

**USO E INTEGRAZIONE DI DATI AMMINISTRATIVI E DA *SURVEYS*
NEL PROGETTO MIUR
“DINAMICHE E PERSISTENZE NEL MERCATO DEL LAVORO ITALIANO”**

Ugo Trivellato

Università di Padova

Note introduttive

1. Questo intervento prende le mosse da esperienze di ricerca condotte prevalentemente nell’ambito di progetti cofinanziati dal MIUR dal 2000 al 2005, incentrati sulla mobilità dei lavoratori e sulla valutazione di politiche del lavoro¹.

Mi propongo innanzitutto di dare conto di aspetti attinenti all’utilizzo e all’integrazione, a questi fini, di dati amministrativi e da *surveys*.

Sulla scorta di queste esperienze, presento poi alcune riflessioni di indole generale sulla realizzazione di un sistema informativo del lavoro rispondente, tra l’altro, alle esigenze conoscitive e di supporto a politiche a diverse scale territoriali.

2. In via preliminare, segnalo che l’intervento non avrà come fuoco analisi sui – ed esigenze informative dei – mercati del lavoro locali. E ciò per due ragioni:

- (a) la prima è fattuale, ed è presto detta: le esperienze di ricerca svolte non avevano come ambito di interesse prevalente i mercati del lavoro locali;
- (b) la seconda sta nella convinzione che gli interrogativi attinenti ai mercati del lavoro locali – vuoi in termini di conoscenze di base, vuoi in chiave di informazioni e analisi di supporto a politiche – non siano scindibili da analoghi interrogativi a diverse scale territoriali, fino a quella nazionale – a sua volta attenta alle interrelazioni con resto del mondo –, per le forti interdipendenze del processo economico e per la crescente mobilità territoriale dei lavoratori.

Detto altrimenti, e anticipando alcune considerazioni conclusive, sono dell’opinione che le esigenze informative dei mercati del lavoro locali siano importanti tessere di un mosaico, che spesso necessitano di specifici approfondimenti, ma che possono trovare risposte soddisfacenti soltanto nel contesto dell’intero mosaico. Fuori di metafora, in un sistema informativo nazionale del lavoro, che per la sua natura di ‘sistema’ non può che essere integrato e coerente.

3. Come anticipato, le ricerche svolte si sono incentrate sulla mobilità dei lavoratori e sulla valutazione dell’impatto di selezionate politiche del lavoro. Sul terreno dei fabbisogni informativi c’è un tratto che le accomuna, e che è cruciale mettere in luce: questi temi richiedono di osservare (segmenti di) storie lavorative, quindi di disporre di basi di dati longitudinali, possibilmente *linked employer-employee*.

Tali ricerche possono dunque essere considerate esemplari di problemi e difficoltà che si incontrano in materia di informazioni statistiche disponibili, così come di strategie per cercare di utilizzarle al meglio. Per un altro verso, va tuttavia detto chiaramente che esse sono esemplificative, perché non toccano – o toccano in misura modesta – altri aspetti

¹ Vi hanno concorso, inoltre, convenzioni di ricerca con l’Inps (da parte del Dipartimento di Economia Cognetti De Martiis dell’Università di Torino), con l’Istat (da parte del Dipartimento di Scienze Statistiche dell’Università di Padova), con Veneto Lavoro (ancora da parte del Dipartimento di Scienze Statistiche patavino) e con il Ministero del Lavoro e delle Politiche Sociali (da parte di Veneto Lavoro, con un gruppo di lavoro che includeva anche ricercatori del Dipartimento di Scienze Statistiche dell’Università di Padova).

dell'eterogeneità che contraddistingue il mercato del lavoro (in tema di forme di partecipazione, orario, retribuzioni, carriere, ecc.)².

Evidenze salienti dalle esperienze di ricerca

4. Buona parte dei risultati delle ricerche svolte sono raccolti in Contini e Trivellato (2005a).

Nella prospettiva dell'utilizzo/integrazione di dati da fonti diverse a fini di studio della mobilità dei lavoratori, essi ruotano attorno a quattro nodi: (a) l'utilizzo di basi di dati di fonte Inps; (b) l'utilizzo di basi di dati prodotti a partire dalle rilevazioni dei Centri per l'impiego; (c) l'utilizzo di dati da *surveys*, prima fra tutte la rilevazione sulle forze di lavoro (RFL); (d) l'integrazione di basi di dati – quelle menzionate e altre *ad hoc* –.

5. Abbiamo esplorato le potenzialità nell'uso di dati amministrativi di fonte Inps – e ci si siamo confrontati con i loro limiti –, impiegando (i) il *database* campionario WHIP (*Work Histories Italian Panel*) sviluppato dal LABORatorio R. Revelli e (ii) il *database* VWH (*Veneto Work Histories*), realizzato in via sperimentale dal Dipartimento di Scienze economiche dell'Università Ca' Foscari di Venezia, riferito a una scala territoriale ridotta – due province del Veneto, per il periodo 1975-1997 – ma con dati censuari.

Le potenzialità sono notevoli, e sono state documentate da tempo dalle ricerche condotte soprattutto da Bruno Contini e collaboratori (a partire da Contini e Revelli, 1987 e 1992). In Contini e Trivellato (2005a) esse sono testimoniate da saggi: (i) sulla mobilità dei lavoratori: nel complesso, disaggregata per vari caratteri, riferita a sottopopolazioni di particolare interesse – giovani, extracomunitari, lavoratori in transizione verso la pensione –; (ii) sugli effetti di alcuni istituti contrattuali (lavoratori co.co.co., lavoratori interinali e art. 18 dello Statuto dei lavoratori); (iii) sulle retribuzioni e sui collegamenti fra cambi di lavoro e dinamica salariale.

In chiave esemplificativa, presento sinteticamente due ordini di risultati: le stime del diverso grado di mobilità per dimensione d'impresa e per ripartizione geografica (Leombruni e Quaranta, 2005); la tendenza al dualismo nella mobilità dei lavoratori, documentata da Tattara e Valentini (2005) in un'analisi che contrappone gli *stayers* ai *permanent movers*: detto altrimenti i lavoratori con carriere lunghe prevalentemente nella stessa impresa e i lavoratori con numerosi rapporti di lavoro brevi.

La rilevanza che hanno età dei lavoratori e dimensione di impresa sulla mobilità suggerisce di considerare innanzitutto l'effetto congiunto di questi due fattori. La Figura 1 presenta i valori medi del *gross worker turnover* per il periodo 1987-1999, disaggregati per classi quinquennali di età e cinque classi dimensionali di impresa.

Come atteso, le curve che descrivono il *gross worker turnover* in funzione dell'età mostrano il tipico *pattern* a U. Prime evidenze degne di nota stanno nel fatto che esse si collocano via via più in alto al diminuire della dimensione di impresa, e che per le piccole imprese hanno un andamento più piatto, segnalando minori differenze nel *gross worker turnover* tra giovani e meno giovani. Il punto che merita sottolineare è comunque l'entità delle differenze fra grandi e piccole imprese. I lavoratori nel pieno della loro maturità lavorativa, o prossimi alla pensione, che svolgono la loro carriera in piccole imprese sono caratterizzati da una mobilità addirittura superiore a quella che si osserva per i 25-30enni impiegati in imprese sopra i 200 addetti. In sintesi, al crescere della dimensione di impresa diminuisce la mobilità da/verso l'esterno, mentre il *gross worker turnover* nella mano d'opera impiegata avviene sostanzialmente tramite le entrate dei giovani e le uscite per pensionamento. Viceversa, nelle piccole imprese i tassi di entrata e di uscita sono

² Ad es., sull'orario di lavoro vedi Istat (2005) e soprattutto Istat (2006, cap. 4).

decisamente elevati per tutte le età, e la maggiore esperienza maturata dai lavoratori meno giovani non si traduce in rapporti di lavoro sensibilmente più stabili. Questo fatto si può spiegare, in parte, con scelte del lavoratore: le possibilità di carriera all'interno di una piccola impresa sono minori, e più frequentemente il lavoratore cercherà progressioni di carriera cambiando impresa. Ma esso chiama in causa anche eventi 'traumatici' quali gli elevati tassi di nati-mortalità delle imprese sotto i 20 addetti, che generano mobilità indipendentemente dall'età del lavoratore.

Figure 1 e 2 circa qui

Con la Figura 2 l'attenzione è portata su differenze territoriali nella mobilità, misurata ora dal tasso di riallocazione – cioè a dire dalla frazione di lavoratori mobili –. Essa mostra che i livelli più alti di mobilità sono nel Mezzogiorno. Ad un esame più articolato, emerge poi che nel Mezzogiorno la frazione di movimenti *job-to-job* è relativamente bassa, dell'ordine del 19%; all'opposto, i tempi di rientro nell'occupazione dipendente privata extra-agricola sono dilatati, interrotti da lunghi episodi di non occupazione – la frazione di rientri entro l'anno supera appena il 40% –. Ciò suggerisce che all'elevata mobilità si accompagna una marcata segmentazione: e tutto ciò avendo come sfondo un basso tasso di attività. Emerge dunque una connotazione negativa della *performance* del mercato del lavoro meridionale, alla quale concorre appunto anche una sorta di 'cattiva' mobilità, ben diversa da quella relativamente 'buona' che si osserva nel Nord-Est (dove una percentuale notevole di lavoratori – il 39% – transita *job-to-job* o comunque ha prevalentemente successioni di lavori interrotti da periodi piuttosto brevi di non occupazione, sicché la frazione di rientri entro l'anno sfiora il 61%).

In merito all'eterogeneità dei lavoratori quanto a *pattern* di storie lavorative, è interessante guardare alla Figura 3. Essa si riferisce a un ambito settoriale e territoriale abbastanza ristretto – l'industria manifatturiera di due province venete – e riguarda la coorte di uomini nati nel 1950-60 e con episodi lavorativi nel periodo 1982-90. La figura contrappone la fase iniziale della carriera nel lavoro dipendente privato di un tipico lavoratore *stayer* a quella di un tipico *permanent mover*. Per lo *stayer* i primi quattro episodi di occupazione coprono 176 mesi su 202; il tempo di attesa fra un'occupazione e la successiva rispetto al tempo lavorato è dunque dell'ordine di 1 a 7, e inoltre si riduce progressivamente. Al polo opposto, per il *permanent mover* i primi cinque episodi di occupazione coprono 84 mesi su 140, con un rapporto fra tempo di attesa e tempo lavorato che è di 2 a 3, pressoché stabile mano a mano che si passa a successivi lavori. I *permanent movers* sono circa il 10% dei lavoratori, in prevalenza donne, con basso livello di istruzione, addetti mansioni mediamente poco remunerate. La polarizzazione è chiara e stabile per tutto il periodo – si fa anzi più marcata negli anni Novanta – e segnala una segmentazione strutturale del mercato del lavoro manifatturiero veneto³.

Figura 3 circa qui

Difficoltà e limiti nell'utilizzo della fonte Inps rimangono peraltro severi. In sintesi, essi vengono da:

³ Altre evidenze, riassunte in Contini e Trivellato (2005b, pp. 48-49), corroborano l'ipotesi che le recenti trasformazioni nella composizione e nelle caratteristiche dell'occupazione – e le 'riforme al margine' nella regolazione del mercato del lavoro – abbiano operato nel senso di accentuare i tratti di dualismo dell'occupazione, con gli incrementi di mobilità concentrati su una frazione contenuta di lavoratori, in prevalenza giovani, donne, immigrati.

- la copertura parziale dell’universo degli occupati, anche se piuttosto larga (e in progressiva espansione, segnatamente con la realizzazione del “Casellario delle posizioni previdenziali attive”);
- problemi di qualità, o comunque di pertinenza, delle informazioni rispetto alle esigenze conoscitive (il caso di scuola è quello dei lavoratori co.co.co.: vedi Berton, Pacelli e Segre, 2005);
- la mediocre tempestività nell’aggiornamento della fonte;
- i vincoli a una marcata disaggregazione territoriale (e/o secondo altri caratteri) e ad un’analisi delle politiche – retributive, di mobilità, ecc. – delle imprese⁴ conseguenti alla natura campionaria di WHIP.

6. In materia di *database* statistici costruiti a partire dalle informazioni raccolte dai Centri per l’impiego, ci si confronta con esperienze variegata nelle diverse regioni. Di una tra le più mature, il *database* GIOVE riguardante il Veneto, prodotto da Veneto Lavoro, dà conto l’intervento di Anastasia, Disarò, Gambuzza, Maurizio e Rasera (2006), al quale rinvio.

In Contini e Trivellato (2005a) le potenzialità di GIOVE sono documentate da analisi sulla partecipazione al lavoro e la mobilità degli extracomunitari e dei giovani. Con riguardo all’universo di riferimento, i dati sono censuari e sono resi disponibili con notevole tempestività (anche se a prezzo di qualche incompletezza nella copertura spaziale). Consentono quindi analisi convenientemente disaggregate e aggiornate.

Elementi critici nell’utilizzo dei dati dei Centri per l’impiego sono:

- per quanto attiene agli occupati, la copertura parziale, seppure larga – relativa essenzialmente al lavoro dipendente privato – e la rilevazione incentrata sui movimenti (assunzioni, trasformazioni, cessazioni di rapporti di lavoro), mentre manca quella degli *stock*;
- per quanto riguarda la disoccupazione, la perdurante, mediocre affidabilità dei dati sui disoccupati registrati. Anastasia e Disarò (2005) e Perrone e Sestito (2006) documentano vistose divergenze fra disoccupati registrati nei Servizi Pubblici dell’Impiego e disoccupati stimati dalla RFL, che giustificano il permanere dello scetticismo sulla possibilità di usare i disoccupati registrati per confronti di *stock* nello spazio e nel tempo⁵;
- la notevole variabilità nel processo di *data entry* da parte dei Centri per l’impiego – con ovvi riflessi negativi sulla qualità, completezza e tempestività del *database* prodotto a livello regionale –, e ancor più la marcata diversità delle esperienze delle varie regioni.

7. L’impiego di dati da *surveys* a fini di analisi della mobilità dei lavoratori riguarda innanzitutto, e soprattutto, la rilevazione sulle forze di lavoro nella versione trimestrale

⁴ Sulla necessità di disporre di dati su un considerevole numero – se non sull’intero insieme – di lavoratori di una data impresa per un ampio numero di imprese, per studi di tali politiche e per analisi della produttività delle imprese, vedi, ad es., Lazear e Shaw (2006) e Bryson e Forth (2006).

⁵ Le discrepanze fra disoccupati registrati nei servizi pubblici per l’impiego e disoccupati stimati dalle rilevazioni sulle forze di lavoro non sono peculiari dell’Italia, ma in altri paesi dell’Unione Europea hanno una dimensione ben più contenuta e trovano spiegazioni ragionevoli (Melis e Lüdeke, 2006). Va aggiunto che stime della disoccupazione coerenti con i criteri dell’*International Labour Office*, prodotte dalla rilevazione sulle forze di lavoro, non sono indiscutibili, proprio perché poggiano su criteri – segnatamente quello dell’azione attiva di ricerca effettuata nell’ultimo mese – in sostanza convenzionali. Recenti ricerche, condotte con approcci di analisi diversi (Brandolini, Cipollone e Viviano, 2006; Battistin, Rettore e Trivellato, 2006), sono concordi nel concludere in favore di una misura della disoccupazione ‘larga’, estesa alle persone non occupate che dichiarano di essere alla ricerca di lavoro, immediatamente disponibili a lavorare e di aver compiuto un’azione di ricerca, senza alcun vincolo di prossimità temporale della stessa.

(RTFL), e in prospettiva nella nuova versione di indagine continua (Istat, 2004). In proposito, è stato sviluppato il metodo dell'“abbinamento a tre occasioni” (Baretta e Trivellato, 2004; Trivellato, Paggiaro, Leombruni e Rosati, 2005), che sfrutta appieno la dimensione longitudinale del disegno campionario della RTFL per abbinare i record pertinenti allo stesso individuo in tre occasioni d'indagine entro un anno di calendario e per combinare l'informazione sulle storie lavorative risultante dagli stessi.

In tal modo si dispone di un piano di osservazione che consente di identificare fino a due associazioni e due separazioni, e fino a tre episodi di occupazione, per ogni individuo nell'anno (vedi la Figura 4): esso documenta dunque i segmenti di storie lavorative in maniera abbastanza dettagliata, seppur incompleta. Considerato poi che per un dato anno si generano due panel a tre occasioni (rispettivamente, gennaio.*t*-aprile.*t*-gennaio.*t*+1 e gennaio.*t*-ottobre.*t*-gennaio.*t*+1) i quali sono campioni disgiunti e indipendenti dalla stessa popolazione, li si aggrega e si ottiene un campione *pooled* di numerosità grosso modo doppia, che fornisce stime della mobilità dei lavoratori più precise.

Figura 4 circa qui

Il confronto dell'informazione sulla mobilità dei lavoratori che in tal modo si trae dalla RTFL con quella documentata da WHIP, presa come *benchmark*, per gli anni 1995 e 1998, suggerisce che i dati di panel a tre occasioni della RTFL forniscono un'informazione ragionevolmente affidabile, in particolare per quanto riguarda la stima della frazione di lavoratori mobili (vedi dopo la Figura 5).

Il vantaggio che ne discende è duplice: il metodo dell'abbinamento a tre occasioni sulla RTFL consente loro di condurre analisi (*i*) estese all'insieme degli occupati delle famiglie residenti e, soprattutto, (*ii*) aggiornate (in Contini e Trivellato, 2005a, le analisi su WHIP si arrestano al 1999, mentre quelle sulla RTFL giungono al 2003).

La contenuta dimensione campionaria del panel a tre occasioni comporta, d'altra parte, vincoli severi nella disaggregazione territoriale delle analisi di mobilità, che non può spingersi oltre le grandi ripartizioni.

8. L'esperienza di utilizzo integrato di dati da varie fonti – di origine amministrativa e da *surveys* – è stata condotta a due livelli: (*i*) confrontando stime di aggregati e/o indicatori da fonti diverse, soprattutto a fini di raccordo di serie (per ovviare alla scarsa tempestività di una delle due fonti); (*ii*) procedendo all'abbinamento dei microdati e conducendo poi l'analisi su un *database* integrato.

Un interessante esempio nella prima direzione è rappresentato proprio dal confronto e dalla combinazione delle stime di mobilità dei lavoratori dipendenti del settore privato extra-agricolo tratte da WHIP e dalla RTFL (vedi la Figura 5).

Figure 5 circa qui

La Figura 5 accosta appunto – sulla scorta di un positivo vaglio di comparabilità – le stime del tasso di riallocazione fornite, per gli occupati dipendenti dell'industria e dei servizi privati, da WHIP e dalla RTFL. Le due serie di indicatori di riallocazione sono presentate anche a popolazione costante, mantenendo invariata la struttura per genere e classe di età della popolazione del 1998. Il principale tratto che emerge è una certa stabilità del *pattern* della mobilità dei lavoratori, in parte dovuta al processo di invecchiamento della forza lavoro (per il procedere verso le età centrali delle generazioni dei *baby-boomers* e per la riduzione dei giovani in ingresso nel mercato del lavoro), che ‘oscura’ il trend di moderata crescita della mobilità desumibile dalle serie a popolazione costante.

Un esempio di integrazione in chiave di *linkage* di microdati è in recenti studi sulla valutazione del programma “Liste di mobilità”. La possibilità di valutare l’impatto del programma è, ovviamente, condizionata dalla base informativa disponibile. In generale, precedenti studi (vedi, ad es., Brunello e Miniaci, 1997, e Paggiaro e Trivellato, 2002) si sono basati sulle informazioni dagli archivi regionali risultanti dalla gestione delle Liste. Tali informazioni, peraltro, sono insufficienti per uno studio di valutazione d’impatto per tre principali ragioni: (i) mancano informazioni sulla storia lavorativa degli iscritti durante la loro permanenza in Lista, storia che può includere episodi di occupazione a tempo determinato, la cui individuazione è cruciale per la valutazione dell’impatto⁶; (ii) le informazioni disponibili cessano con l’uscita del lavoratore dalla Lista e quindi non permettono l’analisi degli effetti a medio termine; (iii) sono carenti le informazioni relative alla storia lavorativa dell’individuo prima del suo ingresso in Lista, il che rende problematico un adeguato controllo dell’eterogeneità fra gli iscritti.

Tali problemi possono essere (in larga parte) risolti mediante lo sviluppo di una base dati integrata risultante dal *linkage* tra l’archivio regionale delle Liste e un archivio del tipo GIOVE risultante dai dati dei Centri per l’impiego. Esso consente, infatti, di osservare tutti gli episodi di occupazione alle dipendenze (entro l’universo di riferimento⁷), sia a tempo determinato che a tempo indeterminato, sperimentati dai lavoratori inclusi nel campione utilizzato per la valutazione d’impatto in un conveniente arco temporale prima e dopo l’iscrizione alla Lista (Paggiaro, Rettore e Trivellato, 2006). Ciò permette sensibili miglioramenti nello studio dell’impatto del programma – tipicamente degli effetti differenziali dello stesso, cioè a dire della conseguenza dei maggiori benefici monetari per i lavoratori tra 40 e 49 anni rispetto a quelli per i lavoratori con meno di 40 anni sulla probabilità di ri-occupazione –, in termini di generalità dei risultati e di loro robustezza.

Va aggiunto, peraltro, che ulteriori possibilità di analisi (ad es., sugli effetti differenziali in termini di reddito da lavoro) potrebbero venire da una più ampia integrazione di informazioni alla scala micro, estesa agli archivi dell’Inps.

I ritardi e le difficoltà che si incontrano su questi fronti (si noti che le “Liste di mobilità” restano un programma per molti aspetti esemplare, sul quale si è svolta un’attività di valutazione degli effetti da parte di parecchi ricercatori) sono il sintomo di una carenza basilare: l’assenza di un approccio maturo al disegno e alla realizzazione delle politiche – e degli interventi di regolazione del mercato del lavoro –, orientato quindi (anche) alla valutazione dei loro effetti e alla conseguente predisposizione della rilevazione delle informazioni necessarie (Rettore, Trivellato e Martini, 2003).

Verso un sistema informativo del lavoro: ritardi e prospettive

9. La ricognizione di un’esperienza di ricerca – sulla mobilità dei lavoratori e sugli effetti di politiche del lavoro – è, per sua natura, parziale. Allo stesso tempo, è un interessante studio di caso, che ben si presta per evidenziare le implicazioni per l’informazione statistica di cui serve disporre, su due piani, distinti ma interconnessi:

- quello di *affidabili insieme di indicatori*, convenientemente disaggregati per aspetti tematici e territoriali, prodotti correntemente e comparabili nel tempo;

⁶ Gli episodi di occupazione a tempo determinato (ciascuno con contratto al più di un anno), infatti, ‘fermano l’orologio’ che misura il tempo di permanenza nella Lista per un periodo pari, complessivamente, alla durata massima ammissibile (consentono, cioè, di raddoppiare la durata ammissibile di permanenza nella Lista).

⁷ Al solito, l’occupazione dipendente privata. È peraltro plausibile ritenere che le transizioni al pubblico impiego o al lavoro autonomo siano trascurabili.

- quello di *adeguate basi di microdati*: longitudinali, *linked employer-employee*⁸, multi-livello.

10. L'indicazione generale suggerita dalle buone pratiche di altri paesi è nel senso della realizzazione di un *sistema statistico*, inteso nell'accezione tecnico-funzionale di organizzazione e gestione dell'insieme delle attività attinenti a *input/processing/output* dell'informazione statistica.

In via preliminare, è essenziale considerare le implicazioni che ne vengono per i tre grandi momenti nei quali si articola il processo di generazione dell'informazione statistica sul lavoro:

- saper cogliere le domande, e tradurle in fabbisogni di informazione statistica,
- saper produrre l'informazione,
- saperla diffondere.

La distinzione è importante. Il primo e il terzo momento – saper cogliere le domande e saper diffondere l'informazione – poggiano, infatti, sui rapporti con gli utilizzatori, in primo luogo Regioni ed enti locali. Per il secondo momento, invece, quello propriamente della produzione dell'informazione statistica, assumono un rilievo preminente considerazioni di indole interna, che attengono alla qualità del processo produttivo e portano in primo piano le nozioni di integrazione, modularità, flessibilità del sistema (Trivellato, 2003).

Nel guardare alle buone pratiche nel panorama internazionale, l'esempio di *Denmark Statistics* è particolarmente istruttivo (una sua breve presentazione è in Anastasia e Trivellato, 2001). Il sistema statistico si configura come organizzazione e gestione integrata di insiemi di basi di dati, alimentate da una varietà di fonti: censimenti, *surveys*, soprattutto archivi di origine amministrativa, loro convenienti combinazioni. Una sua semplificata, schematica rappresentazione è nella Figura 6. Le esigenze di assetto coerente integrato del sistema stesso diventano i termini di riferimento essenziali e dettano le prescrizioni lungo tutto l'asse del processo produttivo: dalle definizioni, nomenclature e classificazioni, all'organizzazione dei dati elementari (comprensiva della descrizione con metadati), fino alla produzione dei risultati statistici.

 Figura 6 circa qui

11. Per il lavoro, i principali soggetti ai quali confluiscono – e che elaborano – le informazioni elementari sono ben noti: l'Istat; l'Inps, e per l'occupazione nel settore pubblico l'Inpdap; l'Inail; i Centri per l'impiego, e i conseguenti sistemi informativi provinciali e/o regionali; gli archivi fiscali.

Le grandi tessere del mosaico informativo sul lavoro, dunque, ci sono. Ma slabbrate – di qualità molto variabile, con ridondanze e insieme con vistose falle – e soprattutto largamente sconnesse. L'esigenza cruciale è appunto di ricondurle a sistema statistico.

12. Sul terreno della realizzazione di un sistema statistico del lavoro, se è agevole riconoscere che nell'ultimo decennio sono stati compiuti apprezzabili passi, è altrettanto difficile negare sensibili ritardi⁹. In sintesi:

- Non siamo completamente allineati agli standard europei per quanto attiene alle batterie di indicatori connessi al monitoraggio delle *Employment Guidelines* dell'Unione Europea e alla loro qualità.

⁸ Sulla rilevanza di *employer-employee data* per un vasto insieme degli studi sul mercato del lavoro e sulle politiche delle imprese, a fini analitici e/o di *policy analysis*, vedi Bryson, Forth e Barber (2006).

⁹ Traggo parte delle riflessioni che seguono da Trivellato (2006).

- Registriamo vistose carenze nel monitoraggio dei nuovi istituti contrattuali – a cominciare dalla loro semplice quantificazione (valga, per tutti, il già menzionato esempio dei lavoratori co.co.co.) – e dei nuovi servizi, pur essendo questa attività spesso prevista come obbligatoria dalla normativa.
- Abbiamo un grave ritardo nell'approntamento e nella messa a disposizione di basi di dati adeguate alla valutazione degli effetti di politiche attive del lavoro e di misure di regolazione.
- In generale, i progressi compiuti nella direzione di un sistema statistico del lavoro sono stati meno che mediocri. È paradossale – e grave – che le potenzialità dei grandi archivi amministrativi e dei flussi informativi dei Centri per l'impiego siano utilizzate poco, in maniera non integrata, al più a 'a macchia di leopardo', e che il Sistema Informativo (nazionale) del Lavoro resti una sorta di fata Morgana, che dal 1998 pare di intravedere all'orizzonte, ma mai si materializza.

13. I fattori che sono alla base di queste carenze sono numerosi. Vi sono problemi di risorse. Ma vi sono anche *cruciali problemi di 'governance'*, acuiti da un'enfasi impropria, distorta sugli aspetti istituzionali. Mi riferisco ad atteggiamenti – di Regioni, di singole amministrazioni statali, di enti – che pongono l'accento sulla dimensione 'patrimoniale', proprietaria, nell'acquisizione e nell'uso dei dati, sulla costruzione di sistemi statistici regionali o settoriali venati da pretese di autosufficienza¹⁰. Mi auguro – e confido – che questo convegno segni un deciso superamento di tali atteggiamenti.

14. La questione assume particolare rilievo per i dati di fonte amministrativa. I *dati amministrativi* assumono, infatti, un'importanza crescente, per una varietà di ragioni. Sono riferiti a tutta la popolazione pertinente e, di massima, correntemente aggiornati. Il loro utilizzo, per generare corrispondenti archivi statistici, comporta grandi vantaggi operativi in termini di riduzione dei costi e di eliminazione di un gravame aggiuntivo per i rispondenti. Inoltre, amplia in maniera decisiva l'offerta informativa, perché consente di (i) produrre stime riferite a piccole aree, in particolare a domini funzionali quali sono i mercati del lavoro locali, (ii) generare basi di dati longitudinali e (iii) monitorare in maniera adeguata le politiche.

Tali archivi, peraltro, devono soddisfare l'ovvia condizione preliminare di essere *buoni di per sé, rispetto alle finalità proprie* (tre soli esempi, non incoraggianti, ai quali ho già fatto riferimento: i lavoratori co.co.co., i disoccupati registrati nei Centri per l'impiego, l'informazione sul programma Liste di mobilità). Sotto questo profilo, l'impiego a fini statistici e di ricerca dei dati amministrativi costituisce un'importante cartina di tornasole della trasparenza e della qualità dell'azione amministrativa.

Inoltre, essi richiedono di essere meglio regolati per quanto attiene alla definizione di modelli e formati di raccolta delle informazioni, in modo che ne siano assicurate *ex-ante* essenziali caratteristiche di *completezza, affidabilità e integrabilità*, anche ai fini di una loro utilizzazione statistica¹¹. Affidare all'Istat compiti incisivi in materia – di indirizzo, di assistenza metodologica e tecnica, di controllo –, e insieme assicurare un efficace

¹⁰ Né vanno sottaciuti altrettanto impropri slittamenti delle priorità sulla dimensione informatica. Le potenzialità dei moderni mezzi di trattamento computerizzato dei dati sono notevolissime, anche per aspetti attinenti al vaglio della qualità e all'estrazione di informazioni rilevanti da vaste masse di dati 'imprecisi' – ad esempio, con tecniche di *data mining* –. Tuttavia, è bene non scordare una vecchia massima: "*garbage in, garbage out*". Essa mantiene una buona dose di verità, ed è un'utile bussola nel disegnare sistemi informativi, sia a fini gestionali che a scopi statistici.

¹¹ Si noti che ciò non vale ad escludere la necessità di una altrettanto impegnativo lavoro *ex-post*: di riconduzione a concetti, definizioni e classificazioni rilevanti a fini statistici (e, più in generale, conoscitivi) e di verifica/integrazione con altre fonti. Semplicemente, pone le condizioni perché questa attività *ex-post* si svolga contando su basi più solide e in maniera più agevole.

coordinamento fra l'Istituto e le amministrazioni e gli enti interessati, è un'esigenza pressante, che è essenziale sia sollecitamente soddisfatta.

15. Un'ultima considerazione attiene all'accesso alle basi di microdati: vitale in generale per la ricerca scientifica, essenziale per studi sui mercati del lavoro locali.

Pur in un contesto segnato da ostacoli e ritardi – la normativa sulla tutela della *privacy* non è certo di grande aiuto –, stanno emergendo buone pratiche di produzione e diffusione di basi di microdati, sovente longitudinali. Oggi WHIP, CLAP (il Campione Longitudinale degli Attivi e dei Pensionati, dagli archivi Inps, prodotto e distribuito dalla Divisione Coordinamento attività statistiche del Ministero del Lavoro) e GIOVE sono *public use files*. Ai quali è affiancata la possibilità di comunicazione di *files* più ricchi, con dati identificativi, per progetti di ricerca che ne necessitino, previa argomentata richiesta e l'impegno a rispettare la normativa sulla *privacy*¹².

Una domanda analoga, ma ben più impegnativa perché generale, si pone nei confronti dell'informazione statistica ufficiale, in primo luogo dell'Istat. Essa attiene appunto alla possibilità per i ricercatori – in generale per gli *analytical users* – di accedere a basi di microdati sì privi di identificativi diretti, ma riferiti a individui identificabili. Un'eventualità, questa, tutt'altro che peregrina. A ben vedere, essa si dà non appena i *files* abbiano informazioni individuali sufficientemente ricche: il che è pressoché inevitabile quando l'analisi coinvolga storie lavorative e storie di imprese, in particolare alla scala di mercati del lavoro locali.

Non è forse questa la sede – e per certo non sta a me – addentrarmi sul terreno delle modalità per procedere su questa via. Concludo prospettando, sommariamente, quelli che a me appaiono l'obiettivo e la strada da imboccare.

- (a) L'obiettivo è realizzare condizioni positive di effettiva disponibilità di basi di microdati, riferiti a individui potenzialmente identificabili, per la ricerca: prodotte in via corrente, mantenute e distribuite entro una cornice organizzativa adeguata, rese accessibili secondo il principio dell'"*universal access to all 'bona fide' analytical users*".
- (b) La strada maestra è, ragionevolmente, nella realizzazione di un "sito virtuale sicuro" sulla rete web¹³. Il CEIES – il Comitato Europeo consultivo sull'Informazione statistica nei campi Economico e Sociale – l'ha raccomandata a livello dell'Unione Europea (CEIES, 2002), e gli istituti statistici di alcuni paesi europei vi si sono già incamminati.

Riferimenti bibliografici

Anastasia B. e M. Disarò (2005), "Il profilo del disoccupato a partire dai dati amministrativi veneti", Workshop *Metodi e applicazioni per la valutazione delle politiche del lavoro e di aiuto alle imprese*, Università di Salerno in collaborazione con l'Istituto Italiano di Studi Filosofici, Napoli 29 settembre-1 ottobre 2005 (mimeo).

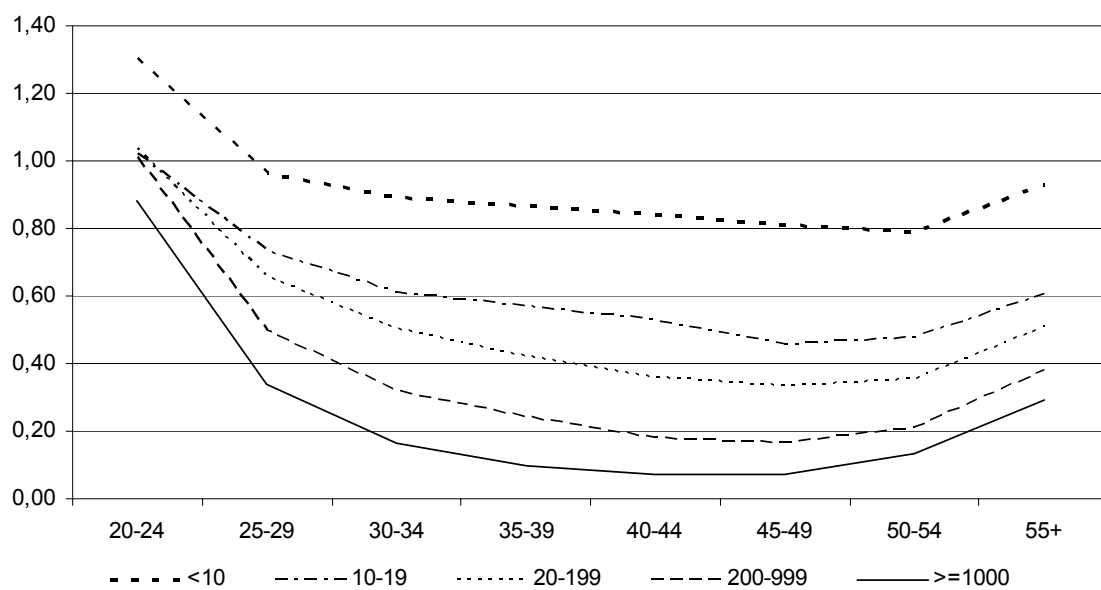
¹² Per il CLAP, che è prodotto e distribuito dalla Divisione Coordinamento attività statistiche del Ministero del Lavoro – un soggetto del Sistan –, occorre inoltre sottoscrivere una convenzione, per la quale è peraltro disponibile un agile formato-tipo.

¹³ Questa soluzione, qui prospettata per i ricercatori, palesemente potrebbe valere anche per la comunicazione di microdati fra i soggetti del Sistan, per la quale il Comstat ha adottato una direttiva la cui attuazione si presenta piuttosto laboriosa (vedi Istat, "Direttiva 20 aprile 2004. Criteri e modalità per la comunicazione di dati personali nell'ambito del Sistema statistico nazionale (Direttiva n. 9)", *Gazzetta Ufficiale*, n. 300, 23 dicembre 2004).

- Anastasia B., M. Disarò, M. Gambuzza, D. Maurizio e M. Rasera (2006), “‘Giove’: il *public use database* prodotto da Veneto Lavoro a partire dai dati dei Centri per l’impiego”, relazione a questo convegno.
- Baretta, P. e U. Trivellato (2004), “La mobilità dei lavoratori da fonti amministrative e da *surveys* sulle famiglie: un’analisi comparata”, *Statistica*, XLIV, n. 1, pp. 23-55.
- Battistin E., E. Rettore e U. Trivellato (2006), “Choosing among alternative classification criteria to measure the labour force state”, *Journal of the Royal Statistical Society - Series A*, 169 (forthcoming).
- Berton F., L. Pacelli e G. Segre (2005), “Il lavoro parasubordinato in Italia: tra autonomia del lavoratore e precarietà del lavoro”, *Rivista Italiana degli Economisti*, 10, n. 1, pp. 57-99.
- Brandolini A., P. Cipollone e E. Viviano (2006), “Does the ILO definition capture all unemployment?”, *Journal of the European Economic Association*, 4, n. 1, pp. 153-179.
- Bryson A. e J. Forth (2006), “The policy relevance of linked employer-employee data”, in A. Bryson, J. Forth e C. Barber (ed.) (2006), cit..
- Bryson A., J. Forth e C. Barber (ed.) (2006), *Making linked employer-employee data relevant to policy*, DTI Occasional Paper No. 4, London, Department of Trade and Industry.
- CEIES (2002), *CEIES opinion: Access to microdata*, adopted at the 12th meeting of the CEIES, Brussels, 22 November 2002 (CEIES 2002/12/6.2/EN).
- Contini, B. e R. Revelli (1987), “Job creation and job destruction in the Italian economy”, *Labour*, 1, n. 3, pp. 121-144.
- Contini, B. e R. Revelli (1992), *Imprese, occupazione e retribuzioni al microscopio*, Bologna, Il Mulino.
- Contini B. e U. Trivellato (a cura di) (2005a), *Eppur si muove. Dinamiche e persistenze nel mercato del lavoro italiano*, Bologna, Il Mulino.
- Contini B. e U. Trivellato (2005b), “Dinamiche e persistenze nel mercato del lavoro italiano: una sintesi”, in B. Contini e U. Trivellato (a cura di) (2005a), cit..
- Istat (2004), *La nuova rilevazione sulle forze di lavoro - Contenuti, metodologie, organizzazione*, Roma, 3 giugno 2004.
- Istat (2005), *Flessibilità dell’orario e organizzazione del lavoro*, «Statistiche in breve», Roma, 29 dicembre 2005.
- Istat (2006), *Rapporto annuale. La situazione del Paese nel 2005*, Roma, 24 maggio 2006.
- Lazear E.P. e K.L. Shaw (2006), “Wage structure, raises and mobility”, in A. Bryson, J. Forth e C. Barber (ed.) (2006), cit..
- Leombruni R. e R. Quaranta (2005), “Eppure si muoveva già. Una breve storia della mobilità dei lavoratori in Italia”, in B. Contini e U. Trivellato (a cura di) (2005a), cit..
- Melis A. e B. Lüdeke (2006), *Registered unemployed (RU) compared with harmonised unemployed (LFS)*, «Working Papers and Studies», Luxembourg, Eurostat.
- Paggiaro A., E. Rettore e U. Trivellato (2006), *The effect of extended duration of eligibility in an Italian labour market programme for dismissed workers*, paper presented at the AEA Conference “Policies against unemployment”, Naples, June 1-2, 2006 (mimeo).
- Pirrone S. e P. Sestito (2006), *Disoccupati in Italia. Tra Stato, Regioni e cacciatori di teste*, Bologna, Il Mulino.
- Rettore E., U. Trivellato e A. Martini (2003), “La valutazione delle politiche del lavoro in presenza di selezione: migliorare la teoria, i metodi o i dati?”, *Politica Economica*, 19, n. 3, pp. 301-341.

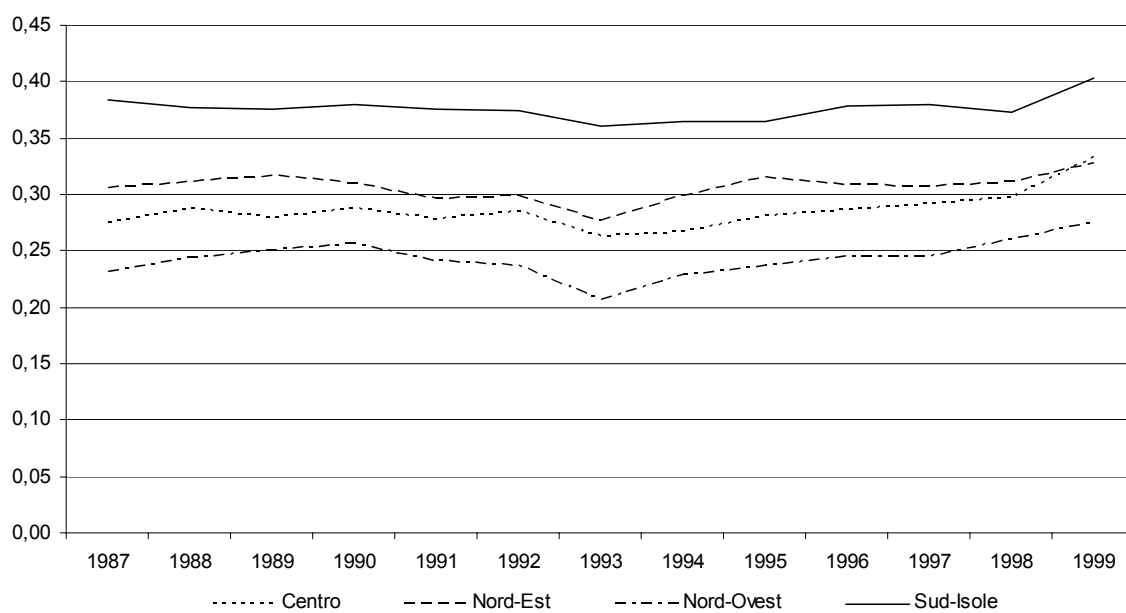
- Tattara G. e M. Valentini (2005), “La mobilità dei lavoratori dell’industria nel Veneto: dinamica di lungo periodo e aspetti differenziali”, in B. Contini e U. Trivellato (a cura di) (2005a), cit..
- Trivellato U. (2003), “Qualità dell’informazione statistica ufficiale e esigenze informative di regioni e città”, *Giornale del Sistan*, 8, n. 22, pp. 10/All.-19/All.
- Trivellato U. (2006), “Trasformazioni del mercato del lavoro italiano, politiche del lavoro e disponibilità/fabbisogni informativi: un percorso a metà del guado”, *Economia e Società Regionale*, n. 2 (in corso di pubblicazione).
- Trivellato U., A. Paggiaro, R. Leombruni e S. Rosati (2005), “La dinamica recente della mobilità dei lavoratori, 1998-2003”, in B. Contini e U. Trivellato (a cura di) (2005b), cit..

Fig. 1. *Gross worker turnover* per età del lavoratore e dimensione d'impresa, 1987-1999.



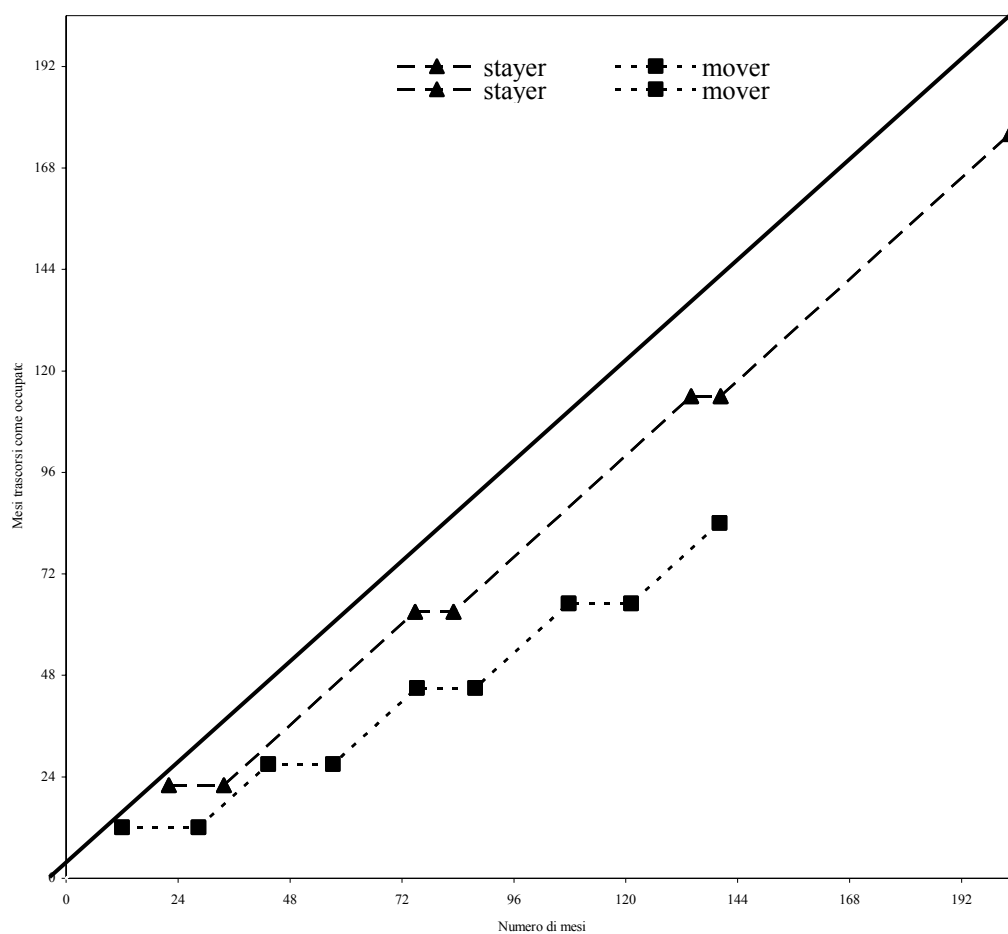
Fonte: Leombruni e Quaranta (2005), pag. 224.. Elaborazioni su dati da WHIP.

Fig. 2. Tasso di riallocazione per area geografica, 1987-1999



Fonte: Leombruni e Quaranta (2005), pag. 227. Elaborazioni su dati da WHIP.

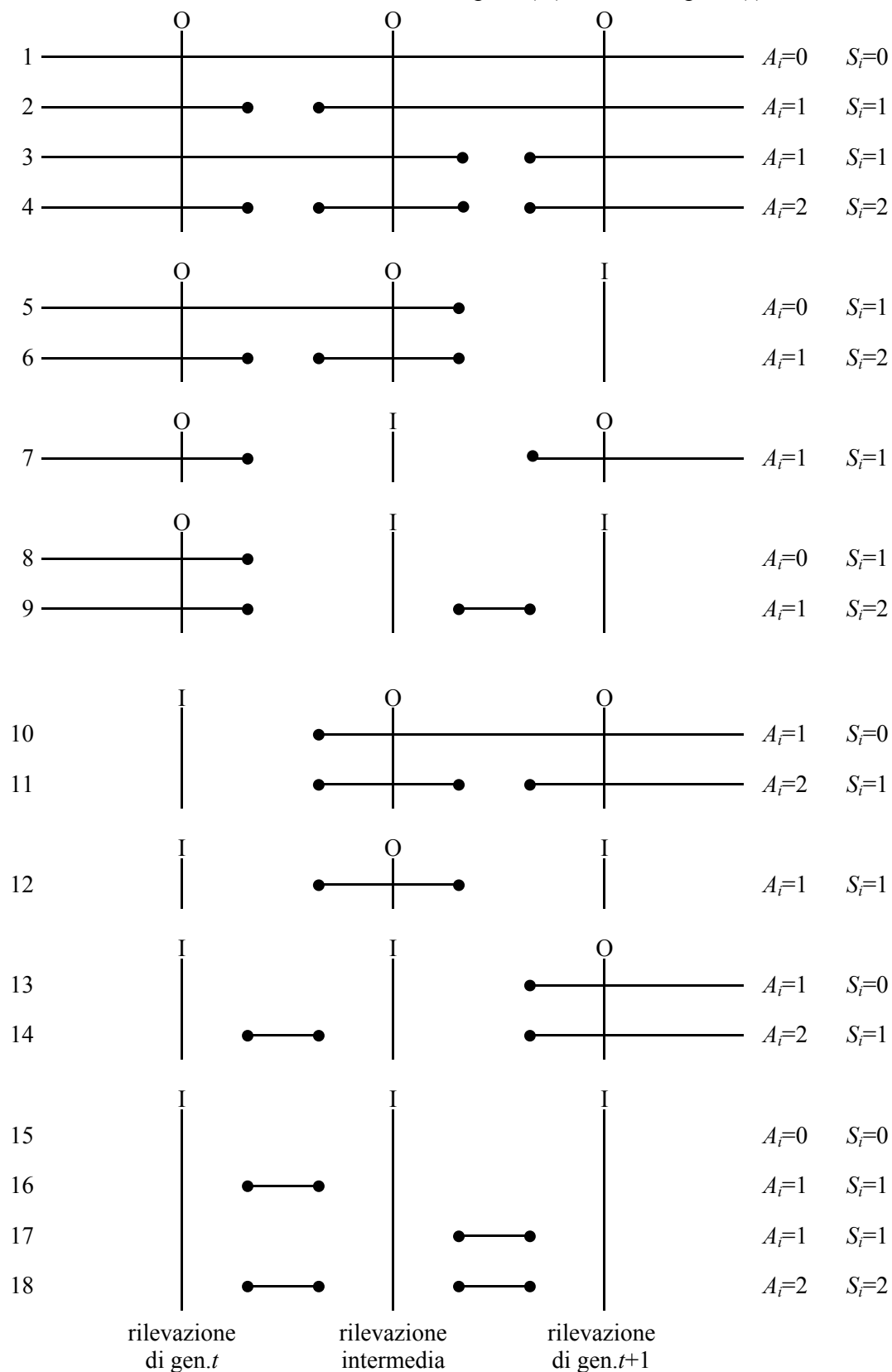
Fig. 3. La carriera del *permanent mover* e dello *stayer* rappresentativo nel periodo 1982-1997: uomini (media degli uomini della coorte di nascita 1950-60 che hanno lavorato nel periodo 1982-1990)



Nota:: la successione degli episodi di lavoro-non lavoro per lo *stayer* e per il *permanent mover* rappresentativo è limitata rispettivamente ai 4 e ai 5 episodi di occupazione iniziali della carriera lavorativa.

Fonte: adattato da Tattara e Valentini (2005), pag. 266. Elaborazioni su dati da VWH.

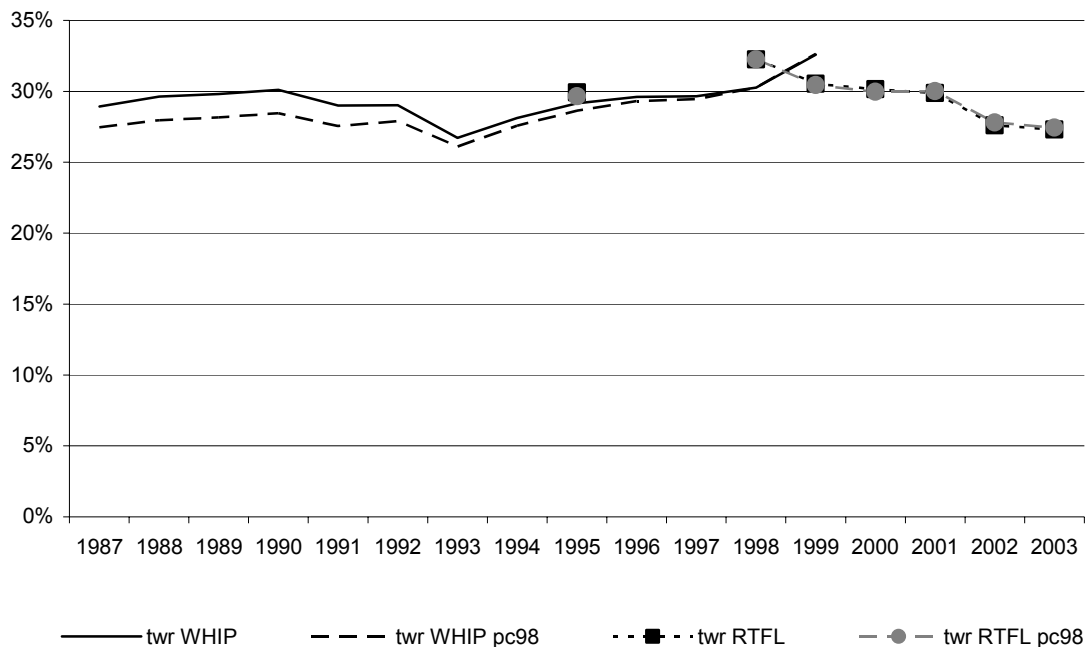
Fig. 4. Schema per l'identificazione degli episodi lavorativi e delle transizioni dall'abbinamento a tre occasioni di RTFL: occupato (O) e non occupato (I).



Legenda: Ogni rapporto di lavoro è rappresentato da un segmento orizzontale, che viene intersecato dalle occasioni d'indagine. La condizione lavorativa di un individuo rilevata correntemente in ogni occasione è indicata con O (occupato) o I (non occupato). Una transizione (associazione, A_i , o separazione, S_i) entro l'intervallo gen.t-gen.t+1 è indicata con '•'.

Fonte: Trivellato, Paggiaro, Leombruni e Rosati (2005), p. 278.

Fig. 5. Serie storiche dei tassi di riallocazione degli occupati dipendenti nell'industria e nei servizi privati, popolazione corrente e popolazione costante (1998), rispettivamente dalla fonte WHIP per il periodo 1987-1999 e dalla RFL per il 1995 e 1998-2003.



Nota: Le serie a popolazione costante sono ottenute mantenendo invariata la struttura della popolazione del 1998 per genere e per le classi di età 15-29, 30-49, 50-64.

Fonte: Contini e Trivellato (2005b), p. 42.

Fig. 6. Schema-tipo di sistema statistico che integra basi di dati da una pluralità di fonti

