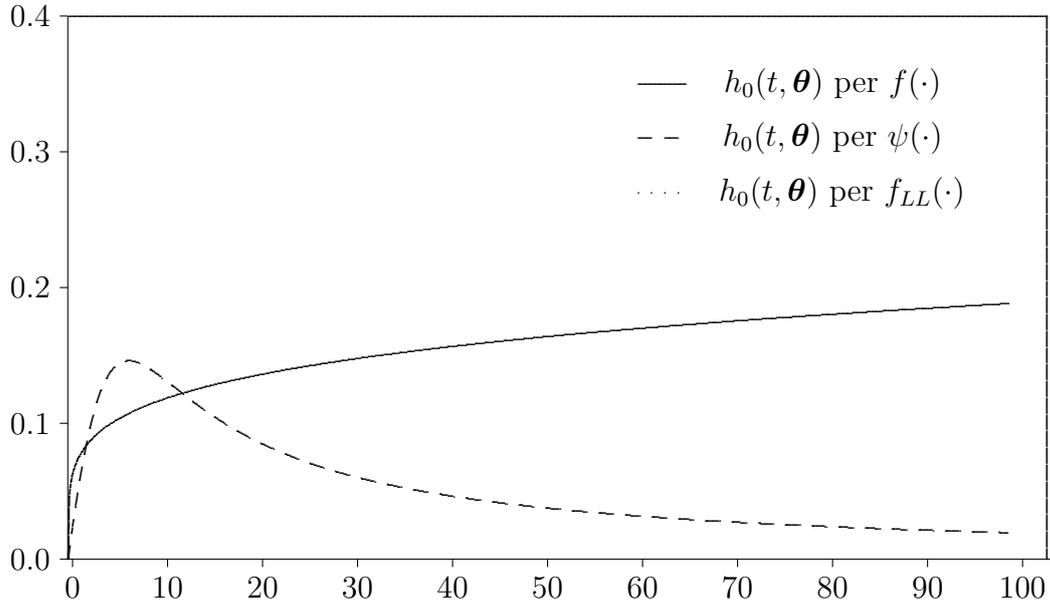


Figura 11. Funzioni di rischio stimate nei modelli per i periodi completi



Si noti, infine, che gli errori standard dei coefficienti sono stati ottenuti dall'inversa della matrice di informazione di Fisher, stimata tramite la matrice di covarianza delle derivate prime della funzione di verosimiglianza, valutate nel punto di massimo. Tale stimatore, detto BHHH o prodotto esterno dei gradienti (OPG), è conveniente perché richiede soltanto il calcolo delle derivate prime e la matrice risulta sempre definita positiva (Greene, 1993: 138-139). Tuttavia, in alcuni casi può produrre valori tali che la matrice di informazione di Fisher non sia invertibile anche nel punto di massimo della funzione di verosimiglianza. Gli errori standard delle stime dei parametri sono stati calcolati (soltanto per i periodi incompleti) anche con l'inversa della matrice di informazione di Fisher costruita con le derivate parziali seconde, cambiate di segno e valutate nel punto di massimo, per verificare le differenze con le stime conseguite tramite il prodotto esterno dei gradienti: le significatività dei coefficienti restavano immutate, ma i valori degli errori standard BHHH oscillavano tra il 5% e il 30% in più di quelli ottenuti con le derivate seconde. Il coefficiente di determinazione, R^2 , riportato nella Tabelle 6-9 è stato calcolato tramite la relazione: $R^2 = (1 - L_R^{2/n})/[1 - L(\theta)^{2/n}]$, dove L_R indica il rapporto tra le funzioni di verosimiglianza dei modelli senza variabili esplicative e con variabili esplicative, ossia $L_R = L(\theta)/L(\theta, \beta)$ (Maddala, 1983; Greene, 1993). I modelli che includono l'eterogeneità non osservata o assumono un rischio non proporzionale presentano un coefficiente di determinazione più basso da cui si può desumere che forse si ha già un elevato grado di adattamento delle funzioni di ripartizione stimate con le funzioni di ripartizione empiriche, anche in assenza di covariate.

Tabella 6 — *Stime dei parametri delle covariate, della distribuzione esponenziale per l'effetto telescopico, e della distribuzione di Weibull per i periodi incompleti (in mesi) dal primo trimestre 1993 al primo trimestre 1995*

θ, x	93:1	93:2	93:3	93:4	94:1	94:2	94:3	94:4	95:1
$\hat{\lambda}$	0,100 (0,092)	0,107 (0,088)	0,143 (0,123)	0,377 (0,348)	0,259 (0,199)	0,524 (0,485)	0,119 (0,103)	1,356 (1,306)	0,563 (0,424)
$\hat{\alpha}$	0,886 (0,119)	1,022 (0,145)	0,978 (0,141)	0,751 (0,105)	0,881 (0,117)	0,960 (0,116)	1,118 (0,148)	0,760 (0,100)	0,876 (0,117)
$\hat{\gamma}$	0,031 (0,005)	0,025 (0,005)	0,029 (0,005)	0,032 (0,005)	0,029 (0,005)	0,020 (0,003)	0,018 (0,005)	0,020 (0,004)	0,016 (0,003)
Femmine	0,039 (0,160)	-0,101 (0,131)	-0,050 (0,163)	-0,165 (0,132)	-0,231 (0,165)	-0,224 (0,162)	-0,444* (0,254)	-0,106 (0,128)	-0,176 (0,139)
E/10	-0,068 (0,385)	0,084 (0,417)	-0,471 (0,491)	-0,228 (0,350)	-0,423 (0,390)	-0,894* (0,517)	-0,026 (0,538)	-0,527 (0,415)	-0,691* (0,367)
(E/10) ²	0,028 (0,049)	-0,008 (0,060)	0,068 (0,073)	0,032 (0,047)	0,077 (0,057)	0,123 (0,075)	-0,016 (0,079)	0,033 (0,059)	0,089* (0,051)
Diploma	0,338 (0,164)	0,277 (0,182)	0,103 (0,191)	0,079 (0,139)	0,233* (0,130)	0,180 (0,171)	0,089 (0,215)	0,173 (0,122)	0,273 (0,128)
Laurea	1,370 (0,431)	0,467 (0,298)	0,393 (0,297)	0,136 (0,222)	0,104 (0,278)	0,629 (0,265)	0,511 (0,430)	0,673 (0,250)	0,364 (0,304)
C.F.	-0,416 (0,259)	-0,292 (0,214)	-0,164 (0,287)	-0,325 (0,246)	-0,553 (0,255)	-0,456 (0,231)	-0,170 (0,291)	-0,042 (0,170)	-0,202 (0,198)
P.O.	-0,080 (0,325)	-0,409 (0,349)	0,104 (0,366)	0,371 (0,267)	0,456* (0,241)	-0,300 (0,301)	0,859* (0,448)	0,078 (0,257)	0,212 (0,268)
N.I.	-0,778 (0,213)	-0,562 (0,183)	-0,370 (0,187)	-0,547 (0,168)	-0,773 (0,214)	-0,914 (0,247)	-0,351 (0,324)	-0,765 (0,189)	-0,492 (0,194)
Rientri	-0,006 (0,183)	-0,375 (0,164)	-0,114 (0,200)	-0,266* (0,147)	-0,536 (0,157)	-0,498 (0,147)	-0,041 (0,234)	-0,254 (0,156)	-0,297* (0,153)
Azioni	0,039 (0,052)	-0,038 (0,054)	0,013 (0,052)	-0,077* (0,044)	-0,053 (0,047)	-0,003 (0,049)	-0,089 (0,079)	-0,059 (0,038)	-0,035 (0,037)
C.T.I.	-0,333 (0,157)	-0,041 (0,141)	0,109 (0,198)	-0,279* (0,150)	0,140 (0,138)	-0,187 (0,152)	-0,208 (0,207)	-0,377 (0,168)	-0,053 (0,140)
C.T.e	-0,575* (0,320)	0,225 (0,330)	0,662* (0,344)	0,401 (0,437)	-0,173 (0,294)	0,137 (0,326)	-0,157 (0,491)	0,085 (0,321)	0,787 (0,339)
T.P.o	0,450 (0,161)	0,083 (0,156)	0,111 (0,176)	0,447 (0,160)	-0,076 (0,131)	0,098 (0,159)	0,205 (0,196)	0,229 (0,153)	-0,002 (0,137)
S.C.Ra.	-0,067 (0,126)	0,029 (0,129)	0,206 (0,146)	0,112 (0,108)	0,201* (0,117)	0,386 (0,139)	0,108 (0,209)	-0,083 (0,130)	-0,194 (0,130)
Sussidio	0,630 (0,230)	0,355 (0,263)	0,778* (0,455)	0,723 (0,323)	0,452 (0,280)	0,024 (0,266)	1,318 (0,355)	0,535 (0,223)	0,509 (0,202)
Log(L)	-793,0	-936,9	-793,5	-861,1	-960,8	-972,5	-721,7	-879,6	-1004,5
R ²	0,228	0,103	0,122	0,175	0,169	0,157	0,172	0,209	0,136
n	260	290	248	291	304	285	222	278	304
$\lambda_{K-S; oss}$	0,934	1,221	1,083	1,650	1,191	0,961	1,033	1,397	1,446

Grassetto : p (probabilità di significatività) $\leq 0,05$. / Asterisco : $0,05 < p \leq 0,10$.

Tabella 7 — *Stime dei parametri delle covariate, della distribuzione esponenziale per l'effetto telescopico, e della distribuzione di Weibull con eterogeneità non osservata per i periodi incompleti (in mesi) dal primo trimestre 1993 al primo trimestre 1995*

θ, x	93:1	93:2	93:3	93:4	94:1	94:2	94:3	94:4	95:1
$\hat{\vartheta}$	12,58 (9,82)	4,757 (3,351)	9,813 (7,075)	2,399 (2,048)	8,501 (6,478)	3,451 (2,851)	8,919 (6,195)	0,901 (0,812)	1,601 (1,244)
$\hat{\alpha}$	2,073 (0,243)	2,522 (0,264)	2,313 (0,284)	1,674 (0,192)	1,967 (0,223)	2,245 (0,234)	2,922 (0,359)	1,704 (0,188)	1,847 (0,216)
$\hat{\gamma}$	0,030 (0,005)	0,025 (0,005)	0,029 (0,005)	0,032 (0,005)	0,030 (0,005)	0,020 (0,003)	0,018 (0,005)	0,020 (0,004)	0,016 (0,003)
Femmine	0,081 (0,350)	-0,345 (0,334)	-0,322 (0,371)	-0,347 (0,283)	-0,378 (0,334)	-0,459 (0,362)	-1,082 (0,488)	-0,175 (0,279)	-0,356 (0,287)
E/10	0,015 (0,828)	-0,664 (0,951)	-0,770 (0,968)	-0,911 (0,754)	-0,240 (0,845)	-1,552 (1,121)	-0,543 (1,066)	-1,358 (0,864)	-1,796 (0,822)
(E/10) ²	0,050 (0,109)	0,084 (0,138)	0,107 (0,142)	0,132 (0,103)	0,075 (0,122)	0,211 (0,164)	0,002 (0,149)	0,089 (0,120)	0,237 (0,116)
Diploma	0,948 (0,346)	0,215 (0,364)	-0,071 (0,387)	0,272 (0,294)	0,414 (0,302)	0,215 (0,363)	-0,072 (0,453)	0,304 (0,280)	0,535* (0,282)
Laurea	3,178 (0,880)	0,991 (0,628)	0,831 (0,604)	0,464 (0,499)	-0,146 (0,615)	1,591 (0,656)	1,148 (0,860)	1,398 (0,568)	0,555 (0,607)
C.F.	-0,952* (0,564)	-0,558 (0,503)	-0,498 (0,609)	-0,844* (0,503)	-1,161 (0,538)	-0,278 (0,510)	-0,001 (0,595)	0,203 (0,384)	0,033 (0,395)
P.O.	-0,074 (0,710)	-0,788 (0,806)	0,366 (0,764)	1,109* (0,605)	0,973 (0,619)	-0,435 (0,656)	1,898 (0,913)	0,084 (0,601)	0,313 (0,569)
N.I.	-2,423 (0,508)	-1,818 (0,462)	-0,708 (0,444)	-1,321 (0,353)	-1,682 (0,431)	-2,227 (0,545)	-0,539 (0,649)	-1,870 (0,401)	-1,338 (0,402)
Rientri	-0,582 (0,391)	-1,192 (0,402)	-0,436 (0,432)	-0,626* (0,329)	-1,130 (0,344)	-1,315 (0,358)	0,148 (0,487)	-0,855 (0,330)	-0,804 (0,328)
Azioni	0,078 (0,113)	-0,084 (0,115)	-0,033 (0,117)	-0,268 (0,098)	-0,151 (0,098)	-0,004 (0,102)	-0,135 (0,153)	-0,124 (0,077)	-0,064 (0,074)
C.T.I.	-0,915 (0,361)	-0,597 (0,365)	0,337 (0,446)	-0,684 (0,322)	0,192 (0,332)	-0,531 (0,343)	-0,411 (0,454)	-0,834 (0,351)	-0,102 (0,300)
C.T.e	-1,573 (0,686)	-0,155 (0,726)	1,803 (0,735)	0,612 (0,913)	-0,084 (0,662)	0,620 (0,716)	-0,595 (0,988)	0,075 (0,692)	1,375* (0,715)
T.P.o	1,084 (0,344)	0,316 (0,363)	0,060 (0,391)	0,819 (0,334)	-0,018 (0,297)	0,239 (0,332)	0,490 (0,415)	0,581* (0,313)	-112 (0,293)
S.C.Ra.	-0,211 (0,308)	0,086 (0,327)	0,716 (0,325)	0,389 (0,251)	0,431 (0,272)	0,781 (0,310)	0,303 (0,433)	-0,050 (0,275)	-0,249 (0,267)
Sussidio	1,066 (0,501)	0,950 (0,580)	1,643 (0,842)	1,497 (0,697)	0,699 (0,580)	-0,037 (0,562)	4,301 (0,816)	1,077 (0,457)	0,906 (0,448)
Log(L)	-785,9	-924,1	-794,7	-860,6	-957,4	-966,6	-718,2	-878,6	-1004,9
R ²	0,234	0,112	0,118	0,173	0,150	0,137	0,189	0,195	0,112
n	260	290	248	291	304	285	222	278	304
$\lambda_{K-S; oss}$	1,079	0,961	1,516	1,422	0,947	0,585	1,406	1,292	1,153

Grassetto : p (probabilità di significatività) $\leq 0,05$. / Asterisco : $0,05 < p \leq 0,10$.

Tabella 8 — *Stime dei parametri delle covariate, della distribuzione esponenziale per l'effetto telescopico, e della distribuzione log-logistica con rischi non proporzionali per i periodi incompleti (in mesi) dal primo trimestre 1993 al primo trimestre 1995*

θ, x	93:1	93:2	93:3	93:4	94:1	94:2	94:3	94:4	95:1
$\hat{\omega}$	0,095 (0,072)	0,262 (0,182)	0,114 (0,078)	0,745 (0,627)	0,147 (0,111)	0,310 (0,248)	0,123 (0,082)	1,562 (1,388)	0,893 (0,705)
$\hat{\alpha}$	2,699 (0,237)	3,158 (0,259)	3,049 (0,300)	2,317 (0,185)	2,596 (0,218)	2,904 (0,232)	3,606 (0,367)	2,353 (0,183)	2,463 (0,208)
$\hat{\gamma}$	0,030 (0,005)	0,025 (0,004)	0,029 (0,005)	0,032 (0,005)	0,030 (0,004)	0,020 (0,003)	0,019 (0,005)	0,020 (0,004)	0,016 (0,003)
Femmine	0,086 (0,450)	-0,435 (0,416)	-0,490 (0,467)	-0,517 (0,376)	-0,423 (0,419)	-0,532 (0,450)	-1,312 (0,565)	-0,241 (0,371)	-0,499 (0,372)
E/10	0,098 (1,063)	-0,960 (1,174)	-0,939 (1,201)	-1,554 (1,012)	-0,262 (1,100)	-1,862 (1,397)	-0,673 (1,249)	-1,911* (1,149)	-2,537 (1,095)
(E/10) ²	0,062 (0,141)	0,121 (0,171)	0,127 (0,176)	0,223 (0,139)	0,091 (0,158)	0,253 (0,206)	0,001 (0,173)	0,121 (0,160)	0,338 (0,155)
Diploma	1,327 (0,435)	0,218 (0,442)	-0,154 (0,483)	0,416 (0,393)	0,504 (0,390)	0,208 (0,445)	-0,112 (0,540)	0,355 (0,377)	0,646* (0,371)
Laurea	4,173 (1,068)	1,286 (0,784)	1,095 (0,764)	0,742 (0,668)	-0,297 (0,804)	2,019 (0,827)	1,410 (1,015)	1,829 (0,745)	0,609 (0,771)
C.F.	-1,380* (0,726)	-0,614 (0,627)	-0,631 (0,765)	-1,243* (0,663)	-1,517 (0,678)	-0,180 (0,646)	-0,003 (0,707)	0,401 (0,523)	0,228 (0,517)
P.O.	-0,123 (0,908)	-0,973 (0,979)	0,551 (0,955)	1,551* (0,810)	1,146 (0,809)	-0,461 (0,819)	2,325 (1,071)	0,051 (0,813)	0,477 (0,746)
N.I.	-3,336 (0,622)	-2,299 (0,572)	-0,907 (0,558)	-1,870 (0,457)	-2,281 (0,531)	-2,845 (0,660)	-0,636 (0,767)	-2,681 (0,510)	-1,918 (0,507)
Rientri	-0,891* (0,498)	-1,517 (0,489)	-0,631 (0,544)	-0,900 (0,442)	-1,467 (0,437)	-1,712 (0,449)	0,221 (0,577)	-1,347 (0,436)	-1,133 (0,427)
Azioni	0,095 (0,142)	-0,098 (0,140)	-0,051 (0,147)	-0,425 (0,133)	-0,223 (0,124)	-0,009 (0,127)	-0,155 (0,178)	-0,162 (0,102)	-0,080 (0,097)
C.T.I.	-1,240 (0,465)	-0,806* (0,455)	0,470 (0,562)	-0,985 (0,428)	0,255 (0,438)	-0,714* (0,433)	-0,495 (0,543)	-1,155 (0,458)	-0,167 (0,394)
C.T.e	-2,014 (0,871)	-0,279 (0,884)	2,467 (0,920)	0,757 (1,159)	0,165 (0,883)	0,946 (0,922)	-0,744 (1,170)	0,050 (0,906)	1,645* (0,911)
T.P.o	1,471 (0,430)	0,441 (0,446)	0,031 (0,493)	1,022 (0,439)	-0,023 (0,390)	0,299 (0,417)	0,570 (0,493)	0,841 (0,415)	-177 (0,381)
S.C.Ra.	-0,296 (0,393)	0,083 (0,405)	0,961 (0,409)	0,577* (0,341)	0,564 (0,352)	0,994 (0,385)	0,360 (0,512)	0,057 (0,365)	-0,215 (0,348)
Sussidio	1,257 (0,629)	1,253* (0,723)	2,159 (1,022)	2,087 (0,928)	0,775 (0,732)	-0,030 (0,705)	5,617 (0,965)	1,427 (0,594)	1,114* (0,594)
Log(L)	-785,0	-922,0	-795,4	-862,4	-957,8	-966,4	-718,2	-880,2	-1006,3
R ²	0,234	0,112	0,118	0,173	0,150	0,137	0,189	0,195	0,112
n	260	290	248	291	304	285	222	278	304
$\lambda_{K-S; oss}$	1,068	0,888	1,668	1,581	0,974	0,631	1,454	1,330	1,292

Grassetto: p (probabilità di significatività) $\leq 0,05$. / Asterisco: $0,05 < p \leq 0,10$.

Tabella 9 — *Stime dei parametri delle covariate, della distribuzione esponenziale per l'effetto telescopico, e della distribuzioni $f(t, \mathbf{x}; \cdot)$, $\psi(t, \mathbf{x}; \cdot)$, $f_{LL}(t, \mathbf{x}; \cdot)$, per i periodi completi (in mesi) ricostruiti nell'intervallo che va dal primo trimestre 1993 al primo trimestre 1995*

θ, \mathbf{x}	$f(t, \mathbf{x}; \cdot)$			$\psi(t, \mathbf{x}; \cdot)$			$f_{LL}(t, \mathbf{x}; \cdot)$		
	M	F	T	M	F	T	M	F	T
$\widehat{\lambda}, \widehat{\vartheta}, \widehat{\omega}$	0,117 (0,076)	0,091 (0,047)	0,098 (0,036)	4,927 (3,181)	7,781* (4,051)	6,599 (2,531)	0,203 (0,131)	0,128* (0,066)	0,151 (0,058)
$\widehat{\alpha}$	1,239 (0,070)	1,217 (0,053)	1,205 (0,040)	2,008 (0,118)	1,911 (0,082)	1,927 (0,067)	2,008 (0,118)	1,910 (0,082)	1,927 (0,067)
$\widehat{\gamma}$	0,005* (0,003)	0,007 (0,002)	0,007 (0,002)	0,005* (0,003)	0,007 (0,002)	0,007 (0,001)	0,005 (0,003)	0,007 (0,002)	0,007 (0,001)
Femmine	—	—	-0,148* (0,086)	—	—	-0,175 (0,158)	—	—	-0,175 (0,158)
E/10	-0,352 (0,425)	-0,385 (0,366)	-0,264 (0,237)	-0,428 (0,734)	-0,352 (0,596)	-0,275 (0,424)	-0,428 (0,734)	-0,352 (0,596)	-0,275 (0,424)
(E/10) ²	0,035 (0,058)	0,057 (0,055)	0,031 (0,034)	0,035 (0,097)	0,048 (0,088)	0,029 (0,061)	0,035 (0,097)	0,048 (0,088)	0,029 (0,061)
Diploma	0,255 (0,185)	-0,016 (0,102)	0,043 (0,084)	0,303 (0,282)	-0,148 (0,187)	-0,001 (0,153)	0,303 (0,282)	-0,148 (0,187)	-0,001 (0,153)
Laurea	0,651* (0,372)	0,360 (0,236)	0,460 (0,178)	0,816 (0,536)	0,569 (0,368)	0,659 (0,299)	0,816 (0,536)	0,569 (0,368)	0,659 (0,299)
C.F.	-0,169 (0,246)	0,131 (0,217)	-0,154 (0,148)	-0,134 (0,407)	0,126 (0,362)	-0,157 (0,237)	-0,135 (0,407)	0,126 (0,362)	-0,157 (0,237)
P.O.	-0,167 (0,326)	0,629 (0,237)	0,308 (0,180)	-0,428 (0,521)	0,594 (0,407)	0,219 (0,312)	-0,428 (0,521)	0,594 (0,407)	0,219 (0,312)
N.I.	-0,443* (0,241)	-0,554 (0,144)	-0,505 (0,115)	-0,484 (0,372)	-0,662 (0,250)	-0,593 (0,203)	-0,484 (0,372)	-0,662 (0,250)	-0,592 (0,203)
Rientri	-0,181 (0,211)	-0,345 (0,124)	-0,270 (0,101)	-0,315 (0,315)	-0,517 (0,210)	-0,439 (0,170)	-0,315 (0,315)	-0,517 (0,210)	-0,439 (0,170)
Azioni	0,005 (0,056)	0,021 (0,037)	-0,002 (0,029)	-0,098 (0,092)	-0,002 (0,059)	-0,047 (0,049)	-0,098 (0,092)	-0,002 (0,059)	-0,047 (0,049)
C.T.I.	-0,149 (0,193)	-0,075 (0,121)	-0,103 (0,097)	-0,300 (0,270)	-0,096 (0,201)	-0,192 (0,160)	-0,300 (0,270)	-0,096 (0,201)	-0,192 (0,160)
C.T.e	-0,253 (0,390)	-0,128 (0,194)	-0,098 (0,162)	-0,680 (0,740)	-0,127 (0,330)	-0,216 (0,297)	-0,680 (0,740)	-0,127 (0,330)	-0,216 (0,297)
T.P.o	0,319 (0,209)	0,197* (0,120)	0,275 (0,093)	0,401 (0,292)	0,408 (0,188)	0,441 (0,154)	0,401 (0,292)	0,408 (0,188)	0,441 (0,154)
S.C.Ra.	0,344 (0,168)	229 (0,102)	0,264 (0,078)	0,391 (0,247)	0,266 (0,179)	0,290 (0,141)	0,391 (0,247)	0,266 (0,179)	0,290 (0,141)
Sussidio	0,368 (0,343)	0,776 (0,269)	0,655 (0,204)	0,352 (0,457)	0,764 (0,352)	0,664 (0,276)	0,352 (0,457)	0,764 (0,352)	0,664 (0,276)
Log(L)	-777,2	-1500,1	-2285,4	-770,2	-1496,8	-2272,8	-770,2	-1496,8	-2272,8
R ²	0,125	0,141	0,124	0,067	0,091	0,075	0,067	0,091	0,075
$\lambda_{K-S; oss}$	1,228	1,641	1,925	0,532	1,085	1,157	0,532	1,085	1,157

Grassetto: p (probabilità di significatività) $\leq 0,05$. / Asterisco: $0,05 < p \leq 0,10$.

4. Conclusioni

I metodi utilizzati hanno permesso di accertare alcune tendenze: le donne aspettano piú degli uomini e presentano una struttura di dipendenza dalle covariate piú complessa, assenza di o lieve dipendenza dallo stato, effetto telescopico marcato e significativo, miglioramento dell'aderenza dei dati osservati a quelli stimati con l'inclusione dell'eterogeneità non osservata, un comportamento variabile nel tempo del parametro di forma (α) della distribuzione dei periodi di disoccupazione. Le variazioni dei valori di $\hat{\alpha}$ e β_t possono essere attribuite, non inequivocabilmente, a possibili effetti stagionali sui modelli a causa di ingressi e uscite dovuti a posti di lavoro stagionali che potrebbero indurre spostamenti repentini e locali nelle medie delle durate di attesa. Tuttavia, occorrono ulteriori approfondimenti per verificare quanto ciò dipenda da variazioni stagionali di flussi di ingresso e uscita dal mercato del lavoro e quanto dipenda, invece, da altri fattori: come la distorsione campionaria delle lunghezze.

Ulteriori sviluppi dell'analisi sono stati trascurati per brevità. Per esempio, si sarebbe anche potuto indagare su metodi piú efficaci per verificare gli assunti sulle distribuzioni adottate sia con procedure statistiche specifiche (Chesher, Lancaster, Irish, 1985), sia con strategie di simulazione. L'analisi dei periodi completi poteva essere migliorata per i vari sottogruppi di persone in cerca di lavoro e i parametri si potevano stimare separatamente o con un modello a rischi concorrenti; in particolare, si poteva ricorrere anche a una formulazione piú ampia e aderente alle teorie economiche. L'unione di tutti i periodi di disoccupazione, completi e incompleti, in un'unico insieme non è stato effettuato perché il numero dei periodi completi diventerebbe relativamente piccolo rispetto a quello degli incompleti e potrebbe distorcere le stime. In alternativa, si potevano includere tra i periodi completi quelli incompleti delle ultime due rilevazioni, che non sono state utilizzati nell'abbinamento per la loro ricostruzione, al fine di accertare quale modifica inducono nei parametri del modello. Si può comunque presumere che non debbano emergere differenze notevoli perché la coerenza tra le diverse stime ottenute è molto elevata: anche se ciò non esclude che qualche nuova covariata diventi significativa e qualche altra non lo sia piú. Per limitare l'ampiezza dell'analisi si è trascurato quanto sopra.

Un altro aspetto da esaminare in profondità concerne la valutazione degli effetti del metodo di stima degli errori standard perché il fenomeno che si è osservato richiede una verifica di "stabilità" tramite un processo di simulazione.

L'andamento di α per i periodi completi potrebbe essere indagato con altri metodi piú flessibili e orientati ai dati per accertare quale sia la forma piú idonea per i periodi disponibili. In questo caso, la metodologia delle reti neurali potrebbe offrire spunti interessanti: tra gli altri si veda Liestøl, Andersen, Andersen (1994); tuttavia, non è possibile analizzare i periodi trascorsi con questi metodi, senza incorrere negli stessi limiti.

Il metodo utilizzato presenta, comunque, diverse difficoltà: i risultati non sono immediati e intuitivi; il formalismo è complesso; la stima dei parametri richiede

calcoli laboriosi, e l'uso "obbligato" del metodo di massima verosimiglianza. Infine, la Weibull potrebbe essere sostituita con altre distribuzioni anche se essa risulta assai flessibile, e consente di ottenere espressioni senza l'operatore di integrazione. Nonostante le difficoltà e le limitazioni menzionate, i risultati delle elaborazioni sembrano soddisfacenti perché, tra l'altro, indicano che quando dallo studio dei completi è possibile determinare una distribuzione delle durate, allora si può eseguire l'analisi dei periodi incompleti con un discreto successo.

Riferimenti bibliografici

- Allison P.D. (1985), Survival Analysis of Backward Recurrence Times, *Journal of the American Statistical Association*, 80, pp. 315–322.
- Attanasio O., Paoda-Schioppa F. (1991), Regional Inequalities, Migration and Mismatch in Italy, 1960–1986, in Padoa-Schioppa F. (ed.), *Mismatch and labour mobility*, Cambridge University Press, Cambridge, pp. 237–324.
- Biagi M., Tiraboschi M. (1998), Nuovi lavori e nuove forme di utilizzazione della forza-lavoro nella provincia di Modena, in AA. VV., *Rapporto sulla situazione economica e sociale della Provincia di Modena, 1998*, Associazione Mario Del Monte–Università degli Studi di Modena e Reggio Emilia, pp. 11–33.
- Burr I.W. (1942), Cumulative Frequency Functions, *The Annals of Mathematical Statistics*, 13, pp. 215–223.
- Campiglio L. (1996), *Il costo del vivere: Nord e Sud a confronto*, il Mulino, Bologna.
- Capecchi V. (a cura di) (1995), *Osservatorio domanda-offerta. Qualità delle assunzioni e domanda di profili professionali da parte delle imprese in Emilia-Romagna. Rapporto 1994*, Quaderni di ricerca N. 18, Osservatorio del mercato del lavoro, Regione Emilia-Romagna, Bologna.
- Chesher A., Lancaster T., Irish M. (1985), On Detecting the Failure of Distributional Assumptions, *Annales des l'INSEE*, 59/60: pp. 7-45.
- Cox D.R. (1972), Regression Models and Life Tables, *Journal of the Royal Statistical Society*, B, 34, pp. 187-220.
- Cox D.R., Oakes D. (1984), *Analysis of Survival Data*, Chapman and Hall, London.
- Di Nicola P. (1995), Quale flessibilità? Lavoro atipico e part-time in Italia e in Europa, *Economia & Lavoro*, XXIX (3), pp. 21–38.
- Faini R., Galli G., Rossi F. (1996), *Mobilità disoccupazione in Italia: un'analisi dell'offerta di lavoro*, CSC Working Paper, Roma.
- Falorsi S., Falorsi P.D. (1994), Indagine sulle forze di lavoro: descrizione della procedura di stima attualmente utilizzata ed analisi di metodi alternativi, *Quaderni di ricerca*, n. 4, ISTAT, Roma.

- Favro-Paris M.M., Gennari P., Oneto G.P. (1996), La durata della disoccupazione in Italia: un'applicazione della struttura longitudinale dell'indagine sulle forze di lavoro, ISTAT, *Quaderni di Ricerca*, nuova serie, 4, pp. 1-79.
- Flinn J.C. (1986). Econometric Analysis of CPS-Type Unemployment Data, *The Journal of Human Resources*, XXI, pp. 456-484.
- Grace A. (1994), *Optimization Toolbox User's Guide*, The Math Works, Natick, MA.
- Greene W.H. (1993), *Econometric Analysis*, 3.rd ed., Macmillan, New York.
- Groot W. (1990), The Effects of Benefits and Duration Dependence on Re-Employment Probabilities, *Economics Letters*, 32, 4, pp. 371-376.
- Ham J.C., Rea S.A. Jr. (1987), Unemployment Insurance and Male Unemployment Duration in Canada, *Journal of Labour Economics*, 5, 3, pp. 325-353.
- Hannan M., Carroll G. (1992), *Dynamics of Organizational Populations: Density, Legitimation, and Competition*, Oxford University Press, New York.
- Heckman J.J., Singer B. (1984a), Econometric Duration Analysis, *Journal of Econometrics*, 24, pp. 63-132.
- Heckman J.J., Singer B. (1984b), A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data, *Econometrica*, 52, pp. 271-320.
- Hujer R., Schneider H. (1989), The Analysis of Labor Market Mobility Using Panel Data, *European Economic Review*, 33, 2/3, pp. 530-536.
- ISTAT (1994), *Annuario statistico italiano*, ISTAT, Roma.
- ISTAT (1995a), *Annuario statistico italiano*, ISTAT, Roma.
- ISTAT (1995b), *Rilevazione delle forze di lavoro. Media 1993*, Collana d'informazione n. 17, ISTAT, Roma.
- ISTAT (1995c), *Rilevazione delle forze di lavoro. Media 1994*, Collana d'informazione n. 18, ISTAT, Roma.
- Johnson N. L., Kotz S., Balakrishnan N. (1995), *Continuous Univariate Distributions*, 2.nd edition, vol. 2, John Wiley & Sons, New York.
- Lalla M., Pattarin F. (1998), Analisi della durata della disoccupazione su dati longitudinali e trasversali: il caso Emilia-Romagna, Atti della XXXIX Riunione Scientifica della SIS, Sorrento, 14-17 aprile.
- Lancaster T. (1990), *The Econometric Analysis of Transition Data*, Cambridge University Press, New York.
- Listøl K., Andersen P.K., Andersen U. (1994), Survival Analysis and Neural Nets, *Statistics in Medicine*, 13, pp. 1189-1200.
- Lynch L. (1989), The Youth Labour Market in the Eighties: Determinants of Re-Employment Probabilities for Young Men and Women, *The Review of Economics and Statistics*, 71, 1, pp. 37-45.
- Maddala G.S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.

- McCall B.P. (1996), Unemployment Insurance Rules, Joblessness, and Part-Time Work, *Econometrica*, 64 (3), pp. 647–682.
- Narendranathan W., Nickell S., Stern J. (1985), Unemployment Benefits Revisited, *The Economic Journal*, 95, 307–329.
- Narendranathan W., Stewart M.B. (1993), Modelling the Probability of Leaving Unemployment: Competing Risks Models with Flexible Base-line Hazards, *Journal of the Royal Statistical Society, Series C*, 42, 63–83.
- Nickell S. (1979), The Effect of Unemployment and Related Benefits on the Duration of Unemployment, *Economic Journal*, 89, 34–49.
- Osservatorio del mercato del lavoro (1995), *Rapporto annuale sul mercato del lavoro*, Regione Emilia-Romagna, Bologna.
- Osservatorio del mercato del lavoro (1997), *Il mercato del lavoro in Emilia-Romagna. Andamento e prospettive*, Regione Emilia-Romagna e Ministero del Lavoro e della Previdenza Sociale, Bologna.
- Salant S.W. (1977), Search Theory and Duration Data: a Theory of Sorts, *Quarterly Journal of Economic*, 91, pp. 39–57.
- SAS Institute (1990), *SAS/STAT User's Guide*, Version 6, Fourth Edition, SAS Institute, Cary, NC.
- Sestito P. (1995), Mobilità territoriale e mercato del lavoro, *Economia & Lavoro*, XXIX (4), pp. 3–19.
- Sikkel D. (1985), Models for Memory Effects, *Journal of the American Statistical Association*, 80, pp. 835–841.
- Sudman S., Bradburn N.M. (1973), Effects of Time and Memory Factors on Response in Surveys, *Journal of the American Statistical Association*, 68, pp. 805–815.
- Torelli N., Trivellato U. (1988), Modelling Job-Search Duration from the Italian Labour Force Data, *LABOUR*, 2, 1, pp. 117–134.
- Torelli N., Trivellato U. (1993), Modelling Inaccuracies in Job-Search Duration Data, *Journal of Econometrics*, 59, pp. 187–211.

Summary

Some Models to Analyse Complete and Incomplete Spells of Unemployment: The Case of Emilia-Romagna

The analysis is based on data collected by ISTAT in Emilia Romagna from the first quarter of 1993 to the first quarter of 1995. Exploiting the 2-2-2 rotating scheme of the survey, a set of completed spells was obtained from the nine originally available samples of incomplete ones. Both data sets were modelled through a hazard function based on a Weibull distribution of duration. Therefore, referring to the incomplete spells, it was assumed that the unobserved completed spells were Weibull-distributed. The empirical distributions of spell durations showed pronounced spikes which suggests the presence of a considerable telescopic effect. This effect was modelled by an exponential heaping function. Heterogeneity was controlled for by adding a suitable set of covariates. The hazard function was also based on a log-logistic distribution of duration in a non-proportional formulation to compare the different behaviour of the models, the baseline hazard, and the parameter estimates of the covariates. The full models were estimated separately on the completed spell data and for each quarter, on the incomplete spell data, to make the comparisons stated above and to envisage the potential seasonal effect. The results showed that the shape parameters changed over time. The parameters of each covariate proved to be statistically unchanged over time (*i.e.*, quarterly models) and were also equal to the parameters of the completed spell model. Therefore, the analysis of incomplete spells is fairly feasible when an even (closed form) baseline hazard function is found in a completed spell analysis.

Keywords

Unemployment spell, Telescoping effect, Heaping effect, Heterogeneity.

Working Papers già pubblicati

1. E. Battistin, A. Gavosto e E. Rettore, *Why do subsidized firms survive longer? An evaluation of a program promoting youth entrepreneurship in Italy*, Agosto 1998.
2. N. Rosati, E. Rettore e G. Masarotto, *A lower bound on asymptotic variance of repeated cross-sections estimators in fixed-effects models*, Agosto 1998.
3. U. Trivellato, *Il monitoraggio della povertà e della sua dinamica: questioni di misura e evidenze empiriche*, Settembre 1998.
4. F. Bassi, *Un modello per la stima di flussi nel mercato del lavoro affetti da errori di classificazione in rilevazioni retrospettive*, Ottobre 1998.
5. Ginzburg, M. Scaltriti, G. Solinas e R. Zoboli, *Un nuovo autunno caldo nel Mezzogiorno? Note in margine al dibattito sui differenziali salariali territoriali*, Ottobre 1998.
6. M. Forni e S. Paba, *Industrial districts, social environment and local growth. Evidence from Italy*, Novembre 1998.
7. B. Contini, *Wage structures in Europe and in the USA: are they rigid, are they flexible?*, Gennaio 1999.
8. B. Contini, L. Pacelli e C. Villosio, *Short employment spell in Italy, Germany and Great Britain: testing the "Port-of-entry" hypothesis*, Gennaio 1999
9. B. Contini, M. Filippi, L. Pacelli e C. Villosio, *Working careers of skilled vs. unskilled workers*, Gennaio 1999
10. F. Bassi, M. Gambuzza e M. Rasera, *Il sistema informatizzato NETLABOR. Caratteristiche di una nuova fonte sul mercato del lavoro*, Maggio 1999.
11. M. Lalla e F. Pattarin, *Alcuni modelli per l'analisi delle durate complete e incomplete della disoccupazione: il caso Emilia Romagna*, Maggio 1999.