

VOLUME LXIV – NN. 1-2

GENNAIO- GIUGNO 2010

**RIVISTA ITALIANA
DI ECONOMIA DEMOGRAFIA
E STATISTICA**

COMITATO SCIENTIFICO

Prof. LUIGI DI COMITE , Prof. FRANCESCO FORTE, Prof. GIOVANNI MARIA GIORGI
Prof. VINCENZO LO JACONO, Prof. MARCELLO NATALE, Prof. ALBERTO QUADRIO CURZIO
Prof. GIOVANNI SOMOGYI

COMITATO DI DIREZIONE

Prof. PIETRO Busetta, Prof. CATELLO COSENZA †, Prof.ssa SILVANA SCHIFINI D'ANDREA
Prof. SALVATORE STROZZA, Prof. SILIO RIGATTI LUCHINI

DIRETTORE

Prof. ENRICO DEL COLLE

REDAZIONE

Dott. GIOVANNI CARIANI, *Redattore capo*

Dott. CLAUDIO CECCARELLI, Dott. ANDREA CICCARELLI, Dott.ssa PAOLA GIACOMELLO
Prof.ssa ANNA PATERNO, Dott.ssa ANGELA SILVESTRINI
GABRIELLA BERNABEI, *Segretaria di Redazione*



Direzione, Redazione e Amministrazione

Piazza Tommaso de Cristoforis, 6

00159 ROMA

TEL. e FAX 06-43589008

E-mail: sieds@tin.it

Stampato da CLEUP sc
“Coop. Libreria Editrice Università di Padova”
Via G. Belzoni, 118/3 – Padova (Tel. 049/650261)
www.cleup.it

IN QUESTO NUMERO

Questo volume accoglie, in ordine alfabetico, una selezione effettuata in sede di referaggio, delle comunicazioni libere dei Soci presentate in occasione della XLVI Riunione Scientifica della Società Italiana di Economia, Demografia e Statistica tenutasi a Firenze dal 28 al 30 maggio 2009, sul tema “Povertà ed esclusione sociale”. La Riunione Scientifica è stata organizzata in collaborazione con il Dipartimento di Statistica “Giuseppe Parenti” dell’Università degli Studi di Firenze.

Enrico Del Colle

INDICE

Anna Maria Altavilla, Angelo Mazza, Antonio Punzo, <i>Sull’impiego di un indice di dissimilarità nello studio della disposizione di popolazioni straniere su un territorio urbano</i>	7
Francesco Antonio Anselmi, <i>Ecoturismo: caratteristiche e potenzialità di sviluppo in Italia</i>	15
Francesco Antonio Anselmi, <i>Una nuova politica regionale dell’ocse e dell’unione europea</i>	23
Vincenzo Asero, Simona Gozzo, <i>La qualità nel settore turistico: un’analisi empirica</i>	31
Giuseppe Avena, Emilia De Luca, <i>Interpretazione della collaborazione familiare in giovani coppie attraverso l’analisi testuale di interviste soggettive</i>	39
Odo Barsotti, Federico Benassi, Linda Porciani, Alessandro Valentini, Silvia Venturi, <i>Percorsi scolastici e previsioni della popolazione scolastica. Un caso di studio</i>	47

Odoardo Bussini, Donatella Lanari, <i>Condizione socio-economica e disuguaglianze di salute nella popolazione anziana in Europa</i>	55
Oliviero Casacchia, Lorenzo Cassata, Piero Giorgi, Cecilia Reynaud, Salvatore Strozza, <i>La mobilità interna in Italia: le principali caratteristiche dei cambiamenti di residenza</i>	63
Cinzia Castagnaro, Antonella Guarnieri, <i>Le nascite fuori dal matrimonio: tendenze in atto e principali caratteristiche</i>	71
Cinzia Conti, Antonella Guarnieri, Francesca Licari, Enrico Tucci, <i>La mobilità interna degli stranieri in Italia: uno studio attraverso il record linkage tra archivi</i>	79
Rosario D'Agata, Venera Tomaselli, <i>La struttura dei sistemi turistici locali in Sicilia: analisi della stagionalità incoming</i>	87
Elena De Filippo, Paolo Diana, Raffaele Ferrara, Linda Forcellati, <i>Alcuni aspetti dell'integrazione degli immigrati nella Provincia di Napoli</i>	95
Anna Di Bartolomeo, Antonella Guarnieri, Paola Muccitelli, <i>Performance scolastiche e presenza di alunni stranieri: un'analisi spaziale nel Comune di Roma</i>	103
Marta Disegna, Christine Mauracher, Isabella Procidano, <i>Segmentazione della domanda di vino in Italia mediante mappe di Kohonen</i>	111
Raffaele Ferrara, Piero Giorgi, Marija Mamolo, Salvatore Strozza, <i>Il ruolo della fecondità degli stranieri in due paesi di recente immigrazione: il caso dell'Italia e della Spagna</i>	119
Romana Gargano, <i>Modelli gerarchici per lo studio dell'inquinamento atmosferico nello stretto di Messina</i>	127
Romana Gargano, Giuseppe Avena, <i>Modelli variabili latenti per la valutazione della soddisfazione degli utenti del servizio di trasporto pubblico</i>	135
Rosa Giaimo, Dario Corso, Giovanni Luca Lo Magno, <i>Mercato del lavoro femminile e contesto familiare nelle Regioni Italiane</i>	143

Pietro Iaquina, Francesca Di Lazzaro, <i>Fattori di attrattività delle migrazioni: le minoranze etniche. Il caso degli arbëreshë</i>	151
Letizia La Tona, Angela Alibrandi, <i>Statistical implicative analysis and principal components method: a comparison between two approaches for the analysis of social data</i>	159
Antonio Lucadamo, Giovanni Portoso, <i>La rasch analysis associata ai modelli multilevel nella valutazione della customer satisfaction</i>	167
Silvia Meggiolaro, <i>Low birth weight and parental resources</i>	175
Giovanni Portoso, Antonio Lucadamo, <i>Soglie inferiori e superiori di taluni indici di connessione</i>	183
Mauro Reginato, Tiziana Barugola, <i>Immigrati stranieri in Piemonte e fecondità una prima analisi descrittiva</i>	191
Giuseppe Riccardo Lamonica, <i>La Performance dei modelli di previsione della volatilità</i>	199
Angela Silvestrini, Alessandro Valentini, <i>Evidenze poco note dell'incremento della popolazione straniera nei Comuni italiani</i>	207
Valeria Stragapede, Annamaria Stramaglia, <i>Industria e territorio tra aspettative locali e esigenze globali. Prime evidenze di un caso studio</i>	215

SULL'IMPIEGO DI UN INDICE DI DISSIMILARITA' NELLO STUDIO DELLA DISPOSIZIONE DI POPOLAZIONI STRANIERE SU UN TERRITORIO URBANO

Anna Maria Altavilla, Angelo Mazza, Antonio Punzo

1. Introduzione

I differenti pattern insediativi che caratterizzano i diversi gruppi demografici vengono spesso associati, implicitamente o esplicitamente, alle specificità dei processi comportamentali tipiche di ciascun gruppo le quali, sia con riferimento ai diversi modi di reagire alle caratteristiche ambientali, sia con riferimento all'interazione tra ed all'interno dei gruppi, possono talora favorire una reciproca attrazione tra gli individui e talaltra un allontanamento fra gli stessi.

E' possibile considerare una data disposizione osservata come una singola manifestazione di un processo di collocazione e tale processo può presentare caratteristiche differenti per i diversi gruppi demografici. In generale, l'interesse del ricercatore si concentra sulle caratteristiche "sistemiche" di tali processi, prescindendo da fluttuazioni di origine casuale che potrebbero coinvolgere le singole osservazioni.

In letteratura sono stati proposti numerosi indici di segregazione, ciascuno dei quali ne sottintende una specifica definizione; Massey e Denton (1988) individuano cinque dimensioni della segregazione. Tra queste, la dissimilarità è quella tradizionalmente più studiata e l'indice D di Duncan e Duncan (1955), del quale si tratta in questo lavoro, ne è indubbiamente la misura più impiegata.

E' noto che una problematica connessa con l'impiego dell'indice D è che questo tende a sovrastimare la dissimilarità sistemica quando il gruppo di minoranza rappresenta una proporzione piccola della popolazione o quando le subaree geografiche oggetto dell'analisi sono costituite da un numero basso di individui (Carrington e Troske, 1997). Inoltre, come verrà chiarito attraverso l'impiego di simulazioni Monte Carlo, la distorsione di D tende a crescere al diminuire del livello di dissimilarità del processo (Allen, Burgess e Windmeijer, 2009).

Da quanto detto, si può comprendere facilmente come l'interpretazione di tale indice possa essere fuorviante, soprattutto se si confrontano gruppi demografici con

caratteristiche molto diverse tra loro in termini dei fattori sopramenzionati da cui dipende la distorsione di D .

In questo lavoro, dopo una breve disamina delle proposte più recenti fornite in letteratura per migliorare il comportamento di D , viene presentata una correzione *bootstrap* di tale indice, che consente di ridurre in maniera notevole l'entità della distorsione. Infine, viene presentata un'applicazione ai processi di collocazione realizzati nella città di Catania da parte di alcuni gruppi d'immigrati stranieri appartenenti a differenti nazionalità, che hanno cominciato ad insediarsi stabilmente nella città da tempi più o meno recenti.

2. Una correzione *bootstrap* dell'indice di dissimilarità

Il processo di collocazione di n individui, caratterizzati dal possedere o meno un certo attributo dicotomico $c=0,1$ (ad esempio l'essere italiani o stranieri), in un territorio suddiviso in J subaree indicate con $j=1, \dots, J$, è governato dalle probabilità condizionate

$$p_j^c = P(\text{subarea di appartenenza} = j \mid c), \quad j = 1, \dots, J, \quad c = 0, 1. \quad (1)$$

Si ha dissimilarità "sistemica" se vi è almeno una subarea nella quale gli individui afferenti ad uno dei due gruppi hanno una diversa probabilità di collocarsi, cioè se:

$$\exists j: p_j^0 \neq p_j^1.$$

A partire dalla (1), la probabilità di osservare un dato pattern insediativo può essere così modellata mediante il prodotto di due distribuzioni multinomiali indipendenti, una per $c=0$ ed una per $c=1$ ¹ (Ransom, M. R. 2000; Allen, Burgess e Windmeijer, 2009):

$$P(n_1^0, n_2^0, \dots, n_J^0, n_1^1, n_2^1, \dots, n_J^1) = \prod_{j=1}^J \prod_{c=0}^1 n_j^c! \frac{(p_j^c)^{n_j^c}}{n_j^c!},$$

dove n_j^c rappresenta il numero di individui nella subarea j con attributo c , ed n_c indica il numero totale di individui con attributo c nella popolazione.

¹ Il fondamento di tale formula può essere compreso facilmente se si interpreta la collocazione di ciascun individuo, in una delle J subaree, come il risultato di un esperimento casuale in cui ognuno dei J possibili esiti ha probabilità p_j^c , $j=1, \dots, J$ (modello multinomiale).

Data una particolare suddivisione in subaree del territorio, l'indice di dissimilarità che si considererà in questo lavoro, riferito al processo di collocazione degli individui, è dato da (Duncan and Duncan, 1955):

$$D_{pop} = \frac{1}{2} \sum_{j=1}^J |p_j^1 - p_j^0|. \quad (2)$$

Naturalmente l'indice assume valori nell'intervallo $[0,1]$ ed il suo valore aumenta all'aumentare della dissimilarità sistemica. Inoltre, è facile constatare come il caso $D_{pop}=0$ (assenza di dissimilarità sistemica) sia ottenibile se, e solo se, $p_j^0 = p_j^1$ per ogni $j=1, \dots, J$.

Dal punto di vista inferenziale, data una collocazione osservata, è possibile calcolare l'equivalente campionario (stimatore naturale) di D_{pop} come:

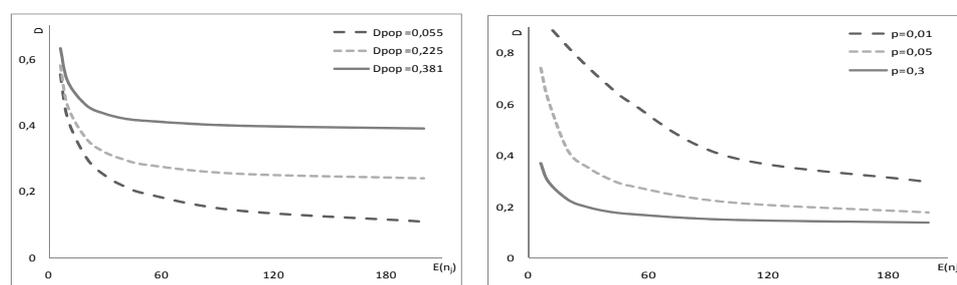
$$D = \frac{1}{2} \sum_{j=1}^J \left| \frac{n_j^1}{n^1} - \frac{n_j^0}{n^0} \right| = \frac{1}{2} \sum_{j=1}^J |\hat{p}_j^1 - \hat{p}_j^0|. \quad (3)$$

Sfortunatamente, come ben noto in letteratura (Carrington e Troske, 1997), D tende a sovrastimare il reale ma incognito valore di D_{pop} per valori non elevati di una o più delle seguenti quantità: p , $E(n_j)$ e D_{pop} . In questi casi, l'impatto della cosiddetta allocazione "casuale" fa sì che l'entità della distorsione

$$bias(D) = E(D) - D_{pop}$$

possa risultare, anche considerevolmente, elevata. In particolare, come si evince dai grafici in Figura 1 e Figura 2, al variare di una delle quantità sopra citate, tenute fisse le altre due, $bias(D)$ si presenta come una funzione decrescente e convergente a zero.

Figura 1 – Valori di D in funzione di $E(n_j)$, $J=50$. In entrambi i grafici, mediante procedura Monte Carlo, sono stati estratti 1.000 campioni in corrispondenza di 8 diversi valori di $E(n_j)$ compresi tra 6 e 200. Per ognuno di questi valori, D è ottenuto come semplice media aritmetica dei 1.000 valori dell'indice di dissimilarità calcolato sul singolo campione.



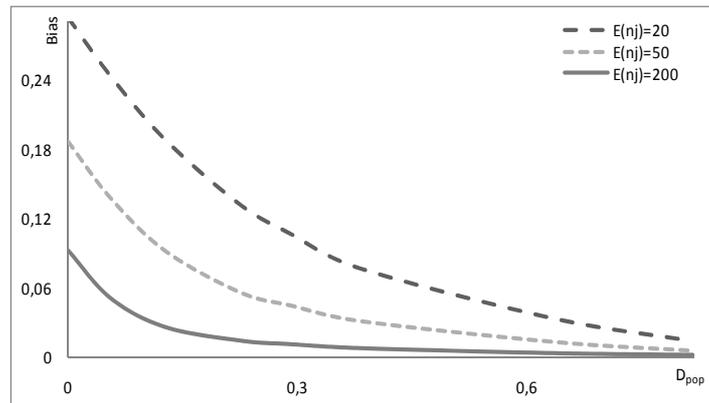
Nel tentativo di eliminare, o quantomeno ridurre, tale distorsione, Allen, Burgess e Windmeijer (2009) hanno proposto una correzione di tipo *bootstrap*. Utilizzando le probabilità osservate \hat{p}_j^c , $c=0,1$, e $j=1,\dots,J$, tramite modello multinomiale, è possibile generare B campioni *bootstrap* e, per ognuno di essi, calcolare l'indice di dissimilarità denotato con D_b . Sia allora

$$\bar{D}_b = \frac{1}{B} \sum_{b=1}^B D_b$$

l'indice di dissimilarità medio riferito ai B campioni *bootstrap*. L'idea degli autori potrebbe "informalmente" essere riassunta dalla proporzione:

$$D_{pop} : D = D : \bar{D}_b$$

Figura 2 – $Bias(D)$ in funzione di D_{pop} , $J=50$ e $p=0.1$. Mediante procedura Monte Carlo, sono stati estratti 1.000 campioni in corrispondenza di 8 diversi valori di $E(n_j)$ compresi tra 0 e 0.818. Per ognuno di questi valori, il valore di D nella formula di $bias(D)$ è ottenuto come semplice media aritmetica dei 1.000 valori dell'indice di dissimilarità calcolato sul singolo campione.



stante a significare che la relazione che intercorre tra D_{pop} e D è, in linea teorica, proporzionale a quella intercorrente tra D e D_{pop} ; in questi termini, uno stimatore "bootstrap" di $bias(D)$ dovrebbe essere $D_{pop} - D$. È facile a questo punto ottenere uno stimatore, con correzione bootstrap della distorsione, nel seguente modo:

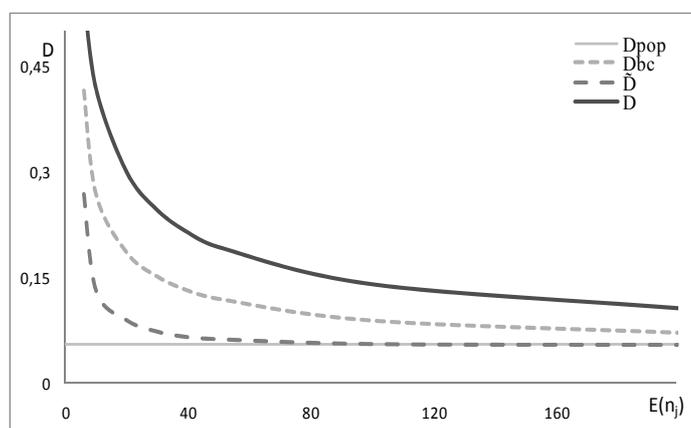
$$D_{bc} = D - (\bar{D}_b - D) = 2D - \bar{D}_b. \quad (4)$$

Allen, Burgess e Windmeijer (2009) mostrano attraverso simulazioni Monte Carlo, facendo variare congiuntamente le quantità $E(n_j)$, D_{pop} , e p , come tale stimatore attenui la distorsione presente in D . La correzione (4) proposta degli autori non

elimina del tutto la distorsione poiché questa non è costante rispetto a D_{pop} ma ne è funzione decrescente (vedi Figura 3).

Nel seguito, si propone uno stimatore di D_{pop} che in generale è meno distorto di D_{bc} . La proposta, forse ancora più intuitiva di quella avanzata da Allen, Burgess e Windmeijer (2009), sfrutta la “filosofia” della massima verosimiglianza ed utilizza come stimatore di D_{pop} quel valore \tilde{D} , sintesi delle probabilità \tilde{p}_j^c , tale che ricampionando con tecniche bootstrap da un processo con tali probabilità, si ottenga in media un valore $\bar{\tilde{D}}$, il più vicino possibile al D osservato². Tale metodo è stato implementato tramite una procedura in linguaggio R che può essere richiesta agli autori.

Figura 3 – Comportamento degli indici D , D_{bc} e \tilde{D} in funzione di $E(n_j)$, $p=0.1$, $J=50$.



3. La segregazione residenziale negli stranieri residenti a Catania

Dalla fine degli anni settanta la città di Catania è meta di flussi migratori che arrivano da tutti i continenti e che, in linea con il processo di frammentazione in atto in tutto il territorio italiano, appartengono a numerose etnie. Alcune di queste, provenienti dai PVS del Mediterraneo, continuano a dare vita, come nel passato, a limitati flussi stagionali, ma la maggior parte ha cominciato a realizzare progetti migratori di lunga durata insediandosi stabilmente nella città. Uno studio, relativo

² E' evidente che esistono infinite coppie di sequenze di \tilde{p}_j^c cui corrisponde uno stesso valore di \tilde{D} . Il criterio che è stato adottato è quello di distribuire la quantità $(D - \tilde{D})$ tra le J coppie $(\hat{p}_j^0, \hat{p}_j^1)$ in proporzione al corrispondente scarto $|\hat{p}_j^0 - \hat{p}_j^1|$.

alla fine degli anni ottanta, sui processi di collocamento degli immigrati stranieri che avevano fissato stabile dimora nella città (Altavilla, 2001) ha evidenziato l'esistenza di un'alta concentrazione di questi all'interno di alcune subaree identificate con i quartieri più poveri. Da studi successivi (Altavilla, 2002; Altavilla e Mazza, 2008) è emerso, tuttavia, che le varie nazionalità presenti nella città hanno mostrato e continuano a mostrare una diversa evoluzione non solo numerica ma anche occupazionale, la quale ha determinato nel tempo la realizzazione di processi d'insediamento e d'integrazione nel territorio molto diversificati tra le varie comunità, con la conseguente trasformazione dei relativi pattern insediativi.

Con l'ausilio dell'indice proposto, si è misurato il grado di segregazione dei gruppi di stranieri numericamente più consistenti residenti nel territorio catanese³. La Tabella 1 mostra la proporzione p degli immigrati stranieri sulla popolazione residente totale, i valori dei tre indici D , D_{bc} e \bar{D} , gli scostamenti tra di essi ed il rango nelle graduatorie secondo i tre indici.

Tabella 1 – Indici di dissimilarità D , D_{bc} e \bar{D} relativi agli stranieri residenti delle dieci nazionalità più numerose. Il territorio urbano del comune di Catania è stato suddiviso in subaree quadrate di lato pari a m.1.000 e $E(n_j)=1470$.

Nazionalità	p	D	D_{bc}	\bar{D}	$D - \bar{D}$	$D_{bc} - \bar{D}$	Graduatoria		
							D	D_{bc}	\bar{D}
Bangladese	0,05%	0,87	0,86	0,86	0,01	0,00	1	1	1
Senegalese	0,17%	0,84	0,84	0,84	0,00	0,00	2	2	2
Cinese	0,09%	0,82	0,81	0,81	0,01	0,00	3	3	3
Marocchina	0,07%	0,68	0,65	0,64	0,04	0,01	4	4	4
Tunisina	0,10%	0,58	0,56	0,55	0,03	0,00	6	5	5
Greca	0,05%	0,59	0,54	0,52	0,07	0,02	5	6	6
Srilankese	0,46%	0,51	0,50	0,50	0,01	0,00	9	8	7
Colombiana	0,04%	0,57	0,51	0,49	0,08	0,03	7	7	8
Mauriziana	0,72%	0,46	0,45	0,45	0,01	0,00	10	10	9
Rumena	0,03%	0,56	0,46	0,40	0,16	0,06	8	9	10

I valori dei tre indici riportati nella Tabella 1 sembrano, in effetti, evidenziare un differenziato percorso d'inserimento all'interno del tessuto urbano da parte dei dieci gruppi di stranieri analizzati. I valori più bassi fanno riferimento ai due gruppi provenienti dalle isole Mauritius e dallo Sri Lanka che rappresentano a Catania le due comunità più numerose e di più antica immigrazione (Altavilla, 2002). Il livello di segregazione per questi due gruppi si presenta abbastanza vicino a quello

³ I dati cui si fa riferimento sono il data base del registro anagrafico del comune di Catania, aggiornato all'agosto 2005. I dati relativi alle residenze sono stati georeferenziati ed integrati in un sistema informativo geografico.

delle due comunità europee analizzate e sembra coerente con un buon livello d'integrazione. Dediti prevalentemente al lavoro domestico, i singoli membri delle due comunità hanno, infatti, rispetto agli altri gruppi, maggiori possibilità d'interagire positivamente con la popolazione locale e di ben inserirsi nella città. L'elevata concentrazione rilevata per le comunità di nazionalità bangladesi, senegalese e cinese sembra denunciare, invece, la persistenza di un basso livello d'integrazione.

Con riferimento alle differenze nei valori assunti dai tre indici, si può osservare che $D \leq D_{bc} \leq \tilde{D}$. Gli scostamenti di D e D_{bc} da \tilde{D} sono molto bassi o nulli per i gruppi di stranieri maggiormente numerosi (Mauriziani e Srilankesi) o caratterizzati da livelli più alti di dissimilarità (Bangladesi, Senegalesi, Cinesi) mentre sono più elevati per quei gruppi al contempo meno numerosi e caratterizzati da una minore dissimilarità (Colombiani, Greci, Marocchini e, in particolare, Rumeni); tali differenze hanno un ordine di grandezza tale da incidere anche sull'ordinamento dei gruppi. Quanto osservato è in linea con la presenza di una distorsione positiva di D in corrispondenza di valori bassi di p e di D_{pop} che, come già visto tramite le simulazioni Monte Carlo, \tilde{D} riesce a correggere meglio di D_{bc} .

4. Conclusioni

L'indice di dissimilarità D di Duncan e Duncan (1955) è uno stimatore positivamente distorto della dissimilarità sistemica; l'entità di tale distorsione cresce se il gruppo di minoranza rappresenta una frazione piccola della popolazione totale, se le subaree oggetto di analisi contengono un numero basso di individui e quando la stessa dissimilarità sistemica è bassa. Tali circostanze rendono difficile l'interpretazione dell'indice, sia in termini assoluti che relativi.

In questo lavoro si è proposta una correzione *bootstrap* dell'indice D ; le simulazioni Monte Carlo effettuate hanno mostrato che tale correzione, pur non eliminando del tutto la distorsione, riesce a limitarla considerevolmente, comunque in maniera superiore a quanto consentito da altre correzioni proposte in letteratura di recente.

L'applicazione presentata relativa ai pattern insediativi degli stranieri residenti nell'area urbana del comune di Catania ha mostrato un livello elevato di eterogeneità tra i diversi gruppi, sia in termini di consistenza numerica che di dissimilarità. In corrispondenza dei gruppi meno numerosi e diffusi in maniera più omogenea sul territorio, dove cioè la distorsione di D era più elevata, l'effetto della correzione proposta è stato più marcato e tale da incidere anche sulla graduatoria dei diversi gruppi in funzione della dissimilarità.

Riferimenti bibliografici

- Allen R., Burgess S., Windmeijer F. 2009. *More Reliable Inference for Segregation Indices*, in Working Paper No. 09/216, Centre for Market and Public Organization, University of Bristol.
- Altavilla A. M., 2001. *Modelli di dispersione della popolazione straniera a Catania*, in Annali della Facoltà di Economia dell'Università di Catania, vol. XLVII, pp. 497-527.
- Altavilla A.M. 2002. *Indagine sulle famiglie di extracomunitari a Catania*, in L. Di Comite, M.C. Miccoli (a cura di), Cooperazione, multietnicità e mobilità territoriale delle popolazioni, Cacucci, pp. 117-154.
- Altavilla A. M. Mazza A. 2008 *Sull'analisi dei pattern insediativi degli immigrati in un territorio urbano*, Rivista italiana di Economia, Demografia e Statistica, vol. LXII, 2008.
- Carrington W.J., Troske K.R., 1997. *On Measuring Segregation in Samples with Small Units*, in Journal of Business & Economic Statistics, Vol. 15, No. 4, pp. 402-409.
- Duncan, O. and Duncan, B., 1955, A Methodological Analysis of Segregation Indexes, American Sociological Review, 20, 210-217.
- Massey, D.S. and Denton, N.A. 1988. *The Dimensions of Residential Segregation*, Social Forces, 67, 281-315.
- Ransom, M. R. 2000. *Sampling Distributions of Segregation Indexes*, Sociological Methods & Research, 28, 454-475.

SUMMARY

Duncan and Duncan dissimilarity index is one of the most commonly used measures of segregation. It is well known that this index reflects random allocation as well as systematic group segregation; this leads to potentially large values of the segregation index when unit sizes or minority proportions are small, even if there is little or no systematic segregation. To overcome this bias issue a few methods have been proposed. In this paper it is presented a new bootstrap correction and Monte Carlo simulations prove it to perform better than the existing ones. This new correction is illustrated by comparing dissimilarity levels across the settlement patterns of foreign immigrants in the city of Catania.

Anna Maria ALTAVILLA, Ordinario di Demografia, Università di Catania.

Angelo MAZZA, Ricercatore di Statistica, Università di Catania.

Antonio PUNZO, Assegnista di Ricerca, Università di Catania.

ECOTURISMO: CARATTERISTICHE E POTENZIALITA' DI SVILUPPO IN ITALIA

Francesco Antonio Anselmi

1. Introduzione

L'obiettivo di questa ricerca è di esplorare le metodologie che definiscono il fenomeno dell'ecoturismo, in forte espansione negli ultimi venticinque anni nel mondo e il suo sviluppo attuale e potenziale in Italia dal punto di vista della domanda e dell'offerta. Nella prima parte, infatti, si analizza l'ideologia dell'ecoturismo, che propugna uno stile di vita che rimanda a due correnti storiche del turismo esperenziale ed esistenziale; l'importanza del ritorno alla natura e all'autenticità dell'esistenza umana. Nella seconda parte della ricerca si analizza il mercato dell'ecoturismo in Italia, L'ecoturismo in Italia è rilevante con la capacità di crescita annua che arriva al 20%, mentre la crescita annua del turismo tradizionale è pari al 4,3%.

In sintesi, questo contributo vuole: a) diffondere i valori dell'ecoturismo, dando un valore economico riconoscibile alla natura, senza distruggerla; b) sviluppare un turismo che rinforzi la protezione della natura e della cultura, che crei meno impatti ambientali; c) dare delle possibilità di raffinare il turismo di natura tradizionale; d) indirizzare le autorità locali, gli operatori privati di questo settore e sottolineare le problematiche più emergenti nel mercato italiano rispetto agli altri mercati.

2. Ideologie di Ecoturismo

La crescita vertiginosa del turismo in generale ha, da qualche tempo, aperto un dibattito per un processo di revisione dello sviluppo turistico basato sui principi della sostenibilità. È soprattutto questo il senso delle Nazioni Unite che hanno proclamato il 2002 Anno Internazionale dell'Ecoturismo nel Quebec (Canada).

La dichiarazione di Quebec contiene un ampio scacchiere di interessi per una tematica di crescente importanza non solo come settore di notevole potenzialità per lo sviluppo economico, ma anche come potente strumento per la conservazione dell'ambiente naturale a condizione che esso venga adeguatamente gestito.

L'ecoturismo in questa prospettiva è caratterizzato da alcuni aspetti peculiari:

- 1) Mira alla promozione di uno sviluppo sostenibile del settore turistico.
- 2) Non determina il degrado o l'esaurimento delle risorse.
- 3) Concentra l'attenzione sul valore intrinseco delle risorse naturali rispondendo ad una filosofia più biocentrica che antropocentrica.
- 4) Richiede all'ecoturista di accettare l'ambiente nella sua realtà senza pretendere di modificarlo o adattarlo a sua convenienza.
- 5) Si fonda sull'incontro diretto con l'ambiente e si ispira ad una dimensione cognitiva diretta.

L'ecoturismo pone l'accento sulle esperienze nelle aree naturalistiche, favorendo preservazione, apprezzamento e comprensione ambientale e culturale.

Una delle definizioni di ecoturismo maggiormente condivise è quella dell'International Ecotourism Society (TIES, 1990) che recita: *"Ecotourism is responsible travel to natural areas that conserves the environmental and improves the well-being of local people"* (ecoturismo è un modo responsabile di viaggiare in aree naturali, conservando l'ambiente e migliorando il benessere delle popolazioni locali).

In questa definizione, l'ecoturismo ha una forte componente programmatica e descrive anche un insieme di risultati lungamente desiderabili, riguardanti:

- a) la compatibilità ambientale e socio-culturale come condizione fondamentale; b) l'apporto di benefici per i progetti di protezione dell'ambiente e per la popolazione locale (partecipazione, creazione e ampia distribuzione di reddito); c) l'accrescimento della consapevolezza ambientale e maggiore accettazione della conservazione della natura come uso del territorio proficuo e adeguato.

L'associazione Ecoturismo Italia, referente italiano dell'International Ecotourism Society, propone una sua definizione nei termini seguenti: un modo di viaggiare responsabile in aree naturali, conservando l'ambiente in cui la comunità locale ospitante è direttamente coinvolta nel suo sviluppo e nella sua gestione, ed in cui la maggior parte dei benefici restano alla comunità stessa

Le varie definizioni di ecoturismo vogliono aprire a nuove prospettive di vita, che si misurano con visibili comportamenti dei soggetti, con un insieme di pratiche continuative, basate su quattro principi: 1) minimo impatto ambientale; 2) massimo rispetto per le culture ospitanti; 3) sostanziali benefici socioeconomici alle popolazioni locali; 4) massimo beneficio per il tempo libero degli ospiti.

Al presente, il mercato dell'ecoturismo mostra i seguenti elementi:

- 1) Un segmento in rapida crescita. L'Organizzazione mondiale del turismo (OMT) stima che l'ecoturismo rappresenta il 2-4% dell'intero mercato turistico

mondiale e con un potenziale di crescita annua intorno al 20% (The International Ecotourism Society, www.ecotourism.org).

2) In molti dei Paesi in via di sviluppo con una consolidata industria ecoturistica, il segmento ha visto una forte crescita del numero di visitatori con un aumento dei ricavi generati da questi ultimi.

3) Molti Paesi in via di sviluppo, con ecosistemi tropicali vergini, spiagge incontaminate, resti archeologici, vogliono potenziare questo mercato di nicchia e molti Paesi stanno investendo nell'ecoturismo.

4) Molte comunità locali hanno acquisito la consapevolezza dei benefici economici dell'ecoturismo e sono motivate a proteggere le loro risorse naturali ed a adottare i dovuti comportamenti [(vale il motto: Wildlife pays so wildlife stays: (se paghi la natura questa rimane)].

5) Costituisce un valido elemento di spinta alla formazione e tutela ambientale ed interpretazione della natura, che tuttavia necessita di fondi pubblici e privati.

Una delle peculiarità dell'ecoturismo è di essere capace di lasciare nell'economie locali fino al 95% della spesa turistica sostenuta, al contrario di quello che succede nelle altre tipologie di turismo organizzato nelle quali l'80% della spesa sostenuta dal turista si distribuisce tra compagnie aeree, hotel ed altri operatori turistici (M. Honey, TIES, 2004).

3. Ecoturismo in Italia

In base alle definizioni sopra specificate l'ecoturismo in Italia è: a) turismo natura, ossia *osservazione, apprezzamento e conservazione della natura*; b) turismo di *protezione della natura e benessere delle popolazioni locali*.

L'ecoturismo come turismo natura ha come motivazioni principali di considerare la vacanza l'occasione per osservare e godere della natura e della cultura tradizionale, cioè per sintonizzarsi con la natura e la comunità locale.

Nell'ecoturismo come seconda impostazione c'è una mentalità, un atteggiamento, un agire degli operatori e fruitori di rispetto della natura quale condizione necessaria per assicurare nel tempo la redditività delle attività turistiche e la comunità di turismo stesso.

In Italia prevale la prima impostazione, cioè l'ecoturismo è inteso come turismo natura e rappresenta circa il 20% del mercato turistico, ma è rilevante la capacità di crescita annua che arriva al 20%, mentre la crescita annua del turismo tradizionale è pari al 4,3%. In Italia si rilevano un gran numero di piccoli operatori (associazioni, agenzie, ONG), che si stanno impegnando nell'ecoturismo sia come promotori di outgoing che di incoming, in particolare nei parchi, nell'agriturismo (5° Rapporto Ecotur, 2007).

3.1 *La domanda di ecoturismo (turismo natura)*

I segmenti più rappresentativi dell'ecoturismo sulla base della analisi fatte dai tour operator italiani che operano su questa tipologia di turismo rilevano una assoluta univocità nell'individuare "Parchi Aree Protette". Gli altri segmenti sono l'agriturismo, il turismo rurale, montagna, mare e riserve marine e turismo lacuale

Per quanto concerne l'agriturismo, esso non è solo una forma di ricettività: rappresenta il bisogno di evasione dalla vita stressante della città e dei luoghi massificati e standardizzati della vacanza, con un contatto con la natura.

Per quanto concerne le attività preferite dall'ecoturista, soprattutto italiano, in vacanza nei parchi e nelle aree protette, si rilevano le attività sportive per il 54%, tra le quali l'escursionismo che si pratica in montagna, il trekking, il mountain-bike, lo sci di fondo, l'equitazione, la canoa, e il climbing. Successivamente si trova il relax con il 24,3%. Inoltre, c'è la riscoperta delle antiche tradizioni per il 12% e a seguire i prodotti tipici della gastronomia per il 10%. Infine, si rileva l'interesse per la storia, la cultura, i musei naturalistici, ecc. compresi nella categoria "altri" che fanno registrare un valore del 4,4% nel 2006 (5° Rapporto Ecotour, 2007).

In base all'indagine effettuata presso i tour operator nazionali che vendono nei loro cataloghi il turismo natura, si è potuto rilevare negli anni 2005-2006 che i "giovani" sono quelli che fanno registrare il maggiore indice di preferenza (39,9%), seguiti dalle "famiglie" (37,4%), e dagli "over65", scesi dal 24,7% al 22,7%, confermando ancora una volta che il turismo natura è ricercato da tutti.

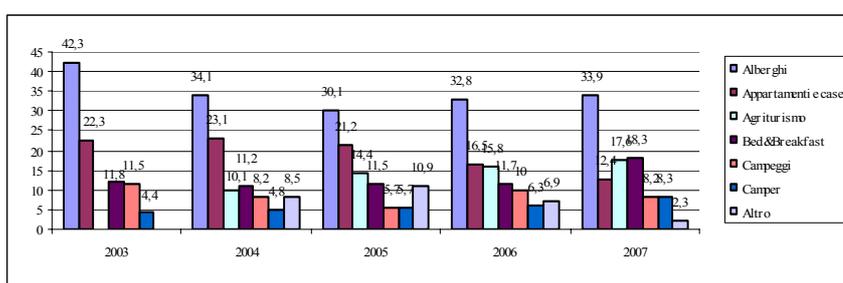
Per quanto attiene al livello di istruzione posseduto, i fruitori del turismo natura hanno un diploma di scuola media superiore nel 46% dei casi, una laurea nel 37,2% e la licenza elementare o media inferiore nel 16,8%. Percentuali che segnano un leggero cambio di indirizzo: diminuiscono i laureati ed aumentano coloro che hanno un grado d'istruzione più basso, sfatando la convinzione che il turismo natura sia una tipologia di élite.

3.2 *Strutture ricettive utilizzate dagli ecoturisti, provenienza e permanenza*

Interessanti sono i dati relativi alle strutture ricettive utilizzate dal turista-natura. Gli alberghi e le pensioni rappresentano, nel 2007, la soluzione ricettiva preferita con il 33,9%, in leggero aumento rispetto al 2006, seguita dai Bed & Breakfast che registrano una crescita notevole rispetto ai valori degli anni precedenti raggiungendo il 18,3%. Anche gli agriturismi segnano una crescita continua raggiungendo il 17,6%. A notevole distanza seguono gli appartamenti/case private, 12,4%, in forte diminuzione rispetto al 2005 (21,6%).i

B&B, il cui valore è standardizzato intorno all' 11-12%, Infine, seguono i campeggi, col 9,2%, nel 2007, mantenendosi costante col 2006, i camper col 6,3% e la voce "altro" il cui valore nel 2007 è stato del 2,3%, livello molto più basso rispetto al 2006, pari all'8%.(comprende foresterie, rifugi, ostelli, convitti, centri didattici, ecc) (Figura 1).

Figura 1 — Strutture ricettive utilizzate dagli ecoturisti dal 2003 al 2007 (Valori percentuali).



Per quanto concerne la provenienza degli ecoturisti, l'indagine evidenzia che gli ecoturisti provengono in misura predominante dalla regione col 46,1%, nel 2007; molti sono di provenienza nazionale, in misura del 36,1%, di provenienza europea sono circa il 13,8%, e del resto del mondo, nel 4%. (6° Rapporto Turismo Natura).

Per quanto riguarda la permanenza dell'ecoturista nei parchi e nelle aree protette si rilevano le seguenti caratteristiche. Nelle aree protette vige il fenomeno dell'escursionismo, che è la categoria più frequente col 40,1% nel 2007, segue quella con "1 giorno" col 12,1%, poi quella degli weekend con il 17%, ed ancora quella dei "3 giorni" col 9,6% e da "1 settimana" col 13,8%, e infine, quella con "più di una settimana" col 7,3%. Tale diversità di risultati deriva dal fatto che i turisti dei parchi nazionali provengono da molto lontano rispetto ai turisti dei parchi regionali e delle altre aree protette; spesso da Paesi europei e da Paesi extraeuropei i quali soggiornano per un tempo più lungo.

3.3 Presenze nelle strutture ricettive e composizione del fatturato dell'ecoturismo

L'ecoturismo italiano (turismo natura) si svolge fundamentalmente nei parchi nazionali e regionali, e nelle aree protette. Le presenze turistiche nelle strutture ricettive ufficiali dei parchi nazionali e regionali sono rilevate dall'ISTAT, mentre nelle aree protette diverse dai parchi le presenze non sono rilevate dall'ISTAT e

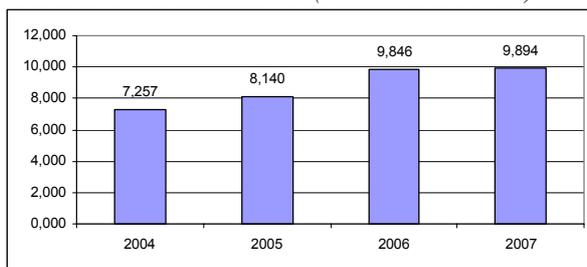
vengono stimate con opportune procedure. Le presenze negli esercizi alberghieri ammontavano nel 2007 a 59.593.788 e quelle degli esercizi complementari a 27.682.966, che determinano 87.276.754 (con una permanenza media di 4,48 giorni). Aggiungendo una stima delle presenze nelle strutture ricettive ufficiali nelle aree protette diverse dai parchi nazionali e regionali quantificate in 8.640.399 presenze nel 2007, si determinano 95.917.153 di presenze totali negli esercizi ricettivi ufficiali¹ delle aree protette in generale, che si possono suddividere, dopo averle ponderate sulla base delle presenze ufficiali registrate nei parchi nazionali e regionali, in 59.756386 di presenze italiane (62,3%) e 36.160.767 di presenze straniere (37,7%).

Per un dimensionamento quantitativo del fatturato dell'ecoturismo italiano (turismo natura) si considera la spesa giornaliera procapite domestica che per il 2007 è di 64,3 euro (il valore è stato determinato sulla base di un'indagine svolta dalla SWG nel 2005, rivalutando opportunamente la spesa originaria nel tempo) e per quello straniero la spesa giornaliera considerata è di 88,61 euro (stime UIC).

Il fatturato per la componente nazionale nelle strutture ricettive ammonta a 3,842 miliardi nel 2007, mentre il fatturato per la componente estera nella strutture ricettive ufficiali è di 3,204 miliardi di euro. Il fatturato nel comparto dell'agriturismo ammonta a 831 milioni di euro; il fatturato dei camperisti ammonta a 667 milioni. Infine, occorre aggiungere il fatturato delle presenze italiane pari a 12.905933 e straniere 7.116.614 negli esercizi non classificati ufficialmente, sulla base di un parametro 19,16% delle presenze negli esercizi ricettivi ufficiali (6° Rapporto Ecotur 2008) che ammonta a 1,350 miliardi di euro.

Le presenze sopra specificate sulla base dei calcoli precedenti hanno generato un fatturato globale di 9,894 miliardi di euro, 2006 (figura 3).

Figura 2 – Fatturato dell'ecoturismo in Italia (in miliardi di euro).



4. L'offerta di ecoturismo in Italia: un'analisi SWOT

L'ecoturismo (turismo natura) in Italia è un segmento in forte evoluzione ed espansione e rappresenta per il turismo una importante occasione di sviluppo.

Un'analisi SWOT realizzata sui maggiori mercati dell'incoming internazionali, sottolinea per l'Italia punti di forza e di opportunità ma anche problematiche che sono rappresentate da punti di debolezza e minacce sintetizzati nello schema seguente (ENIT – Agenzia nazionale del turismo):

Figura 3 – Analisi Swot dell'ecoturismo (turismo natura) in Italia.

Punti di forza	Punti di debolezza
Patrimonio naturalistico diversificato ed unico	Mancanza di promozione unitaria e di commercializzazione
Turismo natura integrabile con altri turismi e in particolare con quello culturale	Carenza di materiale informativo e di guide specifiche
Clima mite o piacevole	Prezzi poco competitivi e rapporto qualità/prezzo insoddisfacente
Crescente interesse per l'ambiente e il contatto con la natura	Carenza di pacchetti All Inclusive
Abbinamento con l'enogastronomia	Difficoltà nella mobilità sul territorio
Facilità di raggiungere l'Italia da parte dei Paesi europei limitrofi	
Interesse crescente per gli operatori per il turismo natura (ecoturismo)	
Opportunità	Minacce
Crescente interesse per l'ambiente e il contatto con la natura	Offerte a prezzi vantaggiosi di altre destinazioni
Una ricerca crescente del benessere	Crescente interesse verso destinaz. dell'Est europeo ed extra europee (Corea, Cina)
Voli low cost verso località minori	Uso di internet crescente per molte offerte della concorrenza e commercializzazione
Siti internet di promozione	Notizie negative relative ad ambiente ed ecologia in Italia
Espansione del turismo della terza età	
Possibilità di sfruttare i parchi marini protetti	
Crescita della richiesta di strutture ricettive alternative (agriturismo, B&B)	
Interesse crescente per gli operatori per il turismo natura (ecoturismo)	

5. Conclusioni

L'ecoturismo si caratterizza generalmente per l'elevata propensione dei soggetti (turisti) attenti all'ambiente, che amano l'aria aperta e praticano vacanza attiva.

L'ecoturista psicologicamente, avverte l'esigenza di staccarsi radicalmente dalla società urbanizzata e da ritmi di vita frenetici e ripetitivi ed è spinto dal desiderio di scoprire luoghi incontaminati dalla presenza dell'uomo della modernità, per scoprire se stesso e rigenerarsi.

Con riferimento all'ecoturismo in Italia, nel forma prevalente del turismo natura si riscontra una crescita del segmento più elevata rispetto a tipologie di turismo consolidate (sole, mare, itinerari culturali, ecc.) con una serie molto elevata di opportunità, tuttavia si rilevano dei punti di debolezza e minacce, per i quali occorre intervenire con tempestività per un'offerta fruibile, che faccia sistema per raggiungere in una forma organica il mercato con modalità chiare e trasparenti altrimenti i flussi turistici si indirizzeranno verso destinazioni che in misura maggiore avranno soddisfatto tali condizioni.

Riferimenti bibliografici

Boccella N. Paziienza P. Rinaldi A. 2008. *Economia e marketing del turismo*, McGraw-Hill, Milano.

Contino A. 2006. *Opportunità di turismo nel territorio* in Nuove Ipotesi Dall'Ara G.

Morandi E. 2006. *I sistemi turistici locali. Normativa, progetti e opportunità*, Halley Editrice, Milano.

- Gentili G. 2008. *L'Italia "verde" sui mercati internazionali*, 6° Rapporto Ecotur sul Turismo Natura, Edizione Agricole de Il Sole 24ORE, Bologna.
- Granata S. 2007. 2007. *Economia del turismo sostenibile. Analisi teorica e casi studio*. Economia dei Servizi, n° 3, settembre- dicembre.
- Honey M. TIES (2004), *Ecotourism: A tool for Sustainable Development*", Workshop on Ecotourism , Colombo (Sri Lanka) 22 ottobre.
- Osservatorio Permanente sul Turismo Natura (a cura) (2007), *5° Rapporto Ecotur sul Turismo Natura*, Edizione Agricole de Il Sole 24ORE, Bologna.
- Osservatorio Permanente sul Turismo Natura (a cura) (2008), *6° Rapporto Ecotur sul Turismo Natura*, Edizione Agricole de Il Sole 24ORE, Bologna.
- Leoni L. E. Morelli P. 2006. *L'offerta ricettiva nei Parchi nazionali e regionali in Italia*, Quarto Rapporto sul Turismo Natura, ECOTUR, Bologna.

SUMMARY

Ecotourism. Characteristics and his potentiality of development in Italy

This research focus attention, at first, on the ideology of ecotourism, that has became one of the fastest growing sectors of the tourism industry, with a growing annually by 10-15% worldwide.

The ideology of ecotourism advocates a style of life that postpones to two current historians of experiential and existential tourism; the importance of the return to the nature, that gives us insight into our impacts as human beings. In the second part of the search the ecotourism market is analyzed in Italy. in the prevailing form of tourism nature. It finds a most elevated growth of the segment in comparison to consolidated typologies of tourism (sun, sea, cultural itineraries, etc.). Besides it is underlined, that the possibilities of growing of the Italian ecotouristic offer are very elevated and with a series of opportunity, nevertheless some points of weakness and threats have been perceived. Insofar it needs to act with timeliness for an enjoyable offer, that makes system to reach in an organic form the market with clear formality and transparency, otherwise the tourist flows will be addressed toward destinations.

UNA NUOVA POLITICA REGIONALE DELL'OCSE E DELL'UNIONE EUROPEA

Francesco Antonio Anselmi

1. Introduzione

Negli anni settanta nasce la politica di sviluppo regionale con l'intento di correggere le disparità economiche (tassi di crescita del Pil e del reddito disequilibrati), tecniche, di redistribuzione delle opportunità tra regioni e di insufficiente dotazione di risorse (umane, finanziarie e infrastrutturali). Essa è basata su massicce assistenze finanziarie (per infrastrutture e per sviluppare i servizi pubblici) alle regioni più povere, alla creazione artificiale di poli tecnologici ed economici di sviluppo a livello regionale, e alle istanze di mantenere in vita settori industriali in declino per proteggere l'occupazione locale.

Questo tipo di interventi è proseguito fino alla fine degli anni ottanta, rilevando risultati generalmente fallimentari e molti Paesi dell'OCSE ed europei si sono interrogati sull'utilità di continuare ad adottarli; tra l'altro emergevano nuovi problemi non solo economici (emarginazione, crescita della povertà assoluta e relativa) che toccavano pure le regioni centrali e i centri urbani. In considerazione di ciò, sia l'OCSE che l'Europa si sono progressivamente orientate verso un nuovo paradigma, anche alla luce del fenomeno della globalizzazione, che poggia su cinque pilastri, che consistono di politiche regionali indirizzate a tutte le regioni (dalle più povere alle più ricche), e volte a mobilitare tutti mezzi per massimizzare le opportunità di sviluppo endogeno, attraverso la formazione del capitale territoriale che permette di sviluppare un moltiplicatore dell'investimento nell'ambito di azioni dello Stato centrale volte a vegliare al mantenimento della qualità delle infrastrutture, dei servizi pubblici; a garantire un ambiente favorevole alle sviluppo delle aziende, ad interventi di diffusione delle conoscenze (istruzione, formazione), delle tecnologie e dell'innovazione; alle misure che mirano allo sviluppo del capitale sociale, alla protezione dell'ambiente,

L'obiettivo di questa ricerca è di analizzare degli aspetti di questi nuovi pilastri (orientamenti) volti al superamento delle disparità territoriali (economiche, tecniche e infrastrutturali) e a garantire la sostenibilità dello sviluppo, promuovendo metodi e pratiche nuovi nell'ambito di nuove tecniche strategiche di nuove tematiche strategiche orientate alla coesione e alla competitività territoriale regionale, attraverso un processo di integrazione di attività economiche, sociali, culturali ed ambientali.

2. Caratteri e risultati di un ventennio di politiche regionali

Le politiche di sviluppo economico regionale negli anni settanta sono poste in essere in termini di correzione delle disparità attraverso misure redistributive (attraverso sovvenzioni particolari, attraverso misure atte a promuovere la localizzazione industriale in zone svantaggiate o economicamente minacciate ed era quindi vissuto come un problema di redistribuzione delle opportunità tra regioni ricche e regioni “economicamente sottosviluppate” e come problema di insufficiente dotazione di risorse (umane, finanziarie, tecniche e infrastrutturali).

Le politiche regionali di quel periodo si possono riassumere, come si è già specificato nell’introduzione, nei seguenti interventi:

- Massicce assistenze finanziarie (per infrastrutture e per sviluppare i servizi pubblici) alle regioni più povere;
- Creazione artificiale di poli economici di sviluppo a livello regionale;
- Creazione ex-nihilo di tecnopoli;
- Tentativi di mantenere in vita settori industriali in declino per proteggere l’impiego locale.

Dopo un ventennio di interventi i risultati raggiunti, come si è già detto prima, sono stati fallimentari. Infatti, la valutazione dei risultati può essere riassunta nei seguenti termini:

a) le massicce azioni finanziarie per costruire le infrastrutture e sviluppare servizi pubblici nelle regioni più povere hanno soprattutto determinato distorsioni dei mercati e una cultura della dipendenza, che ha danneggiato lo sviluppo delle stesse regioni;

b) la creazione artificiale di poli di sviluppo economico delle regioni ha condotto a progetti molto costosi, di cui molto non hanno avuto un futuro;

c) la creazione dal nulla di poli tecnologici non è riuscita, non avendo avuto come interlocutore credibile la regione dove l’investimento è stato realizzato;

d) gli interventi in difesa dei settori industriali in declino per proteggere l’occupazione locale, nella generalità dei casi ha condotto a spese inutili.

Agli inizi degli anni '90, accanto a questi insuccessi si è cominciato ad osservare che il problema regionale doveva essere risolto considerando più angolazioni in relazione al fatto che all’insuccesso degli interventi delle regioni periferiche si rilevava la perdita di competitività delle regioni centrali e dei centri urbani nei quali si concentravano una serie di problemi importanti di natura sociale, emarginazione, crescita della povertà relativa che assoluta, attività delinquenziale e molti Paesi a livello dell’OCSE ed europeo furono progressivamente orientati, anche alla luce del fenomeno della globalizzazione, verso un nuovo paradigma, che poggia su cinque pilastri, che sono illustrati nel paragrafo che segue.

3. I nuovi pilastri della politica regionale nell'ambito dell'OCSE

Gli insuccessi della politica di riequilibrio regionale e i grandi cambiamenti prodotti dalla globalizzazione hanno fatto perdere competitività anche alle regioni più ricche e ai centri urbani e hanno spinto l'OCSE a promuovere nuove strategie di posizionamento competitivo di regioni e di nazioni orientate verso un nuovo paradigma che poggia su cinque pilastri:

- 1) Politiche regionali indirizzate a tutte le regioni, dalle più povere alle più ricche.
- 2) Politiche volte a mobilitare tutti mezzi per mettere in grado di massimizzare le opportunità di sviluppo endogeno di tutte le regioni, attraverso la formazione del capitale territoriale.
- 3) Azioni dello Stato centrale volte a vegliare al mantenimento della qualità delle infrastrutture, dei servizi pubblici e a garantire un ambiente favorevole allo sviluppo delle aziende.
- 4) Interventi volti alla diffusione delle conoscenze (istruzione, ecc.), delle tecnologie dell'innovazione, alle misure che mirino allo sviluppo del capitale sociale, alla protezione dell'ambiente e alla creazione di un quadro di qualità della vita.
- 5) Una politica territoriale esercitata a livello nazionale, che deve essere compatibile con le politiche poste in essere nelle regioni e nelle città, con processi di decentralizzazione di responsabilità supportate da risorse fiscali che ne permettano la realizzazione.

Si tratta di porre azioni che mobilitino tutti i mezzi affinché le regioni siano in grado di massimizzare le loro opportunità di sviluppo endogeno. Gli investimenti vengono attivati in una regione fondamentalmente affinché si costituisca un capitale territoriale, che permetta di sviluppare un moltiplicatore degli investimenti.

Un ruolo importante deve essere svolto dallo Stato centrale che deve vegliare nelle varie regioni, affinché ci sia un mantenimento della qualità delle infrastrutture e dei servizi pubblici e lo sviluppo di mezzi moderni di trasporto e di comunicazione che garantiscano a tutte le regioni un livello di accessibilità.

Un ruolo importante deve essere offerto dalle infrastrutture di natura immateriale. Si tratta di azioni capaci di rinforzare i vantaggi competitivi delle regioni attraverso la diffusione delle conoscenze, delle tecnologie e delle innovazioni; tutte le misure che mirano allo sviluppo del capitale sociale; le misure volte a garantire la pace, la coesione sociale.

Le azioni strategiche dello sviluppo regionale richiedono la valorizzazione e creazione di risorse specifiche a livello locale idonee a competere con l'esterno, col globale e si delineano misure di politica economica e regionale centrate sul comportamento strategico degli attori sulla costituzione di *milieux innovateurs*, che sappiano creare risorse (umane, finanziarie e tecniche, uniche in cui

riconoscersi e farsi riconoscere) (*Unique Selling Propositions*) di sviluppo, che hanno alla base la competitività territoriale che presuppone: le risorse del territorio; il coinvolgimento dei vari soggetti delle istituzioni; l'integrazione dei settori di attività in un'ottica di innovazione; la cooperazione con gli altri settori e l'articolazione con le politiche regionali, nazionali ed europee e il contesto globale.

La competitività territoriale si compone di quattro dimensioni, che si combinano in modo specifico. Esse sono:

1) La competitività sociale: attiene alla capacità dei soggetti (operatori e istituzioni a vari livelli) di concertare efficacemente una visione del futuro. Essa è fiducia reciproca, volontà, e capacità che strutturano in modo articolato interessi individuali e collettivi.

2) La competitività ambientale indica la capacità dei soggetti di valorizzare l'ambiente come elemento distintivo del loro territorio (patrimonio da tutelare). L'ambiente comprende la vita della popolazione sul territorio e quindi anche il patrimonio edificato.

3) La competitività economica: attiene alla produzione e al mantenimento all'interno del territorio il massimo valore aggiunto, consolidando i punti di contatto tra i vari settori e combinando efficacemente le risorse per valorizzare le specificità locali.

4) Il posizionamento rispetto al contesto globale. I vari soggetti devono trovare una propria posizione rispetto agli altri territori per realizzare appieno il loro progetto territoriale in simbiosi con la globalizzazione.

In tale ambito il capitale territoriale (risorse fisiche, risorse umane, culturali, identità, know-how, risorse finanziarie, imprese, istituzioni pubbliche, mercati e relazioni) merita un approfondimento.

4. Il capitale territoriale

Il capitale territoriale può essere definito una ricchezza che integra dimensioni geografiche, economiche, sociali, culturali, politiche, ecc. Non è un'entità, né universale, né immutabile.

Il capitale territoriale si basa su un approccio globale con concertazione del territorio e sulla ricerca di soluzioni per integrare le risorse locali (endogene).

L'approccio globale deve considerare la realtà del territorio nelle sue diverse componenti: ambientali, economiche, sociali, culturali, politiche, al fine di esplorare tutte le risorse esistenti e potenziali del territorio.

L'approccio territoriale spinge gli operatori pubblici e privati, organizzati nell'ambito di una partnership locale a ideare un programma di sviluppo per il

proprio territorio, definito di concerto con la popolazione e negoziato globalmente con i poteri pubblici, regionali o nazionali, competenti sul territorio.

L'approccio territoriale propone, inoltre, un nuovo concetto del rapporto della popolazione rispetto al progetto di sviluppo favorendo il coinvolgimento, la concertazione.

La scelta dell'approccio territoriale è legata all'importanza delle risorse locali ("endogene") per il raggiungimento di uno sviluppo sostenibile.

Le risorse endogene possono essere fisiche, ambientali, culturali, umane, economiche e finanziarie, nonché istituzionali e amministrative. Esse devono essere affinate in base a una molteplicità di componenti che identificano il profilo del territorio. Esse sono:

1. Le risorse fisiche e la loro gestione [(capitale naturale (acqua, suolo, aria, biodiversità vegetale ed animale)] impianti e le infrastrutture, il patrimonio storico e architettonico, ecc.).

2. Il capitale umano che comprende le conoscenze e le competenze presenti in un territorio legate agli individui (istruzione, formazione, ricerca e sviluppo) ed anche le caratteristiche demografiche e la struttura sociale della popolazione.

3. Il capitale sociale identificato con le strutture relazionali in grado di agevolare la realizzazione di progetti collettivi, con l'abilità di fare le cose in maniera collettiva. È contenuto nell'abilità degli individui, gruppi, organizzazioni e istituzioni a lavorare insieme per raggiungere obiettivi e benefici comuni e nella fiducia e reciprocità che si sviluppano attraverso un processo di interazione continua.

4. Il know-how e le competenze (impliciti ed espliciti, la padronanza delle tecnologie, le capacità di risolvere i problemi con tempestività, ecc.).

5. Il capitale culturale e di identità del territorio è rappresentato dalle conoscenze, dalla condivisione di valori non solo economici dei soggetti che operano sul territorio (mentalità, attitudini, forme di riconoscimento, usi e costumi, ecc.).

6. Il capitale istituzionale, che attiene alle forme specifiche e modalità di interazione tra istituzioni e amministrazioni locali, alle regole dell'azione politica, al grado di autonomia nella gestione dello sviluppo, incluse le risorse finanziarie, alla fiducia nelle istituzioni, e alle forme di partecipazione degli attori.

7. L'immagine e la percezione del territorio (sia tra gli stessi abitanti, sia all'esterno), la comunicazione sul territorio.

8. La posizione del territorio nei diversi mercati, i contatti con altre aree, le reti di scambio, ecc.).

Tutte queste componenti costituiscono il capitale territoriale, che non è una nozione statica, ma dinamica e rappresenta il complesso degli elementi materiali ed

immateriali a disposizione del territorio. Essi possono costituire punti di forza o di debolezza a seconda degli aspetti considerati.

Il territorio è un'entità dinamica, dalle molteplici sfaccettature, che si evolve nel tempo, essendo il risultato di un legame tra passato, presente e futuro. Esso si arricchisce confrontando il suo stato attuale con quanto avviene altrove, analizzando le esperienze positive e gli insuccessi, proiettandosi nel futuro in funzione di un'analisi e di una volontà condivisa dagli operatori locali.

Il capitale territoriale può pertanto essere raffigurato come una sfera posta nel punto di intersezione di due assi (un asse passato-futuro e un asse interno-esterno).

Questa sfera è in continua evoluzione, si arricchisce, si definisce in modo sempre più preciso con ad elementi attinti dal passato (la storia), dal futuro (il progetto), dalle componenti interne del territorio (interazioni tra i vari soggetti, istituzioni, reti locali) e dalle sue relazioni con il mondo esterno (scambi con i mercati, le istituzioni e le reti esterne).

5. Novità della politica regionale a livello europeo

Dopo aver illustrato le caratteristiche del capitale territoriale nell'ambito della nuova politica di sviluppo dell'OCSE, è opportuno soffermarsi sulla politica regionale europea, che presenta qualche novità rispetto alla nuova politica regionale dell'OCSE. La nuova politica regionale europea ricerca la competitività nell'innovazione, che richiede il passaggio verso una società basata sulla conoscenza, migliorando le politiche in materia di società dell'informazione e di R&S (Consiglio europeo di Lisbona, 2001).

In tale contesto le azioni innovative di tale politica europea devono promuovere metodi e pratiche nuovi nell'ambito delle seguenti tematiche strategiche:

- a) Un'economia regionale fondata sulla conoscenza e sull'innovazione tecnologica. Essa basa il vantaggio concorrenziale sull'innovazione piuttosto che su una riduzione di costi, che può annullarsi facilmente in un'economia della globalizzazione; conseguentemente essa cerca di costruire un quadro istituzionale regionale che rafforzi il capitale umano, la formazione e l'apprendimento permanente, che diventano fattori determinanti per la creazione e la diffusione delle conoscenze nel tessuto produttivo fonte principale di innovazione e di vantaggio competitivo regionale se si saprà costruire un ambiente istituzionale regionale, che promuova la creazione, la divulgazione e l'integrazione delle conoscenze nel tessuto produttivo, i nessi tra le imprese, con le università, i mercati finanziari e altri partner tecnologici.
- b) Una società dell'informazione al servizio dello sviluppo regionale. La società dell'informazione può contribuire a far uscire le regioni dell'Unione

Europea svantaggiate dal loro isolamento (geografico ed economico) disponendo di un'enorme capacità di collegamento a rete per consentire di valorizzare i punti di forza, per aumentare la competitività, in particolare, delle PMI.

c) Una coesione e una competitività regionale da realizzare mediante l'acquisizione integrata delle attività economiche, ambientali, sociali e culturali (identità regionale e sviluppo sostenibile). Le regioni devono saper migliorare le condizioni di vita e di lavoro dei loro abitanti attraverso i loro punti di forza in termini competitivi e di sostenibilità che possono essere determinati da competenze specifiche, da una ubicazione strategica, da una situazione ambientale eccezionale, da infrastrutture culturali, da tradizioni (prodotti regionali, gastronomia, ecc.).

In conclusione, una nuova politica regionale europea deve rispondere efficacemente alle nuove sfide del futuro (globalizzazione, mutamenti tecnologici, coesione sociale) e valorizzare il capitale territoriale (l'identità regionale), che ha alla base il capitale materiale e umano della regione.

6. Conclusioni

L'obiettivo di questa ricerca ha riguardato la nuova politica regionale dei paesi dell'OCSE e dell'Europa volta a garantire la sostenibilità dello sviluppo, promuovendo nuovi pilastri (metodi e pratiche nuovi), nuove tecniche strategiche di nuove tematiche strategiche orientate alla coesione e alla competitività regionale, attraverso un processo di integrazione di attività economiche, sociali, culturali ed ambientali.

Questi pilastri fanno nascere nuovi concetti come il management territoriale e la competitività territoriale, che pongono l'accento sulla necessità per un territorio di sapere creare un insieme di risorse unico in cui riconoscersi e farsi riconoscere (creare delle *Unique Selling Propositions*).

Nell'economia del ventunesimo secolo a competere non è più solo l'impresa ma il Territorio che deve essere considerato *come un grande e complesso "Organismo"* composto da Attori e Attività il cui Valore Aggiunto è 1) il Valore Organizzato; 2) la capacità di fare sistema; 3) la coesione sociale; 4) la sicurezza a competere.

In conclusione la nuova politica regionale dell'OCSE, ed anche dei Paesi dell'UE, è di natura territoriale, che si deve fondare essenzialmente su tre principi: lo sviluppo endogeno: la messa in valore delle opportunità d'espansione proprie ai territori; lo sviluppo sostenibile: che concilia efficienza economica, coesione sociale e equilibrio ecologico; una governance più responsabile.

Riferimenti bibliografici

- Camagni R. 1991. *From the local milieu to innovation through cooperation networks*, in Camagni R. (ed). *Innovation Networks, Spatial Perspectives*, GREMI, Belhaven Press, p.1-9.
- Camagni R., Capello R. 2002. “*Milieux Innovateurs and Collective Learning: From Concepts to Measurement*”, in Acs Z., de Groot H. e Nijkamp P. (eds.), *The Emergence of the Knowledge Economy: A Regional Perspective*, Springer Verlag, Berlino, pp. 15-45
- Camagni R., Capello R. 2002. *Apprendimento collettivo e competitività territoriale*, Franco Angeli, Milano
- Camagni R., Maillat D. 2006, *Milieux innovateurs: théorie et politiques*, Economica, Antropos, Paris
- Camagni R. (2007), *Towards a Concept of Territorial Capital*, ERSA, Cergy-Paris.
- Camagni R., Maillat D., Matteaccioli A. 2004, *Ressources naturelles et culturelles, milieux et développement local*, Editions EDES, Neuchâtel
- Maskell P. (1999), *Social Capital, Innovation and Competitiveness*, in J. Stefen Baron Field and T. Schuller, eds., *Contribution to the “ Social Capital Collection”*, Oxford University Press, Oxford.
- <http://www.socialcapitalgateway.org/>

SUMMARY

A New regional policy in the OECD and European

In the seventies the policies of regional economic development were joined by the correction of economic (growth rates of GDP and income unbalanced) and technique differences and by inadequate allocations of human, financial, and infrastructural resources.

After twenty years of operations the results have been disastrous, and the OECD and EU have become oriented towards a new paradigm, even in light of the phenomenon of globalization, that rests on five pillars.

The objective of this research has been of analyzing aspects of these new pillars aimed at overcoming regional economic, technical and infrastructural disparities and to ensure sustainable development promoting new methods and practices of strategic techniques in innovations oriented in cohesion and in regional competitiveness, through a process of integration of economic, social, cultural and environmental activity.

Francesco Antonio ANSELMI, Professore Aggregato di Economia Applicata e di Economia dell'Industria Turistica – Facoltà di Economia – Università di Palermo.

LA QUALITÀ NEL SETTORE TURISTICO: UN'ANALISI EMPIRICA

Vincenzo Asero, Simona Gozzo¹

1. Introduzione

La qualità rappresenta una delle modalità adottate dal mercato per risolvere il problema dell'asimmetria di informazioni nelle relazioni che si instaurano tra venditori e compratori. Le imprese, infatti, investono in qualità utilizzandola come meccanismo di segnalazione per rivelare informazioni circa i beni/servizi offerti. I consumatori, invece, affidano al suo riconoscimento un particolare valore che, sebbene non quantificabile, è in grado di guidare e, in alcuni casi condizionare, le loro decisioni di acquisto. In generale, è possibile individuare almeno tre dimensioni della qualità: quella *certificata*, riconosciuta e documentata attraverso un sistema di marchi e protocolli; quella *attesa*, direttamente legata alle aspettative del consumatore; e quella *percepita*, individuabile tanto dal lato del consumo, come risultato dell'esperienza del consumatore, quanto da quello della produzione, in quanto valore attribuito dal produttore ai beni e servizi da lui offerti, in assenza di apposite certificazioni poste a tutela del mercato.

Anche in ambito turistico la qualità riveste particolare rilevanza, sebbene la sua definizione e misurazione presentino diversi problemi, sia a causa della matrice composita del prodotto turistico, per cui di volta in volta occorre definirne ambiti e criteri di identificazione, sia perché beni e servizi turistici sono prevalentemente *experience good* (Candela e Figini, 2003), per i quali la possibilità di verificarne gli attributi è costituita dal loro consumo, per cui anche la qualità può essere riconosciuta soltanto *ex post*, sulla base di una struttura delle preferenze che si rafforza e consolida nel tempo attraverso l'esperienza (Becker e Stigler, 1977).

¹ Il presente lavoro è stato svolto congiuntamente dagli autori. Più in particolare, i paragrafi 1, 2 e 6 sono stati redatti da Vincenzo Asero ed i paragrafi 3, 4 e 5 da Simona Gozzo.

Questo lavoro affronta il tema della qualità dei servizi offerti al turista attraverso i risultati di una ricerca empirica condotta in una delle province della Sicilia, quella di Enna². I servizi considerati sono quelli ricettivi e di ristorazione, poiché ad essi si indirizza una parte consistente della domanda turistica che si rivolge verso una destinazione. Per ciascuno di questi, l'analisi è stata effettuata sia tenendo conto dei sistemi di certificazione, sia ricorrendo alla costruzione di indicatori di qualità che si riferiscono a pratiche informali più che al rispetto di disposizioni formali, distinti in relazione alla tipologia di servizi e alla percezione del loro standard da parte degli operatori, posto che la qualità assume valenze diverse a secondo degli ambiti cui si riferisce.

2. Qualità e competitività

Nel mercato turistico la qualità è utilizzata con significati diversi, associati ad ambiti e contesti differenti. In ogni circostanza, però, l'attenzione costantemente rivolta alla soddisfazione dei turisti, sia per ciò che attiene lo standard qualitativo dei beni e servizi offerti dai singoli operatori, sia relativamente alla dimensione olistica del prodotto turistico, costituiscono importanti fattori di competitività e di immagine sul mercato di una data destinazione. Mantenere e migliorare la qualità dell'offerta è, infatti, fondamentale ai fini della difesa o dell'incremento della propria posizione di mercato. A tal fine, la qualità richiede un approccio integrato da parte di tutti gli *stakeholders*, che includa le politiche di offerta, la conoscenza dei bisogni e delle aspettative dei turisti, le valutazioni da questi espresse.

Sebbene la decisione delle imprese di offrire prodotti/servizi di qualità è legata alle reazioni del consumatore, il che costituisce per loro un incentivo sufficiente, queste sono libere di scegliere se investire o non investire in qualità. E se decidono di farlo, possono operare un'ulteriore scelta tra livelli diversi di qualità, preferendo produrre con bassa o con alta qualità sulla base di un confronto tra il maggior profitto derivante da un determinato livello di qualità ed i maggiori risparmi di costo associati a una bassa qualità. Nel mercato turistico questo comportamento è tipico delle imprese che vogliono conquistare o mantenere fette di mercato basando la loro competitività sul prezzo, ma, naturalmente, non è valido nel caso di imprese che preferiscono investire in qualità per ragioni di immagine della destinazione turistica in cui sono localizzate o per scelte di posizionamento (Asero, Gozzo e

² Indagine svolta nell'ambito del Progetto CONACT-Interreg III B Archimed. Il campione è stratificato in: territorialità; tipologia di struttura; tipologia e categoria di alloggio. Il disegno di campionamento ha previsto il coinvolgimento di 62 strutture, di cui è stato raggiungibile il 62%, mantenendo, comunque, le proporzioni adeguate.

Patti, 2008). E' evidente, dunque, che anche il contesto produttivo in cui operano le imprese gioca un ruolo importante nelle loro scelte di investimento in qualità.

La valutazione della qualità può essere effettuata facendo ricorso a metodologie differenti, tutte riconducibili allo studio delle relazioni tra le aspettative e la soddisfazione dei consumatori (Parasuraman, Zeithaml, e Berry, 1985). In ambito turistico, ciò implica che gli operatori puntano a fornire servizi di qualità e ad adottare strategie per il miglioramento dei loro standard sulla base di aspettative ed opinioni fornite dai turisti (Gustafsson e Johnson, 2003). Tuttavia, occorre ricordare che nel turismo i servizi sono suddivisi per categorie, in genere associati a standard qualitativi differenti. Ciò comporta che spesso è possibile riconoscere la qualità non solo in servizi di categoria superiore, ma anche in quelli di categorie più basse. Anzi, tale caratteristica appare sempre più evidente in questo settore, dove è in continua crescita la ricerca di esperienze di qualità da parte dei turisti, a prescindere dalle categorie di servizi domandati. Come l'evidenza empirica conferma, il turista potrebbe parimenti essere soddisfatto nella misura in cui valuti come adeguato il rapporto tra la qualità da lui percepita dei beni e servizi consumati e la spesa sostenuta per ottenerli, a prescindere dal livello del reddito impiegato, e viceversa non esserlo (Asero e Gozzo, 2007).

3. La qualità certificata ad Enna

Secondo l'approccio sociologico neoistituzionalista le organizzazioni operano normalmente in un tessuto strutturato e riconosciuto di vincoli e sostegni spesso sollecitati dalle stesse, che ne vedono un fattore di legittimazione al loro agire. Riconducendo la questione al piano in analisi, la diffusione o meno di un sistema di qualità certificato entro il territorio dipenderebbe dal costituirsi di un contesto istituzionale in cui l'azione della singola impresa è tenuta ad inserirsi, per la propria reputazione. Se il settore turistico è caratterizzato da un processo di isomorfismo, quindi, questo deriva prevalentemente dalla percezione che gli operatori hanno delle aspettative di qualità dei servizi da parte del turista e dalla rilevanza attribuita in tal senso al sistema di certificazione della qualità.

Nel caso di Enna, difficilmente è emersa la pratica di adozione di una certificazione di qualità, che perlomeno non è ritenuta "necessaria" quando si fa riferimento alla propria struttura, sebbene siano riconosciute le potenzialità dello strumento in relazione ad una migliore gestione complessiva dei servizi. Il riconoscimento di un potenziale aggiuntivo in termini di immagine è d'altra parte maggiore tra chi non utilizza la certificazione, ad indicare la percezione "positiva" dello strumento, sebbene non tanto in relazione alla sua "funzione manifesta" – riconducibile alla gestione della struttura – quanto piuttosto a quella "latente" (Merton, 2000), di strumento da utilizzare per "farsi pubblicità" ed "attrarre i clienti". Il fatto che le

potenzialità come meccanismo di segnalazione siano rilevate soprattutto da chi non si serve della certificazione indica che la stessa è comunque percepita positivamente, come uno strumento utile e diffuso tra le organizzazioni “di successo”, dinamica solitamente riconducibile all’emergere di processi di isomorfismo per adattamento nel lungo periodo, a prescindere dal settore preso in considerazione (Powell e DiMaggio, 2000).

Solo il 24% degli operatori presi in considerazione fanno riferimento ad un ente di certificazione per la qualità dei servizi ed il 75% di questi attribuisce ad essa una valenza soprattutto pratica, per i vantaggi che comporta in relazione al miglioramento nella gestione dei servizi. Coloro che non hanno adottato la certificazione sono, invece, più propensi ad attribuirvi una valenza simbolica per una maggiore capacità di attrazione dei clienti. Ne emerge, quindi, una differenza nella percezione della pratica e nella rilevanza attribuita alla certificazione di qualità. I soggetti che non aderiscono ad un sistema di certificazione della qualità investono, in effetti, maggiormente in pubblicità, adottando un numero maggiore di canali di comunicazione rispetto a chi investe in qualità. Si evince, inoltre, dall’analisi condotta un diverso modo di considerare la certificazione di qualità, riconducibile agli specifici servizi. La maggior parte degli operatori del settore alberghiero contattati non usufruisce di una certificazione di qualità, che tuttavia percepisce come strumento utile, di cui non ci si serve soprattutto per mancanza di informazioni su come procedere o per i costi elevati, più che per scarsa “motivazione” o disinteresse. Gli intervistati ritengono che l’impiego della certificazione non costituisca un appesantimento burocratico ed anzi permetta di migliorare la gestione della struttura e garantisca un vantaggio in termini di pubblicità e attrattiva. La categoria dei ristoratori è, di contro, più critica nei confronti della certificazione e difficilmente adotta una procedura formalizzata, ritenendo la certificazione di qualità semplicemente come un altro canale da adottare per “farsi notare” da potenziali clienti. Complessivamente, la certificazione di qualità è considerata uno strumento utile, ma a patto di accelerare i tempi necessari per ottenerla, semplificare l’iter previsto limitando il dispendio di risorse sia economiche che gestionali e prevedere dei canali efficienti per l’informazione dell’utenza. Gli operatori sembrano, in tal senso, lamentare la mancanza, sul territorio, di una rete strutturata cui fare riferimento per ottenere/scambiare informazioni, nonché di un referente cui rivolgersi per ottenere chiarimenti³.

³ In effetti, nel periodo di indagine, l’ISNART ha portato avanti una campagna tesa ad introdurre pratiche di certificazione di qualità per le strutture sul territorio, di concerto con la Provincia di Enna.

Peculiare sembra essere la posizione di alberghi e ristoranti, tipologie per definizione direttamente interessate all'adozione di certificazioni di qualità ma che, nel caso di Enna, sembrano invece essere poco inclini al coinvolgimento. Ci si chiede, in questo caso, quali vincoli assumano rilevanza nel definire una struttura "di qualità", dato il limitato interesse a disporre di una certificazione in tal senso. Sembra importante, innanzitutto, fare riferimento alle "buone prassi" indicate dagli stessi operatori, ovvero a quelle prestazioni che si ritengono necessarie per adeguare il servizio alle esigenze della clientela, differenziando in base al settore.

4. Indicatori di qualità e pratiche informali nelle strutture alberghiere

Prendendo in considerazione i servizi forniti in quanto indicatori indiretti degli standard qualitativi diffusi nelle strutture alberghiere di Enna, si possono distinguere alcuni servizi "fondamentali" (rapporti con la clientela e cura delle camere), indicati dalla totalità degli operatori. L'offerta è caratterizzata anche dalla cura nell'aspetto della struttura (71%), l'attenzione alle richieste della clientela (48%), presentazione dei cibi (48%) e mantenimento di prezzi competitivi (48%)⁴. Sono abbastanza sentite, ma in misura inferiore (solo il 29% degli intervistati seleziona le relative categorie), anche la necessità di un servizio di informazione e la necessità di rispondere immediatamente alle richieste della clientela, due indicatori direttamente riconducibili alla dimensione della funzionalità del servizio. Meno diffusi sono alcuni servizi accessori, per l'infanzia e per gli animali, oltre che la necessità di garantire elevati standard qualitativi dell'offerta (14%), mentre nessuna attenzione è rivolta all'impatto sull'ambiente ed al legame con il territorio. Considerando l'insieme dei servizi elencati, prevale sul territorio un'offerta economica (il 43% delle strutture mantiene prezzi competitivi) e servizi funzionali (29%), più che una d'alta qualità (solo il 14% dichiara di mirare a garantire elevati standard di qualità).

Si è già detto che la maggior parte degli albergatori non dispone di una certificazione di qualità o per mancanza di informazioni o a causa dei costi eccessivi. Le due motivazioni elencate sembrano incidere parimenti sulla scelta di non adottare una certificazione di qualità da parte delle strutture che privilegiano l'alta qualità del servizio, senza che emergano particolari differenze in relazione al numero medio di servizi disponibili. Prendendo in considerazione la totalità dei servizi o i servizi legati alla funzionalità, invece, le strutture che lamentano scarsa disponibilità di informazioni sono anche quelle in cui è presente un minor numero medio di servizi (riconducibili ad un più basso livello di qualità rilevata), a

⁴ Si noti che, a conferma di quanto rilevato, i costi sono considerati dai gestori uno dei limiti principali all'adozione della certificazione.

prescindere dalla tipologia degli stessi, mentre il limite legato ai costi eccessivi caratterizza quelle strutture che presentano un più alto numero di servizi.

5. Indicatori di qualità e pratiche informali nelle strutture di ristorazione

Già in relazione alla percezione della certificazione di qualità, i ristoratori si distinguono in quanto decisamente più critici rispetto all'adozione della stessa. Questo atteggiamento potrebbe dipendere dalla necessità, prevista dalla legge 155/1997, di adottare un manuale di autocontrollo HACCP e rispettare, di conseguenza, una serie di vincoli e presupposti abbastanza rigidi, adeguando la struttura. Come valutare, quindi, la qualità del servizio di ristorazione?

Diversi fattori sono stati analizzati a tal fine, valutandone la rilevanza sulla base della tipologia di struttura (ristorante/pizzeria o locanda/trattoria) e l'ambiente della struttura (indicato dagli operatori, che distinguono tra ambiente "elegante", "rustico" o "caratteristico"). Il rapporto tra tipologia ed ambiente della struttura non è tale da permettere una totale sovrapposizione tra le classi delle due tipologie, legate rispettivamente ad un criterio oggettivo e soggettivo di attribuzione. Una relazione, tuttavia, si mantiene in quanto dall'incrocio tra variabili emerge che nessuno degli operatori di locande e/o trattorie definisce l'ambiente della propria struttura come "elegante", mentre prevale la categoria "rustico" o "caratteristico". Il criterio soggettivo di auto-definizione della categoria di locale gestito sembra, in particolare, avere un elevato potere discriminante in relazione agli standard qualitativi diffusi per tipologia indicata, configurando differenti nicchie di mercato. L'offerta, in generale, è economica e non tipizzante. Scarso il consumo di alta qualità o la presenza di locali "caratteristici"⁵. La distinzione individuata sembra rispondere effettivamente a differenti criteri di gestione e tipologie di prestazioni, come emerge già considerando l'impiego della certificazione volontaria. Tutte le strutture dall'ambientazione elegante hanno, infatti, adottato una certificazione di qualità, contro solo il 33% di ristoranti/pizzerie e nessun'altra tipologia di ristorazione. Di contro, circa l'87% degli operatori delle strutture di ristorazione dichiarano di servire prodotti controllati, che dispongano di certificazione di qualità e/o provenienza, delineando l'immagine di un'offerta comunque attenta a garantire, di fatto, qualità e controllo dei prodotti.

Le informazioni relative alla limitata presenza di un servizio di alta qualità sul territorio sono confermate prendendo in considerazione, più analiticamente, i servizi accessori, considerati come indicatori indiretti di qualità nell'offerta. La

⁵ Il 64% di strutture sono ristoranti/pizzerie; il 54% offrono servizio da asporto; il 37% ambiente rustico e il 37% caratteristico. Meno diffusa la locanda/trattoria e l'ambiente elegante, rispettivamente il 36% e 27% delle strutture.

ristorazione sul territorio è caratterizzata da servizi che non configurano necessariamente un'offerta di qualità superiore⁶, più diffusi entro ristoranti/pizzerie piuttosto che entro locande/trattorie, con una differenza particolarmente elevata considerando gli accessori legati alla qualità superiore o all'ambiente elegante.

I servizi accessori più diffusi nei locali dall'ambiente elegante sono per la maggior parte riconducibili ad un'offerta di qualità superiore⁷, nonostante la limitata presenza di quest'ultima tipologia di servizi sul territorio⁸. Gli ambienti rustici, d'altra parte, si presentano anch'essi come caratterizzati da alcuni servizi di "alta qualità", ma privilegiano la dimensione del confort e della funzionalità⁹.

Un altro fattore che incide in modo significativo sulla qualità dell'offerta attiene all'ambito gestionale ed, in particolare, alla conoscenza normativa/informazione degli operatori ed alla funzionalità del servizio offerto. Sono considerati indicatori di funzionalità la presenza di personale qualificato, di software gestionali, la possibilità di utilizzare differenti mezzi di pagamento, la comunicazione dei tempi di attesa, il periodo di apertura del locale ed i tipi di specialità disponibili, mentre sul piano della conoscenza normativa si distinguono le norme europee e l'attribuzione di responsabilità personale in caso di furti e/o reclami. La dimensione della funzionalità pesa, complessivamente, molto sulla qualità totale del servizio ed è quella che mostra una realtà locale propulsiva di risultati positivi, mentre dovrebbe essere incrementata la conoscenza normativa, in particolare in relazione alla normativa comunitaria, soprattutto considerando locande e trattorie dall'ambiente rustico.

6. Conclusioni

Nel turismo l'investimento in qualità può essere visto come elemento per migliorare la competitività di una destinazione e/o delle singole imprese. I produttori, in genere, adeguano la loro offerta ai bisogni e alle aspettative dei consumatori. Tuttavia, la loro scelta di investimento può riguardare livelli diversi di qualità o, in alcuni casi, semplicemente limitarsi all'adeguamento delle loro

⁶ Sono considerati accessori legati ad alta qualità del servizio la presenza di aree all'aperto, vista panoramica, cantina vini, locale climatizzato, piano-bar e parcheggio privato.

⁷ I servizi maggiormente diffusi entro questa categoria sono: aree all'aperto (100%), vista panoramica (67%), cantina vini (100%), locale climatizzato (67%), piano-bar (67%).

⁸ Si noti che il 100% dei locali dall'ambiente elegante dichiara di aver adottato una certificazione di qualità, ad indicare lo strutturarsi di un processo di isomorfismo specifico legato a questa nicchia di mercato, caratterizzato dal mantenimento di elevati standard di qualità, sia informale che certificata.

⁹ I servizi maggiormente diffusi entro questa categoria sono: 50% sala meeting, 75% organizzazione banchetti, 50% parcheggio privato.

strutture agli standard richiesti dalla categoria di appartenenza. Questa decisione è certamente influenzata dal contesto in cui operano le imprese e dalla posizione occupata da una destinazione turistica sul mercato.

La bassa competitività di Enna sul mercato turistico sembra confermare questi assunti, per cui la dimensione dell'alta qualità non sembra caratterizzare l'offerta di questo territorio.

Riferimenti bibliografici

Asero V. e Gozzo S. 2007. *Comportamento di spesa e soddisfazione del turista: un'analisi di segmentazione*. Turistica, vol. 3, pp. 67-85.

Asero V., Gozzo S., Patti S. 2008. *Qualità percepita e qualità attesa nel mercato turistico*. Rassegna economica, vol. 1, pp. 235-257.

Becker G. S. e Stigler G. 1977. *De gustibus non est disputandum*. American Economic Review, vol. 67, pp. 76-90.

Candela G. e Figini P. 2003. *Economia del turismo*, McGraw-Hill.

Gustafsson A. e Johnson M.D. 2003. *Competing in a service economy: how to create a competitive advantage through service development and innovation*, Jossey-Bass, San Francisco.

Merton R. K. 2000. *Teoria e struttura sociale: studi sulla struttura sociale e culturale*, vol. 2, Il Mulino, Bologna.

Parasuraman A., Zeithaml V.M. e Berry L. 1985. *A conceptual model of service quality and its implications for future research*. Journal of Marketing, 49, pp. 41-50.

Powell W. W. e DiMaggio P. J. 2000. *Il neoistituzionalismo nell'analisi organizzativa*, Einaudi, Torino.

SUMMARY

Quality management and assurance systems of tourism facilities and services are fundamental to a healthy and growing tourism industry. Tourism operators play an important role in ensuring overall quality of a destination. Nevertheless their investment in quality is effected by effort and resources required as well as by the behaviour of the other tourism firms of that destination and the development of tourism in that area. The paper considers this issue presenting the result of a survey made in the Province of Enna, Sicily.

Vincenzo ASERO, Ricercatore in Economia Politica, Facoltà di Scienze Politiche, Università di Catania.

Simona GOZZO, Ricercatore in Sociologia, Facoltà di Scienze Politiche, Università di Catania.

INTERPRETAZIONE DELLA COLLABORAZIONE FAMILIARE IN GIOVANI COPPIE ATTRAVERSO L'ANALISI TESTUALE DI INTERVISTE SOGGETTIVE*

Giuseppe Avena, Emilia De Luca¹

1. Introduzione

Questo lavoro si ispira alle politiche di sensibilizzazione verso la condizione femminile, vincolata tra gli impegni derivanti dai lavori di casa, dall'assistenza familiare e dalla necessità di percepire un reddito e di realizzarsi come individuo.

Tradizionalmente si accusano gli uomini di poca collaborazione nella gestione del menage casalingo, delegando ogni compito alla donna. Ci siamo posti l'obiettivo di valutare la percezione del problema in un gruppo reale di individui.

Per fare ciò si è attinto ad uno studio che ha esplorato i fattori che influenzano la formazione di scelte che comportano una variazione dello stato demografico (eventuale formazione o accrescimento di un nuovo nucleo familiare) e dei comportamenti a tali scelte collegati, tra cui la gestione dei tempi e dei ruoli casalinghi. Il metodo d'indagine usato è quello delle interviste multitematiche aperte, basate su spunti di discussione, l'approfondimento dei quali è dipeso dall'andamento della singola rilevazione, rivolte a coppie giovani di lavoratori², in cui ogni narrante ha rappresentato sé e il partner.

Questo genere di intervista si presta ad essere interpretata principalmente in due modi: il primo, che sta prendendo piede rapidamente, a causa della potenziale capacità informativa, è l'analisi testuale, sia intesa come utilizzo lessicale che come analisi del discorso, che adotta strumenti quantitativi, prevalentemente associativi e probabilistici; il secondo, più comune, si basa sull'interpretazione diretta.

* Lavoro svolto nell'ambito del progetto di Ricerca "Ordinario" 2006/2007 dell'Università degli Studi di Messina su: *Le determinanti delle scelte della vita adulta nell'Italia meridionale: interpretazioni non "rational choice"*

¹ Il lavoro è frutto della collaborazione dei due autori. Tuttavia, per la stesura del testo, Giuseppe Avena è autore dei paragrafi 1, 2 e 5, Emilia De Luca è autore dei paragrafi 3 e 4.

² Tra gli intervistati, alcuni non facevano parte di coppie stabili eterosessuali (riconosciute o meno): a loro non è stata posto lo spunto della "collaborazione".

Questi metodi d'analisi sono stati utilizzati solo sui frammenti d'intervista collegati direttamente o derivatamente agli spunti tematici della collaborazione familiare, per identificare solo gli schemi verbali tipici, ma mantenendo il legame con le caratteristiche demografiche dell'intervistato e del suo nucleo familiare.

L'analisi testuale automatica³ richiede che le variabili siano identificate con rigore su tutte le interviste e di integrare tutte le altre narrazioni disponibili nel metodo qualitativo con l'analisi per selezione di nuclei tematici.

2. Background di riferimento

I soggetti rilevati attraverso le interviste in profondità appartengono ad un campione selezionato dalle popolazioni delle province di Milano e di Messina, con criteri specifici di età del rispondente (range centrato sui 35 anni), del suo stato lavorativo (in condizione lavorativa attiva, anche se instabile), con titolo di studio di scuola media superiore o più elevato; ciascuno di questi caratteri è stato identificato in maniera non rigida, indulgendo per protocollo, in una minima proporzione dei soggetti, sui limiti imposti in origine.

Tra questi individui si è creata una suddivisione in 3 gruppi: single, in coppia stabile senza figli, in coppia stabile con figli coabitanti. Il primo dei 3 gruppi è stato escluso da questo studio, perché non coerente con lo scopo dell'indagine. In origine si è previsto anche una ripartizione di tipo zonale, per identificare differenze tra centro, periferie e località importanti della provincia.

I soggetti sono stati intervistati sull'argomento delle scelte di formazione o di non-formazione della famiglia e portati ad esprimere una percezione della propria esperienza e delle proprie opinioni sui temi proposti, tra i quali vi era la condizione organizzativa della propria casa. Hanno risposto a domande aperte, inserite in un'intervista la cui durata prevista era tra i 40 ed i 90 minuti, esprimendosi nel modo più ampio ed approfondito possibile sugli input del rilevatore, in un faccia a faccia che ha fatto gioco sulla creazione di un rapporto di confidenza temporaneo.

Dal testo sono poi stati estratti i caratteri dell'attività lavorativa dei partner, sia per tipo che per intensità, intesa come tempo lavorativo, organizzata in 3 fasce (lieve <30 h/sett; normale = 30–40 h/sett; intensa >40 h/sett). La scelta di considerare il tempo del viaggio come lavorativo è stata lasciata agli intervistati.

L'estensione dell'intervista, a causa del numero di questioni affrontate per avere una visione completa del tema, ha reso necessaria una dimensione campionaria ridotta e di seguire un filo conduttore ed un livello di approfondimento diverso per ogni colloquio. I rispondenti considerati per il tema dell'organizzazione familiare

³ Il software utilizzato è il T-LAB.

sono solo quelli in unione stabile. La ridotta dimensione campionaria può non essere rappresentativa della popolazione (nazionale), lo è di più della realtà locale.

Questa indagine, derivando da una così ampia, si focalizza, per brevi brani e con un medio livello di approfondimento, sulla ripartizione dei compiti familiari. I testi derivano dall'interazione tra due persone, che si integrano vicendevolmente nel discorso: il peso delle domande è stato quindi analizzato, anche per isolarlo.

3. Cenni di metodo

Si è scelto di valutare i dati principalmente attraverso l'analisi testuale, che si fonda sulla verifica del rapporto tra le parole: gli indici di associazione ci danno una misura della relazione di un vocabolo con ciascuno degli altri. Lo stesso procedimento si può estendere a raggruppamenti di parole più significativi. Questo tipo di analisi richiede una reiterazione di calcoli per ogni "unità lessicale" (parola differenziata), effettuata con uno specifico software, per contenere i tempi. L'associazione così ottenuta si esplicita attraverso tabelle e grafici che identificano relazioni concettuali e strutture linguistiche, che possono consentire di identificare o escludere alcuni schemi espressivi.

Solitamente, a questa prima analisi quantitativa segue una valutazione delle concordanze, cioè la lettura della frasi che contengono una o più parole cardine, che permette di approfondire il senso specifico dell'utilizzo di un termine ed i concetti complessi, ad esempio le frasi fatte o a doppio senso, a cui si lega, creando nuclei argomentativi nel testo/soggetto di provenienza.

Ogni soggetto ed il relativo testo, inoltre, sono uniti dai caratteri rappresentativi sul tema, identificati con una struttura organizzata per variabili, su cui si crea una intelaiatura, anche in questo caso basata sull'associazione, che consenta di identificare chi dice cosa, attraverso la rispettiva modalità di quella variabile. Questa struttura crea cluster (raggruppamenti) tematici e matrici dette di prossimità (di due elementi in quel gruppo di persone), con cui costruire spazi fattoriali, che useremo per "disegnare" queste relazioni tra parole. In questo spazio la relazione tra parole è tanto più stretta quanto più vicini sono i vocaboli rappresentati nel grafico e ciò consente di identificare il linguaggio tipico in probabilità di ogni modalità di una variabile o di ogni gruppo tematico in quello spazio.

Il confronto tra tali grafici permette di ottenere deduzioni, da confermare attraverso l'extrapolazione di frammenti e la lettura selettiva e ripetitiva dei testi, fondata su stadi interpretativi consecutivi, sempre più specifici.

Il raffronto tra gli schemi ottenuti con la creazione di spazi e con la lettura diretta hanno portato ai risultati del paragrafo seguente.

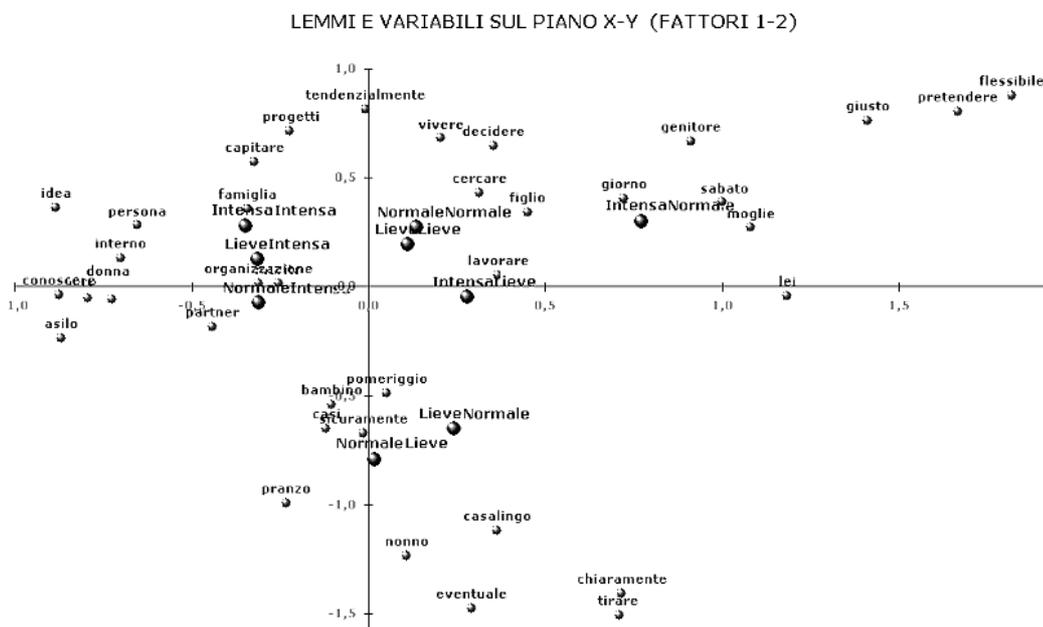
4. Risultati più significativi

Dalla lettura di queste interviste, emerge, come primo elemento rilevante, la frequenza di alcune classi di vocaboli: la rarità di quelli tipici delle attività da svolgere (stirare, lavare, spesa...), l'utilizzo circoscritto di quelli di identificazione individuale (lui, lei, marito, moglie,...), l'ampia presenza di quelli specifici dell'organizzazione generale (decidere, insieme, collaborare) e temporale (mattina, pomeriggio, orario...).

I testi sono poi stati messi in relazione con le caratteristiche dei rispondenti: le zone d'abitazione, centro urbano, periferia, centro suburbano, la presenza o meno di figli, il tipo di occupazione, il ritmo-tempo richiesto dal lavoro di ciascun partner, altre variabili derivate da quelle considerate. I risultati più interessanti sono prodotti dalla variabile che sintetizza i tempi lavorativi d'entrambi i coniugi.

Il Grafico 1 mostra il piano dello spazio fattoriale, creato su tale variabile, che spiega la maggior parte della variabilità.

Grafico 1– Localizzazione delle unità lessicali rispetto alle categorie “rapporto tra le intensità del lavoro dei partner” nelle tre categorie intensa, normale, lieve.



Questo tipo di grafico richiede una certa familiarità con le variabili e le relative denominazioni, per interpretare il sistema di prossimità. In questo caso, si può

leggere l'aggregazione per "proporzione del tempo lavorativo", per cui l'idea di disponibilità, collegata a problemi pratici, si riconduce a coppie con un lavoratore con attività intensa; l'organizzazione del tempo è tipica di un ritmo di lavoro normale, che pare richieda di giocare sui punti di forza dei giorni liberi (ad esempio i fine settimana). Nei casi in cui c'è disponibilità di tempo, con uno dei due partner con attività lievi, s'affronta l'argomento secondo un criterio più teorico, meno organizzativo.

Questo schema si ripete nella composizione dei cluster (Tabella 1).

Tabella 1 – Descrizione dei componenti dei cluster, ottenuti automaticamente (esclusi quelli i cui componenti sono specifici dello schema di domanda)

CLUSTER N. 1	CLUSTER N. 2	CLUSTER N. 3
LEMMA	LEMMA	LEMMA
mattina	pagare	famiglia
pomeriggio	bisogno	problemi
Lei	Io	periodo
lavoro	situazione	cambiare
Sera	disponibilità	caso
lavorare	occupare	presentare
giornata	Spesa	vivere
uscire	Uomo	esperienza
*Rilevatore_GA	faccenda	nostro
*genere_uomo	Tu	grande
Settimana	*Rilevatore_SD	*XX_LieveLieve
Bambino	sbrigare	*lavoro_Atipico
Scuola	parlare	*IntensitàLavoro_Liev
Orario	ambito	bhe
Casa	incombenza	collaborazione
Noi	bolletta	collaborare
Preparare	trovare	*XX_LieveIntensa
Mangiare	*zona_Milazzo	passare
*IntLavoroPar_Normale	piacere	responsabilità
Dedicare	*IntLavoroPar_Intensa	problema

Si nota anche, tornando al Grafico 1, un'aggregazione per gruppi di ritmi lavorativi: un gruppo è formato dalle coppie con ritmi uguali, che si distribuiscono intorno al 2° asse fattoriale; il secondo gruppo comprende le coppie "normale – lieve", isolato dagli altri gruppi ed agli antipodi rispetto al primo; un terzo gruppo identifica coppie con almeno un lavoratore con attività intensa. Queste variabili si sviluppano su più di due assi fattoriali, ma per semplicità si rappresentano con i due più significativi. Ciò comporta che, guardando in profondità questa immagine, le aggregazioni possono cambiare e quindi anche le affinità di comportamento o espressione. Esplorando queste altre dimensioni, i primi due gruppi si confermano nell'aggregazione, evidenziando legami rispettivamente con la collaborazione e con il "caso", mentre il terzo si distribuisce in profondità sul terzo asse.

Considerato che le interviste riguardavano sia uomini che donne, si è provato, attraverso le tecniche delle concordanze e della interpretazione diretta, a verificare i

risultati ponendo una distinzione di genere tra i coniugi lavoratori e cercando una quantificazione numerica del problema. Per semplicità si crea una nuova variabile con solo tre modalità, più esplicite, che si chiarisce nei valori in Tabella 2.

Tabella 2 – Distribuzione degli intervistati per ritmi di lavoro di coppia/impegno casalingo

Confronto ore lavoro di coppia	% per categoria	Impegno casalingo paritario	Impegno casalingo collaborativo	Cooperazione casalinga assente
Lui > Lei	27 (59%)	7	13	7
Lui = Lei	13 (28%)	6	4	3
Lui < Lei	6 (13%)	1	3	2
Valori totali	46	14 (30%)	20 (43%)	12 (26%)

Considerando solo questa nuova variabile, si sono potute interpretare alcune delle interviste precedentemente escluse dall'analisi perché non fornivano informazioni sufficienti alla classificazione delle altre variabili. Si è ottenuto un piccolo campione aggiuntivo, i cui risultati sono riassunti in Tabella 3.

Tabella 3 – Distribuzione degli intervistati per ritmi di lavoro di coppia/impegno casalingo

Confronto ore lavoro di coppia	% per categoria	Impegno casalingo paritario	Impegno casalingo collaborativo	Cooperazione casalinga assente
Lui > Lei	8	2	2	4
Lui = Lei	4	2	2	0
Lui < Lei	0	0	0	0
Valori totali	12	4	4	4

La proporzione di individui che non collaborano alla gestione delle attività domestiche è di circa un quarto dei rilevati totali; tra questi, una percentuale elevatissima (22% del totale, 81% dei non collaborativi) si giustifica tramite uno squilibrio nei ritmi lavorativi dei partner. La distinzione diventa più stringente nei casi in cui *lui* ha un ritmo lavorativo *intenso*, rispetto ad uno *lieve* di lei (rispettivamente il 23% ed il 69%), legittimando questo tipo d'organizzazione domestica.

5. Considerazioni conclusive

Questa serie di risposte mostra gli individui del campione come fortemente partecipativi, percependo la collaborazione, non come aiuto, ma come necessità di condividere l'impegno, in particolare in relazione agli eventuali figli. I cluster e gli schemi linguistici descrivono concetti di cooperazione e organizzazione, collegati

al tempo disponibile dal lavoro, che discordano con la tradizione culturale che delega alla donna la gestione esclusiva delle attività casalinghe, compresa la cura dei figli e degli imprevisti. La maggior parte dei rispondenti ha mostrato uno spirito partecipativo, inteso come reciproca responsabilità e personale interesse nello svolgimento dei compiti necessari alla buona gestione casalinga.

Necessità e disponibilità guidano la partecipazione (dichiarata), molto più di idealismi e prese di posizione. Gli intervistati si preoccupano principalmente di garantire la cura dei figli e della casa, nonostante i sacrifici che ne conseguono. Sono ancora le donne che dedicano il maggior tempo alla cura della casa, come confermato da uno studio dell'Eurostat del 2004, tuttavia non ne sono chiare le ragioni: il primo dubbio è se la collaborazione espressa sia reale o solo percepita; il secondo è se la cura della casa abbia, per le donne, una funzione di immagine, anche sociale; il terzo dubbio è se si instauri un circolo vizioso tra impegno lavorativo e domestico ed elemento culturale, che riproduca l'effetto di delega del ruolo alle donne.

La misura campionaria ridotta e la selettività del campione rappresentano dei limiti che suggeriscono l'utilità di riproporre l'osservazione, ampliando la base e specificando i punti salienti qui rilevati, associandoli ad altri fattori potenzialmente influenti: servizi pubblici disponibili, collaborazioni esterne ed altro.

Riferimenti bibliografici

- Avena G. (2009). Uno studio sulla condivisione maschile nella gestione dell'attività domestica in un ambito territoriale dell'Italia del sud, NIKE, La rivista delle Scienze Politiche, Foggia.
- Avena G. (2007). Corso di vita, famiglie e riproduzione fra scelte e costrizioni a Messina. Materiali e metodi, Messina, Lippolis Editore.
- Bolasco S. (1999). Analisi multidimensionale dei dati. Metodi, strategie e criteri d'interpretazione, Carocci
- Lancia F. (2004). Strumenti per l'analisi di testi, Franco Angeli Editore
- Tuzzi A. (2003). L'analisi del contenuto, introduzione ai metodi e alle tecniche di ricerca, Carocci
- Mencarini L. (2006), I rapporti di genere nell'ambito della coppia e la genitorialità, in Ongaro F., (a cura di, 2006), "Scelte riproduttive tra costi, valori, opportunità" Milano, Franco Angeli Editore.
- Micheli G. A., (a cura di, 2006), Strategie di family formation cosa sta cambiando nella famiglia forte mediterranea, Milano, Franco Angeli Editore.
- Niemi I., Paakkonen H., Hannula U., Lehtikoinen H., Ollila I., (2004), How Europeans spends their time Everyday life of women and men, Luxembourg, Office for Official Publications of the European Communities.

SUMMARY

Interpretations of the cooperation in house management, related to young families, applying text analysis to individual interview

This work explores, using the text analysis and the direct analysis, the aptitude of collaboration in housework, in family management, related to young families, in which one of them is 33-37 years old. Narrations highlight males and females points of view. They are married or in common law marriage, with or without sons and they are going steady.

Looking at the output we can observe that they don't use word connected with traditions and role set and they don't speak about a choice, but talk about the housework and the distribution of duties, pointing on their attitude for cooperation, complete or partial.

They collaborate, sometimes with a task division, otherwise sharing the duties, but this cooperation depends on time used for job by both partners. Looking at the measure of male-female working times, they well adapt own commitments in housework to the external conditions: especially the available time of the partner and the support net.

The key-words evidence the practical vision of the question: they describe how they spend their working-time and their available time; rather than role management and identification of the activity to do. This work shows that difference between genders in housework depend on availability of time, due to job, it isn't felt like an obligation and it isn't important into relationship between partners. This conclusion suggests to study in depth the question, pointing on the characteristics described by the sample and other elements that can influence the matter. We think that the study proposed could produce interesting developments about Italian new generations social structure.

Giuseppe AVENA, Ricercatore di Statistica Sociale, Facoltà di Scienze Statistiche, Università di Messina.

Emilia DE LUCA, Dottore di Ricerca in Statistica e Risorse Ambientali, Università di Messina.

PERCORSI SCOLASTICI E PREVISIONI DELLA POPOLAZIONE SCOLASTICA. UN CASO DI STUDIO

Odo Barsotti, Federico Benassi, Linda Porciani,
Alessandro Valentini, Silvia Venturi

1. Premessa

Il nostro studio si propone, da un lato, di cogliere alcuni fattori significativi che incidono sul fenomeno della dispersione scolastica e, dall'altro, di delineare un possibile scenario futuro della popolazione scolastica attraverso le previsioni demografiche nell'area territoriale della provincia di Pisa, con un'attenzione specifica all'evoluzione della componente straniera. Si tratta, come intuibile, di due aspetti fondamentali di conoscenza nella programmazione dell'offerta formativa a livello locale e nell'adozione di adeguate politiche in grado di contrastare l'allontanamento dal sistema scolastico. Nel nostro caso di studio, ci siamo potuti avvalere della banca dati dell'Osservatorio Scolastico della Provincia di Pisa, che raccoglie le informazioni sugli studenti iscritti nelle scuole della provincia per ogni ordine e grado¹. I dati si possono considerare completi dall'anno scolastico 1993/94 per le scuole secondarie di I grado e dall'anno scolastico 2001/02 per tutti i gradi. Per le previsioni della popolazione scolastica, in quanto previsione derivata da quella demografica, si fa ricorso naturalmente anche ai dati di fonte anagrafica: ammontare e distribuzione secondo il sesso, l'età e la cittadinanza della popolazione residente, matrice dei cambiamenti di residenza tra le subaree territoriali della provincia di Pisa, nel periodo più recente.

2. Successo/insuccesso scolastico: modelli esplicativi

Nella nostra analisi si sono deliberatamente utilizzati i termini “successo” ed “insuccesso” scolastico in luogo di “dispersione scolastica” per due ragioni essenziali: per dare conto della parzialità degli indicatori utilizzati rispetto al più vasto universo degli indicatori che descrivono il fenomeno della dispersione; e,

¹ La ricerca è stata commissionata al Dipartimento di Statistica e Matematica applicata all'Economia dell'Università di Pisa, dalla Provincia di Pisa, Ufficio Istruzione – Area Socio Culturale - Osservatorio Scolastico Provinciale.

soprattutto, per sottolineare l'abbandono della visione della dispersione come imputabile esclusivamente a caratteristiche socioculturali e psicologiche dell'individuo, e l'assunzione della corresponsabilità del sistema scolastico. Il termine "insuccesso", quindi, sottintende sia i problemi degli alunni nell'affrontare il percorso scolastico sia il ruolo delle istituzioni che possono contribuire a provocare o aggravare tali problemi (Provincia di Pisa-OSP, 2008).

Con il termine "successo/insuccesso" scolastico si fa riferimento ad una dimensione temporale entro cui è previsto lo svolgimento dell'iter scolastico nonché alla qualità dell'uscita dal percorso. In particolare, un percorso s'intende concluso con successo quando si compie nei tempi previsti, cioè senza perdita di anni di studio qualunque sia la motivazione: bocciatura, abbandono, trasferimento e rinuncia. La dimensione qualitativa è stata invece misurata attraverso il voto finale di uscita dalla scuola secondaria di II grado.

Lo studio prende in esame due coorti: una composta dalla popolazione studentesca iscritta nell'anno 2001 al primo anno della scuola secondaria di I grado (coorte A) e l'altra composta dalla popolazione studentesca iscritta al primo anno della scuola secondaria di II grado nell'anno 2000 (coorte B). Entrambe le coorti sono state seguite per cinque anni consecutivi attraverso estrazioni successive, escludendo i soggetti che si sono trasferiti o ritirati, perché per loro non è possibile conoscere l'esito del percorso scolastico. La coorte A è stata seguita fino al completamento del primo biennio dell'istruzione secondaria di II grado, considerando questo una tappa fondamentale del percorso scolastico², e la coorte B fino al completamento dell'intero ciclo di istruzione secondaria di II grado, in altre parole fino al conseguimento (o meno) di un diploma³.

Per entrambe le coorti, l'evento successo/insuccesso, inteso nell'accezione parità/ritardo nel percorso scolastico, è stato posto in relazione ad alcune variabili socio-demografiche (genere e cittadinanza), agli esiti scolastici durante il percorso formativo (giudizio finale della scuola secondaria di I grado: sufficiente, buono, distinto, ottimo) e all'area di residenza (Area pisana, Valdarno inferiore, Valdera, Val di Cecina, aree esterne alla provincia). Per la coorte B, è stato possibile, da un lato estendere l'accezione successo/insuccesso anche al voto finale di uscita dalla scuola secondaria e, dall'altro, includere nel sistema di relazioni, oltre alle precedenti, altre variabili quali: la regolarità o il ritardo nel percorso di studio anteriore (età inferiore o uguale a 14 anni o età superiore a 14 anni al momento

² Negli anni più recenti, infatti, tale tappa coincide con il conseguimento dell'obbligo scolastico (L. 296/2006).

³ Sono state scelte due coorti separate e non un'unica da seguire per 8 anni consecutivi poiché, come già detto, i dati raccolti dall'OSP sono completi dal 1993/94 per le scuole secondarie di I grado, ma solamente dal 2001/02 per le scuole secondarie di II grado.

dell'iscrizione al primo anno della scuola secondaria di II grado), il tipo di indirizzo scolastico seguito (professionale, artistico, tecnico, liceale) e il titolo di studio combinato dei genitori (entrambi con al massimo la scuola dell'obbligo, la madre con al massimo l'obbligo ed il padre con diploma o laurea, la madre con diploma o laurea e il padre con al massimo l'obbligo, entrambi con il diploma o la laurea).

E' stato applicato un modello di regressione logistica nel quale la variabile dipendente, espressa dalla probabilità di concludere in pari/non in pari il percorso scolastico per entrambe le coorti, e per la coorte B anche dalla probabilità di ottenere un voto di diploma alto/basso (maggiore di 80 centesimi/non maggiore di 80 centesimi), è posta in funzione delle variabili indipendenti in precedenza specificate⁴. La capacità esplicativa dei modelli non è molto elevata (lo pseudo R² varia da 0,26 a 0,34); il risultato è però tutt'altro che deludente se si considera la natura amministrativa dei dati e la numerosità delle due coorti.

In sintesi, emerge che i fattori associati significativamente al successo/insuccesso scolastico ($p < 0,05$) sono per entrambe le coorti il genere e l'esito finale della scuola secondaria di I grado. In particolare, viene evidenziato che: essere maschio, *cæteris paribus*, ha un effetto negativo sulla probabilità sia di terminare il percorso scolastico in pari sia di ottenere un voto di diploma alto; aver ottenuto un buon giudizio al termine della scuola secondaria di I grado influenza molto positivamente la probabilità di completare in tempo utile il percorso scolastico e di ottenere un voto alto alla maturità. Per la coorte B, quando la variabile dipendente è il voto di diploma, appaiono significative anche le relazioni con: la regolarità e il ritardo nel percorso scolastico anteriore (avere non più di 14 anni al momento dell'iscrizione alla classe prima delle scuole secondarie di II grado ha un impatto positivo sulla probabilità di diplomarsi con un buon voto); il tipo di indirizzo scolastico scelto (l'istruzione tecnica e l'istruzione liceale hanno – e *cæteris paribus* – un impatto negativo, rispetto a quella professionale, sulla probabilità di ottenere un voto alto al diploma).

Seppure statisticamente non significative da un punto di vista inferenziale, meritano tuttavia di essere segnalate le direzioni delle relazioni di altre variabili inserite nel modello di regressione⁵. In particolare viene osservato come: non avere

⁴ Nella coorte A (3.020 unità), gli studenti che hanno superato il primo biennio della scuola secondaria superiore di II grado in pari, e che quindi non hanno mai registrato un evento negativo (bocciatura) nella loro carriera scolastica, sono 2.374, cioè il 78,6% della coorte iniziale. Nella coorte B (2.496 unità), gli studenti che hanno conseguito il diploma senza alcun ritardo sono 1.894, ossia il 75,9% dell'ammontare iniziale della coorte. Tra gli studenti che hanno conseguito in pari il diploma, il 44,3% ha ottenuto una votazione superiore a 80 centesimi.

⁵ Questo è giustificato dal fatto che i collettivi oggetto di analisi non sono due campioni ma due intere coorti di studenti.

la cittadinanza italiana rallenti il percorso scolastico ma non influisca affatto sul voto finale⁶; rispetto all'indirizzo professionale, le altre tipologie di indirizzo (artistico, tecnico e liceale) tendano a ridurre, a parità delle altre condizioni, la probabilità degli studenti di concludere in pari il percorso scolastico; i soggetti i cui genitori hanno un basso titolo di studio (non superiore all'obbligo) o anche quando la sola madre abbia un titolo di studio basso, tendano a sperimentare, rispetto alle "categorie" di genitori con titolo di studio più alto (diploma o laurea), una probabilità inferiore sia di concludere senza ritardi il percorso scolastico sia di ottenere un voto alto al diploma. In conclusione, è opportuno sottolineare come i risultati dei modelli di regressione ridimensionino, o addirittura capovolgano, gli effetti "lordi" delle variabili cittadinanza e tipo di indirizzo scelto nella scuola secondaria di II grado sul successo/insuccesso scolastico delle due coorti considerate. L'essere straniero, una condizione fortemente penalizzante quando si confrontino i tassi lordi di insuccesso tra studenti italiani e non italiani⁷, perde importanza nel segnalare l'incapacità di conseguire in tempi utili sia l'obbligo scolastico che il diploma quando si tengano sotto controllo gli effetti delle altre variabili inserite nel modello: le relazioni nette, benché negative, non sono statisticamente significative né per la scuola superiore di I grado né per quella di II grado. In sostanza, a parità di tutte le altre caratteristiche (il genere, il giudizio finale della scuola secondaria di I grado, la regolarità nel percorso di studio precedente alla iscrizione al primo anno della scuola secondaria di II grado, il titolo di studio combinato dei genitori, l'indirizzo di scuola di II grado frequentato), l'essere straniero non incide in modo significativo sull'insuccesso scolastico. Ed anzi, quando l'insuccesso è considerato in termini di rendimento, la relazione, quantunque statisticamente non significativa, diventa positiva: a segnalare che, *cæteris paribus*, gli studenti stranieri, tendano ad avere un esito non meno favorevole, semmai più favorevole, in termini di voto di diploma (maggiore di 80 centesimi), degli studenti italiani quando gli uni e gli altri abbiano completato il corso di studi senza perdita di anni. Passando a considerare il tipo di scuola secondaria di II grado scelto dalla coorte B, i tassi lordi di insuccesso scolastico degli studenti appaiono nettamente più elevati, in entrambe le accezioni, sia

⁶ Da notare che nelle due coorti considerate la presenza di studenti stranieri non era in effetti ancora molto elevata: il 3,2% nella coorte A e il 2% nella coorte B.

⁷ Nella scuola secondaria di I grado oltre 1/3 degli studenti stranieri non ha concluso in tempo utile il percorso scolastico contro 1/5 degli studenti italiani; nella scuola secondaria di II grado i tassi di insuccesso sono rispettivamente del 36% contro il 24% in termini di ritardo e del 78% contro il 54% in termini di esito finale (voto di diploma inferiore a 80 centesimi).

nell'area dell'istruzione professionale che in quelle tecnica e liceale⁸. Il rovesciamento del segno delle relazioni nette, addirittura statisticamente significative quando nel modello di regressione si assume come variabile dipendente il voto di diploma, sembra allora prospettare che i maggiori insuccessi che sperimentano gli studenti degli indirizzi professionali dipendano non da una maggiore attitudine alla selezione di tale tipo di scuola, ma da una selezione a monte, ossia dal fatto che tale tipo di scuola è scelto da soggetti più deboli per quelle caratteristiche demografiche, socio-culturali e di esiti e regolarità della carriera scolastica che sono state introdotte nei modelli come variabili esplicative⁹.

3. Le previsioni della popolazione scolastica

Il secondo aspetto del nostro studio, ossia le previsioni della popolazione scolastica, ha permesso di valutare le trasformazioni che nel prossimo futuro potrà subire la composizione della struttura della domanda scolastica nella provincia di Pisa e di stimare la crescente importanza della componente straniera già oggi su livelli apprezzabili¹⁰. La popolazione scolastica è composta da due aggregati, tra loro logicamente separati ma in gran parte sovrapposti. Il primo indica la *popolazione scolastica residente in provincia di Pisa*, cioè il numero di studenti

⁸ La proporzione di soggetti in ritardo (che hanno subito almeno una bocciatura) è due volte e mezzo più alta fra gli studenti della coorte iscritti agli Istituti professionali (38,4%) che fra gli studenti iscritti ai Licei (15,7%). Le analoghe percentuali sono pari al 29,4% per gli studenti degli Istituti tecnici e del 27,7% per quelli degli Istituti d'Arte. Ed ancora, il 78% degli studenti iscritti agli Istituti professionali ha ottenuto un voto di diploma inferiore a 80 centesimi, rispetto al 60% degli iscritti agli Istituti tecnici, al 57% degli iscritti agli Istituti d'Arte e a meno della metà (49%) degli iscritti ai Licei.

⁹ Ponendo a confronto le caratteristiche degli studenti della coorte iscritti all'istruzione professionale con quelle degli studenti iscritti all'istruzione liceale (che si collocano al polo opposto) si rileva che: i maschi sono il 62% contro il 30%; i soggetti che hanno avuto un giudizio finale di sufficiente della scuola secondaria di I grado sono l'83% contro l'11%, al contrario nemmeno il 2% ha avuto un giudizio ottimo contro quasi il 30%; coloro che hanno sperimentato ritardi nel percorso scolastico precedente sono il 32% contro appena il 6%; la quota di genitori entrambi con la scuola dell'obbligo è quasi doppia tra gli studenti degli Istituti professionali (74%) rispetto a quelli dei Licei (39%), viceversa tra i secondi è oltre quattro volte più grande la proporzione di genitori entrambi con diploma o laurea (30% contro 7%).

¹⁰ Nell'anno scolastico 2006/07 l'incidenza degli studenti stranieri iscritti nei diversi gradi del sistema scolastico provinciale era pari a: 9,4% nella scuola di infanzia, 9,1% nella primaria, 8,7% nella secondaria di I grado e 5,2% nella secondaria di II grado (Provincia di Pisa-OSP, 2008; pag.56).

residenti nella provincia di Pisa qualunque sia la localizzazione della scuola frequentata, il secondo la *popolazione scolastica presente*, ovvero il numero di studenti che si rivolge ai plessi scolastici della provincia di Pisa qualunque sia la loro residenza. Le previsioni dei due aggregati sono legate al futuro andamento della popolazione e di conseguenza si inquadrano in un modello più ampio di *previsione del numero di residenti* in provincia di Pisa e nelle sue aree. Dal punto di vista metodologico il primo pilastro del modello di riferimento è rappresentato dalle previsioni demografiche di base della popolazione residente in provincia di Pisa (con base 01-01-2007) per classi di età annuali e genere, di recente rilasciate dall'Istat secondo tre scenari: alto, intermedio e basso. Le previsioni provinciali relative allo scenario centrale, quello ritenuto più probabile dallo stesso Istituto di Statistica¹¹, sono state poi disaggregate in base sia alla cittadinanza (italiana e straniera) sia alla zona di residenza (Area pisana, Valdarno, Valdera, Val di Cecina), integrando il classico approccio per componenti di coorte con un sistema di vincoli (*top-down*). In base allo scenario demografico delineato, nei prossimi 20 anni la popolazione residente in provincia salirà dalle attuali 400 mila unità a 439 mila, con un tasso di crescita medio annuo del 4,7%. La popolazione straniera nello stesso periodo triplicherà dalle attuali 22 mila a 63 mila unità, facendo crescere l'incidenza rispetto alla popolazione complessiva dal 5,5% al 14,4%. L'aumento della popolazione straniera e la leggera ripresa della fecondità ipotizzata dallo scenario centrale determinano una crescita della popolazione giovanile in età scolare almeno fino al 2023. Ciò non impedisce tuttavia che prosegua, sia pure con qualche attenuazione, quel processo di invecchiamento che caratterizza ormai da tempo la struttura della popolazione: l'indice di vecchiaia salirà infatti dall'attuale valore di 181,5 fino a 209,9 nei prossimi 20 anni. Applicando opportuni tassi di scolarizzazione, ricavati dai dati dell'Osservatorio, la previsione della popolazione residente si traduce in previsione della popolazione scolastica residente; se a questa si sottrae la quota di studenti pisani che studiano in altre province e si aggiungono i vettori degli studenti che provengono da fuori, si ottiene la previsione della popolazione scolastica presente. I risultati mostrano che la popolazione scolastica residente nella provincia di Pisa, 42.600 unità nel 2007, crescerà sensibilmente nei prossimi anni, fino a superare quota 51.000 unità tra il 2022 e il 2024. Da quel momento in poi si innescherà una fase di declino che la

¹¹ Secondo lo scenario demografico delineato dall'ipotesi centrale per la provincia di Pisa, nei prossimi 20 anni il numero di figli per donna salirebbe dagli attuali 1,29 a 1,50, mentre l'età media al parto crescerebbe da 31,7 a 33,5 anni. Al tempo stesso la speranza di vita alla nascita maschile dovrebbe salire nell'orizzonte di proiezione da 79,3 a 82,2 anni e quella femminile da 84,6 a 87,5 anni. Il numero di immigrazioni dall'estero dovrebbe assestarsi attorno alle 1.400 unità annue.

farà scendere attorno alle 50.000 unità nell'ultimo anno di proiezione. Complessivamente nei 20 anni la crescita sarà dell'8,4% medio annuo. L'incremento è in gran parte determinato dagli studenti stranieri (3.000 unità nel 2007, il 6,9% della popolazione scolastica residente), che nei prossimi anni cresceranno secondo una traiettoria pressoché lineare fino a raggiungere quota 9.600 unità (19,1% del totale) nel 2027. La popolazione scolastica presente è invece pari a 41.600 unità nel 2007 (di cui 2.900 stranieri, il 6,9% del totale). Essa seguirà una traiettoria ascendente nei prossimi quindici anni fino a lambire le 50.000 unità attorno al 2021-2022, e successivamente declinerà fino ad arrivare a quota 49.000 nel 2027 (di cui circa 9.200, il 18,8% del totale, saranno studenti di cittadinanza straniera). Complessivamente nell'orizzonte di proiezione il numero di studenti residenti si manterrà superiore rispetto a quello dei presenti sul territorio provinciale per circa 1.000 unità, un differenziale che tende a crescere nel corso del tempo. Per la popolazione scolastica presente nella provincia sono state infine elaborate anche le proiezioni relative al numero di studenti che frequenteranno le scuole superiori, distinte per tipo di indirizzo: professionale, artistico, tecnico e liceale. Il corpo studentesco in questo tipo di scuole dovrebbe salire nei prossimi 20 anni da 17.400 a circa 20.800 unità. Nelle nostre simulazioni abbiamo ipotizzato che le distribuzioni per indirizzo non cambiassero nel tempo. Ciò significa che l'incremento verrà assorbito per circa 1.700 unità dai Licei, che attualmente attraggono e continueranno ad attrarre quasi la metà degli studenti (48,7% nel 2007); gli Istituti tecnici (32,9% degli studenti nel 2007) incrementeranno di circa 1.000 studenti nei prossimi 20 anni; quelli professionali (2.500 unità nel 2007) di circa 500 unità e infine gli Istituti d'Arte (700 unità) di un centinaio di studenti.

4. Per concludere

Dall'analisi è emerso come alcune caratteristiche degli studenti, espresse da altrettante variabili incluse nel modello, agiscano negativamente sulle *performances* scolastiche: l'essere maschio, talvolta – ma non sempre – l'essere straniero, l'aver concluso la scuola secondaria di I grado con una votazione bassa, l'isciversi non in pari al primo anno della scuola secondaria di II grado, l'aver entrambi i genitori, o comunque la madre, con solo il titolo di studio dell'obbligo sono tutti fattori che determinano una più alta probabilità di insuccesso.

I risultati segnalano, dunque, che sulla continuità e sul rendimento dei percorsi scolastici continuano ad operare le disuguaglianze secondo il livello culturale e socioeconomico dell'ambiente familiare. Queste disuguaglianze potranno essere alimentate ed anche accentuate dalla crescente presenza di famiglie di cittadinanza straniera (le previsioni indicano che gli studenti stranieri triplicheranno nei

prossimi venti anni). Ed il sistema scolastico non appare ancora in grado di contenerle, tende semmai a perpetuarle ove non riesca ad attivare efficaci processi di mobilità sociale. Se, invece, come riconosciuto e promosso a livello nazionale ed europeo, le istituzioni scolastiche intendono essere il motore della mobilità sociale, dovrebbero: agire nella direzione di rimuovere gli ostacoli che si frappongono tra le caratteristiche, le condizioni individuali e le *performances* scolastiche; affrontare la dispersione scolastica, e la perdita di capitale umano che ne deriva, non solo laddove si manifesta (scuola secondaria di II grado) ma dove trova effettivamente origine (scuola secondaria di I grado), come dimostra il forte valore predittivo del giudizio finale della scuola secondaria di I grado sul successo scolastico, a livello del conseguimento sia dell'obbligo sia del diploma; tenere in considerazione nell'impostazione delle *policies* le tendenze demografiche in atto che indicano una diminuzione della popolazione studentesca autoctona e un contemporaneo aumento della popolazione studentesca straniera.

Riferimenti bibliografici

- Istat 2008. *Previsioni demografiche*, Popolazione, Nota informativa 19/06/2008.
Ministero Pubblica Istruzione, Servizio Statistico. 2008. *La dispersione scolastica. Indicatori di base. A.s. 2006/07*.
http://www.pubblica.istruzione.it/mpi/pubblicazioni/2008/allegati/dispersione_2007.pdf
OCSE-PISA 2006. *Rapporto Nazionale Italia*,
http://www.invalsi.it/invalsi/ri/pisa2006.php?page=pisa2006_it_05
Provincia di Pisa – OSP. 2008. *La dispersione scolastica nella provincia di Pisa*, Stylgrafica Cascinese, Cascina.
Valentini A. 2007. *A way to solve the consistency question in a demographic multiregional cohort component model*, Atti della riunione intermedia SIS 2007, Università CA'Foscari, Venezia.

SUMMARY

This paper mainly aims to analyse on one hand the factors related to the leaving school; and on the other hand it shows the results of the school population projections in Province of Pisa, with a particular regard to the foreign population. The results could play a key role for local school policies.

Odo BARSOTTI, Professore di Demografia, Università di Pisa.

Federico BENASSI, Dottore di ricerca in Demografia, Università di Pisa.

Linda PORCIANI, Assegnista di ricerca in Demografia, Università di Pisa.

Alessandro VALENTINI, Ricercatore Istat.

Silvia VENTURI, Ricercatrice in Demografia, Università di Pisa.

CONDIZIONE SOCIO-ECONOMICA E DISUGUAGLIANZE DI SALUTE NELLA POPOLAZIONE ANZIANA IN EUROPA

Odoardo Bussini, Donatella Lanari¹

1. Introduzione

Il sostenuto processo di invecchiamento che sta caratterizzando l'Europa negli ultimi decenni fa sì che l'analisi dello stato di salute della popolazione anziana, suscettibile a deteriorarsi con l'età, diventi un fenomeno di studio di rilevante interesse.

Tra i numerosi fattori che influiscono sulle condizioni di salute degli individui, in particolar modo degli anziani, un posto preminente spetta ai fattori sociali, economici e culturali, oltre ai ben noti meccanismi biologici, genetici e ambientali, il cui contributo, però, nello spiegare le disuguaglianze di salute sarebbe più marginale. Alcuni epidemiologi italiani hanno affermato nel "Manifesto per l'equità" che "nessun fattore di rischio biologico, considerato singolarmente, ha sulla salute un'influenza paragonabile a quella esercitata dallo svantaggio socio-economico e culturale".

Numerosi studi, presenti in letteratura, mettono in evidenza il ruolo importante svolto dalle caratteristiche socio-economiche della popolazione nel determinare livelli differenziali di salute tra popolazioni europee (Mackenbach *et al.*, 2008) o tra diversi gruppi socio-economici nella stessa popolazione (Sihvonen, Kunst, Lahelma *et al.*, 1998). Nonostante i numerosi sforzi a livello internazionale da parte dell'OMS, attraverso campagne volte ad eliminare queste disparità, ancora oggi gli individui con un più basso status socio-economico risultano avere una maggiore probabilità di contrarre le malattie ed essere soggetti più frequentemente a disabilità, oltre ad una più bassa speranza di vita rispetto a chi gode di posizioni sociali privilegiate (Cambois *et al.*, 2001). La constatazione non solo del *gap* in termini di salute esistente tra ricchi e poveri, ma anche del suo accrescersi nel tempo (Mackenbach *et al.*, 2003) ha fatto sì che, negli ultimi anni, un interesse crescente fosse rivolto all'analisi delle disuguaglianze sociali nella salute, tanto da

¹ Il lavoro è frutto della collaborazione congiunta. Tuttavia, il paragrafo 1 è da attribuire a Odoardo Bussini, il paragrafo 2 e l'elaborazione dei dati a Donatella Lanari, il paragrafo 3 ad entrambi.

avere priorità assoluta anche nel quadro delle politiche sanitarie dei vari paesi europei.

Questo lavoro si propone di valutare le condizioni di salute degli anziani (individui con 65 anni e più) in Europa in relazione ai diversi fattori individuali che in letteratura sono considerati fortemente associati allo stato di salute, sia fisico e psichico dell'individuo, in primis la condizione socio-economica, ma anche l'età, il sesso, la struttura familiare e i cosiddetti *living arrangement*. Il modello interpretativo adottato nella nostra analisi considera, tuttavia, oltre alle caratteristiche individuali anche i fattori di contesto che insieme concorrono alla spiegazione del fenomeno. I diversi livelli di reddito e istruzione, la presenza o meno del supporto sociale e familiare, l'ambiente di vita e il paese di residenza, con le diverse politiche pubbliche di sanità adottate, sono tutti elementi che esercitano una forte influenza sulla salute degli individui.

Il nostro interesse rivolto agli anziani si avvale di una duplice considerazione. La prima nasce dall'aumento sostenuto della popolazione anziana e del conseguente invecchiamento demografico in Europa che deriva dall'incremento della durata della vita media, con i progressi maggiori che si concentrano soprattutto nelle età adulte e senili. La seconda deriva dal fatto che molti problemi di salute sopraggiungono proprio nelle età centrali e anziane della vita; in aggiunta al naturale e progressivo peggioramento delle condizioni di salute legato all'età, coloro che appartengono ai ceti sociali più svantaggiati hanno un rischio molto più elevato degli altri di star male, sia in termini di morbosità che di disabilità, anche gravi. Per questo motivo diventa di fondamentale interesse valutare il benessere psico-fisico di questo segmento a rischio della popolazione, privilegiando l'aspetto soggettivo, basato cioè sulla percezione che gli individui hanno del proprio stato di salute. Anche se questo approccio potrebbe sembrare irrazionale rispetto a chi preferisce analizzare la salute in senso stretto, oggettivo – presenza o assenza di malattie o disabilità – l'analisi della percezione negativa del proprio stato di salute da parte degli anziani è altresì importante poiché rappresenta un indicatore sintetico che rispecchia, oltre al valore soggettivo, anche un giudizio più obiettivo sulla effettiva precarietà delle proprie condizioni.

2. Dati e metodi

La banca dati utilizzata è la prima rilevazione del 2004 di SHARE (Survey of Health, Ageing and Retirement in Europe) un database multidisciplinare di microdati delle condizioni di salute, socio-economiche e relative al "network sociale e familiare" di individui con un'età sopra i 50 anni in 11 paesi europei. Il campione selezionato per la nostra analisi si riferisce agli individui con più di 65

anni, per i quali la probabilità del manifestarsi di disabilità o malattie croniche è maggiore.

Per valutare la salute degli anziani sono stati presi in considerazione diversi indicatori tra cui il livello di salute fisica e psichica percepito, la presenza di malattie croniche e di eventuali condizioni di non autonomia funzionale valutati attraverso la scala ADL (Activity of Daily Living). Più precisamente, gli indicatori proposti come variabili risposta sono: a) la percezione individuale della salute, secondo una scala che va da “molto buona” a “pessima”; b) la presenza di malattie croniche²; c) i sintomi depressivi³ riportati dall’intervistato; d) la presenza di eventuali disabilità⁴. Tutti gli indicatori sono stati dicotomizzati in modo da poter essere analizzati con modelli *Probit*. Tra le variabili indipendenti considerate, oltre a quelle già accennate, particolare importanza è attribuita allo status socio-economico e al connubio tra “forme abitative” (*living arrangement*) e la struttura familiare in cui vive l’anziano. Numerosi studi hanno evidenziato come le disuguaglianze nello stato di salute esistenti tra le persone possano essere ricondotte alle diverse condizioni socio-economiche, identificabili nel grado di istruzione e nel generale livello di benessere misurato attraverso il reddito percepito, in modo da costruire un indicatore di “povertà relativa” che tenga conto degli effetti psico-sociali delle disuguaglianze e che, in linea con la teoria avanzata da Marmot (con Wilkinson, 2001), provocherebbe insoddisfazione circa il proprio stato di salute, insicurezza, ansia, depressione. Ciò significa che il rischio di percepire un più basso livello di salute è maggiore nelle classi sociali più svantaggiate, persone meno istruite e più povere, e tra chi vive solo o in situazioni familiari meno protette. Il grado di istruzione influisce sul livello di salute sia in modo diretto che indiretto. Da un lato, chi ha un più alto grado d’istruzione è portatore di conoscenze, abilità, capacità di scelta maggiori mentre, dall’altro, si incrementano le possibilità di svolgere un lavoro più remunerativo che permette uno standard di vita qualitativamente migliore e la possibilità di accedere a prestazioni sanitarie più efficaci e tempestive. Al contrario, l’aumento dell’instabilità familiare si potrebbe accompagnare ad un peggioramento degli indici di salute delle persone interessate. A ciò si aggiungono le differenze

² In questa categoria sono state incluse: ipertensione, colesterolo alto, diabete, malattie croniche di lunga durata, asma, artrite, osteoporosi, cancro, malattia di Parkinson. Gli individui con malattie croniche sono coloro che hanno riferito due o più malattie.

³ I sintomi sono stati classificati secondo la scala EURO-D formata da 12 variabili tra cui pessimismo, pensieri suicidi, senso di colpa, insonnia, irritabilità, perdita di interesse, stanchezza, difficoltà a concentrarsi, pianto, mancanza di appetito. La linea soglia per essere definito depresso è “tre o più sintomi”.

⁴ Necessità di aiuto per svolgere una o più delle seguenti attività elementari: nutrirsi, vestirsi, fare il bagno, spostarsi dal letto alla poltrona, andare in bagno, incontinenza, ecc...

genetiche, di genere, i diversi stili di vita e, non ultimo, il contesto sociale, economico, politico del paese di appartenenza.

La metodologia proposta consiste in un'analisi probabilistica multivariata, in cui sono analizzati gli effetti differenziali nella percezione della salute tra individui ed eventuali differenze tra gli anziani residenti nelle macro-regioni dei paesi europei che hanno preso parte all'indagine SHARE. La presenza di fattori contestuali differenziali, anche a livello più disaggregato di macro-regioni, rappresenta un ulteriore elemento di variabilità che può generare un impatto diverso sulla salute degli individui. Al fine di includere questi aspetti, la robustezza delle analisi probabilistiche concernenti la percezione della salute è stata implementata attraverso l'utilizzo di modelli *multilevel*.

3. Risultati

La tabella 1 dimostra come l'Europa sia attraversata da profonde differenze per quanto riguarda la condizione degli anziani, in tutte le dimensioni analizzate: salute fisica, depressione, morbosità cronica e disabilità. Le differenze più intense si osservano per la salute percepita, dove si passa dal valore minimo del 24% di anziani (Svizzera) che dichiarano di avere una salute "non buona" al caso dell'Italia in cui si registra un dato pari al 63,5%. Oltre al paese elvetico, i residenti in Danimarca, Paesi Bassi e Svezia sembrano godere di un migliore stato di salute rispetto agli altri Stati europei.

Tabella 1 - Percentuale di anziani (65+ anni) con problemi di salute negli 11 Paesi europei

Paese	Salute percepita "non buona"	Depressione	Malattie croniche	ADL
Austria	50,1	35,0	27,4	13,7
Belgio	40,0	40,7	41,9	17,6
Danimarca	39,4	33,1	38,8	17,2
Francia	51,3	53,0	44,2	19,4
Germania	56,4	36,5	36,3	14,4
Grecia	55,6	48,8	42,9	17,0
Italia	63,5	51,8	47,7	17,8
Paesi Bassi	39,5	35,8	26,3	12,5
Spagna	61,6	56,7	46,6	19,2
Svezia	42,7	38,4	29,9	14,9
Svizzera	23,9	33,3	25,8	9,7

Fonte: nostra elaborazione, banca-dati Share, 2004

Note: Salute percepita non buona: discreta, cattiva, molto cattiva; Depressione: tre o più sintomi; Malattie croniche: due o più malattie; ADL: una o più limitazioni

Il profilo che emerge dalle statistiche descrittive sembra confermare l'esistenza di un gradiente nord-sud in Europa che vede i paesi dell'area mediterranea in netto svantaggio, evidenziando la particolare fragilità di tutte le dimensioni di salute considerate. Questa caratteristica è confermata in letteratura dal paradosso secondo cui le persone che vivono nell'Europa del Nord godono di una salute migliore e sono mediamente più ricche, sebbene chi vive nell'Europa del Sud sembra essere avvantaggiato in termini di aspettativa di vita, con una durata media più alta.

I risultati che emergono dai modelli logistici sono riportati nelle tabelle 2 e 3. E' importante precisare che, per vincoli di spazio, la probabilità di percepire un "cattivo stato di salute" è stata analizzata con maggior dettaglio utilizzando tre tipi di modelli (cfr. tabella 2). In particolare: 1) modello benchmark che analizza la relazione tra la probabilità di stare male e la condizione socio-economica (istruzione e risorse economiche percepite), utilizzando come variabili di controllo l'età, il sesso, i *living arrangement*, la rete familiare e la condizione di nativo/immigrato; 2) modello con effetti fissi che include le 51 macro-regioni europee (classificazione NUTS-1); 3) modello con effetti *random* con l'obiettivo di verificare la sensitività dei risultati. L'adozione del modello a effetti variabili si giustifica in considerazione dell'elevato numero di macro-regioni utilizzate e nell'ipotesi di un'accentuata variabilità tra le diverse regioni europee.

L'indagine empirica evidenzia la stretta relazione tra il "cattivo" stato di salute e la condizione socioeconomica delle persone con oltre 65 anni in Europa, confermando il ruolo fortemente discriminante nei confronti del fenomeno "salute percepita". Odds ratio superiori a 1 indicano uno stato di salute peggiore rispetto alla modalità di riferimento. La probabilità di essere in condizioni di salute "non buone" è maggiore tra gli anziani con un basso grado di istruzione (ORs=1.8), ma soprattutto tra chi percepisce che le risorse economiche a disposizione siano non sufficienti, o comunque precarie (ORs=2.2). In particolare, i modelli statistici mostrano l'importanza di fattori quali l'età e il fatto di essere immigrato, nel determinare una più alta probabilità nel percepire un "cattivo" stato di salute, mentre la presenza di una rete familiare stabile si riflette in una migliore soddisfazione. Controllando le differenze di genere, si riscontra una maggiore morbosità tra le donne rispetto alle persone di sesso maschile, a differenza di quanto accade per la mortalità. In particolare, l'interazione tra la variabile di genere e le forme abitative rivela che sono le persone che vivono sole, in particolare le donne, a sperimentare uno stato di salute peggiore rispetto all'altro sesso o a chi vive in coppia. Analizzando, infatti, la variabile di interazione sembrerebbe che gli uomini abbiano una probabilità di stare male inferiore di circa il 20% rispetto alle donne. Per testare la robustezza dei parametri ottenuti, abbiamo stimato un modello ad effetti fissi includendo variabili *dummies* per le 51 regioni Nuts-1 (Modello II).

Tabella 2 - Modello di regressione logistica per valutare la relazione tra la probabilità nel percepire un cattivo stato di salute e le variabili indipendenti selezionate

Variabili	Parametri	Modello I	Modello II	Modello III
Livello di istruzione	Basso	1.79***	1.80***	1.78***
	Intermedio	1.41***	1.28***	1.31***
	Alto (ref)			
Risorse economiche percepite	Basso	2.53***	2.20***	2.22***
	Intermedio	1.42***	1.33***	1.35***
	Alto (ref)			
Età	coorti 1900-1919 (ref)			
	coorti 1920-1929	0.76***	0.74***	0.73***
	coorti 1930-1939	0.47***	0.45***	0.44***
Sesso*Living arrangement	Femmina*vive sola (ref)			
	Maschio*vive solo	0.78***	0.80***	0.81***
	Femmina*vive in coppia	0.98	0.92	0.94
	Maschio*vive in coppia	0.80***	0.78***	0.79***
Figli	nessun figlio(ref)			
	uno	1.20**	1.22**	1.22**
	due	0.98	1.00	1.00
	tre o più	1.06	1.11	1.08
Genitori	Nessun genitore (ref)			
	almeno un genitore vivo	0.65***	0.64***	0.64***
Status nativo/immigrato	nativo (ref)			
	immigrato	1.25***	1.17*	1.20**
NUTS - 1			$\chi^2_{(50)}=273***$	$\rho=0.046***$

Nota: Livelli di significatività: *** < 0.001; ** < 0.05; * < 0.1.

Modello I: benchmark; Modello II: con effetti fissi; Modello III: con effetti random.

In primo luogo, emerge che l'ipotesi nulla di assenza nelle differenze della salute percepita tra le regioni europee, come atteso, è rigettata agli usuali livelli di significatività statistica ($\chi^2(50)=273$ p -value=0.000).

Inoltre, le stime riportate in questa seconda colonna mostrano che molta della variabilità tra macro-regioni è imputabile alle differenze nelle condizioni socioeconomiche. Con riferimento al modello *benchmark*, la risposta della salute percepita di coloro che hanno un livello di istruzione intermedio risulta ridotta di oltre dieci punti percentuali, mentre per i percettori più bassi di reddito, controllando per gli effetti regionali, la probabilità di percepire cattive condizioni di salute è ridotta addirittura di 30 punti percentuali (ORs da 2.53 a 2.2). Nell'ultima colonna della tabella sono riportate, infine, le stime *multilevel*. Lo scopo di questa ulteriore analisi consiste nel valutare la necessità di includere variabili contestuali nello spiegare le differenze nello stato di salute. Il fatto che i parametri stimati siano pressoché identici al modello ad effetti fissi, associato al basso valore della varianza intra-classe (ρ circa 5%), - sebbene statisticamente

significativo – non suggerisce ulteriori approfondimenti nel caratterizzare la variabilità residua attraverso variabili contestuali.

I risultati ottenuti utilizzando come variabile risposta del modello ad effetti fissi gli altri indicatori di salute relativi ai sintomi depressivi, malattie croniche e disabilità (cfr. tabella 3) confermano quanto emerso dall'analisi precedente e, cioè, la correlazione positiva tra le cattive condizioni di salute, sia fisica che mentale, e un basso status socio-economico degli anziani⁵.

Tabella 3 - Modello di regressione logistica per valutare le relazioni significative tra la probabilità di essere depresso, di avere malattie croniche, disabilità e le variabili indipendenti selezionate

		<i>Depressione</i>	<i>Malattie croniche</i>	<i>ADL</i>
<i>Livello di istruzione</i>	Basso	+	+	+
	Intermedio	+		
	Alto (ref)			
<i>Risorse economiche percepite</i>	Basso	+	+	+
	Intermedio	+		
	Alto (ref)			
<i>Età</i>	<i>coorti 1900-1919 (ref)</i>			
	coorti 1920-1929	-	-	-
	coorti 1930-1939	-	-	-
<i>Sesso*Living arrangement</i>	<i>Femmina*vive sola (ref.)</i>			
	Maschio*vive solo	-	-	-
	Femmina*vive in coppia	-		-
	Maschio*vive in coppia	-	-	-
<i>Figli</i>	<i>nessun figlio(ref)</i>			
	uno		+	
	due		+	
	tre o più		+	
<i>Genitori</i>	<i>Nessun genitore (ref.)</i>			
	almeno un genitore vivo		-	-
<i>Nativo/immigrato</i>	<i>nativo (ref)</i>			
	immigrato	+		

Note: è stato riportato il segno della relazione tra il rischio di percepire un cattivo stato di salute e le suddette variabili.

⁵ In questo caso l'impatto delle variabili indipendenti sulla salute è stato sintetizzato riportando il segno della relazione tra il rischio di percepire un cattivo stato di salute e le suddette variabili.

Riferimenti bibliografici

- Cambois, E., Robine J. M., e Hayward M.D. 2001. *Social inequalities in disability-free life expectancy in the French male population, 1980-1991*, Demography 38, pp 513-24.
- Mackenbach, J.P., Stirbu I., Roskam, A-J.R., Schaap, M.M., Menvielle, G., Leinsalu, M., e Kunst A.E. 2008. *Socioeconomic inequalities in health in 22 European Countries*, The new England Journal of medicine, 358, pp 2468-81.
- Mackenbach, J.P., Bos V., Andersen O., Cardano M., Costa G., Harding S., Reid A., Hemström Ö., Valkonen T. e Kunst A.E. 2003. *Widening socioeconomic inequalities in mortality in six Western European countries*. International Journal of Epidemiology, 32, pp 830-7.
- Marmot M., Wilkinson RG. 2001. *Psychosocial and material pathways in the relation between income and health: a response to Lynch*, BMJ, 322, pp 1233-6.
- Sihvonen A-P, Kunst A, Valkonen T, Lahelma E, Mackenbach J. 1998. *Socioeconomic inequalities in health expectancy in Finland and Norway in the late 1980s*, Social Science and Medicine, 47, pp 303-15.

SUMMARY

Socioeconomic conditions have been found to be related to self-rated health, which is an established predictor of morbidity and mortality. In this paper we investigated the associations between socio-economic status and self-perceived health among older people in European countries, who most probably face health problems. We analyzed four indicators of health status based on self-reports concerning physical and mental health, chronic disease and disability, using the SHARE data. Results showed that poor self-rated health was most common among people who experienced economic hardship and with a low-educational level. Women living alone tended to report poor health, confirming the association between social isolation and poor health. Moreover, results from logistic estimations, with fixed-effects included, evidenced the presence of some variability across the European macro-regions considered.

Odoardo BUSSINI, Professore ordinario di Demografia, Università di Perugia, Dipartimento di Economia, Finanza e Statistica, obussini@stat.unipg.it
Donatella LANARI, Dottoranda di Ricerca in Demografia, Università di Roma "La Sapienza", Dipartimento di Studi Sociali, Economici, Attuariali e Demografici, donatella.lanari@uniroma1.it

LA MOBILITÀ INTERNA IN ITALIA: LE PRINCIPALI CARATTERISTICHE DEI CAMBIAMENTI DI RESIDENZA

Oliviero Casacchia, Lorenzo Cassata, Piero Giorgi,
Cecilia Reynaud, Salvatore Strozza

1. Introduzione

Come sottolineato da diversi autori (Pugliese, 2006 e De Santis, 2008) la mobilità interna sembra ormai essere uscita dalla fase di stagnazione degli anni '90. Fattori principali di questa nuova vivacità sono individuabili nel sensibile aumento della propensione allo spostamento dei giovani italiani diplomati e laureati residenti soprattutto nelle regioni del Sud (Piras, 2007), nella ripresa delle migrazioni a corto raggio nelle province del centro-Nord (Bonifazi, 2009), nel contributo crescente portato dalla componente straniera al volume degli spostamenti (ISTAT, 2008).

Considerata l'eterogeneità delle popolazioni che contribuiscono con sempre maggior vigore al fenomeno, diventano particolarmente importanti studi sulle caratteristiche dei flussi migratori interni e quelle degli individui che li compiono. Nel presente lavoro ci si propone un esame delle relazioni tra le caratteristiche attuali dei flussi, le loro tipologie e direzioni, e quelle degli individui, quali il sesso, l'età e la cittadinanza, per arrivare a definire differenti modelli migratori.

2. Antefatto: cresce la popolazione straniera e il suo apporto alla mobilità interna

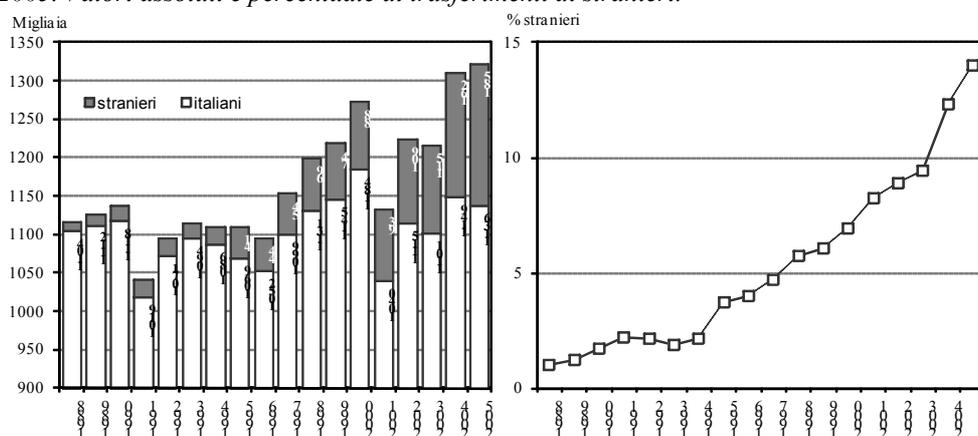
La popolazione straniera ha sperimentato una crescita impetuosa passando da circa 360 mila residenti al censimento del 1991 a quasi un milione e 350 mila a quello successivo, per arrivare alla data più recente (inizio 2008) a quasi 3 milioni e mezzo di residenti in Italia, con un impatto sul totale della popolazione del paese che da meno dell'1% si è avvicinato al 6% in poco più di 15 anni.

Gli spostamenti interni al territorio nazionale hanno ovviamente risentito di questa dinamica tanto che una parte fortemente crescente dei trasferimenti di residenza ha interessato la componente straniera: nel 1991 meno di 23 mila, nel 1995 oltre 41 mila, nel 2000 circa 88 mila, per arrivare nel 2005 a ben 185 mila trasferimenti interni (fig. 1). Il numero degli spostamenti degli italiani non ha fatto registrare modificazioni di rilievo nel periodo 1988-95. Solo nel quinquennio successivo si nota un incremento apprezzabile per tornare negli ultimi anni quasi ai livelli iniziali, tanto che la crescita più recente della mobilità interna complessiva

risulta dovuta proprio al contributo degli stranieri che in tutto il periodo osservato hanno una propensione a muoversi sul territorio da 2 a 3 volte quella degli italiani.

La mobilità degli stranieri in Italia contribuisce notevolmente all'incremento delle migrazioni interne specialmente con riferimento ad alcune direttrici (Istat, 2008). Più rilevante risulta il contributo della componente straniera nel flusso che dal Mezzogiorno si dirige verso il Centro-nord o in quello in uscita dalle aree metropolitane verso i comuni della corona.

Figura 1 – Trasferimenti di residenza tra comuni italiani per cittadinanza. Italia, 1988-2005. Valori assoluti e percentuale di trasferimenti di stranieri.



Fonte: ns. elaborazioni su dati Istat.

L'esame dei percorsi degli stranieri all'interno dell'Italia, nonché il confronto con quanto osservato per gli italiani, appare particolarmente complesso per la notevole eterogeneità territoriale del paese e per la grande variabilità del comportamento migratorio, il quale manifesta intensità e direttrici diverse secondo la composizione per sesso ed età della popolazione e l'articolazione per area di origine della componente straniera. È opportuno ricordare che la popolazione straniera residente in Italia risulta particolarmente articolata, senza la possibilità di individuare poche nazionalità nettamente prevalenti, e che le comunità immigrate hanno caratteristiche e comportamenti non di rado specifici. L'analisi proposta consente di giungere ad alcuni risultati significativi inquadrando la problematica della mobilità territoriale degli stranieri all'interno di quella complessiva.

3. Dati e metodi

Nel lavoro sono stati utilizzati i dati delle iscrizioni e cancellazioni anagrafiche, relativi a tutti i comuni italiani, per l'anno 2001 ed il Censimento della popolazione del 2001. Il dato anagrafico, pur non esente da problemi (Cassata, 2005) è una delle principali fonti per lo studio della mobilità in Italia, anche in virtù della sua esaustività. Gli individui sono stati raggruppati sulla base di sesso, età e cittadinanza; per gli italiani è stata considerata la ripartizione di nascita (nord, centro e sud). Per la cittadinanza sono stati utilizzati 5 gruppi: italiani, stranieri di Paesi a sviluppo avanzato (PSA), Europa dell'Est, Nord Africa e Resto dei Paesi a forte pressione migratoria (PFPM). Allo scopo di distinguere meglio la popolazione italiana, è stato considerato il luogo di nascita, individuando ulteriori 3 gruppi: gli italiani nati in regioni del nord, del centro e del sud. Per l'età sono state utilizzate le classi 0-19, 20-39, 40-59, 60+. Nella classe d'età 60+ la cittadinanza è stata classificata in soli due gruppi: italiani e stranieri dei PFPM, sia per ridurre il numero dei gruppi, sia perché alcuni flussi risultavano troppo esigui. Dalla combinazione delle classificazioni utilizzate si ottengono 50 gruppi, 26 per gli stranieri (4 cittadinanza x 2 sessi x 3 età + 2 età 60+) e 24 per gli italiani (3 ripartizione di nascita x 2 sessi x 4 età), sui quali viene condotta l'analisi.

Per ogni gruppo utilizzato come unità di analisi sono stati calcolati gli indicatori di mobilità. I flussi sono stati caratterizzati sia attraverso la tipologia, distinta in intraprovinciali (diretti all'interno della stessa provincia), diretti (si/no) verso un polo metropolitano e verso un comune di prima o seconda fascia delle aree metropolitane, sia attraverso la direzione, identificata attraverso le 5 ripartizione di origine e di destinazione. Gli indicatori utilizzati sono stati quindi: tasso di mobilità dato dal rapporto tra individui che hanno sperimentato un movimento anagrafico e la popolazione residente al censimento; % flussi intraprovinciali; indice di concentrazione dei flussi rispetto alla provincia di destinazione; distanza media tra i capoluoghi delle province di origine e di destinazione dei flussi; dipendenza dai comuni poli metropolitani (flussi verso i poli delle aree metropolitane sui flussi totali); dipendenza dagli altri comuni metropolitani; quota di flussi tra le ripartizioni (in totale 16 direttrici differenti per origine e destinazione considerando anche i cambiamenti di residenza all'interno della stessa ripartizione).

Su questi indicatori è stata condotta un'analisi fattoriale per verificare le eventuali relazioni tra alcune modalità dei flussi e le caratteristiche degli individui che li compiono. Si è applicata una Analisi in Componenti Principali (ACP) sulla matrice indicatori per unità di analisi (22 x 50). La successiva applicazione della cluster analysis sui fattori derivati dalla ACP ha permesso di raggruppare i trasferimenti anagrafici mediante i fattori che sintetizzano le variabili utilizzate e definire i "modelli migratori" più diffusi.

4. I risultati dell'analisi

L'applicazione dell'analisi in componenti principali ha permesso di estrarre quattro fattori (con autovalore superiore all'unità), che complessivamente spiegano il 91% della variabilità. Già i primi due fattori sembrano condurre a ottimi risultati per l'interpretazione del fenomeno poiché catturano rispettivamente il 42,8 e il 27,3% della varianza complessiva.

Il primo fattore contrappone una forte concentrazione dei flussi alla distanza media: si oppongono così i gruppi di individui che scelgono poche destinazioni e in genere più vicine, a quelli che si disperdono maggiormente compiendo però spostamenti mediamente più lunghi. Alla prima di queste due situazioni si associano i flussi che attraversano le due ripartizioni del Nord, mentre alla seconda tutti i flussi che riguardano il Sud. Il secondo fattore sintetizza la dipendenza dai poli metropolitani che è fortemente associata ai flussi che interessano il Centro, mettendo in luce il forte ruolo attrattivo, ma anche espulsivo, di città come Roma e Firenze; a questo fattore si associa anche la dipendenza dalle fasce metropolitane. L'elevata mobilità si oppone ad un'alta percentuale di flussi intraprovinciali nel terzo fattore, ad indicare che lo spostamento all'interno della stessa provincia non costituisce una migrazione tipica dei gruppi più mobili. Sul semi-asse negativo del quarto fattore, di incerta fisionomia, si collocano le migrazioni tra le ripartizioni del nord (tab. 1).

Sul primo piano fattoriale si può notare come gli italiani nati al sud e al centro si differenziano dagli stranieri e dagli italiani nati al nord (fig. 2a). Ciò accomuna le caratteristiche dei flussi degli stranieri tra di loro e di quelli degli italiani del nord. L'analisi condotta considerando gli italiani non distinti per luogo di nascita faceva giungere ad una netta differenziazione tra italiani e stranieri, mentre l'informazione aggiuntiva della ripartizione di nascita degli italiani ha portato a verificare la nota differenziazione tra le ripartizioni rispetto alla mobilità interna. Si è inoltre evidenziato come la mobilità degli italiani nati al nord sia più simile a quella degli stranieri piuttosto che a quella degli italiani nati nelle altre due ripartizioni. Il primo fattore, infatti, oppone i flussi di individui nati nel Sud Italia, sul semiasse negativo, a quelli dei nati nel Nord, a cui si associano gli individui dei PSA e gli uomini giovani degli altri PFPM. Questi ultimi gruppi di individui compirebbero spostamenti con una distanza più contenuta, prevalentemente tra le due ripartizioni del Nord, e con una elevata concentrazione delle destinazioni; al contrario degli individui nati nel Sud Italia che consumano distanze maggiori e più disperse sul territorio. Il secondo fattore oppone invece gli italiani nati nel centro, caratterizzati da una forte dipendenza dalle aree metropolitane, ai Nord Africani, che quindi sembrano non essere attratti dai comuni metropolitani (fig. 2a). Gli uomini adulti (20-39 e 40-59 anni) degli altri PFPM e dell'Est Europa si collocano sul semiasse positivo del terzo fattore e si contraddistinguono per una elevata mobilità, in contrapposizione con gli italiani, maschi e femmine, nati al nord di tutte le classi di

età considerate, caratterizzati da una più elevata percentuale di flussi intraprovinciali, così come emerge dalla letteratura sulla mobilità interna (fig. 2b). Il quarto fattore oppone, anche se meno nettamente, i gruppi di individui Nord Africani, non interessati dai flussi nelle ripartizioni del nord, ai gruppi di individui di altri PFPM e dei PSA, che quindi gravitano maggiormente in questa ripartizione.

Tabella 1 – Pesi fattoriali.

Indicatori elementari	Fattore 1	Fattore 2	Fattore 3	Fattore 4
Indice di concentrazione	0,83	0,28	0,12	-0,22
% flussi dal Nord-Ovest al Nord-Est	0,65	-0,51	0,25	-0,44
% flussi dal Nord-Est al Nord-Ovest	0,63	-0,52	0,26	-0,45
% flussi dal Centro al Centro	0,16	0,92	0,09	0,30
Dipendenza dai comuni metropolitani	-0,20	0,85	-0,06	-0,32
% flussi dal Centro al Nord-Est	0,52	0,71	0,25	-0,12
% flussi dal Sud al Centro	-0,61	0,68	0,10	-0,10
% flussi dal Centro al Nord-Ovest	0,50	0,64	-0,04	-0,30
Dipendenza dai comuni fasce metropolitane	-0,60	0,61	-0,24	-0,33
% flussi dal Nord-Est al Centro	0,41	0,21	0,80	0,17
Tasso di mobilità	0,29	-0,48	0,70	0,17
% flussi dal Nord-Ovest al Centro	0,49	0,49	0,66	-0,11
% flussi intra-provinciali	0,50	-0,08	-0,74	0,36
% flussi dal Nord-Ovest al Nord-Ovest	0,59	-0,66	-0,34	-0,27
% flussi dal Nord-Est al Nord-Est	0,65	-0,71	-0,10	0,03
% flussi dal Nord-Est al Sud	-0,65	-0,49	0,47	0,17
% flussi dal Nord-Ovest al Sud	-0,81	-0,41	0,36	-0,07
% flussi dal Sud al Nord-Ovest	-0,84	-0,18	-0,23	-0,30
Distanza media	-0,85	-0,25	0,43	-0,11
% flussi dal Sud al Nord-Est	-0,88	-0,24	-0,12	-0,23
% flussi dal Centro al Sud	-0,93	0,13	0,24	0,12
% flussi dal Sud al Sud	-0,95	-0,15	-0,02	-0,12
% variabilità spiegata	42,81	27,26	14,45	6,18

Sono stati individuati sei *cluster*, che in generale rispecchiano la classificazione per cittadinanza degli stranieri e per luogo di nascita degli italiani: le unità statistiche appartenenti ad una stessa area di cittadinanza confluiscono in un unico *cluster* indipendentemente dal sesso e gruppo di età ad eccezione degli stranieri ultrasessantenni.

Il primo *cluster* è costituito dagli stranieri dei PSA a cui si uniscono le straniere donne dei PFPM con più di 60 anni; i flussi più rilevanti di questo *cluster* interessano l'Italia centrale e il Nord, in particolare tra il Centro e il Nord, e privilegiano poche destinazioni. Gli stranieri degli altri PFPM, con meno di 60 anni, costituiscono il secondo *cluster* caratterizzato dai flussi tra le due ripartizioni del Nord e da queste

da una minore quota di spostamenti che dal Nord si dirigono verso il Centro, tipici dei nati al Centro, o verso il Sud, più prevalenti tra i nati al Sud, che quindi sembrano maggiormente essere migrazioni di ritorno. Gli italiani nati al Centro sono attratti dalle aree metropolitane, con spostamenti più ridotti e una minore propensione alla mobilità e fanno registrare una più elevata quota di flussi che interessano la loro ripartizione. Infine gli italiani nati al Sud compiono spostamenti mediamente più lunghi e diretti verso un maggior numero di destinazioni, sono attratti dai comuni appartenenti alle fasce metropolitane e si spostano meno all'interno della stessa provincia; fanno registrare una più elevata quota di tutti i flussi che interessano la loro ripartizione e la minor quota di flussi che avviene tra le altre ripartizioni.

Emerge ovviamente per gli italiani un legame tra la ripartizione di nascita e la quota di flussi che interessa la medesima ripartizione, ma innanzitutto non sempre i gruppi di individui nati in una ripartizione sono caratterizzati da una maggiore quota di tutti i flussi che interessano quella ripartizione – questo è vero solo per i nati al Sud – e in secondo luogo ulteriori analisi condotte senza la distinzione per luogo di nascita o senza le quota di flussi tra le ripartizioni confermano sostanzialmente i risultati ottenuti.

5. Conclusioni

La mobilità interna degli stranieri rappresenta un aspetto importante del processo di adattamento degli immigrati al territorio del paese di accoglimento, consentendo di rimuovere alcune difficoltà incontrate e di cogliere le occasioni che si presentano. La maggiore propensione degli stranieri a spostarsi rispetto agli italiani sembra essere riconducibile al carattere selettivo degli immigrati: concentrati nelle fasce d'età più mobili e con una maggiore attitudine alla mobilità geografica (avendo già sperimentato almeno una migrazione internazionale). Inoltre, nel caso degli stranieri sembrano venir meno dei freni più evidenti per gli autoctoni, come il sostegno della rete familiare e la frequente proprietà dell'abitazione in cui vivono.

Una lettura dettagliata delle caratteristiche dei flussi e degli attori coinvolti ha permesso di aggiungere alla dicotomia Sud-Nord – presente anche nella mobilità delle persone di cittadinanza non italiana – elementi aggiuntivi che differenziano le migrazioni interne degli stranieri da quelle degli autoctoni. Nel 2001 a determinare le differenze non sembrano essere tanto le caratteristiche demografiche (genere ed età) quanto l'origine dei diversi gruppi (cittadinanza e ripartizione di nascita per gli italiani). Appare interessante notare come gli stranieri si caratterizzino non solo per la più elevata propensione a spostarsi, ma anche per la maggiore concentrazione delle destinazioni, la minore dipendenza dalle aree metropolitane e la più marcata rilevanza delle migrazioni che hanno origine e destinazione nelle diverse realtà dell'Italia settentrionale. Anche le differenze all'interno del variegato mondo della

popolazione straniera sembrano di non secondaria importanza, in questo caso confermando strategie e comportamenti differenti tra le molteplici collettività immigrate. Se e quanto la situazione sia cambiata negli ultimi anni e che ruolo svolga il modello distributivo sul territorio italiano sulle diverse traiettorie sono alcuni degli ulteriori aspetti che l'analisi proposta spinge ad investigare.

Riferimenti bibliografici

- Bonifazi C. 2009. "Le migrazioni interne meridionali: vecchi e nuovi ritardi" in *neodemos.it*
- Cassata L. (2005), *Modelli di mobilità degli anziani in Italia*, tesi di dottorato in demografia
- De Santis G. 2008. "La donna è mobile (e l'uomo anche)" in *neodemos.it*
- ISTAT 2008. *Rapporto annuale. La situazione del Paese nel 2007*, Roma.
- Pugliese E. 2006. *L'Italia tra migrazioni internazionali e migrazioni interne*, Il Mulino, Bologna, seconda edizione.
- Piras R., Melis S. 2007. "Evoluzione e tendenze delle migrazioni interne", *Economia Italiana*, n. 2, pp. 437-461.
- Rossi F. e Strozza S. 2007. "Mobilità della popolazione, immigrazione e presenza straniera", in GCD-SIS, *Rapporto sulla popolazione. L'Italia all'inizio del XXI secolo*, il Mulino, Bologna.

SUMMARY

One of the factors of the increase of recent internal mobility in Italy largely due to the high increase of the foreign population. By using 2001 individual data we have defined groups of individuals who changed residence, on the basis of their characteristics: nationality, place of birth for the Italians, sex and age. For each group some indicators were calculated: rate of mobility, percentage of flows intra-province, concentration of flows, dependence on metropolitan areas, percentage of flows between zones. A cluster analysis was applied to the factors extracted by the principal component analysis on the indicators and it resulted in six clusters. In 2001 the differences do not seem to be determined by demographic characteristics but by the origin of the different groups. The foreign population is characterized not only by the highest propensity to travel, but also by the greater concentration of destinations, a reduced dependence on metropolitan areas and a higher propensity of migration as origin and destination, in the north of Italy.

Oliviero CASACCHIA, Professore associato di Demografia, Università di Roma La Sapienza.

Lorenzo CASSATA, Ricercatore, ISTAT.

Piero GIORGI, Professore associato di Demografia, Università di Teramo.

Cecilia REYNAUD, Ricercatore di Demografia, Università di Roma Tre.

Salvatore STROZZA, Professore ordinario di Demografia, Un. di Napoli Federico II.

LE NASCITE FUORI DAL MATRIMONIO: TENDENZE IN ATTO E PRINCIPALI CARATTERISTICHE

Cinzia Castagnaro, Antonella Guarneri

1. Introduzione

Uno dei fenomeni che meglio descrive i cambiamenti avvenuti nei comportamenti familiari e riproduttivi in Italia è rappresentato dall'aumento delle nascite fuori dal matrimonio. Questo aspetto – insieme alla crescente instabilità coniugale e all'aumento delle età sia al primo matrimonio che alla nascita del primo figlio – sembra determinare la lenta convergenza dell'Italia verso modelli familiari meno tradizionali.

I recenti cambiamenti in termini di cadenza e intensità della fecondità richiedono di approfondire il contesto in cui si verifica l'evento nascita, con particolare riferimento alle diverse tipologie familiari. Per definire queste tipologie e per analizzarne le caratteristiche differenziali sono stati utilizzati i dati dell'*Indagine campionaria sulle nascite* condotta dall'Istat.

2. Quadro di riferimento

Il sistema di rilevazioni sulle nascite è stato recentemente modificato per adeguarsi alla normativa in materia di denuncia di nascita¹ che ha imposto la soppressione della rilevazione individuale delle nascite di fonte Stato Civile, effettuata dall'Istat fin dal 1926. Nel nuovo sistema integrato di rilevazioni sulle nascite la *Rilevazione degli Iscritti in Anagrafe per Nascita*, attiva a partire dal 1° gennaio 1999, permette il monitoraggio dell'andamento della fecondità nel nostro Paese per quel che concerne la quantificazione del fenomeno rispetto alle principali variabili anagrafiche. Attraverso questa fonte, è possibile quantificare le nascite al di fuori del vincolo istituzionale del matrimonio. Queste, nel 2007, rappresentano il 20,7% del totale dei nati vivi, risultando quadruplicate rispetto al 1985 (5,4%).

Questi mutamenti sono in parte da mettere in relazione con l'evoluzione normativa che si è verificata in Italia. L'introduzione nel 1975 del nuovo Diritto di famiglia ha consentito il riconoscimento dei bambini nati al di fuori del matrimonio equiparandoli ai nati all'interno del matrimonio. Da quel momento l'incremento

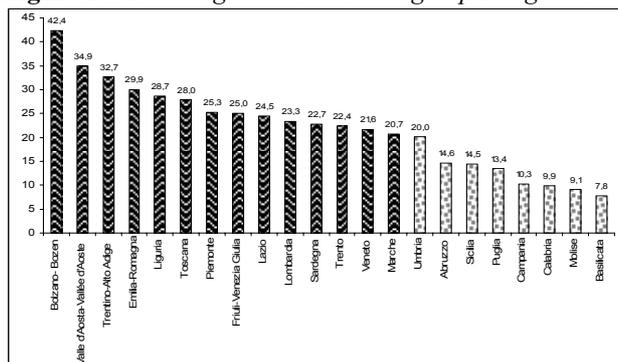
¹ Cfr. legge 127/97 sulla "Semplificazione amministrativa" del 17 Maggio 1997 e dei suoi successivi regolamenti di attuazione.

non si è più arrestato. Nel 1995, il 99% dei nati vivi fuori del matrimonio è stato riconosciuto da almeno un genitore e di questi l'85% è stato riconosciuto da entrambi i genitori. I dati più recenti confermano l'ulteriore consolidamento di questa tendenza. Nel 2005 il 95% dei nati al di fuori del matrimonio, riconosciuti almeno dalla madre, è stato riconosciuto anche dal padre.

Le nascite fuori dal matrimonio sono particolarmente diffuse nell'Italia del Nord (rispettivamente 24,3% nel Nord-ovest e 26,1% nel Nord-est) e del Centro (24,7%). Si nota una netta contrapposizione tra le regioni del Centro-nord e quelle del Mezzogiorno che si posizionano su livelli molto più bassi (il minimo è rappresentato dal 7,8% della Basilicata). Tra le regioni centrali, solo l'Umbria presenta un valore inferiore a quello medio nazionale (Figura 1).

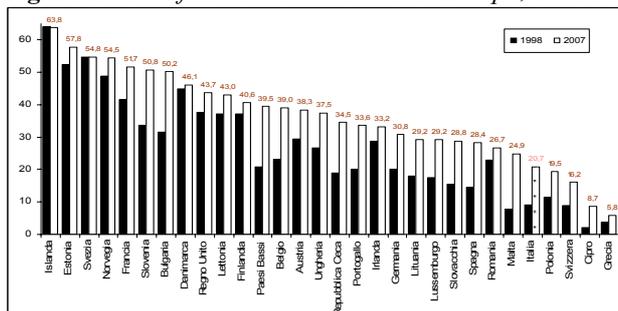
Tuttavia anche le regioni italiane 'meno tradizionaliste' mostrano percentuali piuttosto contenute se confrontate con quanto accade negli altri paesi europei (Figura 2).

Figura 1 – Nati da genitori non coniugati per regione nel 2007. Valori percentuali.



Fonte: elaborazioni su dati Istat

Figura 2 – Nati fuori dal matrimonio in Europa, 1998 e 2007. Valori percentuali (a).



Fonte: dati Eurostat

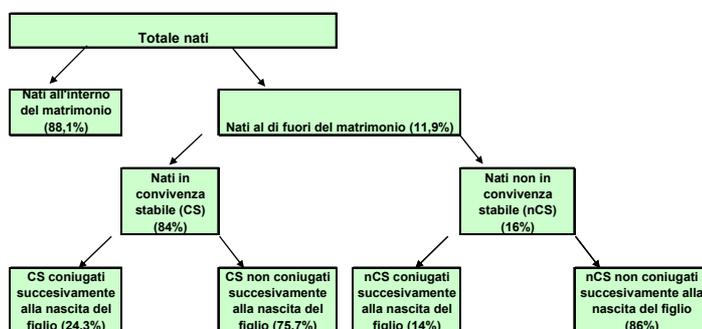
(a) Per Regno Unito, Irlanda e Spagna i dati si riferiscono al 2006 e non al 2007.

Infatti, se i bassissimi livelli di fecondità italiana rientrano a pieno titolo in un andamento di tipo *post-transizionale*, le trasformazioni familiari verso modelli comportamentali meno tradizionali sembrano ancora lontane dal completamento di un processo di convergenza. Anche qui le contrapposizioni sono molto marcate. L'Italia si posiziona agli ultimi posti in Europa per quota di nati fuori dal matrimonio – insieme alla Svizzera, agli altri paesi dell'Europa meridionale e ad alcuni paesi UE di nuova adesione.

3. Principali caratteristiche delle nascite fuori del matrimonio

Nell'ambito dell'attuale sistema integrato sulle nascite all'*Indagine Campionaria sulle Nascite*² spetta il ruolo di rilavare, oltre a notizie di carattere socio-sanitario su gravidanza e parto, informazioni indispensabili per comprendere le dinamiche di fecondità fornendo un utile supporto nella programmazione e nel monitoraggio di interventi di *policy* rivolti alle famiglie con figli al di sotto dei due anni. Sono stati utilizzati i dati del 2005 il cui campione era composto da 50.000 madri estratte dall'archivio degli iscritti in anagrafe per nascita nel 2003 (circa il 10% dell'universo di riferimento). Innanzitutto, appare interessante tracciare un diagramma riepilogativo in cui, a passaggi successivi, si scompone il collettivo dei nati (Prospetto 1).

Prospetto 1 – Diagramma riepilogativo di scomposizione del collettivo dei nati dell'*Indagine Campionaria sulle Nascite* per scelte familiari della madre.

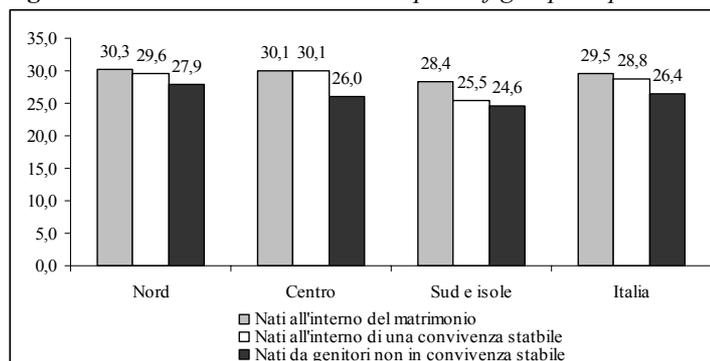


² La tecnica di rilevazione è C.A.T.I. (Computer Assisted Telephone Interview). Il questionario *short form* viene somministrato a tutte le madri intervistate mentre un modulo di approfondimento (*long form*) ha interessato poco più di un terzo del campione. Finora se ne sono realizzate due edizioni, una nel 2002 e una nel 2005. Nel 2010 si svolgerà la terza edizione dell'indagine che abbinerà all'ormai tradizionale approccio trasversale anche una nuova indagine di ritorno su un campione di madri già intervistate nel 2005.

La prima suddivisione mira proprio a distinguere i nati all'interno del matrimonio (88,1%) da quelli nati al di fuori di questo (11,9%)³. Si va poi ad approfondire il collettivo dei nati fuori dal matrimonio distinguendo tra nascite avvenute o meno all'interno di una convivenza stabile. Infine, si è pensato di analizzare una teorica inversione delle tappe tradizionali di formazione della famiglia andando a verificare quante donne tra quelle in convivenza stabile e quante tra quelle non in convivenza stabile si siano sposate nei due anni successivi alla nascita del figlio.

Va osservato come i nati fuori del matrimonio siano più frequentemente primogeniti (il 75% contro il 47,3% dei nati all'interno del matrimonio). Rispetto alla scelta di avere il primo figlio, le coppie che vivono in convivenza stabile presentano interessanti analogie con le coppie coniugate rispetto all'età media delle madri alla nascita del figlio e alla distanza tra la nascita del bambino e l'inizio rispettivamente della convivenza o del matrimonio (Figura 3). Diverso è il caso delle madri non in convivenza stabile che alla nascita del primo figlio hanno in media 3 anni di meno (26,4 anni al primo figlio contro i 29,5 delle coniugate) che diventano 4 anni in meno per le residenti al Sud e nelle isole.

Figura 3 – Età media alla nascita del primo figlio per ripartizione, anno 2005.



Fonte: Istat, Indagine Campionaria sulle Nascite

Per quanto riguarda l'intervallo tra l'inizio della convivenza o la data di matrimonio e la nascita del primo figlio va osservato che rispettivamente il 43% o il 46% dei figli primogeniti nasce a una distanza inferiore ai due anni da questi eventi; circa l'85% entro i primi 5 anni di convivenza o di matrimonio (Figura 4).

³ La percentuale dei nati fuori del matrimonio appare leggermente più bassa rispetto a quella osservata nell'universo di riferimento (pari a 14,5% nell'anno 2003).

Figura 4 – L'intervallo tra inizio della convivenza/ matrimonio e nascita del primo figlio, anno 2005.



Fonte: Istat, Indagine Campionaria sulle Nascite

4. Analisi delle scelte familiari e riproduttive

Per approfondire l'influenza delle caratteristiche individuali e di contesto sull'esperienza della maternità è stato adottato anche un approccio di tipo interpretativo. Sono così stati applicati due diversi modelli logistici⁴ per analizzare le caratteristiche differenziali delle madri che:

- hanno avuto figli in una convivenza stabile rispetto a quelle che li hanno avuti all'interno del matrimonio;
- hanno avuto figli non in convivenza stabile rispetto a quelle che li hanno avuti in una convivenza stabile.

Le variabili esplicative considerate per entrambi i modelli sono: titolo di studio della madre, età della madre al parto, ordine di nascita, titolo di godimento dell'abitazione e ripartizione geografica di residenza.

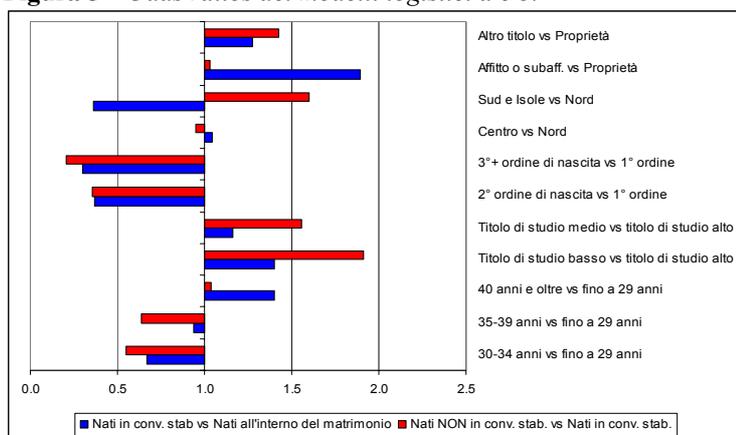
Tra i fattori che maggiormente influenzano la scelta dei genitori di avere un figlio all'interno di una convivenza stabile rispetto ad averlo all'interno del matrimonio emerge il livello di istruzione della madre: la probabilità è maggiore per le madri meno istruite (Figura 5). Rispetto al titolo di godimento dell'abitazione, sembra che la stabilità dell'alloggio sia positivamente correlata alla scelta matrimoniale; infatti, la probabilità di far nascere il proprio figlio all'interno di una convivenza stabile è minore per le madri che dichiarano di possedere una casa di proprietà rispetto ad essere in affitto. Con riferimento alle variabili prettamente demografiche, la probabilità di avere un figlio all'interno di una convivenza stabile e non all'interno del matrimonio decresce all'aumentare dell'età delle madri (con l'eccezione delle ultraquarantenni)⁵ e del numero di figli. La

⁴ Ci si è basati sul criterio logico della parsimonia includendo tutti gli effetti semplici ed escludendo quelli di ordine superiore. L'inclusione delle variabili nei modelli è stata effettuata con la tecnica *stepwise*.

⁵ La probabilità è più elevata per le madri con più di 40 anni; al contrario, rispetto alle più giovani, tale probabilità diventa inferiore a 1 per le classi intermedie.

ripartizione territoriale si conferma fondamentale *proxy* dei comportamenti riproduttivi; infatti, mentre al Centro aumenta, seppur in modo lieve, la probabilità di essere in convivenza stabile alla nascita del figlio rispetto al Nord, nel Mezzogiorno tale probabilità diminuisce notevolmente riportando in primo piano, ancora una volta, la maggiore propensione ad assumere comportamenti di tipo tradizionale.

Figura 5 – Odds ratios dei modelli logistici a e b.

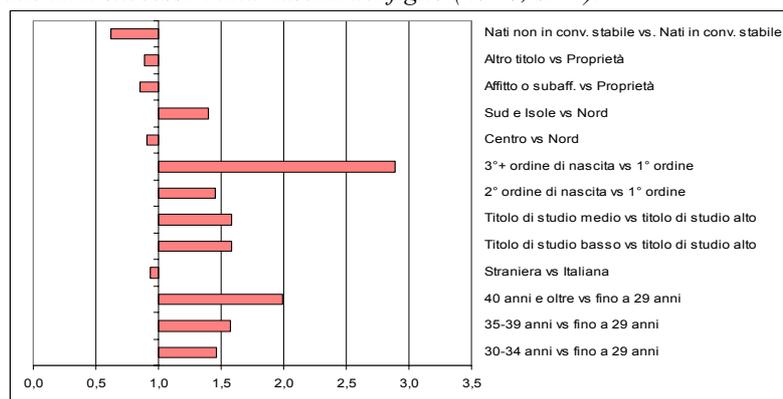


Confrontando la convivenza stabile con quella non stabile, si ripete in buona sostanza quanto già osservato nel precedente modello anche se in questo caso sono le nascite non in convivenza stabile rispetto a quelle in convivenza stabile a mostrare un contesto di maggiore fragilità sia in termini di capitale umano che di stabilità abitativa. La ripartizione territoriale mette in luce, ad esempio, un andamento contrapposto confermando come l'aver un figlio fuori dal matrimonio non sia molto spesso frutto di una scelta legata ad un progetto riproduttivo; nel Sud, come già osservato, avere un figlio all'interno di una convivenza piuttosto che all'interno del matrimonio è una scelta di pochi; inoltre, tra i nati fuori dal matrimonio prevalgono i nati da madri non in convivenza stabile.

Infine, per poter indagare in maggiore profondità su quanto la convivenza rappresenti una scelta definitiva o solo una tappa di un percorso è stato applicato un ulteriore modello logistico per mettere a confronto le madri già in convivenza che si sono sposate in seguito alla nascita di uno o più figli e quelle che, invece, hanno deciso di non contrarre il vincolo matrimoniale (Figura 6). Dai risultati del modello emerge come le madri con una maggiore propensione a sposarsi siano quelle meno istruite, risiedano nel Sud e abbiano figli di ordine successivo al primo. L'età, inoltre, evidenzia una probabilità direttamente proporzionale alla preferenza accordata a scelte di tipo matrimoniale. Sembra essere la stabilità ad avere maggiore influenza sulla propensione a sposarsi dopo la nascita del figlio; la

probabilità di contrarre matrimonio è più elevata per le madri con cittadinanza italiana, che dichiarano di vivere in una casa di proprietà e che vivono in convivenza stabile.

Figura 6 – Odds ratio del modello con variabile risposta: Madri che si sono sposate nei due anni successivi alla nascita del figlio ($no=0$; $si=1$).



5. Conclusioni

Le coppie che decidono di formare una famiglia con figli all'interno di una convivenza stabile presentano generalmente caratteristiche simili a quelle delle coppie coniugate. Le differenze più marcate si colgono confrontando i nati da madri in convivenza stabile con quelli da madri non in convivenza stabile sia a livello di variabili demografiche che di progetti familiari e riproduttivi.

L'analisi della propensione a sposarsi dopo la nascita dei figli mette in luce la complessità dei percorsi che oggi sembrerebbero discostarsi, almeno in parte, dalla sequenza delle tappe che caratterizzano il tradizionale processo di formazione della famiglia. Sono, quindi, numerose le variabili che entrano in gioco quando si va a determinare se per le madri la convivenza all'interno della quale è nato un figlio rappresenti una scelta definitiva o solo una tappa di un progetto di vita più ampio in cui è previsto anche il matrimonio. La centralità della variabile territoriale sembra suscitare la necessità di maggiori approfondimenti proprio per via di differenti comportamenti che contrappongono il Centro-nord al Mezzogiorno.

Considerando la duplice valenza che è possibile attribuire alla convivenza – scelta di emancipazione e indicatore di situazioni di fragilità – le analisi fin qui condotte sembrerebbero far emergere con maggior evidenza questa seconda dimensione. Nuovi modelli interpretativi per ripartizione potrebbero, invece, consentire l'individuazione di profili alternativi. Tuttavia, sulla base dei risultati fin qui discussi, la scelta del matrimonio appare per lo più legata al raggiungimento di una stabilità – *in primis* economica. Nella scala delle priorità la disponibilità

economica sembra rappresentare un vincolo difficile da aggirare. In questo contesto, dunque, il matrimonio costituisce una tappa che è possibile accantonare definitivamente o posporre ad altre tappe del processo di formazione della famiglia. L'analisi delle caratteristiche differenziali di chi compie tali scelte e della sequenza con cui le compie ci aiuta a comprendere il caleidoscopio di percorsi oggi esistenti.

Riferimenti bibliografici

- Castagnaro C., Gualtieri G., Prati S. (2007), Le nascite al di fuori del matrimonio: evoluzione e principali caratteristiche, *Conferenza Nazionale della Famiglia*, maggio 2007.
- Commissione per la Garanzia dell'Informazione statistica (2008), *Indagine su fonti, modalità e strumenti di raccolta delle informazioni sulle famiglie da parte della statistica ufficiale*, Rapporto di indagine, 08.01.
- Istat (2006a), *Avere un figlio in Italia. Approfondimenti tematici dall'indagine campionaria sulle nascite*, Collana "Informazioni", n° 32.
- Istat (2006b), *L'indagine campionaria sulle nascite: obiettivi, metodologia e organizzazione*, Collana "Metodi e Norme", n° 28.
- Terra Abrami V. (2009), Matrimonio e dinamiche sociali in cambiamento, Neodemos, articolo pubblicato il 20/05/2009.
- van de Kaa D. (1988), Europe's second demographic transition, *Population Bulletin*, 42, 1, Population Reference Bureau, Washington, DC, USA.

SUMMARY

One of the elements that best describes changes in Italian family behaviour is the quick increase of births out of marriage. In 2007, these births amount to 112.984, which account for about 21% of total live births. This percentage was quadrupled compared to 1985. This phenomenon is particularly spread in the Northern and Central Italy, while decreasing in southern regions.

The aim of this paper is to analyse changing in family behaviour using individual data provided by the *Italian Birth Sample Survey* carried by Istat. The 2005 sample consists of around 50.000 births enrolled at the Population Registers in 2003, representing about 10% of the total resident births registered.

In order to study the factors influencing family behaviours, different logistic model were applied. In the models, the main socio-demographic characteristics of mother, territorial domain, parity, and type of cohabitation were included as descriptors.

Cinzia CASTAGNARO, Ricercatore Istat.
Antonella GUARNERI, Ricercatore Istat.

LA MOBILITÀ INTERNA DEGLI STRANIERI IN ITALIA: UNO STUDIO ATTRAVERSO IL RECORD LINKAGE TRA ARCHIVI

Cinzia Conti, Antonella Guarneri, Francesca Licari, Enrico Tucci

1. I percorsi degli stranieri regolarizzati

Le varie collettività straniere presenti nel nostro Paese non solo mostrano una differente propensione a spostarsi sul territorio italiano ma anche diversi modelli di mobilità. L'obiettivo del lavoro è descrivere questi modelli attraverso tecniche di *network analysis* e studiare l'influenza di alcune caratteristiche (individuali e relative al percorso migratorio) sugli spostamenti interni.

Il *record linkage* tra l'archivio dei regolarizzati nel 2004 e i permessi di soggiorno a questi rinnovati nel 2007 consente di "fotografare" gli stessi individui in due successivi istanti di tempo e di cogliere il loro eventuale spostamento sul territorio, confrontando il luogo in cui è stato ottenuto il permesso di regolarizzazione nel 2004 con il luogo di rinnovo del permesso nel 2007¹. L'utilizzo di questi dati permette altresì di considerare, oltre alla mobilità, anche il mutamento di altri aspetti nella vita dei singoli individui, come il cambiamento dello stato civile e del motivo del soggiorno.

La mobilità degli stranieri regolarizzati è molto elevata: il 60 per cento si è spostato sul territorio e nella maggior parte dei casi (42 per cento) ha cambiato ripartizione. Come nelle attese, si sono spostati meno frequentemente dalla ripartizione di regolarizzazione coloro che hanno fatto domanda di sanatoria nel Nord-ovest e nel Nord-est; sono rimasti nella stessa ripartizione rispettivamente nel 68 e nel 62 per cento dei casi (Tabella 1). Coloro che si sono regolarizzati nel Mezzogiorno hanno invece più frequentemente rinnovato il permesso in un'altra ripartizione, soprattutto nel Nord-ovest e nel Nord-est.

In generale, comunque, se Sud e Isole fanno registrare una minore stabilità dei cittadini stranieri, la tendenza allo spostamento varia a seconda delle nazionalità. Considerando grandi gruppi di cittadinanze si nota che coloro che provengono dall'America centro-meridionale presentano una tendenza a muoversi molto ridotta rispetto agli altri gruppi. Coloro che si sono spostati in misura maggiore verso altre

¹ I dati considerati permettono, rispetto al collettivo dei regolarizzati, di avere il riferimento relativo alla provincia in cui è stata effettuata la regolarizzazione nel 2004 e quella in cui è stata fatta domanda per il rinnovo del permesso nel 2007.

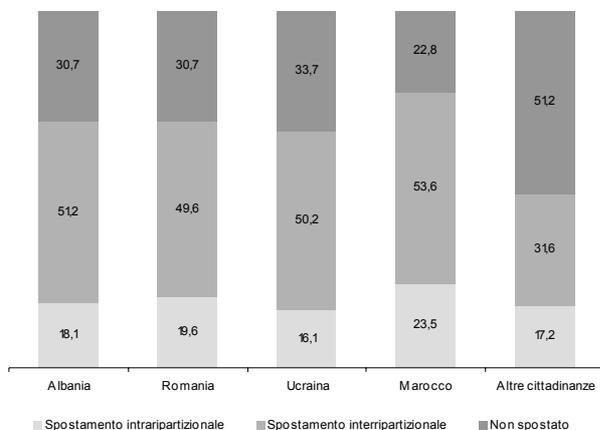
ripartizioni sono gli africani e gli asiatici regolarizzati al Sud: rispettivamente nel 72 e nel 69 per cento dei casi. Gli americani regolarizzati nel Nord-ovest sono coloro che si sono spostati in misura minore (meno dell'11 per cento dei casi). Esaminando le principali cittadinanze, si rileva che i marocchini presentano una mobilità particolarmente elevata sul territorio (77,2 per cento) e che danno luogo alla più alta quota di spostamenti nella stessa ripartizione (Figura 1).

Tabella 1 – *Stranieri regolarizzati per ripartizione di rinnovo del permesso e ripartizione di regolarizzazione (valori assoluti e percentuali).*

RIPARTIZIONE DI REGOLARIZZAZIONE	Ripartizione di rinnovo del permesso					Totale (v.a.)
	Nord-ovest	Nord-est	Centro	Sud	Isole	
Nord-ovest	68,1	13,6	13,2	4,3	0,8	170.996
Nord-est	18,6	61,7	13,0	5,8	0,9	103.037
Centro	21,7	16,8	54,3	6,2	1,0	140.522
Sud	21,1	20,6	15,3	41,6	1,3	77.428
Isole	24,4	20,0	14,1	6,4	35,1	13.475
Italia (v.a.)	185.754	129.089	126.082	55.112	9.421	505.458

Fonte: elaborazioni Istat su dati del Ministero dell'interno

Figura 1 – *Stranieri regolarizzati per principali cittadinanze e tipo di spostamento (valori percentuali).*

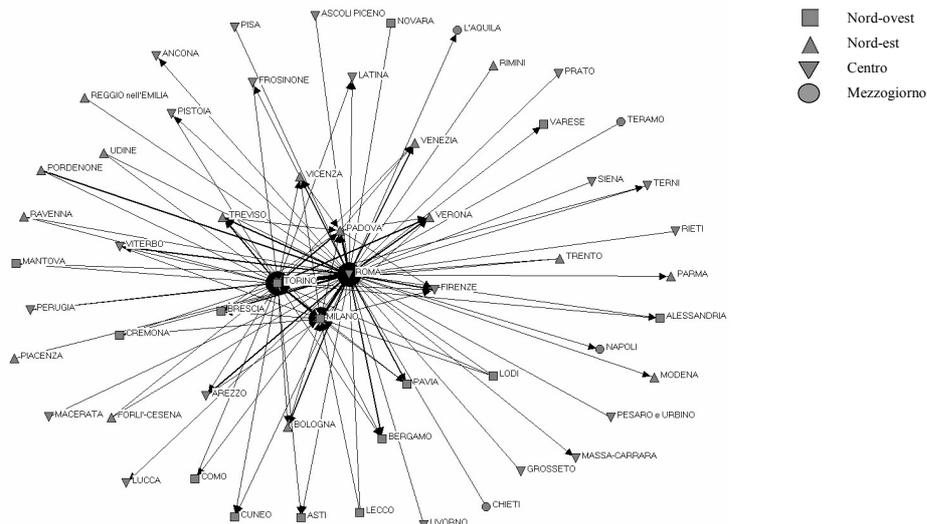


Fonte: elaborazioni Istat su dati del Ministero dell'interno

2. I network dei regolarizzati

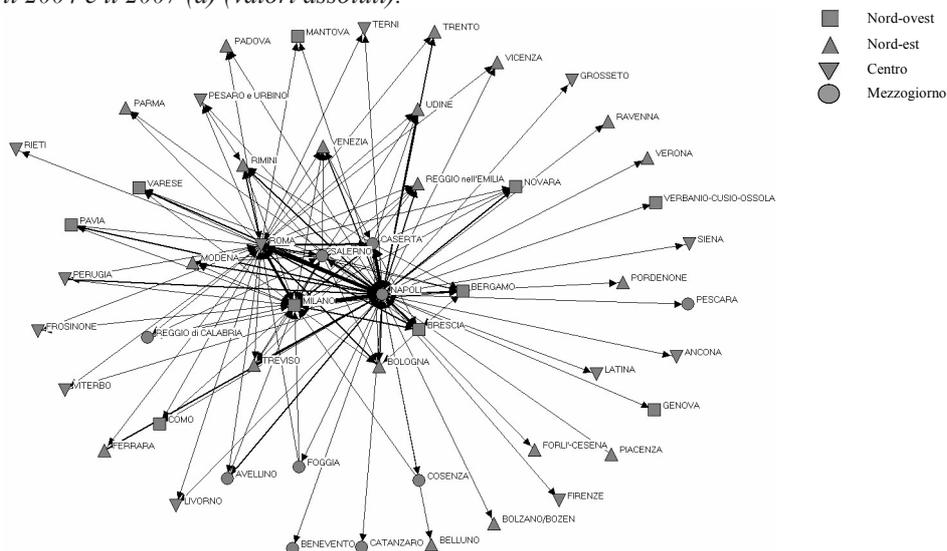
La rappresentazione grafica dei network (grafi) consente di cogliere in maniera sintetica l'esistenza di reti che legano tra loro differenti realtà territoriali, superando l'ottica "a due a due" della matrice origine-destinazione. Di seguito si illustrano i

Figura 3 – Rete degli spostamenti interprovinciali dei regolarizzati di cittadinanza rumena tra il 2004 e il 2007 (a) (valori assoluti).



Fonte: Elaborazione Istat su dati del Ministero dell'interno
(a) Sono stati rappresentati i flussi superiori a 80 spostamenti.

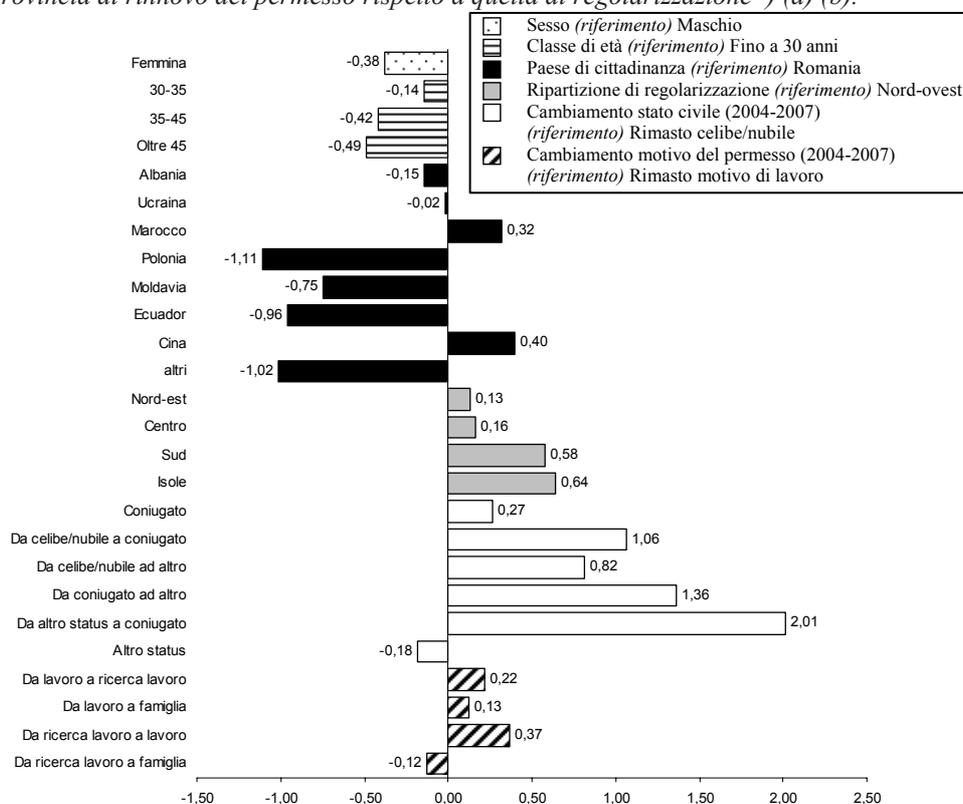
Figura 4 – Rete degli spostamenti interprovinciali dei regolarizzati di cittadinanza ucraina tra il 2004 e il 2007 (a) (valori assoluti).



Fonte: Elaborazione Istat su dati del Ministero dell'interno
(a) Sono stati rappresentati i flussi superiori a 50 spostamenti.

Nel modello sono considerati tutti i regolarizzati con un permesso ancora in vigore al 2007 (Figura 6). Sono soprattutto i maschi e i giovani a trasferirsi. Per l'età si rileva una chiara scala di valori: la probabilità di spostarsi diminuisce man mano che l'età avanza. Con riferimento alla cittadinanza, cinesi e marocchini hanno una maggiore propensione a muoversi sul territorio rispetto ai rumeni³, mentre moldavi ed ecuadoriani mostrano un'attitudine a spostarsi più bassa.

Figura 6 – Stime del parametro del modello di regressione logistica (variabile dipendente: "stessa provincia di regolarizzazione e di rinnovo del permesso/diversa provincia di rinnovo del permesso rispetto a quella di regolarizzazione") (a) (b).



Fonte: Elaborazioni Istat su dati del Ministero dell'interno

(a) Dal confronto tra casi attesi in base al modello e casi osservati risultano correttamente assegnati il 67,1 per cento dei casi. (b) Nel grafico sono riportate solo le modalità significative.

³ I rumeni sono la collettività di riferimento nella specificazione del modello adottato.

Per quanto riguarda la ripartizione, a parità di altre caratteristiche, risulta più elevata la propensione relativa a spostarsi per chi si è regolarizzato in altre ripartizioni rispetto al Nord-ovest. In particolare la propensione allo spostamento è elevata al Sud e nelle Isole. Questo risultato sembrerebbe confermare l'ipotesi che nel Mezzogiorno sia stato richiesto e accettato un numero particolarmente elevato di regolarizzazioni, alle quali successivamente hanno fatto seguito spostamenti sul territorio alla ricerca generalmente di migliori condizioni lavorative⁴.

Infine, sono state inserite nell'analisi due variabili che sintetizzano i principali cambiamenti di status intercorsi nel periodo 2004-2007: modifica dello stato civile⁵ e modifica del motivo del permesso di soggiorno. Una prima considerazione riguarda il cambiamento dello stato civile: chi lo ha modificato si è spostato di più rispetto a coloro per i quali questa caratteristica è rimasta immutata. In particolare, nel periodo considerato, chi si è sposato e chi ha terminato una relazione coniugale (per separazione o morte del coniuge) mostra una propensione relativa a trasferirsi più elevata rispetto chi è rimasto nella condizione di celibe o nubile.

Osservando il cambiamento del motivo del permesso di soggiorno e prendendo come riferimento la categoria di coloro per cui è rimasto invariato il motivo "lavoro", chi ha trasformato il motivo del permesso di soggiorno da "ricerca lavoro" a "lavoro" presenta maggiori possibilità di essersi spostato; mentre, per chi è passato da "ricerca lavoro" a "famiglia", la propensione è meno elevata.

In sintesi

La network analysis mette chiaramente in luce che le differenti collettività, pur condividendo alcune traiettorie, danno vita a reti di spostamenti peculiari. Attraverso l'analisi condotta sui dati del *record linkage* emerge in maniera netta il ruolo chiave avuto da alcune province nei processi di regolarizzazione. In particolare spicca il ruolo di alcuni nodi del Mezzogiorno che risultano, invece, di secondaria importanza quando l'analisi viene condotta considerando i dati dei trasferimenti di residenza (Conti *et al.*, 2009).

L'analisi delle relazioni tra mobilità e altre variabili socio-demografiche evidenzia come i cambiamenti di status (stato civile e motivo del permesso) siano, in maniera più o meno forte, collegati allo spostamento sul territorio. Anche in questo caso è stato possibile cogliere la complessità dei "percorsi" degli stranieri in

⁴ Si veda in proposito: Istat. Rapporto annuale. La situazione del Paese nel 2007. Roma: Istat. 2008.

⁵ Si sottolinea che si lavora sullo "stato civile", quindi sulla condizione familiare formale e non su quella "di fatto" che gli stranieri hanno in Italia. Non necessariamente i coniugati hanno il coniuge in Italia.

Italia, non solamente in senso strettamente territoriale ma anche come sequenza di eventi (ad esempio un matrimonio o un permesso ottenuto per famiglia e non più per lavoro) che condizionano fortemente la vita degli individui e che tra loro si intersecano in molteplici configurazioni.

Riferimenti bibliografici

Conti C., Gabrielli D., Guarneri A., Tucci E. (2009) “Measuring foreigners migration flows: Traditional methods and social network analysis”, *Atti del convegno intermedio SIS - Statistical methods for the analysis of large data-sets*, Pescara, settembre 2009, CLEUP, Padova.

Cordaz D. (a cura di) (2005) Le misure dell’analisi di rete e le procedure per la loro elaborazione mediante UCINET V, Appendice al volume Salvini, A., *L’analisi delle reti sociali. Risorse e meccanismi*, Ed Plus, Pisa University Press, Pisa.

Istat (2009) *Rapporto annuale. La situazione del Paese nel 2007*, Roma.

Maier G., Vyborny M. (2005) Internal Migration between US States - A Social Network Analysis, *SRE-Discussion 2005/04*, Abteilung für Stadt- und Regionalentwicklung, Department of Urban and Regional Development, Vienna University of Economics and Business Administration (WU-Wien).

SUMMARY

In 2007 more than 500 thousand foreigners renewed the residence permit that was previously granted according to the regularization law of 2004. Using the results of record linkage techniques, it is possible to compare the place where the permit was granted in 2004 with the place of renewal in 2007, in order to verify whether a transfer has occurred or not. These data also allow to consider the changes in other aspects of the individual characteristics such as the marital status and the reason of stay.

In a first step, the method of social network analysis (SNA) is applied. The analysis underlines the difference in the networks of the various communities.

In a second step, a logistic model is applied in order to analyze the relations between some characteristics of the foreign population and its propensity to move within the Italian territory. The model takes into account some important events that have an impact on the life of the individuals. Therefore it is possible to better understand the complexity of the “integration route” of the foreigners in Italy.

Cinzia CONTI, Ricercatore Istat.

Antonella GUARNERI, Ricercatore Istat.

Francesca LICARI, Collaboratore tecnico Istat.

Enrico TUCCI, Ricercatore Istat.

LA STRUTTURA DEI SISTEMI TURISTICI LOCALI IN SICILIA: ANALISI DELLA STAGIONALITÀ *INCOMING*

Rosario D'Agata, Venera Tomaselli¹

1. Organizzazione normativa dei sistemi turistici territoriali

Il turismo è un fenomeno a peculiare connotazione geografica per la sua capacità di incidere sull'organizzazione territoriale delle aree di riferimento mediante processi di localizzazione di risorse e di circolazione di tipologie di presenze turistiche.

Con la Legge 135/01, che istituisce i Sistemi Turistici Locali (STL)² in Italia, l'offerta turistica non è più esclusivamente calibrata sulle esigenze della domanda (Costa, 2005). Nell'intenzione del legislatore, i STL dovrebbero svolgere un ruolo decisivo per lo sviluppo socio-economico del territorio, valorizzando le risorse locali e costruendo reti relazionali tra unità amministrative diverse ma complementari. Molti studiosi hanno abbandonato lo studio delle relazioni tra attori economici appartenenti allo stesso settore, focalizzando l'attenzione sull'insieme delle attività produttive comprese all'interno di un sistema territoriale, al fine di far emergere le complementarità tra settori. Il territorio non rappresenta il mero 'contenitore' spaziale in cui operano singole realtà economico-produttive ma diventa esso stesso, in una logica distrettuale, attrattore di domanda turistica e volano per lo sviluppo. Il territorio da semplice insieme di coordinate spaziali e confini amministrativi diviene componente attiva dell'offerta ed esso stesso prodotto turistico (Dall'Ara e Morandi, 2006). In merito alla definizione territoriale dei STL, il legislatore ha concesso totale autonomia alle regioni ed ampia discrezionalità agli attori interessati. Poche regioni, però, hanno reso operativi i STL costituendo forme organizzative estremamente flessibili ed utilizzando criteri d'azione diversificati da regione a regione, anche a causa dell'assenza di riconoscimento giuridico ai STL nella normativa di riferimento.

¹ Il presente lavoro è stato svolto congiuntamente dagli autori. Più in particolare, i paragrafi 4 e 5 sono stati curati da Rosario D'Agata ed i paragrafi 1, 2 e 3 da Venera Tomaselli.

² Dalla definizione normativa infatti, i STL si prefigurano come "contesti turistici omogenei o integrati comprendenti ambiti territoriali anche di Regioni diverse e caratterizzati dall'offerta integrata di beni culturali e ambientali e di attrazioni turistiche compresi i prodotti tipici dell'agricoltura e dell'artigianato locale o dalla presenza di imprese turistiche" (art. 5).

2. Offerta e domanda turistica in Sicilia: proposta di ripartizione in distretti

La Sicilia, con la Legge Regionale 10/2005, ha tentato di individuare criteri di definizione dei STL e di prevedere una programmazione finalizzata alla promozione di strategie di sviluppo comuni per aree territoriali caratterizzate da dimensioni strutturali simili, da vocazioni turistiche complementari e connotate da criticità, tra le quali quella che il programma triennale di sviluppo turistico 2007-09 definisce “paradosso della stagionalità”. Le presenze turistiche italiane nell’Isola, infatti, appaiono concentrarsi quasi esclusivamente nel periodo estivo ed in un periodo un po’ più ampio (da aprile ad ottobre) quelle dei visitatori stranieri.

L’approccio sistemico allo studio delle connessioni territoriali tende ad analizzare la trama delle caratteristiche economiche delle aree in cui si struttura l’offerta turistica e l’andamento delle presenze (Casari, 2008, 23-26). Seguendo tale paradigma, questo lavoro considera una serie storica di dati sulle presenze in arrivo presso i principali porti ed aeroporti, quali punti preferenziali di accesso, per individuare caratteri di stagionalità degli arrivi, evidenziando le peculiarità strutturali del territorio. L’attenzione è rivolta alle dimensioni economiche dei comuni – unità primarie comprese nei STL – tenendo conto degli studi sulle proposte di aggregazione territoriale. La definizione di modelli di organizzazione dell’offerta nella regione è orientata in base ad un approccio metodologico che tiene conto della dimensione territoriale come elemento essenziale per la costituzione di sistemi locali integrati (distretti, sistemi, ecc.) in cui la concentrazione e la specializzazione delle attività economiche nel comparto turistico sia misurabile. Sulla base della legge Quadro nazionale 135/01, infatti, numerosi studi (Censis, Confindustria) hanno affrontato la questione della definizione di aree funzionali per la promozione dello sviluppo di un settore strategico per l’economia italiana e siciliana. Tra questi, nella proposta di Arisco e Caponnetto (2008) si fa riferimento ad “un nuovo approccio allo sviluppo turistico improntato all’ottica territorialista...” (*ibidem*, 210), la cui complessità organizzativa e gestionale ha condotto alla costruzione di un modello sulla base di indicatori di domanda ed offerta turistica di livello comunale³. Dalla sintesi operata

³ Il modello si articola in informazioni e misure relative alla domanda espressiva del target di riferimento, all’accessibilità in termini di offerta infrastrutturale e di strutture e nodi di reti reali e virtuali, all’attrattività, ospitalità e capacità ricettiva, all’ambito di specializzazione delle attività economiche, alla forza e capacità di aggregazione locale in *segue nota a pagina precedente*: base all’effettività di strumenti di programmazione negoziata territoriale (progetti PRUST, PIT, Patti territoriali, Leader e Urban I e II, ecc.), all’ambito paesaggistico e di identità ambientale e storico-culturale delle aree territoriali.

sui diversi aspetti, sono state definite unità omogenee⁴, sistemiche ed integrate con le loro articolazioni gerarchiche rispetto ad aree funzionali alla programmazione di politiche di sviluppo (Figura 3).

3. Materiali e strumenti di analisi

3.1 Analisi della specializzazione dell'offerta turistica nel territorio

L'analisi sulle ripartizioni territoriali ha preso in considerazione una serie di misure relative ad aspetti economici collegati al fenomeno turistico. Sono state selezionate da *data-base* opportunamente integrati (ISTAT, 2008), informazioni trasformate in rapporti di livello comunale, per procedere alla costruzione di un indicatore sintetico territoriale (Tab. 1):

Tabella 1 – Dimensioni economiche del turismo per distretti.

Indicatore	Descrizione	Fonte	Anno
Indice generale di dotazione economica	(Occupati / Popolazione residente) * 100	ISTAT	2001
Indice specifico di dotazione econ.: servizi	(Occupati nei servizi / Popolazione residente) * 100	ISTAT	2001
Indice di imprenditorialità	(Unità locali delle imprese / Popolazione residente)* 1000	ISTAT	2001
Imprese dei servizi	(Imprese registrate nei servizi / Popolazione residente) * 1000	Infocamere	2006
Unità locali dei servizi	(Unità locali dei servizi / Totale unità locali)* 100	Infocamere	2006
Addetti alle unità locali dei servizi	(Addetti alle unità locali di imprese nei servizi / Tot. addetti) *100	ISTAT	2001
Unità locali dei servizi per tipologia: servizi a imprese e consumatore	(Unità locali dei servizi alle imprese ed al consumatore / Totale unità locali dei servizi)* 100	Infocamere	2006
Ricettività turistica	(Posti letto negli esercizi ricettivi / Popolazione residente) * 1000	ISTAT	2005
Grado di turisticità	(Numero medio di turisti -presenze e numero di giorni dell'anno- negli eserc. ricettivi / Pop. resid.) * 1000	Istituto Tagliacarne - ISTAT	2005
Grado di utilizzo delle strutture alberghiere	(Presenze / numero posti letto * giorni dell'anno) * 100	Istituto Tagliacarne	2005

Fonte: ISTAT 2008, Atlante socioeconomico della Sicilia.

⁴ La rappresentazione in figura 3 si riferisce ai “potenziali” distretti turistici ottenuti dall'aggregazione delle unità comunali in 12 aree funzionali rispetto all'articolazione del modello di rilevazione. Nel presente lavoro, tale ripartizione territoriale è stata adottata per l'analisi della specializzazione dell'offerta turistica nel territorio, di cui al successivo paragrafo 4.

Per ottenere un indicatore sintetico a partire dai suddetti indicatori semplici (X_j), è stata applicata una procedura di trasformazione, mediante il calcolo del numero indice a base fissa, costituita dal valore medio dell'indicatore $\mu(X_j)$, di ogni comune (i), per ciascun distretto (D):

$$I_{ij} = \frac{X_{ij}}{\mu(X_j)} \quad (1)$$

dove: X_{ij} = valore dell'indicatore j -esimo nell' i -esimo comune del distretto (D)

$\mu(X_j)$, = media dell'indicatore j -esimo nell' i -esimo comune del distretto (D).

L'indicatore sintetico è ottenuto calcolando la media aritmetica di ciascun numero indice (I_{ij}) per ogni i -esimo comune all'interno di ciascun distretto (D).

3.2 Analisi della domanda: la stagionalità degli arrivi.

L'analisi della domanda turistica ha consentito di individuare e di controllare aspetti significativi in termini di sussistenza e consistenza di andamenti stagionali degli arrivi in Sicilia, presso i principali porti (Palermo, Catania, Messina) ed aeroporti (Palermo, Catania, Trapani), considerati *proxy* delle presenze turistiche.

La stagionalità è strettamente connessa ad una serie di fattori (Vaccaro, 2007), tra cui l'andamento climatico di fondo ed occasionale dell'area territoriale e l'alternarsi delle stagioni, la disponibilità di periodi di tempo non lavorativo, eventi particolari previsti in calendario e anche non prevedibili. Essa, inoltre, si manifesta in termini di aspetti sistematici legati alla variabilità mensile, in base alla quale la domanda di turismo risulta concentrata in alcuni mesi dell'anno; ed alla variabilità settimanale, che a sua volta concentra la domanda di turismo professionale nei giorni feriali della settimana, mentre nei week-end è elevata la concentrazione della domanda di turismo da tempo libero. In riferimento, poi, alle caratteristiche del territorio e dell'offerta, si possono rilevare effetti di stagionalità della domanda turistica in un unico periodo dell'anno oppure periodi plurimi di stagionalità turistica o, infine, tendenze di stagionalità 'inversa' rispetto a quella prevalente proprio nei periodi di stagioni tipicamente turistiche (città di turismo professionale, in cui nei mesi estivi si sospendono le attività).

Per analizzare e valutare la misura di omogeneità interna della serie di dati rilevati, è stato calcolato il rapporto di concentrazione, basato sul calcolo dei pesi percentuali assoluti e cumulati (p_i) e delle distribuzioni cumulate degli arrivi assoluti e relativi (q_i) per i singoli mesi sul totale annuo.

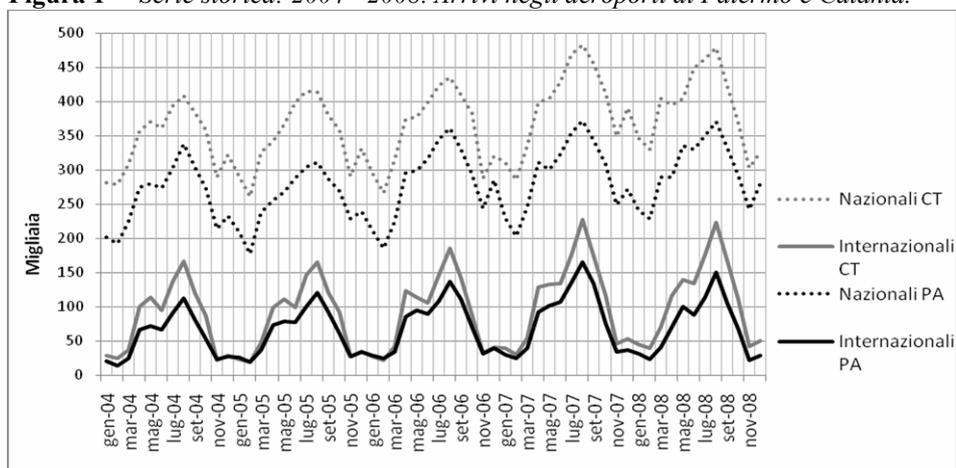
Infine, ai fini dell'analisi dell'andamento delle variazioni, mensili ed annue, relative del valore x_i rispetto alla base, sono stati calcolati i numeri indici degli

arrivi, utilizzando il mese (Agosto) in cui si manifestano effetti di stagionalità come base fissa.

4. Analisi dei risultati

Per evidenziare eventuali caratteristiche stagionali della domanda turistica, dopo aver considerato tutti i principali porti ed aeroporti dell'Isola, l'attenzione si è focalizzata sull'andamento mensile degli arrivi nei due principali aeroporti siciliani⁵: Punta Raisi (PA) e Fontanarossa (CT). In entrambi i casi, è stata analizzata una serie storica (2004-2008) distinguendo gli arrivi nazionali da quelli internazionali (Fig. 1).

Figura 1 – Serie storica: 2004 - 2008. Arrivi negli aeroporti di Palermo e Catania.



Fonte: Assoaeroporti.

Dal confronto tra gli andamenti osservati, emerge come nei 5 anni considerati, è sempre il mese di agosto a far rilevare il numero massimo di arrivi in entrambi gli aeroporti, mentre è a febbraio che si registra il numero minimo di passeggeri. La differenza tra le due aerostazioni, inoltre, appare più marcata rispetto all'ammontare degli arrivi nazionali e non internazionali. L'osservazione di 'picchi' più smussati, nel caso dei flussi nazionali, evidenzia l'esistenza di una distribuzione più omogenea rispetto alla componente internazionale. Per meglio

⁵ In un primo momento era stato preso in considerazione anche l'aeroporto di Trapani Birgi. La scarsa intensità del fenomeno, tuttavia, ha suggerito di escludere la suddetta aerostazione dall'analisi.

comprendere il grado di omogeneità interno a ciascuna distribuzione, è stato calcolato il rapporto di concentrazione (Vaccaro, 2007) considerando il numero degli arrivi mensili per ciascun anno.

L'analisi dei rapporti di concentrazione evidenzia una maggiore omogeneità nella distribuzione degli arrivi nazionali presso l'aeroporto catanese; di contro, per gli arrivi internazionali, componente probabilmente più legata alla dimensione turistica, si registra un valore del rapporto di concentrazione in entrambe le aerostazioni tendenzialmente più elevato. Gli arrivi internazionali, quindi, tendono a concentrarsi prevalentemente in pochi mesi (luglio, agosto e settembre) più di quanto accada per gli arrivi nazionali.

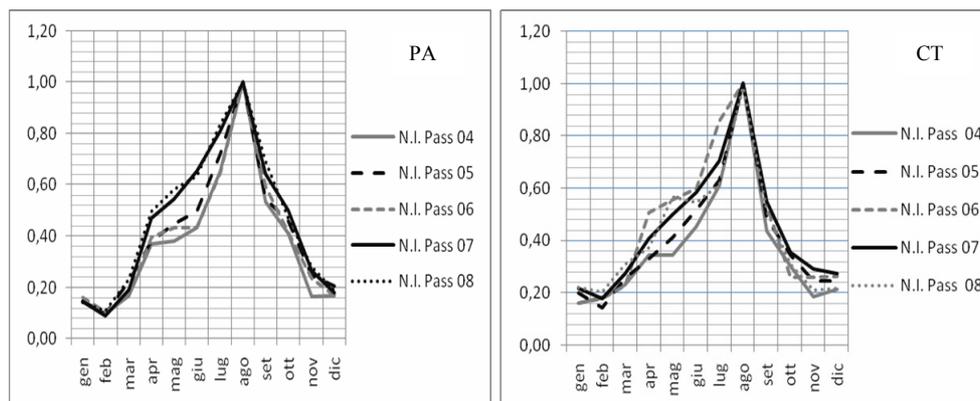
Tabella 2 – Serie storica: 2004 – 2008. Rapporti di concentrazione calcolati sugli arrivi nei due maggiori aeroporti siciliani per nazionalità dei passeggeri.

Anni	Aeroporto CT			Aeroporto PA		
	Naz.	Intern.	TOT.	Naz.	Intern.	TOT.
2004	0,078	0,365	0,132	0,104	0,355	0,148
2005	0,086	0,220	0,137	0,091	0,320	0,135
2006	0,095	0,364	0,147	0,115	0,321	0,157
2007	0,094	0,343	0,148	0,108	0,353	0,159
2008	0,087	0,327	0,137	0,091	0,356	0,141

Fonte: Assoaeroporti.

In un secondo momento, sono stati utilizzati i dati relativi agli arrivi nei porti delle due principali città siciliane. Alla luce delle osservazioni riscontrate in precedenza, si è scelto di analizzare l'andamento degli arrivi attraverso l'impiego dei numeri indici, ponendo come base fissa il valore registrato ad agosto per ciascun anno (Fig. 2).

Figura 2 – Arrivi nei porti di PA e CT. Numeri indici a base fissa (agosto=1).



Fonte: Autorità portuali di Palermo e Catania. Anni: 2004-2008.

L'andamento del fenomeno appare confermare la tendenza osservata in merito agli arrivi via aria. Seppur su dimensioni quantitative diverse⁶, si registra un'elevata concentrazione nel periodo estivo, confermata inoltre dai valori del rapporto di concentrazione, mediamente più elevati a Catania, per quanto concerne i croceristi e a Palermo per quanto riguarda i passeggeri.

Tabella 3 – Serie storica: 2004 – 2008. Rapporti di concentrazione calcolati sugli arrivi nei due maggiori porti siciliani distinti tra passeggeri e croceristi.

Anni	Porto CT			Porto PA		
	Pass.	Croc.	TOT.	Pass.	Croc.	TOT.
2004	0,333	0,458	0,334	0,385	0,496	0,481
2005	0,318	0,515	0,350	0,357	0,418	0,363
2006	0,345	0,556	0,364	0,368	0,462	0,377
2007	0,301	0,535	0,359	0,372	0,303	0,380
2008	0,305	0,473	0,349	0,364	0,441	0,385

Fonte: Autorità portuali di Palermo e Catania. Anni: 2004-2008.

Sul versante dell'offerta, infine, seguendo i criteri esposti nel paragrafo 3 è stato calcolato un indicatore sintetico delle 'capacità' attrattive dei singoli comuni inseriti nei 12 STL proposti. Nella figura 4 sono evidenziati quei comuni, che, all'interno di ciascun STL presentano un valore dell'indicatore sintetico maggiore del 3° quartile.

Figura 3 – Mappa dei potenziali STL.

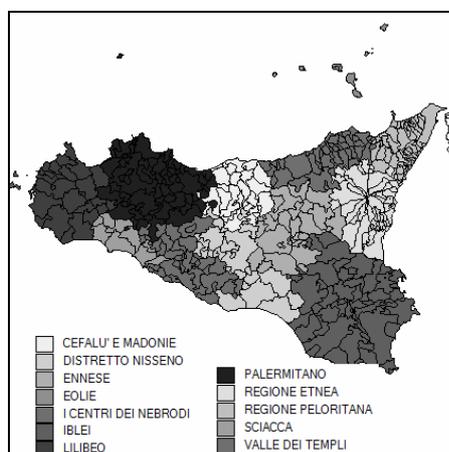
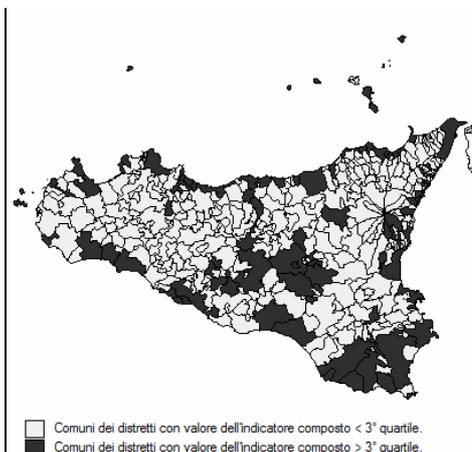


Figura 4 – Mappa dei comuni con un valore dell'indicatore sintetico maggiore del 3° Q.



⁶ In media ogni anno a Palermo sono sbarcati 1.865.648 passeggeri e 370.760 croceristi contro i 218.170 passeggeri e 84.103 croceristi rilevati a Catania.

5. Conclusioni

La vocazione dell'offerta turistica siciliana, come emerge dall'ultimo cartogramma, è sostanzialmente balneare. Se si escludono i STL interni, infatti, si riscontra come i comuni meglio attrezzati a rispondere alla domanda turistica si collochino prettamente in aree marittime. Se da una parte la conformazione geografica della Regione rende prevedibile una siffatta vocazione; dall'altra, la presenza diffusa su tutto il territorio di molteplici e variegati attrattori culturali e paesaggistici, non giustifica l'unidirezionalità dell'offerta turistica alla quale la domanda non può che adattarsi. La stagionalità degli arrivi sull'Isola, rilevata attraverso l'analisi, appare, quindi, la conseguenza di una vocazione turistica incentrata su attrattive inevitabilmente concentrate nei mesi estivi.

Riferimenti bibliografici

- Aci-CENSIS 2002, *I distretti turistici italiana: l'opportunità di innovare l'offerta*.
 Arisco M. N. e Caponnetto D. 2008, "I potenziali distretti turistici siciliani" in: Ruggiero V. e Scrofani L. (a cura di), *Sistemi urbani, reti logistiche e distretti turistici in Sicilia*, Pàtron Editore, Bologna, pp. 209-212.
 Casari M. 2008, *Turismo e geografia*, Hoepli, Milano.
 Confindustria-Comitato Mezzogiorno 2003, *Quaderno 50: I sistemi turistici locali per lo sviluppo di turismo e ospitalità del Mezzogiorno*.
 Costa N. 2005, *I professionisti dello sviluppo turistico locale*, Hoepli, Milano.
 ISTAT 2005, *I Sistemi Locali del Lavoro 2001*, ISTAT, Roma.
 ISTAT 2008, *Atlante socioeconomico della Sicilia*, ISTAT, Palermo.
 Vaccaro G. 2007, *La statistica applicata al turismo*, Hoepli, Milano.

SUMMARY

The structure of the touristic local systems in Sicily: analysis of the *incoming* seasonality

Touristic seasonality in Sicily is analysed according to the different economic features of the Island. We use a proposal of aggregated territorial units (Touristic Local System) in order to define the relationship between touristic demand and supply. The trend of touristic *incoming* shows the exclusive concentration during summer period and only in the seaside townships of the areas.

Rosario D'AGATA, Ricercatore in Statistica Sociale presso la Facoltà di Scienze Politiche dell'Università di Catania.

Venera TOMASELLI, Professore associato in Statistica Sociale presso la Facoltà di Scienze Politiche dell'Università di Catania.

ALCUNI ASPETTI DELL'INTEGRAZIONE DEGLI IMMIGRATI NELLA PROVINCIA DI NAPOLI*

Elena De Filippo, Paolo Diana, Raffaele Ferrara, Linda Forcellati

1. Introduzione

Parlare di integrazione degli immigrati e comprendere il loro livello di inclusione nella società di arrivo è complicato e i rischi di leggere in maniera errata o forzata alcuni indicatori sono tanti. Il termine integrazione esprime peraltro concetti che possono variare nel tempo e nello spazio, che dipendono dal paese, da circostanze storiche e politiche, e soprattutto dalla fase di immigrazione (Strozza, Golini, 2006).

In una realtà come quella della provincia di Napoli le difficoltà nell'interpretare i livelli di integrazione raggiunti dai cittadini immigrati in questi anni sono per certi versi ancora maggiori per le caratteristiche stesse che il fenomeno migratorio ha assunto nel tempo. In particolare per il carattere di area transito che questa realtà ha mantenuto (Orientale Caputo, 2007) e per la sua capacità di "accogliere" e in qualche modo attrarre anche gli immigrati irregolari e quelli che hanno progetti migratori a bassa soglia, cioè orientati al massimo risparmio e con livelli di consumo molto contratti.

Il modello campano dell'immigrazione rimarca ancora oggi in maniera forte il carattere transitorio per una componente dei lavoratori che arrivano e non trovano una adeguata "accoglienza" o che sanno di non poterla trovare o ancora che non sono interessati a rimanere, a cui si è aggiunta in maniera sempre più forte la componente di chi rimane e si "integra" nel contesto locale. D'altronde il processo di stabilizzazione che ha interessato la provincia di Napoli, come del resto l'intera regione, si è presentato già dalla fine degli anni Novanta con una duplice caratterizzazione. Da un lato la componente stanziale dell'immigrazione nell'impatto con la società, con il mercato del lavoro e con i servizi del territorio ha sperimentato percorsi di inclusione e di successo come è possibile evincere, ad esempio, dall'aumento dei nuclei familiari e dalla diversificazione dei lavori svolti, dall'altro ha subito situazioni di esclusione, soprattutto

* L'articolo si basa sui dati dell'indagine campionaria svolta nel comune di Napoli dal Dipartimento di Scienze Statistiche dell'Università di Napoli Federico II e dalla Cooperativa sociale Dedalus, enti che hanno finanziato la rilevazione locale che si colloca all'interno del progetto "La misura del livello di integrazione della popolazione straniera in Italia attraverso dati di indagine" della Fondazione ISMU. Per quel che concerne la stesura del testo Elena de Filippo ha curato i paragrafi 1 e 5, Linda Forcellati il paragrafo 2, Raffaele Ferrara il paragrafo 3 e Paolo Diana il paragrafo 4.

laddove le reti sociali e di comunità non hanno sopperito sufficientemente alle difficoltà nell'accesso ai diritti di cittadinanza per motivi legati alla normativa o alla carenza di servizi. Diverse, e sempre più numerose, sono a questo proposito le categorie di migranti a rischio di ricadere nella marginalità sociale ed economica e questo nella città di Napoli in primo luogo (de Filippo, 2007), dove la condizione di marginalità è spesso condivisa con frange della popolazione locale.

Obiettivo di questo articolo è quello di proporre un primo esame, attraverso i dati di un'indagine "ad hoc" realizzata nel 2008, sul livello di integrazione raggiunto dagli immigrati stranieri nella provincia di Napoli in relazione ad aspetti specifici del loro adattamento nella società di adozione, con la consapevolezza delle difficoltà sopra esposte. Saranno affrontati alcuni temi rilevanti quali il tipo di convivenza e le condizioni abitative, segnalando le differenze per genere, per condizione giuridica di soggiorno e tra i principali gruppi nazionali presenti sul territorio.

2. Caratteristiche generali del campione

In totale sono state realizzate 402 interviste¹ che hanno interessato in oltre il 43% dei casi stranieri provenienti dall'Europa dell'Est e in particolare ucraini (oltre il 23%), i quali anche in base ai dati ufficiali costituiscono oggi il gruppo nettamente più numeroso dell'immigrazione nel napoletano (Strozza, Ferrara, 2007).

Si registra una evidente, per quanto non eccessiva, prevalenza della componente femminile (56%), l'età media supera i 37 anni, la permanenza nella città è mediamente di circa 7 anni e mezzo e il peso dell'irregolarità del soggiorno² risulta di certo non trascurabile (oltre il 34%) (tabella 1). Notevoli sono però le divergenze per cittadinanza. Mentre la collettività ucraina risulta fortemente connotata al femminile (85%) con un'età media elevata (oltre i 42 anni), costituita da persone in età adulta che spesso lasciano nel paese di origine i propri cari o che sono rimaste sole, gli stranieri provenienti dal Senegal sono in maggioranza maschi e con una struttura per età piuttosto giovane (30 anni in media). Entrambe le collettività hanno una durata della presenza sul territorio più breve e una quota di irregolarità più elevata rispetto alle altre comunità rilevanti per numerosità delle presenze. Tra queste ultime, sri-lankesi e polacchi vantano una maggiore durata media della permanenza sul territorio e una minore quota di irregolari, mentre cinesi e sri-lankesi si caratterizzano soprattutto per una struttura di genere molto equilibrata.

¹ Queste le principali caratteristiche della rilevazione: il campionamento è stato realizzato per centri o ambienti di aggregazione; lo strumento di rilevazione è un questionario strutturato a risposte chiuse; i soggetti intervistati sono i cittadini stranieri maggiorenni provenienti dai paesi meno sviluppati (Africa, Asia e America Latina) e dall'Europa centrale e orientale.

² Devono intendersi per irregolari coloro che non sono iscritti in anagrafe e che non hanno un valido permesso di soggiorno.

Tabella 1 – *Età media, durata della presenza e stati civili degli immigrati nel comune di Napoli.*

Aree e principali paesi di cittadinanza	% area o paese di cittadinanza	% Donne	Età Media	Durata Media	% Irregolari
Totale	100,0	56,2	37,5	7,6	34,6
Europa centro-orientale	43,5	78,9	40,1	6,4	36,0
Africa occidentale	12,9	32,7	35,2	8,9	42,3
Resto Africa	8,2	18,2	34,1	6,9	45,5
Asia centro-meridionale	18,7	34,7	35,4	8,7	33,3
Resto Asia	10,7	51,2	37,5	9,3	7,0
America Latina	6,0	70,8	37,1	7,2	45,8
Ucraina	23,1	84,9	41,5	6,3	39,8
Sri Lanka	13,2	45,3	36,9	10,3	24,5
Cina	7,2	44,8	37,0	6,7	3,4
Polonia	7,0	82,1	38,5	10,1	10,7
Senegal	4,7	5,3	30,5	5,6	57,9

Se le prime due collettività citate, ovvero quelle ucraina e senegalese, sembrano adottare un progetto migratorio non definitivo, teso al raggiungimento di un obiettivo meramente economico, le altre tre comunità, quella sri-lankese, cinese e polacca, perseguono probabilmente un trasferimento più stabile.

3. Con chi e dove vivono gli immigrati?

Indicazioni preziose sul livello di integrazione raggiunto dagli immigrati nella nuova società di accogliimento possono derivare dalle informazioni relative al tipo di convivenza e alle condizioni abitative (Strozza, Golini, 2006; de Filippo, 2007). Naturalmente, se vivere con la propria famiglia e in condizioni abitative dignitose può significare un buon inserimento nel paese di arrivo, situazioni differenti possono al contrario delineare uno scenario di disagio o di esclusione sociale.

I dati segnalano che sebbene la maggioranza relativa del campione vive con il proprio nucleo familiare (con coniuge e figli oppure soltanto con il coniuge o soltanto con i figli) si tratta soltanto del 33%³ degli intervistati (tabella 2). Seguono a breve distanza coloro che vivono con amici e conoscenti (28%) o con altri parenti (oltre il 20%), mentre risultano più staccati quelli che vivono da soli (17,5%). Notevoli sono le differenze sia per genere, che per condizione giuridica, nonché per cittadinanza. Vivono più spesso in famiglia soprattutto le donne, i regolari stabili⁴ e le collettività di più “antico”

³ Più in dettaglio: il 21% vive con coniuge e figli; il 12% con coniuge o figli.

⁴ I regolari stabili hanno, al contrario degli irregolari, sia un valido permesso di soggiorno sia l'iscrizione anagrafica; i regolari semi-stabili hanno soltanto il permesso di soggiorno ma non l'iscrizione anagrafica.

trasferimento come quelle provenienti dalla Polonia⁵ e dallo Sri Lanka. Al contrario sperimentano con maggiore frequenza altre forme di convivenza gli uomini, gli irregolari e le comunità provenienti dal Senegal e dall'Ucraina.

Tabella 2 – *Tipo di convivenza dichiarata dagli immigrati nel comune di Napoli.*

Variabili	% per tipo di convivenza in Italia				Totale
	Solo	Con coniuge e/o figli	Con parenti	Con amici e conoscenti	
Totale	17,5	33,1	21,4	28,0	100,0
Sesso					
Uomo	19,3	29,0	21,6	30,1	100,0
Donna	15,8	36,6	21,3	26,2	100,0
Condizione giuridica					
Regolare stabile	13,6	47,2	18,7	20,6	100,0
Regolare semistabile	28,1	12,5	25,0	34,4	100,0
Irregolare	21,4	14,5	25,2	38,9	100,0
Primi paesi di cittadinanza					
Ucraina	21,5	21,5	15,2	41,8	100,0
Sri Lanka	13,2	58,5	13,2	15,1	100,0
Cina	16,7	23,3	46,7	13,3	100,0
Polonia	16,0	68,0	8,0	8,0	100,0
Senegal	21,1	10,5	21,1	47,4	100,0

Per molti studiosi la sfida più importante che gli immigrati devono oggi affrontare è quella della casa, ciò si deve sia ad un mercato degli affitti molto caro che alla grave carenza dell'edilizia pubblica (Reyneri, 2007). Nonostante la qualità degli alloggi sia spesso peggiore per vetustà e dotazione di servizi, gli immigrati pagano affitti normalmente superiori a quelli versati dagli italiani per remunerare i "rischi" legati al maggior affollamento, alla scarsa manutenzione e alla morosità (de Filippo, 2007). Il risultato è che molti immigrati non poveri sono mal alloggiati e immigrati normalmente poveri sono spesso senza casa o in condizione di forte precarietà abitativa (Tosi, 2002).

Le informazioni relative all'area napoletana supportano tale quadro di sfondo e segnalano una difficoltà da parte della maggioranza degli immigrati ad avere una sistemazione abitativa "stabile" certamente maggiore che in altre aree del paese. I proprietari di casa sono infatti quasi inesistenti, meno del 3% contro una media nazionale che nel 2005 risultava pari al 12% (Blangiardo, Farina, 2006). Più in generale solo il 39% degli stranieri intervistati ha una casa di proprietà o un affitto indipendente (non condiviso se non con persone del proprio nucleo familiare) con contratto regolare, mentre oltre il 60% del campione versa in condizione di maggiore precarietà (tabella 3).

⁵ Nel caso dei polacchi è da segnalare che una quota non trascurabile (oltre il 10%) vive soltanto con figli.

In relazione alla stabilità dell'alloggio non si registrano grosse differenze di genere, ma sembrano giocare un ruolo rilevante la condizione giuridica e l'area geografica di provenienza.

Tabella 3 – *Tipo di godimento dell'abitazione dichiarata dagli immigrati nel comune di Napoli.*

Variabili	% per tipologia abitativa						Totale
	Proprietà o affitto indipendente con contratto	Affitto indipendente senza contratto	Affitto con altri immigrati con contratto	Affitto con altri immigrati senza contratto	Abita sul luogo di lavoro	Altra sistemazione	
Totale	39,2	10,2	14,0	11,7	13,7	11,2	100,0
Sesso							
Uomo	40,8	11,4	16,3	15,8	1,6	14,1	100,0
Donna	37,8	9,1	12,0	8,1	24,4	8,6	100,0
Condizione giuridica							
Regolare stabile	58,0	7,2	15,5	3,4	8,7	7,2	100,0
Reg. semistabile	30,3	6,1	12,1	9,1	27,3	15,2	100,0
Irregolare	15,7	15,0	12,4	23,5	17,6	15,7	100,0
Primi paesi di cittadinanza							
Ucraina	25,2	10,8	11,7	11,7	34,2	6,3	100,0
Sri Lanka	64,1	10,3	2,6	10,3	5,1	7,7	100,0
Cina	58,1	3,2	22,6	6,5	3,2	6,5	100,0
Polonia	63,2	10,5	0,0	10,5	15,8	0,0	100,0
Senegal	26,3	15,8	36,8	0,0	0,0	21,1	100,0

Una migliore sistemazione abitativa si riscontra infatti nel caso dei regolari stabili e delle collettività con un progetto migratorio di medio-lungo termine (Polonia, Sri Lanka e Cina). Al contrario, va notato come alla irregolarità giuridica si associ sovente anche quella abitativa e come la maggioranza relativa dei senegalesi e soprattutto degli ucraini (il 34% abita sul luogo del lavoro) si trovi in condizioni di maggiore instabilità.

4. Tipologia familiare e condizione abitativa: un'analisi multidimensionale

Al fine di ottenere un quadro di sintesi delle informazioni inerenti il tipo di convivenza e le condizioni abitative degli immigrati nella provincia di Napoli si è fatto ricorso all'Analisi delle corrispondenze multiple (ACM). Il numero di modalità attive incluse nell'analisi sono 22 provenienti da 6 variabili categoriali: modalità abitativa, tipologia di convivenza, spesa per l'abitazione, spazio a disposizione nell'abitazione, indice di affollamento, presenza di rimesse in denaro⁶.

⁶ Modalità abitativa: proprietà della casa o affitto indipendente con contratto regolare; affitto indipendente senza contratto; affitto con altri immigrati con contratto; affitto con altri immigrati

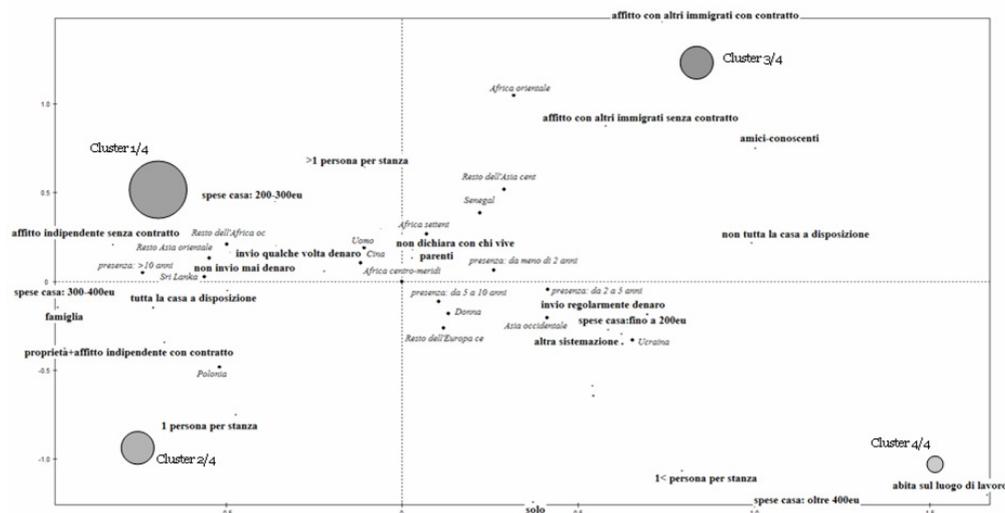
I primi due fattori estratti riproducono il 26,3% dell'inerzia totale. La prima dimensione fattoriale è da ricondurre al “*tipo di convivenza*” degli intervistati, in quanto correlata negativamente al vivere in una famiglia di tipo nucleare (*family*) e positivamente al vivere in una famiglia allargata (*household*) composta da individui che coabitano presso lo stesso indirizzo, consumano pasti preparati insieme e partecipano alla cura della casa (Atkinson, 2000). L'asse configura inoltre un gradiente rispetto ai luoghi e alle forme dell'affettività (Ehrenreich, Hochschild, 2004): dalla famiglia, luogo di relazioni e affetti primari, si passa alle relazioni parentali, a quelle amicali e al vivere soli presso il luogo di lavoro, dove spesso si consuma una traslazione degli affetti verso la famiglia di “accoglienza” (figura 1). In definitiva, il primo fattore potrebbe rappresentare le diverse fasi del ciclo migratorio contrapponendo l'immigrazione più recente (semiasse positivo), giovani in gran parte illegali e senza famiglia, alla più antica immigrazione (semiasse negativo), adulti coniugati che hanno formato o ricostituito la propria famiglia con la quale coabitano e che assorbe la totalità delle risorse economiche disponibili. Rispetto ad esso se le collettività senegalese, cinese e ucraina si pongono sul semiasse positivo, i polacchi e gli sri-lankesi si collocano sul versante opposto.

Il secondo fattore rappresenta invece “*l'affollamento abitativo*” in quanto positivamente correlato alla modalità “più di una persona per stanza” e negativamente legato a quelle “vive solo” e “spende oltre i 400 euro”. Se sul semiasse positivo si collocano in prevalenza gli stranieri provenienti dal continente africano, in particolare i senegalesi, si trovano invece sul semiasse negativo gli stranieri provenienti dall'Ucraina e dalla Polonia.

La *cluster analysis* (fatta utilizzando i primi due fattori dell'ACM) ha infine consentito di individuare quattro gruppi principali per situazione familiare e condizione abitativa: 1) il primo gruppo (36,4%), che si caratterizza per la presenza di immigrati provenienti dallo Sri Lanka, presenta un profilo familiare di tipo tradizionale, vive in affitto indipendente e con un affollamento medio che supera l'unità per stanza; 2) il secondo gruppo (23,8%), si differenzia dal primo per un minore affollamento e si tratta, presumibilmente di famiglie meno numerose; 3) il terzo gruppo (23,8%) proviene in prevalenza dal continente africano e da quello asiatico, è composto soprattutto da maschi che vivono con amici o conoscenti, è in Italia da meno di due anni e mostra una situazione abitativa più precaria; 4) il quarto *cluster*, infine, è composto in prevalenza da donne ucraine che vivono sole presso il luogo di lavoro, dove svolgono sovente il servizio di *colfo* badanti, ed hanno il più basso livello medio di spesa mensile per la casa.

Segue nota a pagina precedente: senza contratto; abita sul luogo di lavoro; altra sistemazione. Tipologia di convivenza: solo; famiglia; amici/conoscenti; parenti. Spesa per l'abitazione in euro: fino a 200; da 200 e 300 euro; da 300 a 400 euro; oltre 400 euro. Spazio a disposizione nell'abitazione: non tutto a disposizione; tutta a disposizione. Indice di affollamento: meno di una persona per stanza; una persona per stanza; più di una persona per stanza. Rimesse di denaro: non invio; invio regolare; invio saltuario.

Figura 1 – Analisi delle corrispondenze multiple: proiezione delle modalità attive e delle modalità illustrative. Assi 1 e 2.



5. Conclusioni

L'analisi del processo di integrazione dei cittadini immigrati nella provincia di Napoli riprende il tema della duplice caratterizzazione del territorio e dell'esistenza di una molteplicità di modelli migratori che non consentono una lettura univoca degli indicatori di integrazione.

Comunità nazionali di più antica immigrazione, come quella senegalese, ad esempio, presente nella regione e nella città di Napoli già a metà degli anni Ottanta, possono presentare caratteri dell'integrazione rispetto alle variabili considerate più simili a quella ucraina il cui arrivo sul territorio campano va ricondotto alla fine degli anni Novanta. Nello specifico caso infatti la maggior parte gli immigrati senegalesi arrivati nel corso di questi 25 anni ha avuto un tempo medio di permanenza di sei/sette anni e la comunità a Napoli è ancora oggi costituita sostanzialmente da giovani maschi commercianti che vivono temporaneamente l'esperienza migratoria nella città partenopea per poi cercare altre soluzioni più o meno definitive nel nord dell'Italia o nel centro Europa. Va inoltre sottolineato che l'evolversi dei percorsi migratori dipende quasi sempre dall'interazione e dall'interagire da un lato di fattori soggettivi di cui è portatore il migrante, dall'altro da alcune caratteristiche del contesto di inserimento, come ad esempio il livello di diffusione e stabilizzazione del sistema di *welfare*, l'accessibilità al mercato del lavoro, le opportunità di inserimento abitativo e di relazione con la comunità ospite. Sono questi gli elementi la cui presenza o meno determina la direzione che imbocca l'immigrato nel suo

sforzo di inclusione nella nostra società. Più aumenta la solitudine sociale e relazionale del migrante, più le complicazioni e gli ostacoli rallentano o determinano il fallimento, del progetto migratorio.

Riferimenti bibliografici

- Atkinson A. B. 2000. *La povertà in Europa*. Il Mulino, Bologna.
- Blangiardo G.C. e Farina P. (cur.) 2006. *Il Mezzogiorno dopo la grande regolarizzazione. Immagini e problematiche dell'immigrazione*. Vol. 3, Franco Angeli, Milano.
- de Filippo E. 2007. *Il modello di stabilizzazione*. In Orientale Caputo (cur.), pp. 146-175.
- Ehrenreich B. e Hochschild A. R. 2004. *Donne globali. Tate, colf e badanti*. Feltrinelli, Milano.
- Orientale Caputo G. (cur.) 2007. *Gli immigrati in Campania. Evoluzione della presenza, inserimento lavorativo e processi di stabilizzazione*. Franco Angeli, Milano.
- Reyneri E. 2007. *La vulnerabilità degli immigrati*. In Saraceno C., Brandolini A. (cur.). *Disuguaglianze economiche e vulnerabilità in Italia*. Il Mulino, Bologna.
- Strozza S e Ferrara R. 2007. *Vecchi e nuovi volti dell'immigrazione: ritratti per aree di origine e di insediamento*. In Orientale Caputo (cur.), pp. 50-92.
- Strozza S. e Pane A. (cur.) 2000. *Gli immigrati in Campania. Una difficile integrazione tra clandestinità e precarietà diffusa*. L'Harmattan Italia, Torino.
- Strozza S. e Golini A. 2006. *Misure e indicatori dell'integrazione degli immigrati*. In Golini A. (cur.). *L'immigrazione straniera: indicatori e misure di integrazione*. Il Mulino, Bologna.
- Tosi A. 2002. *Immigrati stranieri in Italia: dall'accoglienza alla casa*. Rapporto per European Observatory on Homeless, FEANTSA, Bruxelles, (www.lps.polim.it/download).

ABSTRACT

The aim of this paper is to examine specific aspects of immigrants' integration in Naples, focusing attention on family and housing. The data are drawn from a 2008 sample survey on 402 foreigners coming from Less Development Countries and Central and Eastern Europe. The data collected with a structured form are analysed using multidimensional methods.

Elena DE FILIPPO, Docente di Sociologia delle migrazioni, Università di Napoli Federico II, e Presidente della cooperativa sociale DEDALUS.

Paolo DIANA, Ricercatore confermato, Università di Salerno.

Raffaele FERRARA, Assegnista di ricerca, Università di Napoli Federico II.

Linda FORCELLATI, Ricercatrice confermata, Università di Napoli Federico II.

PERFORMANCE SCOLASTICHE E PRESENZA DI ALUNNI STRANIERI: UN'ANALISI SPAZIALE NEL COMUNE DI ROMA

Anna Di Bartolomeo, Antonella Guarneri, Paola Muccitelli

1. Introduzione

Nonostante la crescente sensibilità da parte del mondo politico e degli operatori in merito ai risultati scolastici della componente straniera, i figli degli immigrati mostrano una maggiore fatica nell'inserimento e nel percorso scolastico rispetto agli autoctoni. Questo fenomeno è particolarmente grave se consideriamo le traiettorie della seconda generazione di immigrati, che, in quanto nati o almeno scolarizzati nei paesi di accoglienza, dovrebbero godere delle stesse opportunità contestuali degli autoctoni.

A partire da queste considerazioni, l'obiettivo del lavoro è approfondire le relazioni tra la presenza di alunni stranieri, le *performance* scolastiche e il territorio dove sono localizzate le scuole, considerando al contempo altre determinanti legate al territorio. A tal fine sono stati applicati strumenti di analisi spaziale. Come dettaglio territoriale è stato scelto il comune di Roma sia per l'elevata quota di alunni stranieri nei vari ordini scolastici che per l'eterogeneità delle cittadinanze rappresentate.

Il resto del lavoro è così strutturato: il secondo paragrafo traccia un quadro di riferimento; nella terza parte, vengono presentati i dati utilizzati e la metodologia adottata; la quarta sezione riporta e descrive i risultati dell'analisi; il quinto paragrafo conclude.

2. Quadro di riferimento

Con riferimento alla regolarità dei percorsi scolastici, emergono elevati differenziali tra alunni stranieri e italiani. Per l'anno scolastico 2007/2008 il tasso di ripetenza nelle scuole secondarie di I grado degli alunni stranieri è pari al 6,4% (contro il 2,7% degli italiani) e il tasso di ritardo scolastico¹ degli alunni stranieri è pari a 51,7% (6,8% per gli italiani). Tali differenze sono da attribuire

¹ Il tasso di ritardo scolastico è calcolato mediante il rapporto tra gli alunni con età maggiore (ritardo) a quella teorica di frequenza ed il totale degli alunni. Sono esclusi gli alunni frequentanti le scuole serali.

principalmente a difficoltà legate alla conoscenza della lingua italiana e a problemi di integrazione sociale.

Oggi, questi differenziali sono al centro di un acceso dibattito nel mondo scientifico e politico italiano che coinvolge e divide l'opinione pubblica. Le ragioni di questo interesse sono molteplici e di diversa natura. Fra queste, l'aspetto quantitativo assume un'importanza fondamentale: il crescente fenomeno dell'immigrazione nel nostro Paese si riflette, infatti, nella scuola italiana che in dieci anni ha registrato un incremento medio annuo pari al 71,1%. Nell'anno scolastico 2007/08, gli alunni stranieri presenti nel sistema scolastico nazionale ammontavano a 574.133 (6,4% del totale). D'altra parte, bassi tassi di rendimento scolastico coincidono con più alti tassi di disoccupazione, minori redditi, maggiore devianza e una più forte dipendenza dallo stato sociale (Rumberger, 1987). Dal punto di vista politico, i costi sociali associati a bassi rendimenti scolastici possono perciò essere enormi e difficili da sostenere nel lungo periodo.

3. Dati e Metodi

I dati utilizzati provengono dalla Rilevazione Integrativa sulle Scuole Secondarie di I grado condotta annualmente dal Ministero dell'Istruzione, Università e Ricerca (Miur). Questa rilevazione presenta un ricco patrimonio informativo che permette di descrivere e monitorare le scuole secondarie di I grado sia in termini di risorse (dotazioni organiche e personale) che di informazioni sugli studenti anche in relazione ai risultati conseguiti.

L'anno di riferimento dell'analisi è il 2007/2008; sono state considerate 217 scuole secondarie di I grado situate nel comune di Roma (Figura 1)². Al fine di descrivere le caratteristiche delle scuole sono stati costruiti i seguenti indicatori: tipo di gestione della scuola (scuola pubblica e scuola privata) e presenza/assenza di servizi aggiuntivi (mensa, scuolabus e pre-scuola)³.

Gli indicatori utilizzati per descrivere la popolazione scolastica delle 217 scuole analizzate sono: tasso di ripetenza (percentuale degli alunni ripetenti sul totale degli alunni) e percentuale degli alunni con cittadinanza non UE sul totale degli alunni⁴.

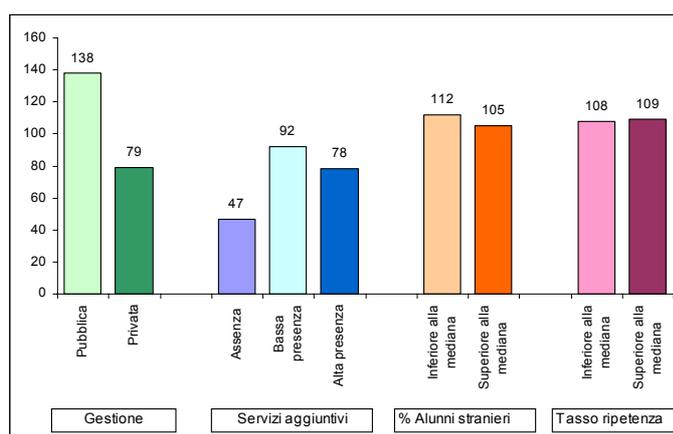
² Sono state escluse 5 scuole per incompletezza delle informazioni a disposizione.

³ Il fenomeno è stato sintetizzato tramite una variabile dicotomica per distinguere le scuole con un alto livello di servizi aggiuntivi (almeno due) dalle scuole con bassa presenza (solo un servizio offerto) o assenza di servizi.

⁴ Questa variabile è stata dicotomizzata in alta presenza (valori superiori alla mediana) e bassa presenza (valori inferiori alla mediana).

Come metodologia di analisi si è fatto ricorso all'analisi spaziale che fa riferimento a dati statistici georeferenziati, dove la variabile geografica non assume solo un ruolo di supporto e/o illustrativo ma diviene un fattore esplicativo del fenomeno analizzato. L'applicazione di tali strumenti permette di studiare le relazioni tra la presenza di alunni stranieri, le *performance* scolastiche e il territorio dove è localizzata la scuola. In questo lavoro, le tecniche di analisi spaziale sono state applicate sulle unità puntuali, ovvero le singole unità scolastiche. In questo caso, poiché la dimensione territoriale è stata codificata in unità spaziali puntuali, l'approccio utilizzato è quello della *Point Pattern Analysis*⁵. Le tecniche di analisi sono state applicate a: dati spaziali semplici (posizione dove l'evento si manifesta, es. ubicazioni delle scuole raggruppate per tipo di gestione) e dati spaziali complessi (posizione dove l'evento si manifesta con diversa intensità, es. % di alunni stranieri nelle singole scuole).

Figura 1 – Numero di scuole per alcune caratteristiche.



Fonte: elaborazione su dati Miur.

Per la georeferenziazione delle unità statistiche, ovvero l'identificazione della posizione dell'unità considerata attraverso l'attribuzione di una coppia di coordinate geografiche, è stato utilizzato il software *Microsoft AutoRoute*. Per l'applicazione delle tecniche di analisi spaziale è stato utilizzato il programma statistico *Crimestat*⁶. I passi dell'analisi spaziale condotta sono stati i seguenti: 1)

⁵ Si utilizza la *Spatially Continuous Data Analysis* nel caso di codifica dello spazio in unità lineari e la *Area Data Analysis* nel caso di unità areali.

⁶ Software sviluppato sotto la direzione di Ned Levine e finanziato dal *National Institute of Justice* degli Stati Uniti per l'analisi della criminalità sul territorio.

Statistiche centrografiche per avere informazioni sulla configurazione spaziale del fenomeno. Il calcolo degli indici centrografici ha permesso di individuare il baricentro spaziale della distribuzione (indici di tendenza centrale) e quantificare il grado di eterogeneità spaziale (indici di dispersione); 2) Analisi della autocorrelazione spaziale per studiare l'influenza della contiguità territoriale sull'intensità delle variabili considerate. Tale analisi tiene conto sia della distanza tra le unità spaziali sia delle intensità del carattere statistico rilevato sulle unità spaziali; 3) Analisi del vicinato per acquisire informazioni sulla struttura (casuale, clusterizzata, uniforme) della distribuzione spaziale.

4. Principali risultati

4.1 *Statistiche centrografiche*

La configurazione spaziale del fenomeno attraverso gli indici di tendenza centrale (baricentro aritmetico) ha permesso di individuare una particolare collocazione dei baricentri delle scuole caratterizzate da una popolazione scolastica multietnica. Infatti, rispetto al centro geografico del totale delle scuole secondarie di I grado del comune di Roma, i baricentri delle scuole con alta e bassa presenza di stranieri sono simmetrici: a nord-est il primo e a sud-ovest il secondo. In effetti, dalla figura 2a sembrerebbe emergere questa contrapposizione: un sud-ovest caratterizzato dalla presenza dei baricentri delle scuole private, di quelle ad alta presenza di servizi aggiuntivi e bassa presenza di stranieri, un nord-est con alta presenza di stranieri e mancanza di servizi aggiuntivi. Gli indicatori di contesto sembrerebbero perciò mostrare un quadro più favorevole per le scuole con un minor peso di alunni stranieri.

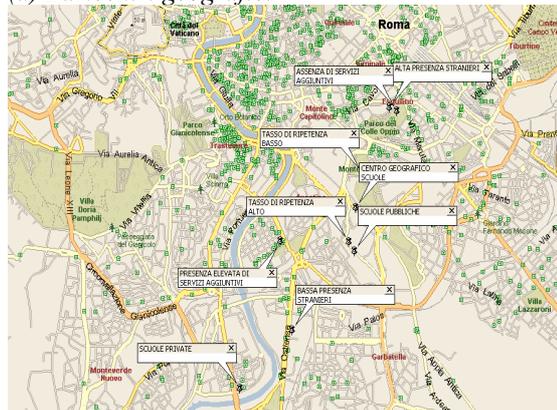
Analizzando le *performance* degli alunni a partire dal tasso di ripetenza, notiamo come il centro medio delle scuole caratterizzate da bassi tassi di ripetenza sia più vicino alla parte geografica dove si collocano le scuole teoricamente più disagiate, mentre il contrario vale per le scuole con *performance* negative.

Infine, è interessante notare come le scuole private sembrano rappresentare un mondo a parte, essendo il loro baricentro notevolmente distante da tutti gli indicatori considerati.

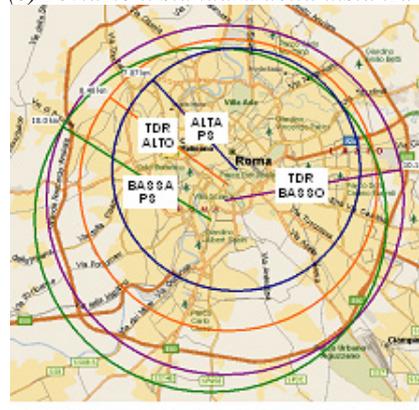
Il grado di eterogeneità territoriale è sintetizzato dalla misura della deviazione standard della distanza (Dds) che registra la dispersione media dei punti-evento rispetto al centro di minima distanza e fornisce informazioni sulla variabilità territoriale. Graficamente la Dds è rappresentata dal raggio del cerchio tracciato sulla mappa. L'analisi della dispersione sul territorio fa emergere considerazioni interessanti (Figura 2b). Le scuole con un'alta presenza di stranieri e un alto tasso di ripetenza si presentano meno disperse sul territorio del comune di Roma.

Figura 2 – Statistiche centrografiche delle scuole per varie caratteristiche.

(a) Baricentri geografici



(b) Deviazioni standard della distanza



Fonte: elaborazione su dati Miur.

4.2 L'analisi della autocorrelazione spaziale

L'autocorrelazione spaziale è basata sulla prima legge di *Tobler* sulla geografia secondo cui “tutte le cose sono collegate tra loro, ma gli oggetti vicini sono sempre più collegati rispetto a quelli lontani” (Tobler W., 1970). L'autocorrelazione spaziale studia l'associazione tra la posizione spaziale delle unità (latitudine e longitudine) e le intensità del fenomeno ad esse riferite, per identificare e misurare la tendenza all'aggregazione/dispersione confrontando i valori di unità vicine. Per l'analisi dell'autocorrelazione spaziale si ipotizza, inoltre, la dipendenza delle osservazioni: “i valori osservati in una certa posizione nello spazio influenzano generalmente i dati in località vicine” (Zani S., 1994).

La misura utilizzata per tale analisi è l'indice di *Moran* che compara i valori di ciascuna area con tutte le altre (enfattizza la co-variabilità rispetto al valore medio)⁷.

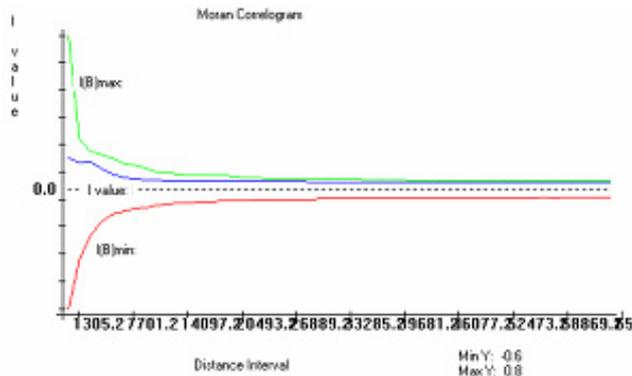
Si rileva un indice di *Moran* statisticamente significativo per l'analisi della autocorrelazione delle scuole con alunni ripetenti e bassa presenza di alunni stranieri (Tabella 1). L'autocorrelazione è positiva: scuole con una bassa presenza di stranieri vicine presentano tassi di ripetenza simili.

⁷ Nel caso di autocorrelazione nulla, ovvero la posizione spaziale non condiziona le modalità con le quali il carattere si manifesta nello spazio esaminato, l'indice assumerà il valore $I = -1/(n-1)$. In presenza di autocorrelazione positiva, unità spaziali vicine presentano modalità simili del carattere statistico esaminato, l'indice sarà $I > -1/(n-1)$. Infine, nel caso di autocorrelazione negativa, unità spaziali vicine presentano modalità dissimili del carattere statistico, l'indice sarà $I < -1/(n-1)$.

Tabella 1 – *Indice di Moran per tasso di ripetenza e presenza straniera nelle scuole.*

	Tasso di ripetenza		
	Scuole secondarie di I grado	Scuole secondarie di I grado con bassa presenza di alunni stranieri	Scuole secondarie di I grado con alta presenza di alunni stranieri
n	217	112	105
I Moran	0,02034	0,03825	0,02932
E (I Moran)	0,00463	-0,00901	-0,00962
Z	0,76121	2,55647	0,33329
P-value (one tail)	n.s.	0,01	n.s.
P-value (two tail)	n.s.	0,05	n.s.

La dipendenza spaziale è evidente anche dall'analisi del correlogramma (Figura 3), dove per *lags* bassi vi sono valori dell'indice di *Moran* elevati: le scuole con una bassa presenza di alunni stranieri e valori simili di tasso di ripetenza tendono ad essere vicine. Tuttavia i valori dell'indice decrescono all'aumentare dei *lags*, pur mantenendosi superiori allo zero, dimostrando l'esistenza di autocorrelazione anche a distanze elevate.

Figura 3 – *Correlogramma di Moran tra scuole con bassa presenza di stranieri e tasso di ripetenza.*

4.3 L'analisi del vicinato

L'analisi del vicinato tende a stabilire se la distribuzione dei punti evento è di tipo casuale, clusterizzata (processo aggregativo) o dispersa (processo repulsivo). I metodi di analisi del vicinato sono vari; in questo lavoro si è utilizzata la *Nearest Neighbour Analysis* che esalta il ruolo della distanza spaziale come fattore esplicativo. Essa consiste nel confrontare la media delle distanze tra vicini,

calcolata sui dati osservati, con la media calcolata su una distribuzione casuale dei punti-evento (dati teorici). Se la media delle distanze tra vicini è maggiore della distanza media teorica, si può affermare che la configurazione spaziale dei punti-evento tende ad essere dispersa (processo repulsivo). Al contrario, se la media delle distanze osservate è minore della teorica il processo è di tipo aggregativo, per cui i punti evento tendono a raggrupparsi. Il *Nearest Neighbour Index* (NNI) è perciò il rapporto tra le due medie (osservate/teoriche)⁸.

Tutti i nostri indicatori mostrano un processo di tipo aggregativo statisticamente significativo (Tabella 2). In particolare, sono le scuole a bassa presenza di stranieri e con un basso tasso di ripetenza a denotare rispetto agli altri una tendenza maggiore alla clusterizzazione. In entrambi i casi, spicca perciò una dimensione fortemente locale del fenomeno.

Tabella 2 – *Nearest Neighbour Index (NNI) su presenza di stranieri, tipo di gestione e tassi di ripetenza delle scuole.*

Scuole	NNI	sign.
Alta presenza di stranieri	0,84	**
Bassa presenza di stranieri	0,69	***
Scuole pubbliche	0,87	**
Scuole private	0,83	**
Alto tasso di ripetenza	0,83	***
Basso tasso di ripetenza	0,58	***

*** p<0.0001; **p<0.01; * p<0.05

5. Conclusioni

Da questo primo tentativo di applicazione di analisi spaziale è emerso come la presenza straniera sia meno dispersa sul territorio del comune di Roma rispetto alle altre variabili considerate. Tra i dati spaziali semplici le scuole private, invece, hanno mostrato un'ampia dispersione territoriale.

L'autocorrelazione positiva tra bassa presenza di alunni stranieri e tasso di ripetenza ribadisce la forte connotazione territoriale delle scuole in cui la presenza di stranieri è inferiore alla mediana. Quante di queste sono vicine tra loro presentano valori simili del tasso di ripetenza. Viceversa l'autocorrelazione tra scuole con alta presenza di stranieri e tasso di ripetenza non è risultata

⁸ NNI < 1 Processo aggregativo: eventi sono più vicini tra loro di quanto ci si aspetterebbe; NNI > 1 Processo repulsivo: eventi tendono a disperdersi.

statisticamente significativa. Dunque, l'associazione tra questi due fenomeni sembra non essere rilevante. In un contesto politico e sociale dove appare frequente l'associazione tra scuole con standard non elevati e scuole dove la presenza di stranieri è cospicua questo risultato si pone come un invito ad adottare una prospettiva di analisi e intervento più articolata e meno stereotipata.

Riferimenti bibliografici

Casacchia O., Natale L., Guarneri A. (2009), *Tra i banchi di scuola. Alunni stranieri e italiani a Roma e nel Lazio*, Franco Angeli, Roma.

Miur (2009), *Alunni con cittadinanza non italiana, Scuole statali e non statali, Anno scolastico 2007-2008*, Miur, Roma.

Portes A., Rumbaut R. G. (2001), *Legacies: the story of the immigrant second generation*, University of California, New York.

Rumberger R.W. (1987), High school dropout: a review of issues and evidence, *Review of Educational Research*, No. 57, 101-121.

Tobler W. (1970), A computer movie simulating urban growth in the Detroit region, *Economic Geography*, 46 (2), 234-240.

Truglia F. G. (2006), La configurazione spaziale del consenso elettorale nel Comune di Roma, *Sociologia e ricerca sociale*, n.81, Franco Angeli, Milano.

Zani S. (1994), *Analisi dei dati statistici*, Giuffrè, Milano.

SUMMARY

This paper aims to deepen the relationships among the presence of immigrants' children per school, scholastic performances and the schools' spatial locations. Data have been taken from the Supplementary Survey on Lower Secondary Schools 2007/2008, held by the Italian Ministry of Education, University and Research. We adopt a spatial analysis approach where geographical variables assume an explicative function abandoning their traditional descriptive role. Our results show the positive autocorrelation between low presence of immigrants' children and repetition rates. Vice versa, the autocorrelation between high presence of immigrants' children and repetition rates is not statistically significant. On this respect our study is relevant as our findings induces to abandon easy and dangerous associations for adopting, instead, a more rational and positive perspective.

Anna DI BARTOLOMEO, Dottoranda in Demografia, Università di Roma La Sapienza.

Antonella GUARNERI, Ricercatore Istat.

Paola MUCCITELLI, Ricercatore Istat.

SEGMENTAZIONE DELLA DOMANDA DI VINO IN ITALIA MEDIANTE MAPPE DI KOHONEN

Marta Disegna, Christine Mauracher, Isabella Procidano

1. Introduzione

Una valida politica di comunicazione deve trovare fondamento nell'analisi del comportamento del consumatore obiettivo. Non è possibile, infatti, comunicare in modo appropriato se non si conosce il destinatario del messaggio.

Al fine di comprendere i consumatori, le loro caratteristiche e per tentare di determinare i loro bisogni e le loro abitudini d'acquisto risulta necessario segmentare il mercato. La segmentazione può essere definita come il processo attraverso il quale le imprese suddividono la domanda in insiemi di clienti potenziali, in modo che gli individui che appartengono allo stesso segmento siano caratterizzati da funzioni di domanda il più possibile simili tra loro e contemporaneamente il più possibile diverse da quelle degli altri insiemi (Valdani, 1986). Le variabili attraverso le quali segmentare il mercato sono innumerevoli e lasciano ampio spazio alla fantasia degli operatori di marketing. Tuttavia la scelta delle variabili più adatte alla segmentazione di un particolare mercato è una decisione molto importante e complessa sulla base della quale si definiscono le strategie d'impresa. I criteri di selezione possono essere costituiti, ad esempio, dalle caratteristiche socio-demografiche, dalle occasioni e dai comportamenti d'uso e d'acquisto del prodotto e dai bisogni, desideri, benefici ricercati.

Nel settore vitivinicolo sempre più si assiste ad un dualismo nelle strategie di identificazione dei vini che vede, da un lato, i vini a forte connotazione territoriale e, dall'altro, i vini di marca (Cesaretti *et al.*, 2006). In particolare l'offerta italiana di vino si ripartisce in due principali categorie: i vini da tavola (nelle sottocategorie dei vini da tavola senza altre qualificazioni e dei vini da tavola con Indicazione Geografica Tipica – IGT) e i vini a denominazione di origine (nelle sottocategorie dei vini a Denominazione di Origine Controllata – DOC e di vini a Denominazione di Origine Controllata e Garantita – DOCG)¹. Attualmente a livello nazionale oltre

¹ La recente riforma dell'Organizzazione Comune di Mercato (OCM) prevede, a partire dalla campagna produttiva del 2009, la distinzione in "vini senza indicazione geografica" e "vini con indicazione geografica". In quest'ultima categoria, ispirandosi alla normativa relativa alle DOP e IGP (Reg. CE 510/2006), sono previsti "vini con Denominazione di Origine Protetta" e "vini con Indicazione Geografica Protetta".

il 30% della produzione totale di vino è costituita da vini a denominazione di origine che deriva dalla produzione di 316 DOC e 42 DOCG.

Considerata la sempre maggiore rilevanza del territorio d'origine nella differenziazione e valorizzazione del prodotto vino e di conseguenza nella comunicazione, risulta interessante una segmentazione che tenga conto anche di questi aspetti.

Obiettivo di questo lavoro è quello di segmentare la domanda di vino in Italia utilizzando una tecnica non supervisionata, le mappe di Kohonen (Kohonen, 1984).

Tale metodologia è stata applicata al fine di individuare il profilo del consumatore di vino Spumante, vino a denominazione di origine e da tavola, distinguendo, per queste ultime tipologie, anche in base al colore del prodotto.

2. Materiali e metodi

2.1 Descrizione del dataset

I dati utilizzati provengono da un'indagine campionaria, condotta dall'Osservatorio AcNielsen, riguardante gli acquisti domestici delle famiglie italiane nel 2003. Dall'indagine sono esclusi gli acquisti extra-domestici riferibili al canale Ho.re.ca. Il campione è costituito da 6.000 famiglie ed è rappresentativo dell'intero universo nazionale. La rilevazione degli acquisti è stata effettuata con il procedimento di *home scanning* che, attraverso il codice a barre, prevede la registrazione da parte della famiglia di tutti i prodotti acquistati presso diversi punti commerciali². Il campione analizzato riguarda le 4244 famiglie italiane che hanno consumato vino nel periodo considerato. Le variabili impiegate per la segmentazione sono state le quantità consumate (in Litri) dalla famiglia per ogni tipologia di vino e la spesa totale di vino. Le variabili utilizzate nella profilazione dei consumatori sono state: l'età, il genere e il titolo di studio della persona di riferimento oltre all'area geografica di residenza e al livello di reddito della famiglia (tabella 1).

² L'indagine non considera il canale di vendita diretto che, per alcune regioni, ricopre un ruolo rilevante.

Tabella 1 – Analisi delle frequenze per le variabili di profilazione.

Variabili	Freq.	%	Variabili	Freq.	%
Età			Reddito		
20 -44	1109	26.13	Alto	1667	39.28
44 -60	1825	43.00	Medio-alto	1240	29.22
Oltre 60	1310	30.87	Medio-basso	972	22.90
Genere			Basso	365	8.60
Maschio	3654	13.9	Area geografica		
Femmina	590	86.1	Nord Est	707	16.66
Tit. di studio			Nord Ovest	1266	29.83
Alto	556	13.10	Centro	840	19.79
Medio-alto	1941	45.74	Sud	1431	33.72
Medio-basso	1295	30.51			
Basso	452	10.65			

2.2 Le mappe di Kohonen

Le mappe auto-organizzate SOM (*Self-Organizing Map*), o mappe di Kohonen, sono il metodo statistico utilizzato in questo lavoro. Tale metodologia è stata proposta in letteratura come strumento in grado di proiettare dei punti osservati in uno spazio di input p -dimensionale in uno di output, o mappa, di dimensioni inferiori. Nello spazio di input la j -esima unità statistica è descritta tramite un vettore p -dimensionale, detto nodo di input, $\mathbf{x}_j = [x_{j1}, x_{j2}, \dots, x_{jp}]$, $j=1, \dots, N$ (dove N rappresenta il numero complessivo di unità statistiche). Lo spazio di output è composto da M nodi (o neuroni) disposti generalmente in una griglia bidimensionale di tipo esagonale o rettangolare. Ad ogni nodo dello spazio di output sono associati due vettori: un vettore p -dimensionale detto prototipo, che rappresenta tutti i punti dello spazio di input associati a tale nodo, $\mathbf{m}_i = [m_{i1}, m_{i2}, \dots, m_{ip}]$, $i=1, \dots, M$; un vettore q -dimensionale (con $q < p$) \mathbf{r}_i di coordinate che ne identifica la posizione sulla griglia. Per la costruzione della mappa si usano algoritmi iterativi che dipendono da un parametro di apprendimento α ($0 < \alpha < 1$) e da un raggio β da fissare a priori. Schematicamente l'algoritmo di base richiede: 1) l'inizializzazione degli M prototipi; 2) il posizionamento (o competizione), ossia l'individuazione del prototipo che nello spazio di input è più vicino alla j -esima unità (la metrica solitamente utilizzata è la distanza euclidea), e quindi l'associazione della j -esima unità al nodo vincitore $c(j)$, o *Best Matching Unit* (BMU), determinato sulla base della seguente regola:

$$c(j) = \arg \min_{1 \leq i \leq M} \{ \|\mathbf{x}_j - \mathbf{m}_i\| \} \quad (1)$$

3) l'aggiornamento (o apprendimento) che prevede la modifica sia del prototipo associato al nodo vincitore sia dei prototipi associati ai nodi adiacenti, secondo la

regola seguente:

$$\mathbf{m}_i(k+1) = \mathbf{m}_i(k) + h_{c(j),i}(k) [\mathbf{x}_j - \mathbf{m}_i(k)] \quad (2)$$

dove $k=0, \dots, K$ identifica il passo di iterazione e $h_{c(j),i}(k)$ indica il vicinato³ del neurone vincente $c(j)$ in corrispondenza dell'iterazione corrente (*neighborhood function*). Le performance dell'algoritmo di base dipendono dal tasso di apprendimento e dal raggio che definisce il vicinato. E' possibile utilizzare funzioni di vicinato, $h_{c(j),i}(k)$, di tipo continue che consentono di far variare nel tempo sia il tasso di apprendimento sia il raggio dell'intorno impiegando. La più utilizzata in letteratura è quella Gaussiana:

$$h_{c(j),i}(k) = \alpha(k) \exp\left(-\frac{\|\mathbf{r}_{c(j)} - \mathbf{r}_i\|^2}{2\sigma^2(k)}\right) \quad (3)$$

dove il tasso di apprendimento $\alpha(k) \in (0, 1)$ e l'ampiezza della funzione di vicinanza $\sigma(k)$, che può essere interpretata come il raggio, sono funzioni decrescenti all'aumentare delle iterazioni, $\mathbf{r}_{c(j)}$ è il vettore contenente le coordinate del nodo vincitore $c(j)$ e \mathbf{r}_i è il vettore delle coordinate dell' i -esimo nodo nello spazio di output. I passi 2) e 3) vengono quindi ripetuti fino a quando i nodi convergono. La versione *Batch* dell'algoritmo di base (Kohonen, 1995) è una variante che migliora l'efficacia e la performance della procedura di apprendimento senza supervisione. In tal caso, nella fase di competizione, ad ogni iterazione tutte le unità di input sono sottoposte contemporaneamente alla rete. Inoltre la regola di apprendimento è la seguente:

$$\mathbf{m}_i^*(k+1) = \frac{\sum_{j=1}^N h_{c(j),i}(k) \mathbf{x}_j}{\sum_{j=1}^N h_{c(j),i}(k)} \quad (4)$$

In questo algoritmo non è necessario specificare il tasso di apprendimento $\alpha(k)$, mentre $\sigma(k)$ rimane una funzione monotona decrescente durante la fase di aggiornamento. I risultati prodotti dalle SOM sono sintetizzabili in rappresentazioni grafiche (*heat maps*) tra le quali le più significative sono la *U*-matrix e il piano delle componenti. La *U*-matrix consente di visualizzare la sintesi completa dei risultati ed è costruita sulla base della matrice delle distanze tra i nodi

³ Con il termine vicinato di un nodo si intende l'insieme dei nodi di output aventi una distanza dal vincitore minore di β , raggio di interazione.

di output. All'interno di ogni nodo di output sono collocate le unità di input per le quali esso è risultato vincitore. Le distanze sono visualizzate usando colori diversi.

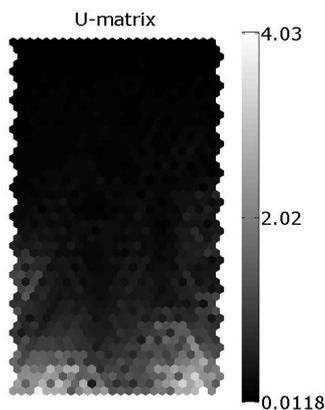
Il piano delle componenti è costituito da tante mappe quante sono le variabili rilevate sulla j -esima unità statistica: ogni mappa può essere vista come una porzione della U -matrix che consente, per ogni singola variabile, di osservare la posizione e la modalità assunte da ciascuna unità. La U -matrix e le mappe che costituiscono il piano delle componenti possono essere sovrapposte, permettendo, attraverso tale lettura, di ricostruire la matrice dei dati iniziali.

3. Risultati e conclusioni

Per l'applicazione abbiamo adottato una mappa di Kohonen utilizzando la versione *Batch* dell'algoritmo e disponendo i nodi in una griglia di tipo esagonale. I valori iniziali dei pesi sono stati scelti seguendo una procedura di inizializzazione lineare (Kohonen, 1984) e il numero di nodi dello spazio di output sono stati calcolati sulla base della seguente formula $M = 5\sqrt{N}$ (Sang *et al.*, 2008).

La U -matrix e il piano delle componenti sono visualizzati utilizzando i colori della scala di grigio. In particolare nella U -matrix i nodi colorati di bianco indicano disomogeneità rispetto ai nodi vicini nella mappa, delineando così i bordi dei possibili segmenti (cluster) individuati dall'algoritmo, mentre sono colorati in nero i neuroni vicini e omogenei tra loro (figura 1).

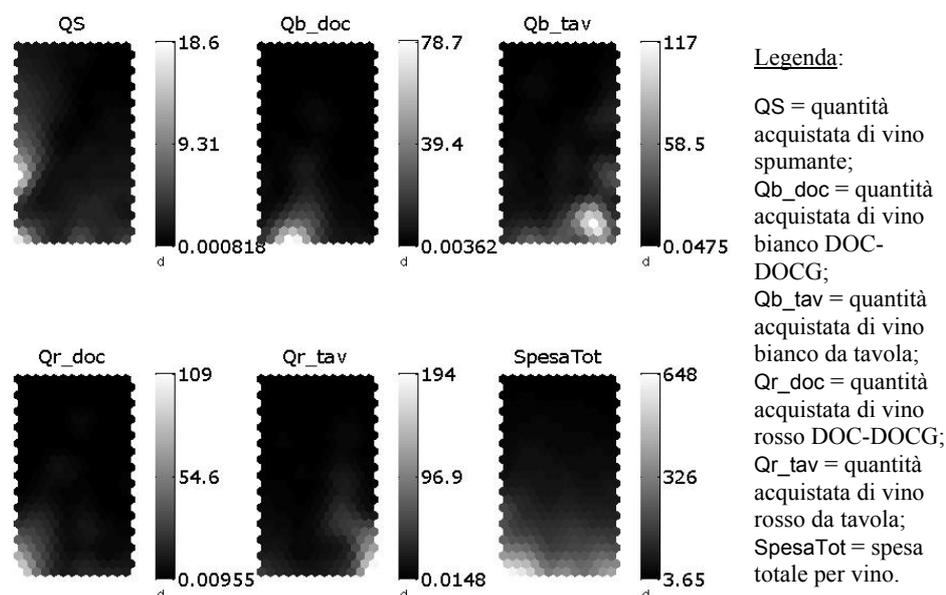
Figura 1 – U -Matrix.



Si può così osservare come i nodi con le distanze maggiori si concentrino nella parte in basso della mappa e non identifichino un contorno ben definito, in base al

quale individuare specifici cluster. In altre parole ciò significa che, considerando congiuntamente le differenti tipologie di vino, non vi sia da parte delle famiglie un comportamento di consumo omogeneo. La disomogeneità rilevata in questa zona può essere spiegata analizzando il piano delle componenti (figura 2).

Figura 2 – Piano delle componenti.



Tale piano è costituito, nel nostro caso, da 6 mappe e, in ciascuna, i nodi colorati di bianco (nero) identificano le unità statistiche dello spazio di input che presentano alti (bassi) consumi. Come si può osservare, i consumi di vino, per le diverse tipologie di prodotto, risultano elevati in zone circoscritte delle mappe, individuate dalle aree più chiare, che sono situate per lo più nelle fasce basse, ma in posizioni differenti. Ciò denota una diversità di comportamento delle famiglie a seconda della tipologia di vino considerata, evidenziando una certa eterogeneità nelle preferenze.

I diversi consumatori identificati possono essere profilati utilizzando le variabili descritte nel paragrafo 2.1. I risultati ottenuti sono sintetizzati in tabella 2. Per quanto riguarda l'età, si può osservare come le famiglie con una persona di riferimento di età inferiore a 44 anni non rientrino nella categoria degli 'alto' consumatori per nessuna tipologia di vino. Ciò presumibilmente è imputabile all'esclusione dall'indagine di specifici canali di acquisto. I consumatori più

anziani (con oltre 60 anni) sembrano prediligere il vino rosso, sia per il vino a denominazione di origine sia per il vino da tavola. Le donne con reddito basso acquistano principalmente vino bianco da tavola, mentre gli uomini con un titolo di studio medio-alto e un alto reddito preferiscono il vino spumante e il vino DOC-DOCG, sia rosso che bianco. Dall'analisi emerge, inoltre, come il comportamento di consumo sia fortemente influenzato dalla variabile territoriale. L'area Nord Ovest presenta il maggiore consumo di spumanti e di vino rosso DOC-DOCG, oltre che la più elevata spesa per l'acquisto di vino in generale. Le regioni del Nord Est primeggiano per il consumo di vino bianco da tavola, mentre il consumatore residente nell'Italia centrale si contraddistingue per un alto consumo sia di vino bianco (DOC-DOCG e da tavola) che di vino rosso da tavola.

Tabella 2 – *Caratteristiche degli 'alto' consumatori per tipologia di vino.*

Tipologia di vino	Età	Genere	Titolo di studio	Area geografica	Reddito
<i>Spumante</i>	Oltre 44 anni	Maschio	Medio-Alto	Nord Ovest	Alto
<i>Rosso DOC</i>	Oltre 60 anni	Maschio	Medio	Nord Ovest	Alto
<i>Bianco Tavola</i>	44 - 59 anni	Femmina	Medio-Alto	Nord Est/Centro	Basso
<i>Bianco DOC</i>	Oltre 44 anni	Maschio	Alto	Centro	Alto
<i>Rosso Tavola</i>	Oltre 60 anni	Maschio	Basso	Centro	Medio-basso

Riferimenti bibliografici

- Cesaretti G. P., Green R., Mariani A., Pomarici E. 2006, *Il mercato del vino. Tendenze e strategie dei concorrenti*, Franco Angeli, Milano.
- Kohonen T. 1984. *Self-organizing and Associative Memory*. (3rd edition 1989) Berlin, Heidelberg, Germany: Springer-Verlag.
- Kohonen T. 1995. *Self-organizing Maps*. Springer Series in Information Sciences, Vol. 30, Springer, Berlin, Heidelberg, New York. (3rd edition 2001)
- Sang H., Gelfand A. E., Lennard C., Hegerl G. e Hewitson B. 2008. *Interpreting Self-Organizing Maps through Space-Time data models*. Institute of Mathematical Statistics in The Annals of Applied Statistics, vol. 2, pp. 1194-1216.
- Valdani E. 1986. *Marketing strategico*. Etas Libri, Milano.

SUMMARY

Demand segmentation of wine in Italy through the Kohonen's Map

The aim of this work is to segment the Italian wine market in order to profile the consumers of sparkling, Controlled Designation of Origin (DOC-DOCG) and table wine. The statistical method adopted is the Kohonen's Map (Kohonen, 1984), an unsupervised learning technique. The dataset is provided by AcNielsen and it regards domestic Italian Household purchases in 2003. The segmentation variables are the quantities (litre) consumed by the family for each type of wine and the total expenditure. The variables used to profile the consumers are: age, gender, educational, geographical area and income level.

The main results show that sparkling and red Controlled Designation of Origin wine are primarily consumed in the North-West of Italy (in this area it is observed the highest total expenditure). The wine-drinking is male, older than 44 years and with an high income level. As for the educational level, the highest is observed for the white Controlled Designation of Origin wine consumer.

Marta DISEGNA, Assegnista di ricerca, Facoltà di Economia, Università Ca' Foscari Venezia.

Christine MAURACHER, Professore associato in Economia Agraria, Facoltà di Economia, Università Ca' Foscari Venezia.

Isabella PROCIDANO, Professore associato in Statistica Economica, Facoltà di Economia, Università Ca' Foscari Venezia.

IL RUOLO DELLA FECONDITÀ DEGLI STRANIERI IN DUE PAESI DI RECENTE IMMIGRAZIONE: IL CASO DELL'ITALIA E DELLA SPAGNA *

Raffaele Ferrara, Piero Giorgi, Marija Mamolo, Salvatore Strozza

1. Introduzione

Crescente è l'attenzione verso il comportamento demografico degli immigrati nei paesi dell'Europa meridionale che, diventati negli ultimi decenni importanti aree di accoglimento, di recente hanno registrato un rapido aumento della popolazione straniera e, di pari passo, un sensibile incremento degli eventi demografici ad essa relativi (Bonifazi, 2007; Cordón, Maldonado, 2006). In particolare, grande interesse assume l'analisi del comportamento riproduttivo degli immigrati in due paesi come l'Italia e la Spagna che hanno sperimentato i livelli più bassi del tasso di fecondità totale (TFT) del momento tra tutti i paesi di una certa dimensione demografica (Livi Bacci, Deldado Perez, 1992).

In questi due paesi gli stranieri sono ormai una realtà di ampie dimensioni assolute (all'inizio del 2007 in Spagna gli stranieri iscritti al *Padrón Municipal* sono quasi 4,5 milioni, in Italia i soli residenti sfiorano i 3 milioni e tenendo conto anche della componente non residente, regolare e irregolare, si arriva appena al di sotto dei 4 milioni), nonché di rilevante peso demografico (nel primo paese quasi il 10% del totale dei residenti; nel secondo poco meno del 7%). Contemporaneamente alla crescita della popolazione straniera si è registrata in entrambi i paesi anche una leggera crescita della fecondità complessiva: nel periodo 1995-2006 il TFT è passato, sia in Italia che in Spagna, da meno di 1,2 a oltre 1,35 figli per donna.

In che misura l'apporto delle donne straniere ha influito sulla recente evoluzione della fecondità in Italia e in Spagna? Quanto la fecondità tardiva è stata contenuta per effetto degli immigrati con comportamenti riproduttivi più intensi nelle fasce d'età più giovani? Tra i due paesi è possibile riscontrare differenze nel contributo delle donne immigrate alla fecondità del momento? E all'interno dei due territori nazionali quanto sono variabili le situazioni tra le diverse aree? A questi quesiti si cercherà di dare risposta ricorrendo a

* Articolo realizzato nell'ambito della ricerca su "La fecondità degli stranieri nei paesi europei di vecchia e nuova immigrazione", diretta da Salvatore Strozza e finanziata dal Dip. di Scienze Statistiche (Un. di Napoli Federico II). Tale ricerca rientra nel PRIN 2008 su "Società in transizione: dai vecchi assetti ai nuovi equilibri" coordinato da Luigi Di Comite. L'articolo è stato ideato e realizzato congiuntamente dagli autori, anche se Ferrara R. ha curato il par. 3.1, Giorgi P. il par. 2.2, Mamolo M. il par. 3.2 e Strozza S. i parr. 1 e 2.1. Le conclusioni sono in comune tra gli autori, mentre Ferrara R. ha predisposto i cartogrammi.

metodi di decomposizione delle variazioni del TFT e della quota di fecondità tardiva nel periodo 2001- 2006, allo scopo di misurare l'apporto dovuto alla componente straniera.

2. Metodi di decomposizione della variazione della fecondità

2.1. Contributo delle donne straniere alla variazione del TFT

Per valutare quanto la variazione negli ultimi anni del TFT del momento sia dovuta al contributo delle donne straniere è stato proposto un modello di decomposizione che consente di misurare l'apporto assoluto e relativo delle donne con cittadinanza del paese considerato (C) e delle donne straniere (S) alla fecondità totale (Strozza *et al.*, 2007). La variazione assoluta del TFT in un dato periodo di tempo (ΔTFT) dipende da tre effetti: la variazione del TFT delle donne autoctone; la variazione del TFT delle donne straniere; la variazione dell'incidenza delle donne straniere nella fascia d'età riproduttiva. In formula:

$$\Delta TFT = \sum \Delta f_x^S \cdot \left({}_0\bar{d}_x^S + \frac{1}{2} \Delta \bar{d}_x^S \right) + \sum \Delta f_x^C \cdot \left({}_0\bar{d}_x^C + \frac{1}{2} \Delta \bar{d}_x^C \right) + \sum \Delta \bar{d}_x^S \left[({}_0f_x^S + \frac{1}{2} \Delta f_x^S) - ({}_0f_x^C + \frac{1}{2} \Delta f_x^C) \right] \quad [1]$$

dove il simbolo Δ indica la variazione nell'intervallo di tempo 0-1, f_x^C e f_x^S il tasso di fecondità all'età x rispettivamente delle autoctone e delle straniere mentre \bar{d}_x^C e \bar{d}_x^S la proporzione tra le donne di età x delle autoctone e delle straniere (anteposto in pedice è il riferimento all'anno).

Seguendo un approccio classico, è stato valutato il contributo di ciascuno dei tre fattori distinguendo gli effetti semplici da quelli dovuti alla variazione congiunta di due o di tutti e tre i fattori e all'effetto interazione. Poiché la variazione della fecondità delle straniere e la variazione della fecondità delle autoctone sono considerate tra loro indipendenti, l'interazione assume valore nullo. Anche l'interazione tripla tra i tre fattori considerati risulta pertanto nulla. Di conseguenza, è possibile sintetizzare le componenti dell'evoluzione nel modo seguente:

$$\begin{aligned} \Delta TFT &= \Delta(TFT^S) + \Delta(TFT^C) + \Delta(\bar{d}^S) + I(TFT^S, \bar{d}^S) + I(TFT^C, \bar{d}^S) = \\ &= \Delta_2(TFT^S) + \Delta_2(TFT^C) + \Delta_2(\bar{d}^S) \end{aligned} \quad [2]$$

in cui gli *effetti semplici* sono dovuti alla variazione di uno dei tre fattori considerati supponendo fissi nel tempo i valori dei restanti due e le *interazioni doppie*, $I(TFT^S, \bar{d}^S)$ e $I(TFT^C, \bar{d}^S)$, sono ottenute sottraendo agli effetti doppi gli effetti semplici (si veda Strozza *et al.*, 2007). Pertanto, il modello di decomposizione della variazione del TFT contiene cinque elementi additivi di cui tre sono gli effetti singoli e due quelli di interazione. Assegnando le interazioni doppie in modo equo tra gli effetti singoli coinvolti, si ritorna però ai tre fattori della formula [1], a cui si farà riferimento nell'articolo. Ciascuna di queste tre componenti non è nient'altro che la somma degli effetti singoli, dovuti alla variazione nel tempo di quella sola componente, e di parte degli effetti delle interazioni doppie di quella componente con ciascuna delle altre due.

2.2. Contributo delle donne straniere alla variazione della fecondità oltre i 30 anni

La proporzione di fecondità tardiva, ovvero realizzata oltre i 30 anni ($\%TFT_{30+}$) fornisce esclusivamente una visione congiunturale del grado di invecchiamento della fecondità complessiva, considerando sempre in maniera implicita i cambiamenti sottostanti nei reali calendari delle coorti (Giorgi, Mamolo, 2007). Così l'incremento della proporzione di fecondità realizzata oltre i 30 anni in una prospettiva di periodo non corrisponde necessariamente ad una crescita della fecondità oltre i 30 anni, per le diverse componenti per nazionalità considerate. Ad esempio, una variazione positiva potrebbe essere dovuta alla dinamica della fecondità nelle età giovani (Giorgi, 1995), come anche cambiamenti nella struttura per ordine di nascita, anche in assenza di modificazioni del calendario nei singoli ordini, possono causare cambiamenti nel calendario complessivo della fecondità

Per valutare il contributo delle donne straniere alla variazione della $\%TFT_{30+}$, si farà riferimento al modello di decomposizione utilizzato in Giorgi e Mamolo (2007). La dimensione relativa all'ordine di nascita viene sostituita dalla dimensione relativa alla cittadinanza della madre, con l'obiettivo di quantificare, pur con i limiti dell'indicatore utilizzato, gli effetti sulla variazione della $\%TFT_{30+}$ della fecondità delle donne con cittadinanza del paese considerato rispetto alla fecondità delle straniere. In questa sede si tenta quindi di spiegare la dinamica della proporzione di fecondità realizzata oltre i 30 anni distinguendo gli effetti dei cambiamenti della fecondità secondo due grandi gruppi di età (15-29 e 30 e più) e secondo la cittadinanza, distinguendo le donne con cittadinanza del paese (C) da quelle straniere (S). Contrariamente a quanto fatto per l'intensità (par. 2.1), in questa parte del contributo si utilizza una misura della fecondità basata sulla popolazione totale.

Secondo il modello generale descritto in Giorgi e Mamolo (2007) che si rifà ai modelli classici di decomposizione additiva con struttura ad effetti singoli, congiunti, di interazione e totali (Kitagawa, 1955; Das Gupta, 1978), la variazione della proporzione di fecondità di periodo realizzata oltre i 30 anni viene decomposta utilizzando un modello a quattro parametri:

$$\Delta \frac{TFT_{30+}}{TFT} = \Delta_{tot} (TFT_{15-29}^C) + \Delta_{tot} (TFT_{30+}^C) + \Delta_{tot} (TFT_{15-29}^S) + \Delta_{tot} (TFT_{30+}^S) \quad [3]$$

I quattro parametri considerati sono la fecondità di periodo alle età 15-29 delle autoctone (TFT_{15-29}^C) e delle straniere (TFT_{15-29}^S), la fecondità di periodo alle età 30 e più delle autoctone (TFT_{30+}^C) e delle straniere (TFT_{30+}^S). Il singolo elemento è dato dalla somma degli effetti semplici, doppi, tripli e dei rispettivi effetti di interazione (Giorgi, Mamolo, 2007).

3. Risultati

L'analisi è stata condotta per l'Italia e la Spagna tra il 2001 ed il 2006. Il livello territoriale considerato, oltre a quello nazionale, è stato quello delle province, seppure l'analisi della fecondità per cittadinanza con tale dettaglio geografico possa porre alcune questioni sul livello qualitativo delle misure ottenute. I dati utilizzati per le nascite e le popolazioni sono stati tratti dalle banche dati online disponibili presso ISTAT e INE.

3.1. Il contributo delle donne straniere alla fecondità del momento

Le stime del TFT del momento per l'Italia e la Spagna evidenziano come nei due paesi si sia verificato tra il 2001 e il 2006 un lieve incremento del livello della fecondità che oggi risulta pari a 1,36 figli per donna in entrambi i paesi (tab. 1). L'incremento assoluto è più ampio in Spagna che in Italia, rispettivamente 138 e 91 figli per 1000 donne, ma è scaturito da determinanti comuni, ovvero dall'accresciuta incidenza delle straniere e dall'aumento della fecondità del momento tra le autoctone. Se nel caso dell'Italia tali componenti sembrano aver agito in egual misura (rispettivamente 48 e 43 figli per 1000 donne), in Spagna risulta prevalente l'effetto dovuto all'incremento della fecondità tra le spagnole (50 contro 106 figli per 1000 donne).

È infine interessante notare come l'effetto dovuto all'intensità del fenomeno tra le straniere sia risultato sempre irrilevante ai fini della ripresa della fecondità complessiva: nullo nel caso dell'Italia e addirittura negativo in Spagna.

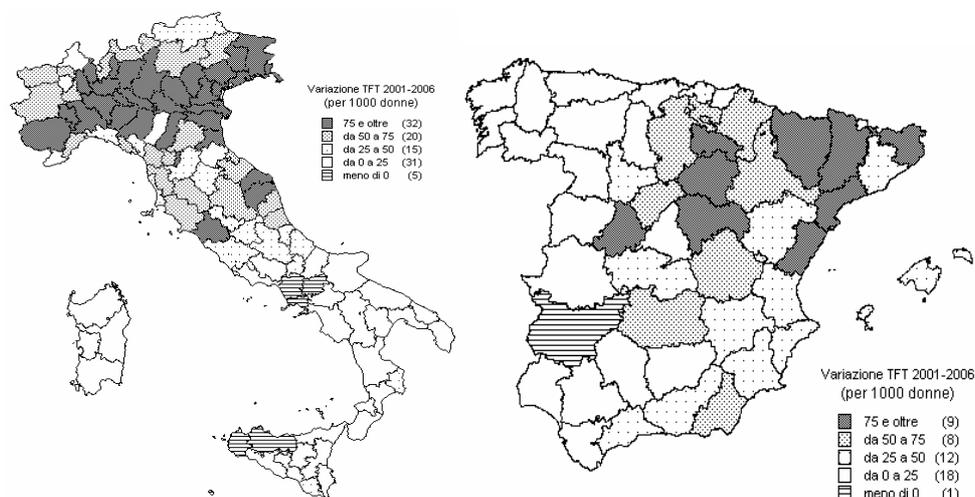
Tabella 1 – *Variazione del TFT del momento tra il 2001 e il 2006 ed effetti dovuti alla variazione dell'incidenza delle donne straniere, del TFT delle straniere e delle autoctone in Italia e in Spagna.*

Paese	TFT (per 1000 donne)		Variazione assoluta TFT 2001-2006	Effetti dovuti alla variazione di:			Effetto complessivo dovuto a stranieri
	2001	2006		Incidenza straniere	TFT straniere	TFT cittadine	
Italia	1265	1356	91	48	0	43	48
Spagna	1219	1357	138	50	-19	106	31

Scendendo dal livello nazionale al dettaglio provinciale i risultati del modello di decomposizione evidenziano le seguenti particolarità: a) mentre in Italia risultano in atto due tendenze differenti, un incremento della fecondità nelle province del Nord e una contrazione della stessa in quelle del Sud, in Spagna la ripresa della fecondità risulta estesa a tutto il territorio nazionale, sebbene con intensità variabile; b) l'apporto complessivo delle straniere alla ripresa della fecondità in Italia appare più consistente nella parte settentrionale del paese, in Spagna invece sembra esistere un gradiente che segue l'asse Ovest-Est; l'ipotesi di fondo è che tali differenze rispecchino sia il diverso peso della popolazione immigrata nelle diverse aree del territorio nazionale, sia i differenti modelli migratori e composizioni per area geografica di provenienza delle collettività in esse presenti; c)

l'effetto incidenza della popolazione immigrata rispecchia chiaramente la geografia insediativa della presenza straniera nei due paesi, segnalando una variazione più marcata nelle province economicamente più forti e dinamiche del Centro-Nord in Italia e della parte orientale della Spagna (fig.1); d) nei due paesi esaminati sono numerose le province nelle quali, durante il periodo 2001-2006, si è registrata una contrazione del TFT tra donne immigrate.

Figura 1 – Contributo delle donne straniere alla variazione della fecondità complessiva del momento in Italia e in Spagna nel periodo 2001-2006.



3.2. L'apporto delle donne straniere alla dinamica della fecondità tardiva

Nel 2001 e nel 2006 la proporzione di fecondità realizzata oltre i 30 anni (%TFT₃₀₊) è pari rispettivamente a 54,11 e 57,96% in Italia e 59,22 e 60,11% in Spagna (tab. 2). In entrambi i paesi la proporzione di fecondità realizzata oltre i 30 anni è aumentata nel periodo 2001-2006, suggerendo, pur con tutti i limiti interpretativi che la misura pone, un invecchiamento della fecondità. Esiste con tutta evidenza un gradiente territoriale Nord-Sud (valori più alti verso valori più bassi) sia nei livelli della %TFT₃₀₊ sia nella sua variazione. Le zone settentrionali di entrambi i paesi mostrano variazioni negative nella %TFT₃₀₊. Il contributo delle autoctone, nei rispettivi paesi, sia al di sopra sia al di sotto dei 30 anni favorisce l'incremento della proporzione di fecondità realizzata oltre i 30 anni: diminuisce la fecondità alle età giovani mentre aumenta la fecondità alle età più elevate.

In Italia il contributo delle donne italiane alla variazione relativa della proporzione di fecondità tardiva è pari a +6,1% al di sotto dei 30 anni e +4,1% al di sopra dei 30 anni. In Spagna il contributo delle spagnole alla variazione relativa della proporzione di fecondità

realizzata oltre i 30 anni è pari a +3,5% al di sotto dei 30 anni e +2,5% al di sopra dei 30 anni. Così il ruolo delle autoctone è, nei rispettivi paesi, sempre positivo ed esse contribuiscono ad accrescere la proporzione di fecondità realizzata oltre i 30 anni. In entrambi i paesi il ruolo maggiore tra le cittadine è ascrivibile alla riduzione della fecondità nella fascia di età al di sotto dei 30 anni (+6,1% e +3,5%), ossia l'invecchiamento della fecondità è maggiormente dovuto alla sua contrazione nelle fasce di età più giovani. In entrambi i paesi le straniere tra 15 e 29 anni fanno sempre ridurre la proporzione di fecondità realizzata oltre i 30 anni, ovvero sembrano contribuire a ringiovanire la fecondità o meglio a frenare il suo invecchiamento imposto dalle donne locali: in Italia le italiane contribuiscono con +10,2% frenato dal -3,1% delle straniere, in Spagna invece le spagnole contribuiscono con +5,95% frenato dal -4,45% delle straniere. Così le autoctone, nei rispettivi paesi, fanno sempre aumentare la %TFT₃₀₊ producendo invecchiamento nella fecondità, mentre le straniere hanno sempre un ruolo contrario.

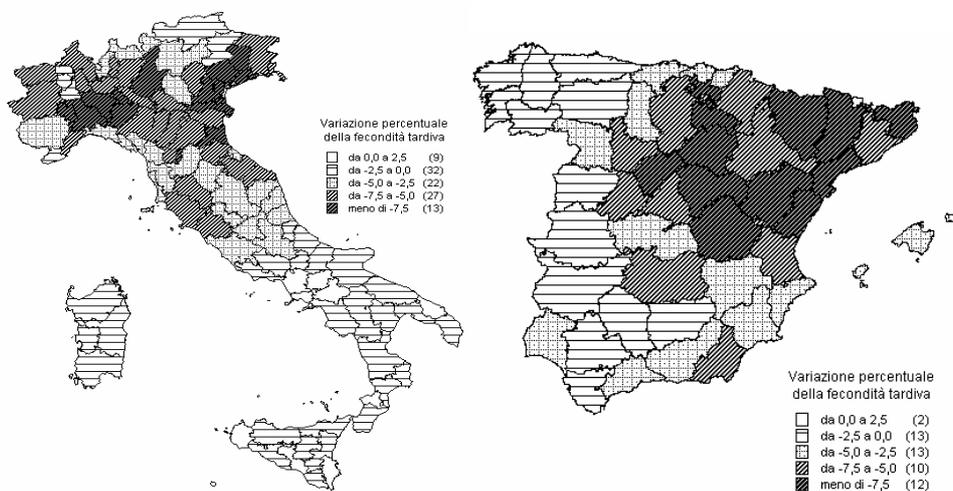
Tabella 2 – *Proporzione di fecondità oltre 30 anni e più, variazioni relative della proporzione di fecondità a 30 anni e più, effetti singoli ed effetti aggregati relativi.*

	Italia			Spagna		
	2001	2001-06	2006	2001	2001-06	2006
%TFT ₃₀₊ / TFT	54,10		57,96	59,22		60,11
%Δ(TFT ₃₀₊ / TFT)		7,1251			1,4987	
<i>Effetti singoli</i>						
Δ _{tot} (TFT ₁₅₋₂₉ ^C)		6,0864			3,4728	
Δ _{tot} (TFT ₃₀₊ ^C)		4,1234			2,4808	
Δ _{tot} (TFT ₁₅₋₂₉ ^S)		-5,2177			-6,8938	
Δ _{tot} (TFT ₃₀₊ ^S)		2,1330			2,4388	
<i>Effetti aggregati</i>						
Δ _{tot} (TFT ^C)		10,2098			5,9536	
Δ _{tot} (TFT ^S)		-3,0848			-4,4549	
Δ _{tot} (TFT ₁₅₋₂₉)		0,8687			-3,4209	
Δ _{tot} (TFT ₃₀₊)		6,2564			4,9196	

A livello provinciale i parametri del modello consentono di evidenziare alcuni aspetti particolari: a) in entrambi i paesi c'è un gradiente territoriale nella proporzione di fecondità realizzata oltre i 30 anni - in particolare in Italia questo gradiente va dai valori più bassi del Sud a quelli più elevati del Nord, mentre in Spagna al gradiente Sud-Nord sembra aggiungersi quello Est-Ovest; b) esiste anche un gradiente nella variazione di %TFT₃₀₊ con la fecondità che tende a ringiovanire nelle regioni che hanno raggiunto i livelli più alti di fecondità tardiva; c) nelle province del Sud Italia è il calo della fecondità nelle età giovani che gioca il ruolo maggiore, mentre nel Nord si manifesta con evidenza anche il ruolo della effettiva fecondità oltre i 30 anni; nel Nord la fecondità sotto i 30 anni fa invecchiare la fecondità complessiva - simili conclusioni valgono anche per la Spagna anche se il gradiente territoriale è meno netto ed evidente; d) le straniere hanno sempre un ruolo di freno all'aumento di %TFT₃₀₊, sebbene

vi siano alcune eccezioni nel Sud Italia, con un evidente gradiente territoriale sia in Italia sia in Spagna (fig. 2); e) il contributo delle donne straniere nelle età inferiori ai 30 anni riduce l'invecchiamento della fecondità, mentre il contributo delle donne straniere al di sopra dei 30 anni, seppure positivo, rimane piuttosto modesto. Il gradiente territoriale molto evidente mette in luce il freno crescente da Sud a Nord in Italia e da Sud-Est al Nord-Ovest in Spagna e mostra come nelle regioni settentrionali di entrambi i paesi il freno delle straniere giovani sia più forte; in Italia ciò è collegabile direttamente all'incidenza delle straniere stesse.

Figura 2 – Contributo delle donne straniere alla variazione della fecondità tardiva in Italia e in Spagna nel periodo 2001-2006.



4. In breve ... e per continuare

La variazione della fecondità totale nel periodo 2001-2006 è risultata chiaramente dovuta, sia in Italia che in Spagna, a due fattori principali quali l'incremento della presenza straniera sul territorio ed il lieve incremento della fecondità tra le donne autoctone. Durante il periodo considerato la proporzione di fecondità tardiva è aumentata in entrambi i paesi, svolgendo però le donne straniere sempre un ruolo di freno all'invecchiamento della fecondità.

Restano a questo punto alcuni aspetti degni di essere approfonditi. Sul piano dei modelli di decomposizione appare necessario superare, per quanto possibile, la separazione netta tra le due soluzioni adottate per giungere ad un modello unico in grado di controllare contemporaneamente le diverse dimensioni considerate. In termini di approccio sembra importante valutare se e come sia possibile far entrare nelle analisi la dimensione relativa alle generazioni: tale aspetto solleva però problemi sia sul piano del metodo sia – soprattutto – sul piano della disponibilità di dati longitudinali relativi alla componente straniera. A livello di

valutazione dei risultati è certamente auspicabile la possibilità di legare quanto emerso alle differenti strutture per area di origine e tipologia familiare degli stranieri per verificare se alcuni gradienti territoriali siano connessi a tali aspetti.

Riferimenti bibliografici

- Bonifazi C. 2007. *L'immigrazione straniera in Italia*. Il Mulino, Bologna.
- Cordón J. A. F. e Maldonado J. L. 2006. *Analisis territorial de la demografia Espanola*. Fundacion Fernando Abril Martorell.
- Das Gupta P. 1978. *A general method for decomposing a difference between two rates into several components*. Demography, vol. 15, n. 1, pp. 99-112.
- Giorgi P. 1995. *La fecondità ad età elevate in Italia*. Continuità e discontinuità nei processi demografici. L'Italia nella transizione demografica, SIS, Rubbettino, pp. 353-360
- Giorgi P. e Mamolo M. 2007, *Decomposition of Late Fertility Dynamics across Italian Regions during the Period 1955-2000*. Genus, vol. LXIII, nn. 1-2, pp. 65-91
- Kitagawa E.M. 1955. *Components of a Difference Between Two Rates*, Journal of the American Statistical Association, vol. 50, n. 272, pp. 1168-1194.
- Livi Bacci M. e Delgado Perez M. 1992. *Fertility in Italy and Spain: the Lowest in the World*. Family Planning, vol. 24, n. 4.
- Strozza S., Labadia C. e Ferrara R. 2007. *Il contributo delle donne straniere all'evoluzione recente della fecondità nelle province italiane*. In Rivista Italiana di Economia, Demografia e Statistica, vol. LXI, nn. 3-4, pp. 419-428.

SUMMARY

By means of decomposition models we investigate the contribution of foreign women to the recently observed increase of period fertility as well as their role in the growth slowdown of the proportion of fertility over age 30 in Italy and Spain during the period 2001-2006. The analysis has been performed at both national and provincial level.

Piero GIORGI, Professore associato di Demografia, Università di Teramo.

Raffaele FERRARA, Assegnista di Ricerca, Dip. Scienze Statistiche, Università di Napoli Federico II.

Marija MAMOLO, Dottore di Ricerca in Demografia e ricercatore ISTAT.

Salvatore STROZZA, Professore ordinario di Demografia, Università di Napoli Federico II.

MODELLI GERARCHICI PER LO STUDIO DELL'INQUINAMENTO ATMOSFERICO NELLO STRETTO DI MESSINA

Romana Gargano

1. Introduzione

Il problema dell'inquinamento atmosferico ha assunto negli ultimi anni un ruolo centrale nelle discussioni che coinvolgono cittadini ed istituzioni con quest'ultime che spesso devono adottare provvedimenti a salvaguardia della salute pubblica in considerazione del fatto che nelle aree urbane le sorgenti dell'inquinamento atmosferico sono riconducibili sia agli impianti termici civili, che soprattutto al traffico veicolare ed alle attività industriali. Nell'area metropolitana dello Stretto di Messina, studiata in questo lavoro, insistono, oltre alle fonti già citate, anche altre possibili sorgenti quali il notevole transito di treni a trazione diesel e di navi. In un recente studio (Corbett et al 2007) è stato, infatti, dimostrato che l'inquinamento prodotto nell'aria dalle navi uccide ogni anno almeno 60000 persone. Annualmente le navi rilasciano in atmosfera 1,2-1,6 milioni di tonnellate di gas di scarico. Più specificatamente sono particelle grandi meno di 10 micron prodotte dalla combustione di carburante delle navi costituite prevalentemente da fuliggine ultrasottile di carbonio, zolfo e biossido di azoto.

In letteratura sono stati proposti numerosi approcci al problema dell'inquinamento atmosferico (Pollice e Lasinio 2009, Cocchi et al. 2007, Riccio et al 2006, Smith et al 2003).

In questo lavoro preliminare è proposto un modello gerarchico basato sulla metodologia *Classification and Regression Tree* (CART) per la modellizzazione delle concentrazioni medie di monossido di carbonio (CO) e polveri sottili (PM₁₀) nell'area metropolitana dello Stretto di Messina. La scelta dell'uso di questa tecnica è da imputarsi all'eterogeneità del fenomeno "concentrazione di CO e PM₁₀", alla presenza di un certo numero di variabili esplicative ed alla facilità di interpretazione in questa prima fase.

La presente ricerca è così strutturata: nel paragrafo 2 verrà introdotto il modello gerarchico utilizzato e analizzato il caso relativo al CO ed al PM₁₀ nell'area dello Stretto di Messina nell'anno 2007, nel paragrafo 3 verrà proposto un approccio alternativo basato sui modelli di mistura di regressione e nel paragrafo finale verranno riportate le principali conclusioni e gli sviluppi in corso.

2. Modello gerarchico per lo studio del CO e del PM₁₀ nello Stretto di Messina

2.1 Area in studio e dati

L'area metropolitana dello Stretto di Messina assorbe giorno e notte il traffico veicolare che, mediante l'attraversamento dello Stretto a bordo di traghetti, dalla Sicilia è diretto al continente europeo e viceversa. Il transito forzato dei centri urbani da parte del flusso veicolare continua a provocare, sia a Messina che a Villa San Giovanni, alti tassi di inquinamento atmosferico, documentati anche dalle crescenti patologie fra gli abitanti, nonché rappresenta un serio pregiudizio per la sicurezza e l'incolumità dei pedoni. Bisogna ricordare che tra le tre città dell'area metropolitana dello Stretto, Messina è l'unica che, considerata la sua particolare conformazione che si estende in lunghezza secondo un'unica direttrice di attraversamento nella quale transita sia il 70% dell'interscambio siciliano, è dichiarata a stato di emergenza ambientale.

In questa ricerca, sono state analizzate le serie storiche dal 1 gennaio al 31 dicembre 2007 delle concentrazioni medie orarie del monossido di carbonio e delle polveri sottili registrate in 12 stazioni di monitoraggio site nelle città di Messina, Villa San Giovanni e Reggio Calabria. La scelta di studiare questi due inquinanti è imputabile esclusivamente al fatto che sono gli unici due inquinanti rilevati in tutte e tre le città (ma non in tutte le centraline).

Le variabili meteorologiche considerate sono: Direzione del vento (gradi nord), Velocità del vento (m/s), Temperatura (°C), Pressione atmosferica (ettopascal), Umidità relativa (%). Il rilevamento di tali variabili è avvenuto a Messina e Reggio Calabria tramite le centraline della Rete Mareografica nazionale dell'Istituto Superiore per la Protezione e la Ricerca Ambientale mentre a Villa San Giovanni direttamente nelle centraline equipaggiate di "palo meteorologico". In figura 1 è riportata la mappa con la dislocazione delle centraline di rilevamento nell'area metropolitana dello Stretto di Messina.

La percentuale di dati mancanti varia dal 10% al 50% (le percentuali più alte riguardano le concentrazioni di PM₁₀ nella città di Villa San Giovanni).

In tabella 1 sono riportati i giorni di superamento del limite previsto dalla legislazione vigente per il PM₁₀. Per il monossido di carbonio gli unici superamenti dei limiti previsti dalla legge sono stati registrati nella centralina V2 (22 superamenti) di Villa San Giovanni sita nei pressi degli imbarchi navali.

Figura 1 – Mappa Stretto di Messina.**Tabella 1 – Numero superamenti PM_{10} .**

Città	N. superamenti
Messina	
Archimede	55
Bocchetta	2
Caronte	7
Reggio Calabria	
Asilo	15
V.S. Giovanni	
Stazione	7
Bergamo	8

2.2 La segmentazione gerarchica

La metodologia *Classification And Regression Tree* (Breimen et al., 1984) è un esempio di ciò che tecnicamente prende il nome di partizionamento binario ricorsivo. Il processo è binario in quanto un nodo padre è bipartito sempre in esattamente due nodi figli ed è ricorsivo perché la bipartizione può essere ripetuta trattando ogni nodo figlio come un nodo padre. Gli elementi chiave di un'analisi CART sono un insieme di regole per:

- splittare (dividere) ogni nodo dell'albero;
- stabilire quando un albero è completo;
- assegnare un valore previsto (o un'etichetta nel caso degli alberi di classificazione) ad ogni nodo terminale.

L'albero di regressione considera:

$$f(x) = \sum_{m=1}^M \beta_m I_{\{x \in R_m\}} = \sum_{m=1}^M \beta_m \phi_m(x; \alpha_m)$$

Dove M è il numero di nodi terminali dell'albero, $I_{\{\cdot\}}$ è la funzione indicatrice che assume valore 1 se si verifica l'argomento $\{\cdot\}$ e zero altrimenti, ϕ una funzione non negativa della stima che $y=j$ dato che x appartiene al nodo, α è un parametro positivo che pesa la complessità del modello.

La metodologia CART è una tecnica non parametrica poiché non richiede la specificazione della forma funzionale e presenta alcuni vantaggi determinanti per la

scelta dell'applicazione nel presente lavoro: non richiede la selezione anticipata delle variabili, è estremamente robusta all'effetto degli *outliers*, può utilizzare congiuntamente variabili categoriche e continue, può processare con dati mancanti, produce alberi ottimali e può essere utilizzata a supporto dei modelli parametrici convenzionali. L'obiettivo è quello di stimare la relazione fra la variabile y (CO o PM_{10}) ed il vettore di variabili indipendenti \underline{x} (giorni della settimana, fasce orarie, variabili meteorologiche) stimando, di conseguenza, la funzione $f(\cdot)$ che lega y a \underline{x} .

2.3 I Risultati della metodologia CART

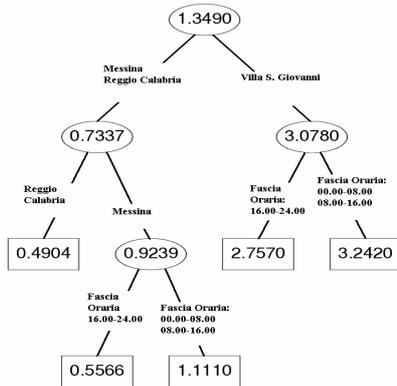
Nella prima parte dello studio si è voluto descrivere il comportamento concentrazioni atmosferiche dei due inquinanti al netto dell'influenza dei fattori meteorologici, considerando, di conseguenza, come covariate le città, i mesi, i giorni della settimana e le fasce orarie (00.00-08.00; 08.00-16.00; 16.00-24.00). In figura 2a e 2b sono riportati i diagrammi ad albero dai quali si evince come per il monossido di carbonio la città di Villa San Giovanni registra valori medi tripli rispetto a Messina e Reggio Calabria, non sembrano avere influenza sui livelli medi di CO il mese ed il giorno di rilevazione. Per le polveri sottili, invece, la situazione è opposta ossia è Messina che registra valori medi cinque volte maggiori delle altre due città. In particolare, è possibile osservare come le concentrazioni medie di PM_{10} tendano ad aumentare nei giorni feriali tra le ore 8:00 e le 16:00.

Sulla scorta di questi risultati, si è ritenuto opportuno ricondurre l'analisi considerando non più le città ma le centraline di rilevamento. Gli alberi di regressione riportati in figura 2c e 2d mostrano come per il CO le due centraline di Villa San Giovanni site in periferia hanno un comportamento medio simile a quelle delle altre città e che, a registrare le concentrazioni medie maggiori è la centralina V2. Per il PM_{10} , invece, è la stazione posta in prossimità dello svincolo autostradale di Messina che registra i valori medi più alti, anche se non bisogna sottovalutare le concentrazioni elevate osservate nelle altre stazioni della città.

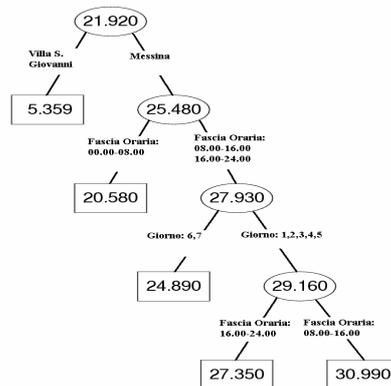
Successivamente, essendo note le relazioni tra gli inquinanti e le variabili meteorologiche ed essendo la zona studiata caratterizzata da un clima particolare (orografia complessa che influenza i movimenti delle masse d'aria), si è voluta verificare, anche in questa zona, l'eventuale dipendenza del CO e PM_{10} dalle covariate meteorologiche. Nelle figure 2e e 2f sono riportati i diagrammi ad albero dai quali si osserva che in caso di alte percentuali di umidità relativa e basse temperature le concentrazioni medie di CO sono significativamente più alte.

Figura 2 – Diagrammi ad albero.

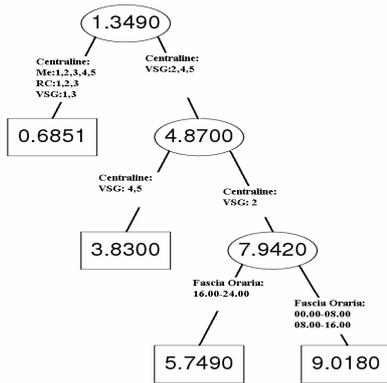
2a - CO~città, fasce orarie, giorni



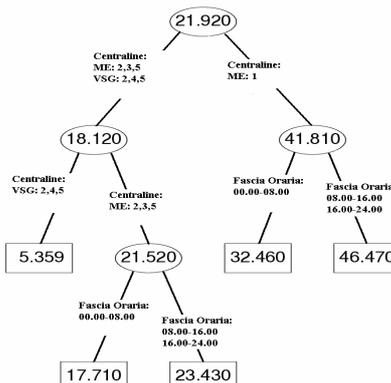
2b - PM₁₀~città, fasce orarie, giorni



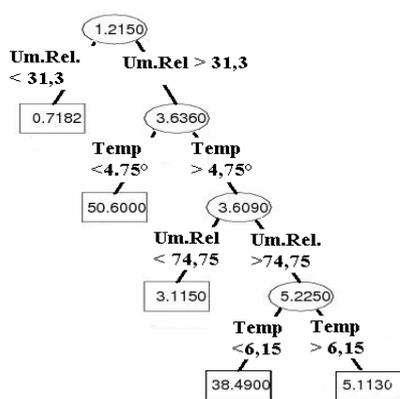
2c - CO~centraline, fasce orarie, giorni



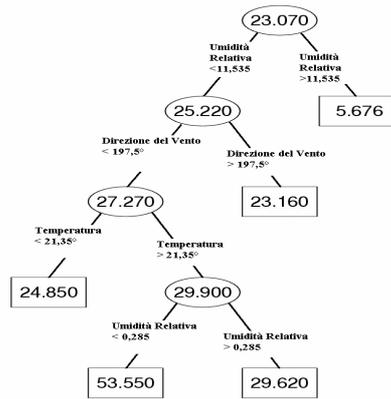
2d - PM₁₀~centraline, fasce orarie, giorni



2e - CO~variabili meteo



2f - PM₁₀~variabili meteo



Riguardo al PM_{10} tra le variabili meteorologiche risultano determinanti per l'aumento delle concentrazioni medie le basse percentuali di umidità relativa, la direzione del vento in particolare venti da SE, E, NO (rispettivamente Libeccio, Ponente e Maestrale) e le temperature superiori a 20 °C.

3. Mixture regression model per lo studio del CO e PM_{10} nello Stretto di Messina

Sulla scorta dei risultati finora ottenuti, volendo, ancora una volta, identificare i parametri delle distribuzioni di probabilità caratteristiche di ciascun segmento stimando, però, contemporaneamente i modelli di regressione relativi a ciascuna variabile in ciascun segmento, si è pensato di proporre un approccio basato sui *mixture regression models*.

3.1 Mixture regression models

I modelli mistura di regressione rappresentano dei popolari metodi applicati ormai in svariati campi (McLachlan e Peel, 2000) per modellare l'eterogeneità inosservata o in caso di *overdispersion* dei dati. Tali modelli si esprimono come la somma pesata di S componenti, ogni componente segue una distribuzione parametrica ed ha assegnati dei pesi che indicano la probabilità a priori per una certa osservazione di provenire da quella componente. In generale un modello mistura di regressione Gaussiana con variabili concomitanti, ipotizzato in questo lavoro, è così definito:

$$H(y | x, \omega, \Theta) = \sum_{s=1}^S \pi_s(\omega, \alpha) N(y | \mu_s(x), \sigma_s^2) \quad \text{con} \quad \sum_{s=1}^S \pi_s(\omega, \alpha) = 1 \quad \text{e} \quad \pi_s(\omega, \alpha) > 0$$

In cui ω rappresenta la variabile concomitante e Θ è il vettore di tutti i parametri, α i parametri della variabile concomitante.

3.2 Risultati del mixture regression model

Al fine di modellare l'eterogeneità eventualmente presente nelle concentrazioni di CO e PM_{10} rilevate nell'area metropolitana dello Stretto di Messina si è proceduto a stimare, attraverso l'algoritmo EM, il *mixture regression model* in

oggetto; in particolare si è fatto ricorso a un modello di regressione che prevede l'utilizzo di una variabile concomitante (le città).

L'ispezione dei test diagnostici (AIC, BIC, CAIC) utilizzati per individuare il numero corretto di componenti della mistura suggerisce l'adozione di un modello a due componenti sia per il monossido di carbonio che per le polveri sottili.

Le probabilità a posteriori, stimate per verificare la corretta classificazione mostrano una parziale correttezza, dovuta, probabilmente, all'effetto delle centraline (in questo caso non considerate).

Nel modello per il CO stimato considerando come covariate le variabili meteorologiche la prima componente ha mostrato l'influenza statisticamente significativa dell'umidità relativa ($p=0.001$) e della temperatura ($p=0.005$). Per la seconda componente nessuna variabile meteo risulta spiegare significativamente i livelli di CO nell'area metropolitana dello Stretto di Messina.

Nel modello stimato per il PM₁₀ le due componenti differiscono per l'intercetta. Nella spiegazione della prima componente hanno un ruolo statisticamente significativo la temperatura ($p=0.009$) e l'umidità relativa ($p=0.010$), mentre per la seconda componente la direzione del vento ($p=0.001$).

5. Conclusioni e sviluppi futuri

Dall'analisi effettuata sembra emergere che nell'area metropolitana dello Stretto di Messina il problema dell'inquinamento atmosferico derivi dalla somma degli effetti dovuti al traffico veicolare, marittimo e ferroviario. Le concentrazioni maggiori di CO sono state registrate a Villa San Giovanni ed in particolare nelle centraline poste nel centro, in prossimità sia alla stazione marittima che ferroviaria e da dove a qualsiasi ora del giorno i mezzi provenienti da e per la Sicilia sono costretti a transitare. A Messina, invece, sono stati registrati i valori maggiori del PM₁₀ nella centralina sita in prossimità dello svincolo autostradale. Le alte concentrazioni registrate nei giorni feriali e nelle ore di punta fanno pensare che, anche in questo caso, siano da attribuirsi al traffico urbano ed extraurbano. Riguardo le relazioni del CO e del PM₁₀ con le variabili meteo entrambi gli approcci hanno mostrato che per il monossido di carbonio esiste un'influenza statisticamente significativa di temperatura e di umidità relativa, mentre per il PM₁₀ è risultata statisticamente significativa anche la direzione del vento.

Tali risultati confermano la bontà delle politiche di salvaguardia adottate dal legislatore che recentemente ha riconfermato lo stato d'emergenza traffico nella città di Messina ove notevoli benefici sono stati apportati dalla realizzazione a sud della città di un nuovo approdo che bypassa il centro cittadino. Sarebbe altresì

opportuna una valutazione globale del problema inquinamento nell'intera area metropolitana dello Stretto e non separatamente così come finora avvenuto.

Nel futuro si prevede di affrontare questo problema con i modelli gerarchici mistura di esperti (*hierarchical mixture-of-experts model*), poiché permettono di combinare aspetti dei modelli mistura finiti e dei GLM, e forniscono una buona generalizzazione per problemi di regressione non lineare, compresa la classificazione. Essi rappresentano un "serio concorrente" delle procedure CART.

Riferimenti bibliografici

- Breiman L., Friedman J.H., Olshen R.A., Stone C.J. (1984), *Classification and Regression Tree*, Chapman & Hall, New York
- Cocchi D., Greco F., Trivisano C. 2007. *Hierarchical space-time modelling of PM₁₀ pollution*, Atmospheric Environment, 41: 532–542
- Corbett J.J., Winebrake J.J., Green E.H., Kasibhatla P., Eyring V., Laurer A. (2007) *Mortality from ship emission: A Global Assessment*, Environmental Science and Technology, 41, 8512-8518.
- McLachlan G., Peel D., 2000. *Finite Mixture Model*. Wiley, New York
- Pollice A., Lasinio G.J., 2009. *A multivariate approach to the analysis of air quality in high environmental risk area*. GRASPA Working paper, 32
- Riccio, A., Barone, G., Chianese, E., Giunta, G., 2006. *A hierarchical Bayesian approach to the spatio-temporal modeling of air quality data*. Atmospheric Environment 40, 554–566
- Smith, R.L., Kolenikov, S., Cox, L.H., 2003. *Spatio-temporal modelling of PM_{2.5} data with missing values*. Journal of Geophysical Research 108, D24.

SUMMARY

Hierarchical models for the air pollution study in the Strait of Messina

The aim of this paper is to describe the air pollution by carbon monoxide and particular matter in the Strait of Messina metropolitan area by hierarchical structure models. The results show high pollution levels throughout the area and in particular in the Villa San Giovanni, where the scale of emissions reductions with no record even on holidays and at night.

Romana GARGANO, Ricercatore di Statistica, Facoltà di Scienze Statistiche, Università di Messina.

MODELLI A VARIABILI LATENTI PER LA VALUTAZIONE DELLA SODDISFAZIONE DEGLI UTENTI DEL SERVIZIO DI TRASPORTO PUBBLICO

Romana Gargano, Giuseppe Avena¹

1. Introduzione

La diffusione dei programmi di certificazione ISO 9001-2000 e l'adozione dei principi definiti nella Carta dei Servizi del settore trasporti hanno portato un sempre maggiore numero di aziende pubbliche ad attivare delle procedure di valutazione non solo in termini di efficacia e di efficienza ma anche in termini più generali di qualità. La qualità di un servizio e, di conseguenza, la soddisfazione espressa dagli utenti sono variabili latenti (v.l.), non osservabili direttamente ma misurabili attraverso una serie di indicatori reali che misurano sia le singole quantità sottostanti che le relazioni causa-effetto tra esse. In tale contesto i modelli ad equazioni strutturali (SEM) rappresentano un'ottima soluzione per descrivere e stimare queste strutture concettuali. Ogni modello ad equazione strutturale si caratterizza per avere un doppio livello di relazioni: quelle tra variabili latenti (modello strutturale) e quelle tra quest'ultime ed il proprio "insieme" di variabili manifeste (modello di misura).

In letteratura esistono numerosi esempi di applicazione dei SEM in processi di *customer satisfaction* (Magnani et al. 2003, Laudisio Di Bonito 2006, Bernini e Lubisco 2005).

Obiettivo del lavoro è valutare l'importanza che gli utenti di un servizio di trasporto pubblico urbano di una città dell'Italia meridionale conferiscono ai diversi aspetti del servizio offerto. A tale scopo verrà proposta un'applicazione dei modelli ad equazioni strutturali e in particolare di due metodi proposti in letteratura per la stima degli stessi: Lisrel al fine di ricostruire la matrice di varianze e covarianze tra le variabili manifeste, osservate attraverso i parametri del modello stimato, e PLS-Path Modeling per stimare le variabili latenti in modo che siano le più rappresentative possibile del proprio blocco di variabili manifeste e che, contemporaneamente, spieghino nel miglior modo le relazioni individuate nel modello strutturale.

¹ Il lavoro è frutto della collaborazione congiunta dei due autori. Tuttavia, per la stesura del testo, Romana Gargano è autore dei paragrafi 3 e 4, Giuseppe Avena è autore dei paragrafi 1 e 2

2. Lo studio

Preliminarmente, è stata condotta un'indagine campionaria allo scopo di realizzare la base di dati necessaria alla costruzione del modello da analizzare. A tal fine, è stato somministrato un questionario ad un campione di 600 utenti nel momento in cui usufruivano del mezzo di trasporto. Il questionario, predisposto per rilevare il livello di soddisfazione dei passeggeri, contiene domande a risposta graduata, con giudizi espressi con un voto da 1 a 10, di carattere generale sugli utenti e sulle abitudini di trasporto (età, sesso, nazionalità, titolo di studio, professione, motivo della presenza nel Comune, tipo di cliente e titolo di viaggio), nonché domande sulle caratteristiche proprie della qualità: "accesso al servizio", "personale", "stato degli automezzi", "trasporto", "sicurezza". La soddisfazione complessiva del servizio è monitorata con due *item*: uno di valutazione complessiva ed uno di soddisfazione rispetto alle proprie aspettative.

Per specificare il modello di soddisfazione dell'utente è necessario individuare gli *item* del questionario rilevanti e le componenti latenti sottese. L'analisi fattoriale esplorativa ha permesso di individuare una variabile latente endogena e quattro componenti latenti esogene relative alla soddisfazione nei confronti dei diversi aspetti del servizio. In seguito alla costruzione del modello, utilizzando il coefficiente *alpha di Cronbach*, è stata condotta un'analisi dell'attendibilità al fine di valutare se, relativamente ad ogni fattore latente, tutti gli *item* risultanti dall'analisi fattoriale misurino coerentemente il costrutto sottostante. Nelle tabelle 1 e 2 si riportano i risultati dell'analisi esplorativa. Si evince che gli *item* relativi ai vari aspetti della soddisfazione si riducono dai 18 iniziali ad 11 e che essi misurano adeguatamente i costrutti latenti sottostanti dando origine a valori elevati del coefficiente *alpha di Cronbach*.

Tabella 1 – Variabili latenti e variabili osservate.

Variabili latenti	Indicatori di soddisfazione
Soddisfazione globale (η)	Giudizio complessivo Giudizio rispetto alle aspettative
Organizzazione delle corse (ξ_1)	Regolarità Orario e percorsi Puntualità
Condizioni di viaggio (ξ_2)	Sicurezza viaggio Confort del viaggio Fiducia negli autisti Accesso ai mezzi
Empatia (ξ_3)	Reperibilità biglietti Cortesia del Personale

Tabella 2 – Valori del coefficiente alpha di Cronbach.

Variabili latenti	Alpha di Cronbach
Soddisfazione globale (η)	0.72
Organizzazione delle corse (ξ_1)	0.81
Condizioni di viaggio (ξ_2)	0.85
Empatia (ξ_3)	0.73

Dopo aver individuato le componenti latenti, è stato specificato il modello di soddisfazione da sottoporre ad analisi confermativa, attraverso l'approccio Lisrel e Partial Least Square Path Modeling, in cui si ipotizza che le componenti latenti di soddisfazione influenzino in maniera casuale la soddisfazione globale del servizio.

3. Il modello strutturale a variabili latenti

Per stimare i parametri di un modello SEM si identificano, generalmente, due diversi approcci: i metodi *covariance-based* e *component-based*.

I metodi *covariance-based* si fondano sull'identificazione dei parametri del modello che meglio ricostruiscono la struttura della matrice di covarianza osservata tra le variabili manifeste. In questo ambito, il metodo LISREL, basato sulla massima verosimiglianza, è sicuramente il più utilizzato. Tale approccio presenta però alcuni limiti che ne possono pregiudicare l'applicazione. Nelle equazioni di misura del modello LISREL non esistono, infatti, le condizioni sufficienti per l'unicità dei punteggi dei fattori, rendendo arbitraria la stima dei punteggi delle v.l..

Le tecniche *component-based*, sono basate sull'identificazione delle variabili latenti che spiegano contemporaneamente il proprio "insieme di indicatori" e le relazioni tra gli "insiemi". Tra queste tecniche, il metodo più usato per la stima degli *scores* delle variabili latenti, è il PLS Path Modelling (Wold, 1985; Tenenhaus et al, 2005) che, al contrario di LISREL, non impone ipotesi forti sui parametri, non ha problemi di identificazione degli stessi parametri e di convergenza degli algoritmi.

3.1 L'approccio LISREL

I modelli ad equazioni strutturali descrivono le relazioni statistiche tra variabili latenti e manifeste. Tali relazioni implicano una struttura della matrice di covarianza empirica che, una volta stimati i parametri, può essere confrontata con la matrice risultante dal modello. Testando l'uguaglianza statistica tra le due

matrici, è possibile verificare se il modello stimato possa essere considerato una possibile struttura esplicativa delle relazioni tra le variabili studiate.

I modelli di misura possono essere formalizzati in:

$$\mathbf{Y} = \mathbf{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad (1)$$

$$\mathbf{X} = \mathbf{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta} \quad (2)$$

con \mathbf{Y} e \mathbf{X} che indicano, rispettivamente, i vettori di dimensione $p \times 1$ e $q \times 1$ delle variabili sottostanti, associate alle variabili ordinali osservate, $\boldsymbol{\eta}$ e $\boldsymbol{\xi}$ i vettori di dimensioni $m \times 1$ e $n \times 1$ delle corrispondenti variabili latenti endogene ed esogene, $\boldsymbol{\varepsilon}$ e $\boldsymbol{\delta}$ i vettori degli errori di misura. È bene ricordare che le variabili sono misurate in termini di scarti dalle loro medie; le variabili indipendenti e gli errori sono tra loro incorrelati così come gli errori delle diverse equazioni.

È possibile definire il modello strutturale come (3):

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \mathbf{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \quad (3)$$

in cui compaiono i tre vettori delle variabili endogene ($\boldsymbol{\eta}$), esogene ($\boldsymbol{\xi}$) e degli errori ($\boldsymbol{\zeta}$); le due matrici dei coefficienti strutturali fra le variabili endogene (\mathbf{B}) e tra le variabili esogene ed endogene ($\mathbf{\Gamma}$) rispettivamente di dimensioni $m \times m$ e $m \times n$. \mathbf{B} è una matrice con tutti gli elementi sulla diagonale principale nulli e $\mathbf{I}-\mathbf{B}$ è non singolare. Anche per la specificazione del modello strutturale bisogna ricordare che le variabili sono misurate in termini di scarti dalle loro medie, le variabili indipendenti e gli errori sono tra loro incorrelati e nessuna equazione strutturale deve essere ridondante.

La stima dei parametri del modello avviene a partire dalla matrice di varianze e covarianze tra le variabili osservate. La matrice incognita di varianze e covarianze $\boldsymbol{\Sigma}$ si può esprimere in funzione dei parametri del modello stimati attraverso un processo iterativo di minimizzazione di una funzione di prossimità tra la matrice \mathbf{S} di varianze e covarianze delle variabili manifeste nel campione e la corrispondente matrice teorica. La funzione, $F(\mathbf{S}, \boldsymbol{\Sigma})$, dipende dal criterio di stima scelto, solitamente assumendo la multinormalità delle variabili osservate, il metodo di stima più utilizzato è quello della massima verosimiglianza. In questo lavoro, poiché tutte le variabili manifeste sono misurate su scala ordinale con valori modali collocati su punteggi alti della scala, si fa riferimento all'*underlying variable approach*, che considera le variabili manifeste come realizzazioni di variabili sottostanti continue, normali standardizzate. La matrice \mathbf{S} , calcolata sulle variabili sottostanti, è la matrice di correlazione policorica. In questo caso, il metodo di

stima che tiene conto della natura dei dati, evitando stime distorte e valori dei test di adattamento errati in caso di dati non normali, è il *Weighted Least Square*, secondo il quale la funzione $F(\mathbf{S}, \mathbf{\Sigma})$ può essere espressa come:

$$F_{WLS} = [\mathbf{s} - \boldsymbol{\sigma}]' \mathbf{W}^{-1} [\mathbf{s} - \boldsymbol{\sigma}] \quad (4)$$

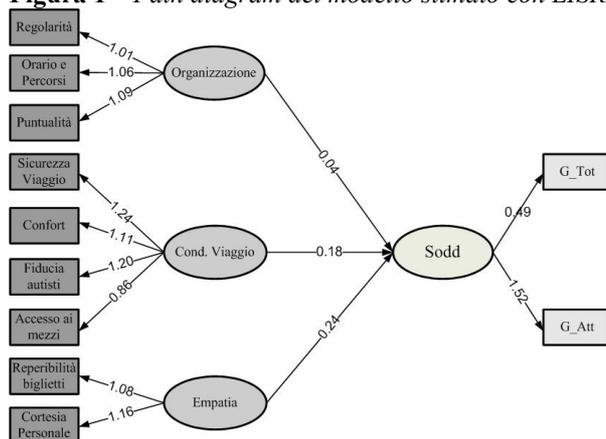
in cui \mathbf{s} indica il vettore degli elementi della matrice di correlazione policorica \mathbf{S} , $\boldsymbol{\sigma}$ il corrispondente vettore di elementi di $\mathbf{\Sigma}$ e \mathbf{W} un'opportuna matrice di pesi.

3.2 I Risultati di LISREL

In figura 1 sono riportate le stime dei parametri del modello di soddisfazione del servizio di trasporto pubblico urbano ottenute con il metodo di stima *Weighted Least Square*. Si evince che le componenti relative al Viaggio e all'Empatia, presentando i valori più alti dei coefficienti strutturali, hanno un'influenza pressoché simile sul livello di soddisfazione globale. In particolare, dall'analisi delle variabili manifeste è possibile osservare che quelle che esercitano maggiore influenza sulle condizioni di viaggio sono la sicurezza del viaggio e la fiducia negli autisti. Reperibilità dei biglietti e cortesia del personale influenzano in maniera simile l'Empatia.

Gli indici di bontà d'adattamento ($R^2=0.67$, $\chi^2=74.45$ ($p=0.000$), Goodness of Fit Index =0.94, Adjusted Goodness of Fit Index =0.89 e Root Mean Square Residual =0.066) confermano un buon livello generale, ossia la corrispondenza del modello teorico ai dati empirici (K. Schermelleh-Engel et al., 2003).

Figura 1 – Path diagram del modello stimato con LISREL.



3.3 L'approccio *Partial Least Square Path Modeling*

L'approccio PLS ai modelli SEM si basa un algoritmo iterativo che consente di stimare le variabili latenti attraverso un sistema interdipendente di elaborazioni alternate, basate su regressioni semplici e multiple. L'idea è di determinare gli *scores* delle variabili latenti attraverso un processo che in modo iterativo determina una stima esterna della variabile latente e successivamente una stima interna della stessa. Formalmente, assumendo che ogni v. l. sia funzione delle proprie variabili manifeste, nella stima esterna ognuna di esse è ottenuta come combinazione lineare delle proprie variabili manifeste. Successivamente, nella stima interna, ogni variabile latente è stimata considerando le relazioni tra le diverse variabili latenti.

L'algoritmo iterativo continua con la stima dei pesi esterni da associare ad ogni variabile manifesta per costruire la variabile latente corrispondente nella sua stima esterna. È possibile ottenere i pesi esterni seguendo due diversi schemi a seconda che i blocchi di variabili siano di tipo riflessivo o formativo.

In particolare, la specificazione dei modelli di misura viene detta di tipo riflessivo, infatti ogni variabile manifesta è considerata essere il riflesso della variabile latente (che ne è la causa). Ogni relazione del modello esterno relativa a quel "insieme" può essere espressa come modello di regressione semplice. I pesi rappresentano il coefficiente di regressione semplice che lega ogni variabile manifesta alla stima interna della propria v.l.

Nello schema di tipo formativo l'idea è che le variabili manifeste formano la variabile latente, ossia si presuppone che la variabile latente sia funzione dei suoi indicatori. In tal caso le relazioni all'interno di ciascun "insieme" di variabili manifeste rappresentano un modello di regressione multipla in cui la variabile risposta è rappresentata dalla variabile latente.

Aggiornato il sistema di pesi esterni, gli stessi pesi sono utilizzati per ottenere una nuova stima esterna delle v.l. L'algoritmo alterna questi passi finché non raggiunge la convergenza tra la stima interna ed esterna delle variabili latenti. Ottenuta una stima definitiva delle v.l. è possibile analizzare la struttura di relazioni che caratterizzano il modello strutturale per mezzo di regressioni semplici e/o multiple che considerano le relazioni tra le variabili latenti. Solitamente per valutare la bontà del modello nell'ambito del PLS-PM vengono utilizzati tre indici: l'indice di comunaltà, l'indice di ridondanza e l'indice di bontà globale (GoF).

3.4 I risultati del *PLS-PM*

Preliminarmente, per uniformità con l'analisi LISREL, si è proceduto alla standardizzazione delle variabili manifeste. La tabella 3 riporta i coefficienti di

correlazione e di regressione di tutte le variabili esogene sull'unica variabile endogena del modello "Soddisfazione globale".

Tabella 3 – Impatto e contributo delle variabili latenti esogene sulla v.l. endogena.

	Org. delle corse	Cond. di viaggio	Empatia
Correlazione	0.434	0.766	0.533
Path coefficient	0.07	0.24	0.15
Contributo all'R ² (%)	5.17	75.22	19.61

Si evince che la soddisfazione dell'utente dipende dalle condizioni di viaggio e, marginalmente, dall'empatia. Non sembra avere influenza l'organizzazione delle corse.

La relazione caratterizzante il modello strutturale stimato si esprime come:

$$\hat{Sodd} = 0.07\hat{Org} + 0.24\hat{Viaggio} + 0.15\hat{Empatia}$$

In relazione alla bontà del modello, l'R², che misura la qualità di ogni relazione del modello strutturale, è risultato essere pari a 0.72. In tabella 4, al fine di misurare la qualità globale del modello, è riportato il valore del GoF.

Tabella 4 – Indici di bontà globale del modello.

	GoF	GoF (Jacknife)	Dev Standard	Critical Ratio	Limite inf (95%)	Limite sup (95%)
Assoluto	0.523	0.522	0.002	543,876	0.521	0.525
Relativo	0.957	0.956	0.001	1754,875	0.956	0.959
Mod. esterno	0.963	0.963	0.000	10342,971	0.962	0.964
Mod. interno	0.946	0.947	0.001	1845,832	0.942	0.948

Si osserva che sono pienamente soddisfacenti sia il valore puntuale ottenuto per il modello stimato, sia la sua validazione ottenuta con tecniche non parametriche.

4. Conclusioni

Sebbene le due soluzioni siano ispirate a due differenti impostazioni di analisi la stima del modello di soddisfazione del servizio di trasporto pubblico urbano, oggetto del presente studio, non sembra evidenziare grosse differenze.

La soddisfazione globale dell'utente risulta influenzata soprattutto dagli aspetti legati alle condizioni di viaggio e, in misura minore, dall'empatia. Influenza particolare è esercitata dagli aspetti legati alla sicurezza del viaggio e alla fiducia

riposta negli autisti, nonché, per quando riguarda la capacità dell'azienda di aprirsi al cliente (empatia), dalla reperibilità dei biglietti e dalla cortesia del personale. Contrariamente a quanto ci si aspettava non risulta determinante ai fini della soddisfazione del cliente l'organizzazione delle corse.

Riferimenti bibliografici

- Bernini C., Subisco A. 2005. *A new extension of dynamic simplex model for the public transport customer satisfaction*, *Statistica*, anno LXV, n. 4.
- Laudisio Di Bonito N. 2006. *Modelli ad equazioni strutturali per la valutazione dell'accessibilità del comune di Napoli*, *Statistica & Società* anno IV n. 2: 17-21.
- Cagnone S., Lubisco A., Mignani S. 2003. *Modelli a variabili latenti per l'analisi della Customer Satisfaction degli utenti dei servizi di trasporto pubblico*, collana *Appunti e ricerche* n. 2, Alma Mater Studiorum Università di Bologna.
- Schermelleh-Engel K., Moosbrugger H., Müller H. 2003. *Evaluating the Fit of Structural equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures*. *Psychol Res Methods Online*, 8: 23-74.
- Tenenhaus M., Esposito Vinzi V., Chatelin Y.M., Lauro N.C. 2005. *PLS Path Modeling*, *Computational Statistics and Data Analysis*, 48, 159-205.
- Wold H. 1985. *Partial Least Square*, *Encyclopedia of Statistical Sciences*, S. Kotz and N.L. Johnson Editors, New York, Wiley, Vol. 6: 581-591.

SUMMARY

Latent variables models for the customer satisfaction of public transport

The aim of this paper is assess the importance that the users of urban public transport to give different aspects of the service offered by referring to two different settings of an analysis that focuses on interpretation (Lisrel) and one in the forecast (Partial Least Square Path Modeling). The model relates to perceived quality of urban public transport service with the offer dimensions.

Romana GARGANO, Ricercatore di Statistica, Facoltà di Scienze Statistiche, Università di Messina.

Giuseppe AVENA, Ricercatore di Statistica Sociale, Facoltà di Scienze Statistiche, Università di Messina.

MERCATO DEL LAVORO FEMMINILE E CONTESTO FAMILIARE NELLE REGIONI ITALIANE

Rosa Giaimo, Dario Corso, Giovanni Luca Lo Magno

1. Introduzione

Le donne italiane sperimentano gravi difficoltà nel conciliare famiglia e lavoro, e incorrono in maggiori difficoltà nell'intraprendere un lavoro retribuito o nel trovare una buona occupazione. Ciò le rende maggiormente vulnerabili e più esposte al rischio povertà, soprattutto quando esse sono disoccupate o con un lavoro mal retribuito e precario (Presidenza del Consiglio dei Ministri, 2000).

Nel 2006 il tasso di attività femminile per le donne in età lavorativa (15-64 anni) risulta del 50.8%, valore che fa collocare l'Italia al penultimo posto tra i paesi dell'Unione Europea a 27, preceduta in questa classifica soltanto da Malta (36.5%). Il modesto risultato italiano nasconde però differenze territoriali rilevanti. Il Nord presenta per il 2006 un tasso di attività femminile del 59.5%, che appare non troppo distante dalla media europea (63%). Il Centro ottiene un po' meno, con un 56% per il tasso di attività femminile, ma molto peggio, distanziandosi notevolmente dal resto dell'Italia, il Sud con un 37.3%.

Osservando le differenze tra i tassi di attività maschili e femminili, ancora per la fascia di età 15-64 anni, si manifestano per l'Unione Europea alcuni *pattern* geografici che vedono i Paesi nordici e i nuovi entranti presentarsi con divari molto contenuti (la Finlandia, con 3.8 punti percentuali di differenza tra il tasso maschile e quello femminile, è la nazione europea con la differenza di genere più contenuta), e i Paesi del Mediterraneo soffrire differenze rilevanti (Malta 41.5, Grecia 24.1, Spagna 21.1). L'Italia presenta una differenza di 23.8 punti percentuali e, ancora una volta, i divari di genere più elevati si verificano al Sud con 32 punti di differenza. Al Nord e al Centro, con differenze rispettivamente di 18.6 e 20.3 punti percentuali, si ottengono risultati migliori, ma con *gap* che bisogna riconoscere essere comunque notevoli.

Questo lavoro si concentra sull'analisi della partecipazione al mercato del lavoro delle donne sposate. Nello studio di tale fenomeno è inevitabile riferirsi al *contesto familiare* di appartenenza, che viene qui inteso come l'insieme dei fattori economici e socio-culturali inerenti la famiglia che possono influire sulla scelta di intraprendere o cercare un lavoro retribuito. Viene quindi analizzata l'influenza di alcuni di questi fattori sulla probabilità che ha la donna sposata di essere attiva,

svolgendo a livello territoriale un confronto tra le tre principali macroaree italiane: il Nord, il Centro e il Sud.

In particolare si è voluto verificare se tra il benessere economico della famiglia, al netto dell'apporto dato dal lavoro retribuito della donna, e la probabilità di essere attiva possa esservi una relazione funzionale che vede tale probabilità diminuire partendo da livelli di benessere bassi per poi risalire una volta superato un certo livello. Tale livello di benessere potrebbe essere inteso come quello in grado di consentire alla donna di liberarsi, grazie alle accresciute risorse economiche a disposizione della famiglia, di parte delle incombenze domestiche.

2. La partecipazione delle donne sposate al mercato del lavoro e le variabili di contesto familiare

Tra i principali fattori che vengono comunemente indicati come cause delle differenze di genere nella partecipazione al mercato del lavoro vi sono la questione della conciliazione lavoro-famiglia e la discriminazione nei confronti delle donne. Per quanto concerne il primo aspetto, c'è da rilevare che il modello più diffuso di divisione dei ruoli all'interno della famiglia italiana è ancora quello dell'uomo *breadwinner*, quello cioè in cui il marito si fa carico di procurare le risorse economiche necessarie per il mantenimento della famiglia, mentre la moglie è confinata prevalentemente nello svolgimento di lavoro domestico non retribuito (Istat, 2008). Come conseguenza di questa differenziazione dei ruoli le donne italiane cercano di conciliare lavoro e famiglia ricorrendo al *part-time* (circa una donna su quattro). D'altra parte il ricorso ai servizi pubblici per l'infanzia risente di alcune carenze del sistema italiano, particolarmente avvertite al Sud (Istat, 2008). La discriminazione invece, in particolare quella relativa alle retribuzioni, è di difficile quantificazione statistica, nonostante i numerosi contributi metodologici che si sono susseguiti a partire dal lavoro seminale di Oaxaca (1973). Inoltre se è vero che l'esistenza della discriminazione scoraggia le donne a presentarsi nel mercato del lavoro, è anche vero che i divari salariali possano essere spiegati dalla ridotta esperienza lavorativa delle donne, dovuta ad una discontinuità temporale di tale offerta nel corso della vita lavorativa (Bono, Giaimo, Lo Magno, 2008).

L'approccio fondamentale adottato in questo lavoro di ricerca è il riconoscimento che le decisioni concernenti il lavoro vengono elaborate dalle donne sposate tenendo in considerazione le esigenze della famiglia di appartenenza. Il contesto familiare è quindi una dimensione sociale di assoluto rilievo, sia per quanto riguarda gli aspetti economici che quelli culturali, e questo studio focalizza la sua attenzione sui fattori che lo descrivono.

La capacità economica del marito è stata considerata una variabile rilevante in

diversi studi sul tema della partecipazione delle donne sposate al mercato del lavoro (Mroz, 1987; Juhn e Murphy, 1997).

Per quanto concerne gli aspetti non monetari, Del Boca et al. (2000) cercano di catturare l'esistenza di una propensione socio-culturale nei confronti del lavoro della donna utilizzando come variabili di *background* familiare il fatto che la madre e la suocera della moglie abbiano lavorato quando avevano la sua stessa età.

In questo studio le variabili di contesto familiare che vengono considerate sono il benessere familiare, la presenza di figli in età pre-scolare e la presenza di una eventuale rete sociale (parenti o amici) su cui la donna può contare in caso di bisogno.

La capacità economica del marito è in genere misurata come la differenza tra il reddito familiare complessivo e il reddito da lavoro della donna. Questo approccio ci è apparso insufficiente, in quanto non tiene in considerazione la dimensione del nucleo familiare e la composizione per età dei suoi componenti, fattori che certamente influiscono sul reale benessere economico che il reddito familiare può effettivamente procurare. La dimensione da utilizzare per lo studio della partecipazione delle donne sposate al mercato del lavoro dovrebbe essere il reale benessere economico di cui la famiglia godrebbe senza contare sull'apporto dato dal lavoro retribuito della donna, dimensione che non può essere colta correttamente sulla base del semplice valore nominale del reddito. Per tenere conto della numerosità e della presenza di adulti o bambini nella famiglia, abbiamo utilizzato una misura di reddito familiare al netto del reddito da lavoro della donna, resa equivalente secondo la scala Oecd modificata (Hagenaars et al., 1994). I redditi familiari equivalenti sono i redditi familiari trasformati per permettere confronti di benessere economico tra famiglie diverse per numerosità dei componenti e per presenza di adulti e bambini. Molti studi e indicatori statistici sulla povertà sono basati sui redditi familiari equivalenti (Commissione Europea, 2004) e il nostro contributo di ricerca rappresenta un collegamento tra studi sulla povertà e studi sul mercato del lavoro. La nostra misura di benessere familiare tiene inoltre in considerazione i fitti imputati per l'abitazione di proprietà, nell'intento di disporre di una migliore misura del benessere effettivamente goduto dalla famiglia.

3. L'analisi svolta

Il *data set* che abbiamo utilizzato è quello dell'Indagine sui Redditi e le Condizioni di Vita, condotta dall'Istat per l'Italia nel 2006 (Eu-Silc 2006). I dati di reddito si riferiscono al 2005, cioè all'anno solare precedente l'intervista. Le altre variabili individuali socio-economiche sono riferite al momento dell'intervista.

Nella nostra analisi utilizziamo un modello *probit* per stimare la probabilità

delle donne sposate di partecipare al mercato del lavoro in funzione della nostra misura di benessere familiare al netto del contributo lavorativo della donna. L'effetto territoriale è stimato utilizzando variabili *dummy* che si riferiscono alle tre macro-ripartizioni italiane. Tra le variabili impiegate includiamo anche una *dummy* che rileva la presenza di una rete parentale o di amicizie sulla quale la moglie può contare in caso di bisogno. Questa variabile ci aiuta a descrivere più compiutamente il contesto familiare nel quale la donna è inserita, in quanto può essere considerata indicativa dell'esistenza di condizioni favorevoli alla conciliazione di famiglia e lavoro. È stato possibile ottenere tale informazione utilizzando uno specifico *item* della batteria di domande sulla partecipazione sociale, batteria che costituisce per l'indagine Eu-Silc una novità, essendo stata inclusa per la prima volta nel questionario dell'indagine 2006.

Il modello *probit* viene utilizzato per verificare l'esistenza di una relazione ad U tra la nostra misura di reddito e la probabilità di essere attiva. Volevamo testare se tale relazione potesse cambiare da decrescente a crescente oltre una certa soglia di benessere, mutamento teoricamente possibile sulla base di una accresciuta disponibilità di beni e servizi per la casa, tale da liberare tempo ed energia alla donna per intraprendere un lavoro retribuito.

Le variabili utilizzate nel nostro modello *probit* per spiegare la partecipazione delle donne sposate al mercato del lavoro sono gli anni di istruzione della donna, gli anni di esperienza lavorativa, l'età, una variabile *dummy* relativa alla disponibilità di una rete parentale o di amicizie su cui poter contare in caso di bisogno, il reddito familiare equivalente al netto del reddito da lavoro della donna (con fitti imputati) e le *dummy* territoriali.

Il modello *probit* può essere così descritto sinteticamente:

$$\begin{aligned} \text{Pr}(\text{attiva} = \text{sì}) = & \Phi(\beta_0 + \beta_1 \text{ISTR} + \beta_2 \text{ESP} + \beta_3 \text{ETA} + \\ & + \beta_4 \text{NBAMBINI} + \beta_5 \text{RETESOCIALE} + \beta_6 \text{REDD} + \beta_7 \text{REDD}^2 + \\ & + \beta_8 \text{NORD} + \beta_9 \text{CENTRO}) \end{aligned}$$

dove il significato delle variabili è facilmente intuibile. L'inclusione del termine quadratico della nostra misura di benessere familiare serve a cogliere l'eventuale relazione non lineare con la probabilità di essere attiva.

Nella Tabella 1 sono riportati i risultati del modello di regressione stimato. I coefficienti presentano tutti il segno atteso. Il termine quadratico del reddito familiare equivalente è risultato positivo e significativo, dando evidenza di una relazione ad U tra tale misura di reddito e la probabilità di essere attiva.

Tabella 1 – Risultati delle stime del modello probit di partecipazione delle donne sposate al mercato del lavoro.

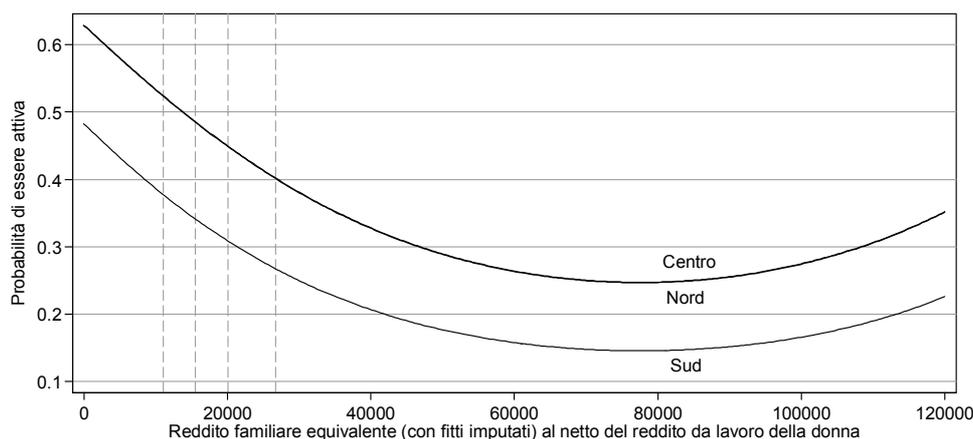
Variabile	Coefficiente		Errore standard
Anni di istruzione	0.1221	***	0.0057
Anni di esperienza	0.0878	***	0.0027
Età	-0.0824	***	0.0033
Numero di figli di età <3	-0.4468	***	0.0561
Può contare su altre persone	0.1881	**	0.0654
Nord	0.3710	***	0.0490
Centro	0.3734	***	0.0576
Reddito familiare equivalente (fitti imputati) al netto del reddito da lavoro della donna	-2.6134e-05	***	3.62e-06
(Quadrato) Reddito familiare equivalente (fitti imputati) al netto del reddito da lavoro della donna	1.6848e-10	***	3.44e-11
Costante	1.2943	***	0.1480

Note: *** significativo allo 0.1%
 ** significativo allo 0.5%.

Per meglio valutare la relazione funzionale tra probabilità di essere attiva e il reddito familiare equivalente al netto del reddito da lavoro della donna si sono rappresentate graficamente tre curve che descrivono tale relazione, ciascuna riferita alle tre ripartizioni territoriali analizzate (Figura 1). Le curve sono state tracciate sulla base dei parametri di regressione stimati e per valori medi di tutti gli altri regressori (il grafico si riferisce quindi ad una generica donna con età, livello di istruzione, esperienza lavorativa a valori medi rilevati nel campione). La *dummy* riferita alla presenza di una rete sociale su cui poter contare è stata considerata uguale a zero.

La Figura 1 mostra un andamento molto simile tra il Nord e il Centro (le due curve sono quasi perfettamente sovrapposte). La curva relativa al Sud si colloca significativamente più in basso rispetto alle altre. La distanza tra le curve dipende dai risultati dei coefficienti relativi alle *dummy* territoriali, che dovrebbero catturare un insieme di fattori di mercato e culturali che differenziano le tre aree. Per facilitare l'interpretazione dei risultati sono state aggiunte quattro linee verticali che rappresentano i quintili della distribuzione dei redditi (quintili ricavati dalla distribuzione che include però, nel reddito complessivo familiare, anche il reddito da lavoro della donna). Questo accorgimento consente di apprezzare come il tratto crescente della curva di probabilità cominci per livelli di benessere elevati.

Figura 1 - Probabilità di essere attiva in funzione del reddito familiare equivalente (con fitti imputati) al netto del reddito da lavoro delle donne (probabilità calcolata per valori medi degli altri regressori e considerando assente la presenza di una rete sociale su cui la donna può contare).



4. Conclusioni

Questo lavoro di ricerca rappresenta un tentativo di collegare logiche e strumenti di misura del benessere familiare a questioni relative al mercato del lavoro.

L'analisi svolta ha evidenziato la rilevanza del reddito del marito nel determinare le scelte lavorative della donna sposata. Tale rilevanza è da ritenersi ancora valida nel mercato del lavoro italiano, nonostante le trasformazioni socio-culturali che vedono la donna portatrice di motivazioni al lavoro sempre più

incentrate su considerazioni di realizzazione personale.

Dai risultati del modello *probit* emerge che la relazione tra probabilità di essere attiva e la misura di benessere familiare da noi impiegata non è lineare. La forma ad U della relazione che ci aspettavamo è stata dimostrata dall'analisi di regressione, ma il tratto ascendente della curva comincia soltanto per livelli di reddito molto elevati.

Sfruttando la novità del questionario Eu-Silc 2006, che ha incluso una batteria di domande sulla partecipazione sociale, abbiamo verificato che l'esistenza di una rete sociale su cui poter fare affidamento influisce positivamente sulla probabilità di essere attiva. Questa informazione completa, sotto il profilo operativo, la concettualizzazione di contesto familiare che proponiamo in questo lavoro. Riteniamo che l'inclusione nel questionario Eu-Silc di tale batteria di domande possa essere proficuamente utilizzata in futuro per studi relativi al mercato del lavoro.

Riferimenti bibliografici

- Bono F., Giaimo R., Lo Magno G. L. 2008. *Un'analisi delle differenze di genere nel mercato del lavoro in Italia* in Atti del Convegno dell'Associazione Italiana di Scienze Regionali "Conoscenza, sviluppo umano e territorio", Bari, 24-26 settembre 2008.
- Commissione Europea 2004. *Common cross-sectional EU indicators based on EU-SILC; the gender pay gap*, Working group on statistics on income and living conditions (EU-SILC).
- Del Boca D., Locatelli M., Pasqua S. 2000. Employment decisions of married women: evidence and explanations. *Labour*, Vol. 14, pp. 35-52.
- Hagenaars A., De Vos K., Zaidi M. A. 1994. *Poverty Statistics in the Late 1980s: Research Based on Micro-data*, Office for Official Publications of the European Communities. Luxembourg.
- Istat 2008. Conciliare lavoro e famiglia. Una sfida quotidiana. *Argomenti* n.33.
- Juhn C., Murphy K. M. 1997. Wage inequality and family labor supply. *Journal of labor economics*, Vol. 15, No. 1, pp. 72-97.
- Mroz T. A. 1987. The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions. *Econometrica*, Vol. 55, No. 4, pp. 765-799.
- Oaxaca R. 1973. Male-female wage differentials in urban labor markets. *International economic review*, Vol. 14, No. 3 (Oct., 1973), pp. 693-709.
- Presidenza del Consiglio dei Ministri, Dipartimento per gli Affari Sociali 2000. *La povertà delle donne in Italia. La ricerca, i dati, le metodologie di analisi*, a cura della Commissione di Indagine sull'Esclusione Sociale (presidente della commissione Saraceno C.).

SUMMARY

Participation rates in the female labor market in Italy are low and considerable regional differences exist. In 2006, 50.8% of the female population in Italy was employed or seeking employment, compared with 37.3% in the south of Italy; the Eu-27 mean was 63%.

In this paper we focus on married women, analysing variables which can explain labor supply. We emphasize household variables which comprise what we define as "family context", i.e. family wellness, the presence of dependent children and the ability to ask relatives, friend or neighbours for help. The aim of this paper is to verify the existence of a "U" relationship between female participation in the labor market and household income. For this reason, we have used equivalent household income, excluding the income derived from female labor as a proxy of family wellness. We wanted to verify whether this relationship is always negative or, departing from a certain level of wellness level, whether it may be positive. This may be due to the presence of economic resources, encouraging women to be gainfully employed out of the home, and thus reducing their work load in the home. Our results revealed the expected "U" form but the ascending part of the curve only starts from high levels of income.

Rosa GIAIMO, Professore ordinario di Statistica per la Ricerca Sperimentale e Tecnologica presso il Dipartimento di Contabilità Nazionale ed Analisi dei Processi Sociali (Dicap) della Facoltà di Economia dell'Università di Palermo.

Dario CORSO, Assegnista di ricerca presso il Centro Interdipartimentale di Ricerca per il Monitoraggio dell'Economia e del Territorio (Cirmet) dell'Università di Palermo.

Giovanni Luca LO MAGNO, Dottorando di ricerca presso il Centro Interdipartimentale di Ricerca per il Monitoraggio dell'Economia e del Territorio (Cirmet) dell'Università di Palermo.

FATTORI DI ATTRATTIVITÀ DELLE MIGRAZIONI: LE MINORANZE ETNICHE. IL CASO DEGLI ARBËRESHË.

Pietro Iaquina, Francesca Di Lazzaro

La Calabria è da secoli terreno d'incontro di correnti migratorie ed esperienze culturali diverse, che continuano a caratterizzare e conformare la vita sociale e culturale. Le numerose minoranze etniche e linguistiche presenti nella nostra regione costituiscono una testimonianza, anche se difficile da cogliere, dato che si tratta di comunità ben integrate nel tessuto sociale e produttivo.

In particolare, il lavoro analizza le caratteristiche delle minoranze etniche – culturali di estrazione Arbëreshë presenti in Calabria e i rapporti socio - culturali che legano a distanza di circa 5 secoli gli Albanesi di ieri e di oggi. E' stato affrontato e descritto il mondo degli albanesi, realizzando un confronto tra gli albanesi arrivati in Italia nel XV secolo, che hanno dato origine alle comunità Arbëreshë, ancora oggi presenti in Calabria e gli albanesi che hanno iniziato ad emigrare dal 1990 fino ai nostri giorni.

Prima di descrivere la minoranza Arbëreshë, appare importante precisare il concetto di minoranza. Secondo la definizione, elaborata nel 1977 da Francesco Capotorti, in conformità con l'art. 27 del Patto sui Diritti civili e politici: “... una minoranza è un gruppo numericamente inferiore al resto della popolazione di uno Stato, in posizione non dominante, i cui membri – che sono cittadini dello Stato – possiedono caratteristiche etniche, religiose o linguistiche che differiscono da quelle del resto della popolazione, e mostrano, anche solo implicitamente, un senso di solidarietà, diretta a preservare la loro cultura, tradizione, religione o lingua ”.

In Italia sono presenti diverse realtà minoritarie. Distinguiamo le *minoranze di antico insediamento* da quelle di *nuova formazione*. Le prime, sparse sul territorio nazionale, hanno conservato nei rapporti familiari e all'interno del gruppo, forme espressive della lingua d'origine, pur avendo piena padronanza della lingua italiana che usano nei rapporti esterni ed ufficiali. Le minoranze di nuova formazione, fanno riferimento agli agglomerati costituiti da immigrati stranieri e da emigrati italiani che dalla loro terra d'origine, si sono trasferiti, per lavoro, in aree di sviluppo industriale, specie nel nord.

Il 15 dicembre 1999 è stata approvata in Parlamento la legge n. 482, che prevede: “ Norme in materia di tutela delle minoranze linguistiche storiche ”.

In particolare, la Repubblica tutela la lingua e la cultura della popolazione: Albanese, Catalana, Croata, Francese, Francoprovenzale, Friulana, Greca, Ladina, Occitana, Sarda, Slovena, Tedesca (art. 6).

Non rientrano tra le minoranze tutelate dalla suddetta legge:

1. Eteroglossie Interne;
2. Minoranze Diffuse;
3. Nuove minoranze.

Secondo la definizione elaborata da Tullio Telmon, nel 1992, le *eteroglossie interne*, sono da intendersi come un "codice linguistico estraneo" ad un certo territorio, un codice dipendente "direttamente dalla mobilità e dai fatti di migrazione". Il caso più vistoso concerne i Tabarchini di Sardegna e i Galloitalici del Meridione (Basilicata e Sicilia).

Non vengono poi contemplate dalla legge le cosiddette '*minoranze diffuse*', ossia quelle comunità che si collocano all'interno di un determinato paese in modo non territoriale, sparse, disseminate a piccoli gruppi sul territorio*. E' intuitivo pensare che il soggetto elettivo di questa categoria siano i Nomadi.

Inoltre il panorama linguistico italiano deve fare i conti con un fenomeno nuovo ed in costante crescita: la presenza di consistenti gruppi di cittadini immigrati provenienti dai più diversi paesi e di parlata diversa da quella italiana che, con espressione coniata da Tullio De Mauro nel 1974, denominiamo '*nuove minoranze*'. Anche se il fenomeno non è così appariscente come in altri paesi, è stato accertato che gli idiomi esogeni praticati nello spazio linguistico italiano, le cosiddette "lingue immigrate", sono almeno 122.

La minoranza linguistica albanese rappresenta una delle principali comunità presenti in Calabria. Gli Albanesi appartengono al gruppo di minoranze di antico insediamento, che non hanno alcuna continuità territoriale con il ceppo d'origine.

Sono, vere e proprie isole linguistiche di antica tradizione, che hanno tramandato, attraverso i secoli e per lo più oralmente, un ricco patrimonio linguistico – culturale – religioso. I due elementi caratterizzanti la cultura arbëreshë sono la lingua e la religione. La cultura albanese si basa su una lingua esclusivamente parlata e non scritta. Gli albanesi sono una popolazione bilingue: l'italiano rappresenta per loro la lingua dell'ufficialità e l'arbëresh, la lingua della famiglia, che ha per la comunità una sorta di funzione aggregativa. Per ciò che riguarda il rito religioso: gli albanesi in Italia seguono il rito bizantino in lingua greca. Molte comunità albanesi hanno però perso, nel corso dei secoli, sia la lingua che il rito. Oggi il rito bizantino sopravvive soprattutto nelle comunità albanesi della provincia di Cosenza e in quelle intorno a Piana degli Albanesi, in Sicilia.

* Giulio Soravia, In un'Europa plurilingue. Culture in transizione, Bologna 1998.

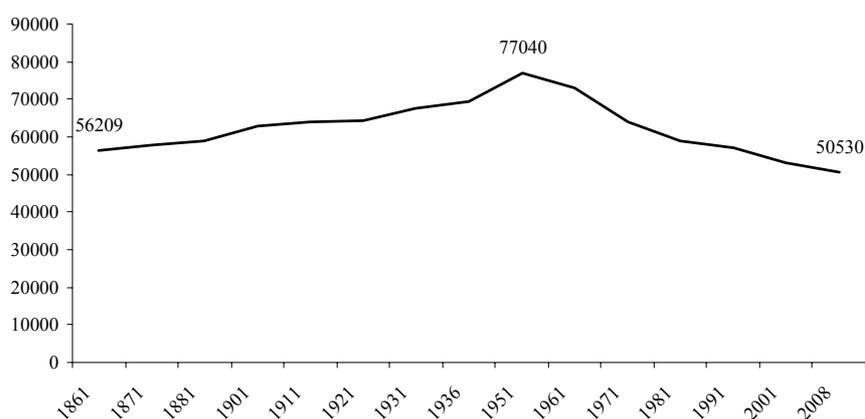
I primi albanesi sono arrivati in Italia intorno al 1400. Non erano né esuli, né profughi. Erano soprattutto soldati, che venivano ad aiutare il re di Napoli contro l'anarchia feudale dei baroni. Le vere migrazioni di popolo, invece, sono cominciate dopo la morte di Scanderbeg nel 1468. Si dirigono verso l'Italia perché qui già vi erano dei possedimenti, che erano già stati dati a Scanderbeg e ad altri condottieri albanesi dai re di Napoli come ricompensa degli aiuti militari ricevuti. A partire erano soprattutto le persone benestanti. Partivano, perché non volevano sottostare al dominio turco e non volevano accettare l'islamizzazione forzata. Infatti, le più importanti migrazioni avvennero proprio a causa dell'invasione turca dal 1470 al 1534. Non è facile stabilire l'esatta provenienza di questi migranti, poiché l'esodo si è perpetrato ad ondate successive. All'inizio poi, per non pagare tasse gravose, come il focatico, spesso queste popolazioni si sono spostate da un posto all'altro, che bruciavano prima di partire. Sono stati per un certo periodo anche nomadi, ma nomadi "atipici", perché erano comunità che cercavano solo il momento buono per fermarsi. E così pian piano hanno fondato nuovi paesi, ma hanno anche ripopolato vecchi paesi abbandonati per la malaria o in seguito a terremoti. In questo modo e in tempi diversi, si sono formate moltissime colonie, molte di più di quelle arbëreshë rimaste fin'ora. Gli ultimi due paesi dove si insediò la minoranza albanese, fu Villa Badessa in Abruzzo (1744) e Brindisi di Montagna in Basilicata (1774).

I 46 comuni di espressione albanofona sono distribuiti in 6 regioni dell'Italia meridionale: Molise, Campania, Puglia, Basilicata, Calabria (27 comuni – 8 frazioni), Sicilia.

In Calabria, le comunità albanesi risultano così distribuite:

- Il 79% si trova in provincia di Cosenza;
- Il 15% in provincia di Catanzaro;
- Il 6% in provincia di Crotona.

Nella provincia di Cosenza vi è il maggior numero di comuni di origine albanese. Sono presenti 21 comuni e 6 frazioni. Le comunità contano pochi abitanti. Il comune più popolato è Spezzano Albanese (con circa 7.000 ab.) e San Demetrio Corone (con circa 4.000 ab.). Il comune più piccolo è Castroregio, che con l'annessa frazione Farneta, conta circa 400 abitanti (Fonte: Istat.)

Figura 1 – *Evoluzione della popolazione arbëreshë in Calabria (1861-2007).*

Fonte: nostra elaborazione su dati Istat

L'attenzione si è soffermata sull'evoluzione della popolazione arbëreshë in Calabria dal 1861 al 2007. In particolare, dal 1861 fino al 1951, si è avuto un incremento della popolazione residente in quasi tutti i comuni arbëreshë, soprattutto grazie al positivo incremento naturale della popolazione che copre il deficit derivante dal movimento migratorio. Da questa data in poi la popolazione inizia a diminuire. Al 2007 i comuni arbëreshë contano circa 53.000 persone.

Dall'analisi, emerge come la comunità Arbëreshë abbia attraversato due crisi demografiche durante il secolo scorso. La prima è avvenuta alla fine del secolo (tra il 1876 e il 1915 circa) come conseguenza della grande emigrazione italiana. Questa crisi è stata in parte compensata dall'elevato tasso di natalità. La seconda crisi, ancora una volta causata da grandi emigrazioni, si è verificata dopo la fine della Seconda Guerra Mondiale e, in quanto ha coinciso con la fine del processo di transizione demografica, ha avuto conseguenze di rilievo:

- Una grande diminuzione della popolazione;
- Un numero crescente di anziani;
- Un grande calo del tasso di natalità.

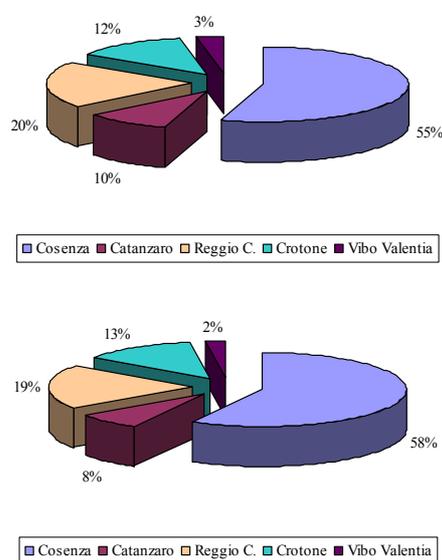
In seguito, per ogni comunità osservata, è stato analizzato l'andamento dei principali indicatori demografici dal 20 ottobre 1991 al 31 dicembre 2007.

Dall'analisi emerge come la diminuzione della popolazione, un numero crescente di anziani e il calo del tasso di natalità, propri della seconda crisi demografica, continuino a caratterizzare quasi tutte le comunità albanesi fino ai

nostri giorni. Con le nuove ondate migratorie prima, ma in particolare con lo scoppio della guerra in Albania (1997), è tornata visibile in Italia la presenza albanese.

L'analisi si conclude, infatti, con la costruzione di alcuni indici, per misurare la capacità che le provincie calabresi hanno di attirare sul proprio territorio la popolazione albanese presente a livello regionale.

Figura 2 – *Popolazione albanese residente in Calabria (31 dicembre 2002 – 2007).*

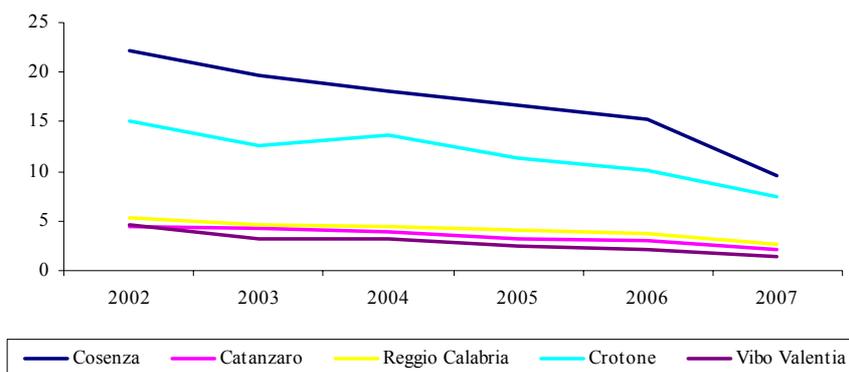


Fonte: nostra elaborazione su dati Istat

Tra le provincie calabresi, la popolazione albanese è concentrata maggiormente nella provincia di Cosenza, dove, tra l'altro, sono già presenti numerose comunità arbërëshe. Possiamo vedere come, al 31 dicembre 2002, l'ammontare della popolazione albanese residente in provincia di Cosenza, equivale al 55% del totale degli albanesi residenti in Calabria. Segue Reggio C., Crotona, Catanzaro e Vibo.

Nel 2007 questo valore aumenta per la provincia di Cosenza, dove risultano residenti circa il 58% del totale degli albanesi residenti in Calabria, e in maniera modesta per la provincia di Crotona.

Figura 3 – Incidenza degli albanesi residenti in Calabria sulla popolazione locale residente.



Fonte: nostra elaborazione su dati Istat

Diminuisce invece, per tutte le provincie calabresi, la quota della popolazione albanese rispetto al totale della popolazione straniera residente. Al 2007 la provincia che mantiene il valore più alto è Cosenza, seguita da Crotona. Le altre provincie presentano valori che oscillano tra l'1,3% e il 2,6%. Mentre al 2002 gli albanesi occupavano il secondo posto tra le comunità straniere presenti in Calabria dopo il Marocco; al 2007 la comunità albanese si colloca al sesto posto.

In particolare, per quanto riguarda la provincia di Cosenza, mentre al 2002 gli albanesi occupavano il primo posto nella graduatoria tra le popolazioni straniere nella provincia di Cosenza; al 2007 vi è stata una presenza maggiore di Rumeni e Marocchini.

Dalle analisi effettuate emerge come la direzione dei flussi migratori non sia assolutamente casuale, ovvero le aree di destinazione delle persone o seguono “catene etniche” e/o si dirigono verso zone particolari in cui le possibilità di lavoro e le condizioni di vita sono indubbiamente migliori rispetto a quelle dei paesi di origine. Gli albanesi hanno scelto l'Italia non semplicemente per la facile raggiungibilità di questa terra, ma per il facile processo di ambientamento grazie alla conoscenza della lingua, della cultura e arte italiana.

A ciò si aggiunge una particolare sintonia, che certamente esiste, dovuta alla minoranza Italo – albanese presente in Italia e in maggior numero in Calabria.

Infatti, abbiamo visto come gli albanesi immigrati in Calabria, sono maggiormente concentrati in provincia di Cosenza, dove è presente il maggior numero di comunità Arbëreshë.

Riferimenti bibliografici

- Barjaba K., Dervishi Z., Perrone L., "L'Emigrazione albanese: spazi, tempi e cause", in Studi Emigrazione, XXIX, n°107, 1992.
- Barjaba K., "L'evoluzione dei modelli migratori albanesi, 1990-1997", in "ISIG Magazine", Istituto di sociologia internazionale, n. 4/97, 1997.
- Barsotti O. , E. Moretti (a cura), Albania, Rimesse e cooperazione allo sviluppo, Franco Angeli, Milano, 2004.
- Di Comite L, Valleri M. A., Problemi demo-economici dell'Albania, «Quaderni mediterranei», n. 1, Ed. Argo, Lecce, 1994.
- Giovannelli Costanza, L'immigrazione in Italia dai paesi balcanici nel corso degli anni '90: caratteristiche e tendenze evolutive, in Popolazione e migrazioni nei Balcani (Di Comite, Pellicani editors), Atti del Convegno "Population at Migrations dans les Balkans", Università degli Studi di Bari, Dipartimento per lo Studio delle Società Mediterranee, Cacucci Editore, Bari, 2002.
- Giulio Soravia, In un'Europa plurilingue. Culture in transizione, Bologna 1998.
- Iaquinta P.- Meksi E., La Mortalità Infantile in Albania; Estratto da Problemi demo - economici dell'Albania, a cura di L.Di Comite e M. A. Valleri; Quaderni Mediterranei 1, pag. 95, Bari, 1994.
- Melchionda U. (a cura di), Gli albanesi in Italia: inserimento lavorativo e sociale, Franco Angeli, Milano, 2003.
- Paterno A. Strozza S., Terzera L., Albanesi in patria e all' estero: migrazioni e popolazione a confronto, in Sospesi tra due rive: migrazioni e insediamenti di marocchini e albanesi, Franco Angeli, Milano, 2006.
- Perrone L. (a cura di), Naufragi albanesi. Studi, riflessioni e ricerche sull' Albania, Ed. Sensibili alle foglie, Roma, 1996.
- Perrone L., «L' Albania tra passato e presente alla ricerca di nuove identità», in «Sociologia urbana e rurale», n. 59, 1999.
- Pittau F., Reggio M., Il caso Albania: immigrazione a due tempi, in «Studi Emigrazione», n. 106, 1992.
- Resta P., Un popolo in cammino: le migrazioni albanesi in Italia, Besa, Lecce, 1996.

Sitografia

www.istat.it

www.instat.gov.al/

www.jemi.it

ABSTRACT

The Italy was for centuries the land set to become the meeting place of migratory flows and different cultural experiences, which have characterized and shaped the social and cultural life.

The numerous ethnic and linguistic minorities still present on Italian soil are an important witness to these phenomena, although difficult to grasp, since it is well-integrated community in the social and productive.

In particular, the work describes the characteristics of ethnic minorities - cultural extraction Arberesh in Calabria and social relations - that bind cultural distance of about five centuries the Albanians arrived in Italy around the fifteenth century, which gave rise to communities and Arberesh Albanians who began to migrate from 1990 to the present day.

The first Albanians arrived in Italy around 1400. The real migration, however, began after the death its national Scanderbeg in 1468, because of 'the Turkish invasion.

The minority Albanian language belongs to the minority group of ancient, which have no territorial continuity with the strain of origin. Are real islands language of ancient tradition, which have passed through the centuries and mostly oral, a rich linguistic heritage, cultural and religious. The common expression albanofona are 46 and are distributed in 6 regions of southern Italy: Molise, Campania, Puglia, Basilicata, Calabria and Sicily.

The work follows the evolution of the population in Calabria Arberesh 1861 to 2008. Then, with new waves of migration, but particularly with the outbreak of war in Albania, is back in Italy the visible presence in Albania.

The choice of Italy is not justified simply by the easy accessibility of this land, but the easy process of adaptation through a particular line, due to the presence of minorities Arberesh. Indeed, from a quantitative analysis on the Albanian presence in Calabria, emerge as the Albanians are more concentrated in the province of Cosenza, where there is a greater number of communities Arberesh.

Albanians are so tied to the population Arberesh, because the settlement of these communities are an important case that integration does not clear, but brings out the differences.

Pietro IAQUINTA, Dipartimento di Economia e Statistica – Facoltà di Economia – Università della Calabria.

Francesca DI LAZZARO, Dipartimento di Economia e Statistica – Facoltà di Economia – Università della Calabria.

STATISTICAL IMPLICATIVE ANALYSIS AND PRINCIPAL COMPONENTS METHOD: A COMPARISON BETWEEN TWO APPROACHES FOR THE ANALYSIS OF SOCIAL DATA

Letizia La Tona, Angela Alibrandi¹

1. Introduction.

In applied statistics a frequent problem that researchers have to solve is the reduction of the statistical variables; in fact in a multivariate context, especially with reference to the social phenomena, statisticians have a great number of variables that must be synthesized and represented through a smaller number of them, providing that the loss of information is minimum.

The classical statistics treats the reduction of variables problem by the principal components analysis (PCA) (Härdle and Simar, 2004), a well known multivariate procedure that carry out transformation of the original variables in some new variables that have to be linear combination of the original ones, have to possess the maximum variance and have to be uncorrelated among them.

The reduction of variables problem has been faced not only within classical statistics but also by the Statistical Implicative Analysis (SIA) (Gras, 1996; Spagnolo, 1997).

The implicative statistical approach is originated from the extension of implication concept, it formalizes particularly structured situations and furnishes information about the implication among classes of variables through the cohesion concept. In the statistical implicative approach Gras (2000) has faced the reduction of variables problem; he underlines that an implication among classes of variables assumes real meaning only if within each class a "cohesion" exists. Given two classes of variables A and B, there is cohesion if an implicative "flow" of the class A on the class B exists. The cohesion is fed by a flow within A and it feeds a flow within B. If we extend the concept to n classes of variables, the global cohesion is obtained by the geometric mean of all couples of classes cohesions. Through the concept of cohesion, in implicative statistics the researchers can distinguish the *principal* variables from the *additional* variables. According to Gras, given a multiplicity of level classes, it is essential to individualize the most pertinent class,

¹ This note, though it is the result of a close collaboration, was specifically elaborated as follows: paragraphs 1 and 3 by L. La Tona and paragraphs 2 and 4 by A. Alibrandi.

with relationship to the classificatory intention of researchers and to the respect of selection criteria.

The purpose of the present paper is to expose the main methodological issues of Statistical Implicative Analysis for the reduction of variables and to compare this approach to principal component methods. In detail the paper is so structured:

- in paragraph 2 the reduction of variables procedures through SIA is exposed;
- in paragraph 3 a methodological comparison between SIA approach and PCA is performed;
- in paragraph 4 an application of two procedure is proposed for the reduction of variables on social data.

2. Reduction of variables procedures through SIA approach

Gras (1996, 2000) faces the reduction of variables according to the SIA approach; the author shows that all variables can be categorized in two sets: *principal* and *additional* variables. Considering the V set of the variables $\{a_1, a_2, \dots, a_m\}$ and the set of the couples (a, b) of $V \times V$ with $a \neq b$, it is possible to form $m(m-1)$ couples to which associate the respective cohesion $(c(a, b))$. Then, it is possible to define the cohesive initial and global preorder on $V \times V$; it is the preorder Ω that is induced by the cohesion application $c(a, b)$ on $V \times V$. For each level of the implicative hierarchy, a class of m_i variables ($2 \leq m_i \leq m$) is formed with cohesion different than zero. The cohesion of these variables can be estimated by the cohesion index for class (a, b) , that can be expressed as follows:

$$c(a, b) = \sqrt{1 - H^2} \quad (1)$$

with

$$H = p \log_2 p - (1 - p) \log_2 (1 - p) \quad (2)$$

where

- $\varphi(a, \bar{b})$ denotes the implication intensity from a on b ,
- $\varphi(b, \bar{a})$ the implication intensity from b on a ;
- $p = \max(\varphi(a, \bar{b}), \varphi(b, \bar{a}))$ denotes the maximum value between two intensities.

H denotes the entropy index, that allows to measure the disorder between two variables. In particular:

- if $p=1$ the cohesion index $c(a, b)=1$;
- if $p \leq 0.5$ the cohesion index $c(a, b)=0$.

Extending the concept to a class of r variables (a_1, a_2, \dots, a_r) , the cohesion $c(a_1, a_2, \dots, a_r)$ is given by the geometric mean of cohesions for classes of two elements, as follows:

$$c(a_1, a_2, \dots, a_r) = \left[\prod_{\substack{i \in \{1, \dots, r-1\} \\ j \in \{2, \dots, r\}, j > i}} c(a_i, a_j) \right]^{\frac{2}{r(r-1)}} \quad (3)$$

The variables that are characterized by a strong cohesion are defined *principal* or *intrinsic*; the ones with low cohesion are *additional* or *extrinsic* variables (and therefore gregarious, eliminable). In this way, it is possible to affirm that the implicative criterion for reducing the number of statistical variables is based on the cohesion evaluation: the principal and the additional variables are identified on the basis of their implicative power and the contribution that each of them gives to the class constitution.

In Counturier et al. (2004) a new method to reduce the number of variables in a large set of data is introduced. SIA approach, which builds association rules with a measure more powerful than conditional probability, is used to detect *quasi-equivalent* variables. The notion of *quasi-equivalence* is related to a given problem and the detection process is strongly dependent on several parameters. The number of variables and the size of data are likewise important clues to take into account. The authors aim is to provide a solution in order to automatically detect some *quasi-equivalent* variables, thus allowing the generation of a new smaller set of data without the equivalent variables in which the semantic should be globally conserved.

After the detection of the set of equivalent variables, it is easier to analyse only one of them (or a new one) that should be used to replace the initial set. Obviously, two variables A and B are equivalent if and only if $A \Rightarrow B$ and $B \Rightarrow A$. Thus, based on the logical definition of the equivalence, the *quasi-equivalence* is so defined:

$$Quasi(A, B) = \sqrt{Relevance(A \Rightarrow B) * Relevance(B \Rightarrow A)} \quad (4)$$

As we can notice, the *quasi-equivalence* index is build by computing the geometric mean of $A \Rightarrow B$ and $A \Rightarrow B$. Given a threshold and using this formula for the *quasi-equivalence*, a list of couples of variables having a coefficient of *quasi-equivalence* greater or equal to the threshold is obtained. Then a *quasi-equivalence* class of n variables A_1, \dots, A_n is measured by the *quasi-equivalence* coefficient of this class, that is defined as follow:

$$Quasi - equivalence - class(A) = \min_{i,j} (quasi(A_i, A_j)) \quad (5)$$

After the determination of a *quasi-equivalence* class, an existing variable or a new one is selected to replace all the variables of the class. This variable can be

called “*leader class*”. It really possesses a leader role as it represents a better class than the others. The choice of the leader class is based on a distance notion. The distance between two variables A and B is defined as:

$$Dist(A, B) = |Relevance(A, B) - Relevance(B, A)| \quad (6)$$

By this distance, it is possible to know the distance between a variable and the other variables of the same class. The distance of a variable A_i with all other variables A_j in a class C with cardinality $|C|$ is defined as:

$$dist(A_i, C) = \frac{1}{|C|} \sum_{A_j \in C; A_i \neq A_j} dist(A_i, A_j) \quad (7)$$

The leader of a class is the variable that minimizes this distance with any other variables. The leader A_i of the class C is so defined:

$$leader(C) = \min_{A_i \in C} dist(A_i, C) \quad (8)$$

The leader of a class with only two elements cannot be determined because the distances are equal. As a consequence, Counturier et al (2004) choose to define the leader as the variable with the greater relevance. In the *quasi-equivalent* class of variables each couple of variables is *quasi-equivalent*. The elected variable minimizes the distance with all the other variables.

Blanchard et al. (2002) try to individualize an implicative statistical criterion that allows to reduce the number of variables to use. They define the notion of *quasi-implication* and of *quasi-equivalence* among variables and build the classes from which it's possible to extract a representative one. Subsequently, they illustrate an algorithm to optimize the choice of the representative variable in every class of *quasi-equivalence*.

Two binary variables A and B are equivalent when the two implications $A \Rightarrow B$ and $B \Rightarrow A$ are simultaneously satisfied. In the experimental situations, the cases in which these two implications are logically and simultaneously valid are rare. The authors propose two criterions to evaluate the *quasi-implication*: the first one is inspired to the principle of the likelihood of Lermann, the second is inspired to the principle of the entropy or “*quasi-inclusion*” that translates the implication between two binary variables.

According to the first criterion, two variables A and B are *quasi-equivalent* when the intensity of implication $A \Rightarrow B$ is “few” different of the intensity of implication $B \Rightarrow A$. Nevertheless, the criterion of the intensity of implication is not more sufficiently discriminating when the populations of the observations are great (about a hundred).

The second criterion is based on "unbalance" between the entropies $A \Rightarrow B$ and $B \Rightarrow A$. The littleness of the difference between these two entropies will characterize the simultaneous satisfaction of the relationship of *quasi-implication*.

A representative subset of variables should be chosen by the number of times in which the variable figures in the given class of equivalence. Nevertheless, this methodology doesn't allow optimizing the reduction since the choice of the representative variable is relatively arbitrary. So Counturier et al. (2004) propose the choice of a subset of variables based on *inertia criterion*. Let's indicate with c_i the *implicative cohesion* of the class C_i , with k_i the coefficients (effects) and with g the barycentre of all the variables; the *inertia index* is given by:

$$I = \frac{1}{\sum_{i=1}^k k_i c_i} \left(\sum_{i=1}^k k_i c_i [a_i g]^2 \right) \quad (9)$$

The inertia represents the implicative intra-class variance and expresses a synthetic measure of the implicative homogeneity inside a class. In this way, the variables that give a greater contribution are chosen.

In order to determine an optimum for the choice of the variables subset, there are different algorithms that are implementable through CHIC software.

SIA approach has been extensively used in literature in some context as teaching of mathematics (Spagnolo, 1998), analysis of teaching experiences (Giambalvo et al., 2003), experimental researches (Spagnolo, 2002).

3. A comparison between SIA and PCA

SIA approach and PCA allow to operate a reduction of variables number in two different ways; however a joint examination of two approaches appears interesting, since it's possible to individualize some similarities.

In PCA in order to choose the suitable number of factors the I_q index has to be calculated by the ratio between the explained variance of the first $q < k$ principal components (where the k is the number of the total components and q is the chosen number of components) and the total variance (Härdle and Simar, 2004). It is also known in literature as *Kaiser-Meyer-Olkin Measure of Sampling Adequacy*. This measure varies between 0 and 1 and values near to 1 denote a better synthesis of data. A value of 0.5 is to be considered a minimum value of adequacy. The I_q index within the PCA and the inertia index within SIA are based on the same principle, because both of them are obtained by the ratio between a limited part of information drawn from data and the whole information. In particular the

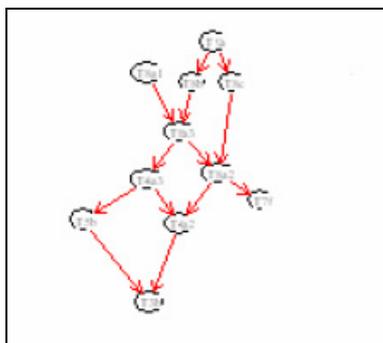
information is constituted by the variance in PCA and by the inertia in SIA. Both index are ranged between 0 and 1.

The reduction of variables through the cohesion index $c(a_1, a_2, \dots, a_r)$ assures the variables orthogonality; also PCA assures this condition and, more, realizes a linear combination of original variables. This allows to obtain adequate new variables that are uncorrelated (corresponding to the orthogonality condition); this is an advantage of the classical method since the new variables express the maximum realizable variance, extracting from dataset the maximum possible information.

4. An application of SIA and PCA to matriculation propensity data in a sample of students attending the last year of high school

We applied both procedures on data, collected in order to analyze the propensity of students of Messina towards the eventual future matriculation of students attending in 2008/2009 the last year of high school in Messina. The research allowed the collection of statistical information about 365 students. A questionnaire allowed to obtain the following variables: age, gender, type of attended school (liceo or not), propensity to university registration, average mark, satisfaction of relational life in the classroom, qualification of father and mother, socio economical welfare of family, number of brothers or sisters attempting university, time spent to study.

SIA approach allows to individuate the implication structure (Figure 1) among the mentioned variables. *Cohesion Index* has furnished a value of 0.658, denoting the existence of a strong cohesive power among the variables. The distance between the university registration and all remaining variables has been calculated. So, we individuate the *principal* variables (type of attended school, propensity to university registration, average mark, satisfaction of relational life in the classroom, qualification of father and mother, socio economical welfare of family, time spent to study) and the *additional* variables (age, gender, number of brothers attempting university).

Figure 1 – Implication structure among variables.

The leader of the variables class is university registration; it is directly explained by the time devoted to the study and the type of attended school. Other variables also indirectly explain their effects, but their cohesive power is less strong; inertia index of SIA assumes value of 0.699; similarly we applied PCA and estimated the I_q index in PCA, obtaining a value of 0.742; both indices denote a good synthesis of data and an high implicative homogeneity inside the class. This application of SIA and PCA for variables reduction on social data shows the concordance between the results obtained by the two different approaches; therefore their combined use may be useful to have a better guarantee of results.

References

- Blanchard J, Kuntz P, Guillet F, Gras R. 2003. *Implication intensity: from the basic statistical definition to the entropic version*. In: Bozdogan H (ed) *Statistical data mining and knowledge discovery*. Chapman & Hall, Boca Raton, pp 473-485.
- Couturier R., Gras R., Guillet F. 2004. *Reducing the number of variables using implicative analysis*. *Classification, Clustering and Data Mining Applications*, 277-285. Springer-Verlag.
- Giambalvo O., Milito A. M., Spagnolo F. 2003. *L'analisi implicative per lo studio di una esperienza didattica in statistica*. *Quaderni di Ricerca in Didattica* n.13, 136-151, Palermo.
- Gras R. 1996. *L'implication statistique. Nouvelle methode de données, Recherches en Didactique des Mathématiques*, La Pensée Sauvage, Grenoble.
- Gras R. 2000. *Les fondements de l'analyse statistique implicative*, *Quaderni di ricerca in Didattica*, n.9, p.187-208, Palermo
- Härdle W., Simar L. 2004. *Applied Multivariate Statistical Analysis*, Springer, Germany.
- Spagnolo F. 1997. *L'analisi a-priori e l'indice di implicazione statistica di Gras*, *Quaderni di Ricerca in Didattica GRIM*, n.7, Palermo.

Spagnolo F. 1998. Insegnare le matematiche nella scuola secondaria, La Nuova Italia, Firenze.

Spagnolo F. 2002. L'analisi quantitativa e qualitativa dei dati sperimentali, Argomentare, congetturare e dimostrare nella scuola di tutti, supplemento al n.10 dei Quaderni di Ricerca in Didattica GRIM, n.7, Palermo.

SUMMARY

In applied statistics, a frequent problem that the researchers have to solve is the reduction of the statistical variables.

The classical statistics faces the reduction of variables by the principal components analysis (PCA), a well known multivariate procedure that carry out transformation of the original variables in some new variables that have to be linear combination of the original ones, have to possess the maximum variance and have to be uncorrelated among them (Härdle and Simar, 2004).

The reduction of variables problem has been faced also by Statistical Implicative Analysis (SIA). SIA is based on the cohesion concepts (Gras 2000) and two kinds of variables are identified as "intrinsic" (principal) and "extrinsic" (additional). Thanks to the contribution that each variable furnishes to the class constitution, the implicative power can be evaluated. The concept of distance between two variables or between a variable A_i with all other variables A_j in a class (Counturier et al, 2004) individualizes the leader of a variables class. Blanchard et al. (2002) propose an implicative statistical criterion, based on inertia index in order to reduce the number of variables.

The purpose of the present paper is to expose the main methodological issues of SIA for the reduction of variables, to compare this approach to PCA and to apply them on the same data related to matriculation propensity in a sample of students attending the last year of high school in Messina.

This application of SIA and PCA for variables reduction on social data shows the concordance between the results obtained by the two different approaches; therefore their combined use may be useful to have a better guarantee of results.

Letizia LA TONA, Professore Ordinario di SECS-S/02 (Statistica per la Ricerca Sperimentale e Tecnologica) presso la Facoltà di Scienze Statistiche dell'Università di Messina.

Angela ALIBRANDI, Ricercatore di SECS-S/02 (Statistica per la Ricerca Sperimentale e Tecnologica) presso la Facoltà di Scienze Statistiche dell'Università di Messina.

LA RASCH ANALYSIS ASSOCIATA AI MODELLI MULTILEVEL NELLA VALUTAZIONE DELLA CUSTOMER SATISFACTION

Antonio Lucadamo, Giovanni Portoso

1. Introduzione

La valutazione di capacità, attitudini, soddisfazione degli individui, relativamente a diversi aspetti della realtà, è da sempre una delle problematiche che si riscontrano nelle scienze sperimentali. Infatti, le qualità sopraelencate non sono direttamente osservabili, ma espresse con modalità di scale politomiche e comunque ordinate. Per poter effettuare la valutazione si rende necessario sostituire alle eventuali modalità categoriche, punteggi, la cui scelta pone problemi di quantificazione in relazione alle distanze fra di essi. Intervengono inoltre difficoltà riguardanti le differenze relative a caratteri stimati, per esempio, in tempi e situazioni diverse o per campioni di individui differenti.

Per ovviare a questi problemi, uno dei modelli più utilizzati è l'analisi di Rasch. A seconda delle peculiarità che assumono i diversi item si avrà a che fare con modelli di Rasch differenti. L'utilizzo della Rasch Analysis sarà il punto di partenza del nostro paper. S'intende infatti valutare la Customer Satisfaction (CS) dei pazienti di un ospedale, ai quali è stato somministrato un questionario volto a rilevare la soddisfazione relativa a diversi aspetti del ricovero. Dopo aver effettuato quindi la quantificazione dei punteggi grazie al modello di Rasch, si passerà a uno studio che cercherà di mettere in risalto come la soddisfazione del paziente possa essere influenzata da una serie di fattori socio-economici o dovuti al contesto in cui l'individuo risulta essere inserito. Per raggiungere questo scopo il modello più opportuno è senz'altro un modello di tipo multilevel.

Alla luce di quanto detto finora si introdurranno di seguito l'analisi di Rasch e i modelli multilevel per poi presentare il caso studio di interesse.

2. La Rasch Analysis

Il modello di Rasch fornisce la probabilità che un individuo assuma una certa modalità nella risposta a un determinato item in funzione delle stime di una serie di parametri quali l'abilità dell'individuo, la difficoltà dell'item e la soglia raggiunta. Si ottiene questo risultato grazie alla costruzione di misure probabilistiche derivate da metodi di calibrazione che partono dagli items unidimensionali. Queste misure

sono di tipo lineare, quantitative, “test free” e “sample free”, cioè libere rispetto al tipo di items considerati e non dipendenti dalle caratteristiche del campione di persone sottoposte alla prova. Questo avviene grazie a tre proprietà fondamentali di cui gode il modello di Rasch, cioè l’additività, la separabilità e l’obiettività specifica. La prima proprietà afferma che le misure costruite tramite la Rasch analysis sono funzione diretta delle abilità delle persone e funzione inversa della difficoltà dell’item e delle difficoltà a raggiungere determinate soglie. Secondo la proprietà della separabilità invece le misure ottenute sono indipendenti dal campione e dalla difficoltà del test, mentre, tenendo conto dell’ultima, possiamo affermare che le probabilità sono esprimibili in punteggi logit, misure che variano lungo tutto l’asse reale e non sono affette da problemi di schiacciamento sui valori estremi (Wright, Masters, 1982). Il modello di Rasch più semplice è il modello dicotomico in cui ci sono soltanto due opzioni, sì o no, che sono in genere codificate come 1 e 0. Il modello è espresso nella seguente forma:

$$\phi_{ni1} = \frac{\exp(\beta_n - \delta_{i1})}{1 + \exp(\beta_n - \delta_{i1})} \quad (1)$$

dove ϕ_{ni1} è la probabilità che la persona n assegni il valore 1 alla domanda i , β_n è l’abilità dell’individuo n e δ_{i1} è la difficoltà ad effettuare il primo (e in questo caso unico) step nella domanda i .

Il modello dicotomico è la base per passare facilmente ai modelli che presentano più di due possibilità di risposta e che sono il modello Partial Credit e il modello Rating Scale. In entrambi i casi bisogna tener conto del fatto che si avrà a che fare con items in cui si hanno p modalità di risposta. Nel modello Partial Credit, fissato che un individuo n abbia superato in un determinato item (i) la modalità $p-1$, la probabilità che lo stesso soggetto superi anche la modalità p sarà data da

$$\phi_{nip} = \frac{\exp(\beta_n - \delta_{ik})}{1 + \exp(\beta_n - \delta_{ik})} \quad (2)$$

Dove ancora una volta β_n è l’abilità del soggetto n -esimo mentre δ_{ik} è la difficoltà nell’item i -esimo a superare il k -esimo step. A questo punto è possibile definire l’espressione generale che calcola la probabilità che il soggetto n risponda col valore p nella domanda i :

$$\pi_{nip} = \frac{\exp \sum_{j=0}^p (\beta_n - \delta_{ij})}{\sum_{k=0}^m \exp \sum_{j=0}^k (\beta_n - \delta_{ij})} \quad p = 1, \dots, m \quad (3)$$

Il numeratore della precedente espressione contiene soltanto le difficoltà relative agli p steps superati, mentre il denominatore è la somma di tutti i possibili numeratori. La differenza fra il modello Partial Credit e il Rating Scale è relativa alla difficoltà che ogni individuo ha nel superare le diverse modalità. Mentre nel Partial Credit Model ogni domanda può presentare difficoltà differenti nel passaggio dallo step $k-1$ allo step k , nel secondo caso il parametro δ_{ik} viene invece suddiviso nel modo seguente:

$$\delta_{ik} = \delta_i + \tau_k \quad (4)$$

Dove δ_i è il valore di scala dell'item mentre τ_k rappresenta la soglia di separazione fra le modalità che, in questo caso, a differenza del Partial Credit Model, non varia al variare delle domande. In questo caso la probabilità che il soggetto n ottenga modalità p nella domanda i sarà data da:

$$\pi_{nip} = \frac{\exp \sum_{j=0}^p [\beta_n - (\delta_i + \tau_j)]}{1 + \exp \sum_{j=0}^k [\beta_n - (\delta_i + \tau_j)]} \quad (5)$$

Una volta definiti i diversi modelli, è ovviamente necessario tener conto di un criterio di stima che consenta di ottenere i valori dei parametri. Esistono diverse procedure da poter seguire quali ad esempio PROX, PAIR, CON o UCON. Nel nostro caso è stata scelta la prima che verrà presentata nel paragrafo relativo al caso studio specifico. Prima di giungere però a tale sezione è opportuno introdurre i modelli statistici multilevel.

3. I modelli multilevel

I modelli multilevel partono dal presupposto che, in una struttura gerarchica, i livelli "più elevati" influenzano il micro livello (Snijders, Bosker, 1999), cioè quello relativo agli individui; ad esempio in un ospedale il reparto può incidere sul livello di soddisfazione dei pazienti. Il modello di base da cui hanno origine tutti gli altri è l'Empty Model, definito come segue :

$$Y_{ng} = \gamma_{00} + \varphi_{0g} + \varepsilon_{ng} \quad (6)$$

In questo modello, che può essere visto come l'ANOVA ad effetti casuali, la variabile dipendente Y è data dalla somma della media generale (γ_{00}), dell'effetto casuale a livello di gruppo (φ_{0g}) e dell'effetto casuale a livello individuale (ε_{ng}). Questo modello è importante perché provvede alla partizione della variabilità fra i due livelli. Infatti, nei modelli multilevel oltre alle ipotesi della regressione classica, si assume che: le variabili casuali φ_{0g} e ε_{ng} siano mutuamente indipendenti, si distribuiscano normalmente con media pari a 0 ed abbiano varianze uguali rispettivamente a ζ^2 e σ^2 .

La varianza totale di Y può essere quindi scomposta come segue:

$$\text{var}(Y_{ng}) = \text{var}(\varphi_{0g}) + \text{var}(\varepsilon_{ng}) = \zeta^2 + \sigma^2 \quad (7)$$

Il coefficiente di correlazione intraclasse, dato da

$$\rho = \zeta^2 / (\zeta^2 + \sigma^2) \quad (8)$$

indica la quota di variabilità totale a livello di gruppo. Se questo è sufficientemente alto ha senso effettuare un'analisi multilevel, poiché si può affermare che buona parte della variabilità è attribuibile ai gruppi (Hox, 2002). Si può ora definire il Random Intercept Model come segue:

$$Y_{ng} = \alpha_{0g} + \alpha_1 x_{ng} + \varepsilon_{ng} \quad \text{con} \quad \alpha_{0g} = \gamma_{00} + \varphi_{0g} \quad (9)$$

con x_{ng} variabili individuali. In realtà se non fosse presente il pedice g l'equazione rappresenterebbe il modello di regressione lineare classico. Se il pedice g si aggiungesse anche a α_1 , oltre che all'intercetta, si avrebbe il Random Slopes Model.

$$Y_{ng} = \alpha_{0g} + \alpha_{1g} x_{ng} + \varepsilon_{ng} \quad \text{con} \quad \alpha_{0g} = \gamma_{00} + \varphi_{0g} \quad \text{e} \quad \alpha_{1g} = \gamma_{10} + \varphi_{1g} \quad (10)$$

4. L'analisi

L'analisi riguarda un'indagine relativa alla soddisfazione dei pazienti di un ospedale napoletano. I pazienti hanno compilato un questionario costituito da 30 domande inerenti i servizi ricevuti durante la degenza in ospedale. Ogni domanda prevedeva una valutazione compresa fra il voto 1 e il 7. Inoltre erano richieste informazioni socio-economiche quali l'età, il sesso, il reddito e l'istruzione. Per questo motivo è stata innanzitutto effettuata un'analisi di Rasch che permettesse di cogliere le propensioni degli individui a dare risposte positive o meno, dall'altro consentisse di evidenziare eventuali differenze fra le valutazioni ottenute dai diversi item. Come accennato precedentemente per poter ottenere le misure desiderate è stata presa in considerazione la procedura PROX, composta da una

serie di fasi che descriviamo. Innanzitutto è necessario effettuare una rimozione dal data-set di tutte quelle persone che hanno risposto con valori estremi a tutte le domande (tutti 1 o 7). La stessa operazione va effettuata per gli item, cioè vengono eliminate tutte quelle domande per cui ci sono state tutte valutazioni estremamente positive o negative. A questo punto si passa ad un'operazione di linearizzazione, effettuata trasformando gli scores di ogni domanda (cioè la somma dei punteggi attribuiti da tutti gli utenti) in una proporzione P_i , rapportandoli al valore massimo ottenibile. Questo valore viene trasformato in una scala logit considerando il logaritmo naturale del rapporto $l_i = \log\left(\frac{1-P_i}{P_i}\right)$ e si effettua poi la centratura,

sottraendo a tutti i punteggi così ottenuti la loro media ($d_i = l_i - \bar{l}_i$). Infine, per eliminare l'effetto del campione, si calcola la dispersione dello stesso nell'ipotesi di distribuzione normale con un fattore di correzione definito da $C_i = \sqrt{[(1+V)/2.89]/(1-UV/8.35)}$ dove V è la varianza corretta degli individui mentre U è la varianza corretta delle attività. Questo fattore viene moltiplicato per i punteggi centrati e si giunge in questo modo alla calibrazione finale: $d_i^* = d_i C_i$. Le operazioni di linearizzazione vengono ripetute anche per gli individui, considerando in questo caso gli scores ottenuti dai singoli soggetti e prendendo ancora una volta in esame la proporzione indicata ora con P_n e il

logaritmo $l_n = \log\left(\frac{P_n}{1-P_n}\right)$. Per eliminare invece la dispersione il fattore da considerare sarà dato da $C_n = \sqrt{[(1+U)/2.89]/(1-UV/8.35)}$ e di conseguenza la stima finale sarà data da: $d_n^* = d_n C_n$ dove $d_n = l_n - \bar{l}_n$.

Attraverso la procedura appena descritta è stato così possibile ottenere le nuove misure relative sia agli individui che alle singole domande. Le misure relative alle domande per le quali ci sono i punteggi più alti o più bassi sono riportate nella Tabella 1 e ci consentono di osservare quali siano gli aspetti del ricovero che hanno prodotto la maggiore soddisfazione nei pazienti. Una prima osservazione da effettuare, in modo da non lasciarsi fuorviare è relativa al fatto che per quanto riguarda gli items, le domande con scores più bassi sono in realtà quelle che hanno una valutazione migliore, dato che lo score per le domande viene calcolato, come visto, tenendo conto del logaritmo della probabilità di "insuccesso" su quella di

“successo” e quindi lo score rappresenta in effetti la difficoltà ad ottenere valutazioni elevate.

Tabella 1 – Risultati relativi ad alcune domande.

Item	Stima
Medici attenti alle problematiche	-0,66
Personale medico garbato e disponibile	-0,55
Informazioni dei medici sulla malattia	-0,44
Buone relazioni umane nel reparto	-0,42
...	...
Spazi per soggiorno e attesa	0,66
Servizi igienici funzionanti e puliti	0,73
Vitto adeguato e tempi rispettati	0,85

Si può quindi osservare come gli aspetti, che abbiano contribuito a una maggiore soddisfazione, sono quelli strettamente legati all’ambito medico e ai rapporti umani, mentre gli aspetti relativi alla struttura risultano essere quelli con valutazione peggiore. Lo stesso tipo di risultato è stato calcolato per gli individui; in questo caso però non si riportano i valori in tabella, per ovvi motivi, ma è importante considerare che i punteggi conseguiti da ogni singolo soggetto sono stati utilizzati come misura di customer satisfaction complessiva per i singoli pazienti. A questo punto era necessario valutare come la CS potesse dipendere da alcune variabili esplicative. L’utilizzo però di un modello di regressione semplice non è parso adeguato in questa circostanza, questo perché i dati a disposizione sono in realtà organizzabili su una struttura gerarchica in cui al primo livello ci sono i pazienti mentre al secondo compaiono i reparti degli ospedali. Su questo livello è stata individuata un’ulteriore variabile esplicativa e cioè gli anni di esperienza dei diversi primari.

Prima di procedere all’analisi, è stata effettuata la standardizzazione delle variabili per meglio valutare l’impatto delle esplicative sulla dipendente. Dopodiché si è verificata la presenza di un effetto di gruppo nell’analisi e, dopo averla accertata, sono stati costruiti i modelli Random Intercept e Random Slopes. Per il secondo modello non ci sono stati risultati significativi mentre il Random Intercept dato dalla formula (11) è quello che ha portato ai risultati migliori.

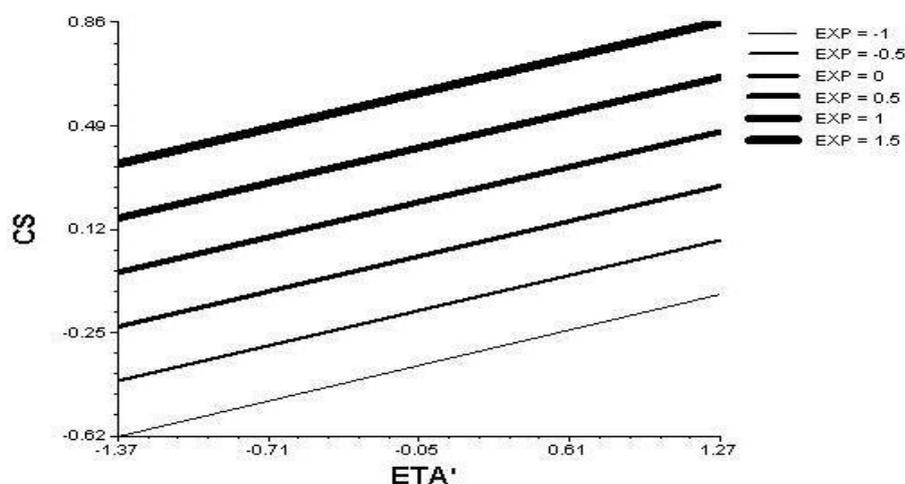
$$CS_{ng} = \alpha_{0g} + \alpha_1(Et\grave{a}_{ng}) + \varepsilon_{ng} \quad (11)$$

L’intercetta della precedente espressione è esprimibile come $\alpha_{0g} = \gamma_{00} + \gamma_{01}(Esp_g) + \varphi_{0g}$, la variabile $Et\grave{a}_{ng}$ è l’età dei pazienti e Esp_g sono gli anni di esperienza dei primari dei g reparti presi in esame. La stima dei parametri ha portato ai risultati che vengono presentati nella Tabella 2.

Tabella 2 – Stima dei parametri del modello Random Intercept.

Parametro	Coefficiente	St. Error	T- ratio	P-value
γ_{00}	0,029342	0,052140	0,563	0,584
γ_{01}	0,391126	0,054041	7,238	0,000
α_1	0,191025	0,055735	3,427	0,001

Come si può facilmente notare la variabile età è caratterizzata da un coefficiente positivo e di conseguenza i pazienti più anziani risultano essere quelli maggiormente soddisfatti. Allo stesso tempo si può notare che un impatto positivo sulla soddisfazione degli individui è dovuto anche al reparto in cui si trovano ricoverati. Maggiore è infatti l'esperienza dei primari più alto sarà il valore della CS. Nella Figura 1 è possibile notare come la CS aumenti all'aumentare dell'età dei pazienti e degli anni di esperienza dei primari.

Figura 1 – Customer Satisfaction in funzione dell'età dei pazienti e dell'esperienza dei primari (valori standardizzati).

5. Conclusioni

Nel presente lavoro è stata effettuata un'analisi di CS grazie all'utilizzo di due strumenti molto diffusi. L'analisi di Rasch ci ha consentito di ottenere punteggi di soddisfazione relativi sia ad ogni singolo paziente, sia ai diversi items considerati. I punteggi ottenuti dagli individui in questa prima fase, sono stati poi utilizzati come base di partenza per effettuare un'analisi multilevel che ha permesso di evidenziare

quali fossero le variabili maggiormente influenti sulla soddisfazione degli utenti. L'analisi realizzata potrebbe essere estesa a più livelli o si potrebbe pensare a una cross section analysis, in cui, per esempio, i pazienti possano essere classificati secondo i diversi reparti o i diversi edifici dell'ospedale, perché non sempre queste due entità coincidono. Un'ulteriore indagine potrebbe essere condotta invece provando ad inserire gli effetti di gruppo, quindi nel nostro caso dei reparti, direttamente nell'analisi di Rasch, prima ancora di effettuare l'analisi multilivello.

Riferimenti bibliografici

- Hox J.J. 2002, *Multilevel Analysis. Techniques and Applications*, Lawrence Erlbaum Associates.
- Snijders T. A. B. e Bosker R. J. 1999, *Multilevel Analysis. An introduction to basic and advanced multilevel modeling*. SAGE Publications.
- Wright B.D. e Masters G.N. 1982, *Rating Scale Analysis. Rasch Measurement*, Mesa, Chicago.

SUMMARY

Rasch Analysis and multilevel models for the evaluation of the Customer Satisfaction

The evaluation of attitudes, capabilities and of individual satisfaction is one of the most important problem of experimental sciences. This qualities in fact, are not observable in direct way, but they are expressed with polycotomous measure scale. To effectuate the evaluation it's necessary to substitute to the qualitative modalities, some scores. These measures can be determined in different way, but problems of quantification or relative to the conditions in which the survey is conduced can arise. For solving this problems one solution can be the Rasch Analysis. We used this technique to quantify the response to an hospital survey about the Customer Satisfaction. In a second step we try to verify if the patient satisfaction can be influenced by socio-economic factors and for this reason we use a Multilevel Model.

Antonio LUCADAMO, Assegnista di Ricerca di Statistica, Dipartimento SEMEQ - Università del Piemonte Orientale "A. Avogadro", alucadam@unina.it
Giovanni PORTOSO, Associato di Statistica, Dipartimento SEMEQ - Università del Piemonte Orientale "A. Avogadro", portoso@eco.unipmn.it

LOW BIRTH WEIGHT AND PARENTAL RESOURCES

Silvia Meggiolaro

1. Introduction

Low birth weight (LBW) is the main cause of perinatal and neonatal mortality and an established determinant of infant mortality. In addition, low birth weight is associated with health problems in childhood (Strauss, 2000) as well as in adulthood (Huxley et al., 2000). Given the negative outcomes associated with LBW, research has examined factors associated with or that predict LBW.

A baby's low weight at birth, commonly defined as less than 2,500 grams, is either the result of preterm birth (before 37 weeks of gestation) or of restricted foetal (intrauterine) growth. In some cases the two processes occur together. As the determinants as well as the consequences of the two causal components of LBW are different (Kramer et al., 2000), it is preferable to consider prematurity and intrauterine growth retardation (IUGR) separately.

Several physiological and behavioural determinants have been associated with the two components of LBW. An unifying framework in the research findings is the large socio-economic disparities: in particular, significant associations have been found between low socio-economic status and LBW, even in those countries where access to prenatal care is universal (Kramer et al., 2000).

The present study aims at examining whether and how socio-economic disparities are present in the two components of LBW with reference to Italy, using data from the *Italian Birth Sample Survey*, conducted in 2005 by the National Statistical Institute (Istat). In this way, the results by a previous study conducted with reference only to the preterm newborns in a tertiary care hospital in Italy (Nobile et al., 2007) can be verified and extended.

The paper is organized as follows: section 2 discusses the socio-economic factors influencing birth weight; section 3 describes data, methods and variables used; section 4 provides the results of a multivariate regression model. The last section (5) contains a discussion of the results and some concluding remarks.

2. Background and hypotheses

Socio-economic disparities in LBW regard both prematurity and IUGR. In particular, preterm births have been reported to be more common among the socially disadvantaged in various countries (Peacock et al., 1995, Koupilova et al., 1998, Olsen et al., 1995). A study referred to Sweden found, instead, small socio-economic differences (Koupilova et al., 1998). Similarly, mixed findings have been showed as regards IUGR (Raum et al., 2001, Tuntiseranee et al., 1999). The contradictions may be, at least partly, explained by the different measures of socio-economic status. Socio-economic status is, indeed, a complex construct; it is usually measured by income, occupation and educational attainment.

People's income is often used to determine material conditions; some studies showed correlations between family income and birth weight (Arntzen et al., 1994). However, estimation of actual family income is problematic and, as a consequence, other measure are often used: usually maternal and paternal occupation. In fact, occupational status represents a social stratification measure, but it is also an indicator of what individuals are exposed to. As a consequence, despite the positive effects of employment on income, maternal strenuous occupations (for example, in those occupations which involve prolonged periods of standing) may affect a foetus in a negative way. Instead, there is consensus that the occupation of the baby's father may be taken as measure of income.

Educational attainment is probably the most often used indicator of socio-economic status in studies on pregnancy outcomes. This may be because, besides being easily available in most surveys, it is closely related both to income and occupation (Pallas 2000, Kerckhoff et al., 2001). The effect of parental education on the duration of gestation and on the rate of foetal growth is confirmed by a large body of literature (see, for example Rodriguez et al., 1995, and Raum et al., 2001).

Lastly, another established determinant of pregnancy outcome, which is strongly connected with socio-economic factors, is marital status. Being an unmarried mother has been identified in a large number of studies as a risk factor for both preterm births and IUGR (Pickering and Deeks, 1991; Rodriguez et al., 1995).

3. Data and methods

Data come from the *Italian Birth Sample Survey*, conducted in Italy in 2005 (Istat, 2006). In this paper we consider information on the birth weight of the

babies born about 18-21 months before the interview and on the duration of pregnancy in weeks as reported by their mothers.

LBW (defined as less than 2,500 grams) infants are separated according to whether they are of less or more than 37 weeks' gestation. A LBW newborn is defined preterm if it is of less than 37 weeks' gestation, whereas it is growth-retarded if it is "term". Normal births are 2,500 grams or above regardless of gestational age.

The sample is composed of 47,118 singleton children: it contains 1,327 preterm births and 1,228 growth-retarded babies.

The different pregnancy outcomes (preterm births, IUGR and normal births) are described through a multinomial logistic regression (Agresti, 1996), in which normal births serve as the reference category.

Several independent variables, with particular attentions to those describing parental socio-economic status, were used to assess the factors influencing the (polytomous) dependent variable. Table 1 lists the covariates and their distributions on the sample by infant outcomes. The variables may be grouped into three categories: mother's physiological characteristics, her socio-economic resources and father's economic characteristics. In addition, the sex of the newborn is considered as control.

Mother's physiological characteristics are represented by age at delivery, parity and prior miscarriages. Table 1 shows that women with preterm births had an older age in comparison with those who had newborns with normal weight and with IUGR. In addition, both preterm and growth-retarded births are more common among first parity. The percentages of prior miscarriages among preterm births are higher in comparison with both those observed among normal newborns and among IUGR.

Mother's socio-economic resources are her education, her employment status at beginning of the pregnancy, and her marital status at delivery. Table 1 shows that both preterm and growth-retarded births are more common among mothers with low educational level. Higher percentages of babies with IUGR are observed among unmarried mothers or among those who are not employed at the beginning of the pregnancy.

Father's economic characteristics are his education and his employment status (they are measured at the interview, and, thus, a causal effect cannot be observed). Table 1 suggests that both preterm births and IUGR are more common among unemployed fathers or with low educational level.

Table 1 – Descriptive statistics of study sample by birth outcome.

Characteristics	Total births	Births > 2,500 gr.	Births < 2,500 gr.	
			Preterm	IUGR
<i>Woman's age at birth</i>				
Under 25	14.1	14.1	12.7	14.4
25-29	27.4	27.5	25.2	28.0
30-34	31.4	31.6	28.0	28.6
35-39	18.7	18.5	21.1	20.9
40 or over	8.4	8.3	13.0	8.1
<i>Parity</i>				
First birth	49.9	49.6	55.2	56.5
Parity 2 +	50.1	50.4	44.8	43.5
<i>Prior miscarriages</i>				
Yes	18.4	18.3	22.5	17.3
No	81.6	81.7	77.5	82.7
<i>Sex of the newborn</i>				
Male	51.2	51.6	49.8	37.2
Female	48.8	48.4	50.2	62.8
<i>Mother's education</i>				
High (university)	17.7	17.9	17.0	14.7
Middle (high school)	44.7	44.8	40.9	43.9
Low (junior school or less)	37.6	37.3	42.1	41.4
<i>Mother's employment status at pregnancy</i>				
Employed	64.3	64.3	65.1	62.7
Not employed	35.7	35.7	34.9	37.3
<i>Marital status of the mother at the birth</i>				
Married	90.1	90.2	89.9	87.0
Unmarried (‡)	9.9	9.8	10.1	13.0
<i>Father's education at the interview</i>				
High (university)	14.3	14.3	13.3	12.7
Middle (high school)	38.5	38.7	34.7	36.7
Low (junior school or less)	47.2	47.0	52.0	50.6
<i>Father's employment status at the interview</i>				
Employed	96.1	96.2	94.3	94.4
Not employed	3.9	3.8	5.7	5.6
<i>N° of cases (Total = 100):</i>	<i>47,118</i>	<i>44,563</i>	<i>1,327</i>	<i>1,228</i>

(‡) Mothers who had their child within a cohabitation. The 1.9% of the unmarried mothers who had their child outside a cohabitation are not considered in this paper.

4. Results

Table 2 lists the results of the multinomial logistic model.

Table 2 – *Coefficients of determinants of preterm birth and IUGR according to a multinomial regression model.*

	Preterm/ Normal Births	IUGR/ Normal Births
<i>Intercept</i>	-3.37***	-3.19***
<i>Age at delivery (ref: under 25)</i>		
25-29	-0.20***	-0.05
30-34	-0.13**	-0.05
35-39	0.18***	0.23***
40 or over	0.49***	0.10
<i>Parity (ref: others)</i>		
First parity	0.44***	0.37***
<i>Prior miscarriages (ref: no)</i>		
Yes	0.20***	-0.06
<i>Newborn sex (ref: female)</i>		
Male	-0.07	-0.59***
<i>Mother's educational level (ref: low)</i>		
High	-0.06	-0.17***
Middle	-0.06	0.03
<i>Mother's employment status at beginning of pregnancy (ref: employed)</i>		
Not employed	-0.01	-0.09
<i>Mother's marital status at delivery (ref: married)</i>		
Unmarried	-0.07	0.24***
<i>Father's educational level (at interview) (ref: low)</i>		
High	-0.08	-0.03
Middle	-0.06	-0.03
<i>Father's employment status (at interview) (ref: employed)</i>		
Not employed	0.40***	0.36***

Maternal age has a strong impact in the preterm/normal birth contrast even controlling for socio-economic and reproductive factors. In particular, an association between young maternal age (under 25) and preterm birth is observed;

the most strong effect is found for older mothers (aged 35-39 and, particularly, 40 or over). As regards the other component of LBW, mothers with the highest risk of IUGR births are those aged 35-39. The variables describing previous birth history are highly significant in determining preterm births. Newborn sex influences only IUGR component: male births have less risk to be IUGR.

Mother's education and her marital status have effects (in the expected direction) in determining births with IUGR. Effects on preterm births are not significant. It is somewhat surprising that employment status has no net effect on either adverse outcome.

As regards the father's economic resources, his employment status is highly significant both in the preterm/normal birth contrast and in the IUGR/normal birth contrast, whereas his education has no effect on either adverse outcome.

5. Discussion and conclusions

Given the strong influence of LBW on the risk of infant mortality and other severe health problems during childhood and adulthood, it is worthwhile examining factors associated with LBW. In fact, studying LBW independently of gestational age may be unsatisfactory. In this paper, preterm birth and intrauterine growth retardation are considered separately. The focus is whether and how socio-economic disparities are present in the two components of LBW in Italy.

Descriptive analyses show that both preterm births and IUGR are more common among mothers with low educational level. In addition, higher percentages of babies with IUGR are observed among unmarried mothers and women who are not employed at the beginning of the pregnancy. Similarly, both components of LBW are more common among unemployed fathers or among those with low educational level.

In fact, these results must be considered with cautions and examined through a multivariate approach. A multinomial regression analysis was used to study the net effects of socio-economic factors. Results show that disparities in adverse outcome by maternal socio-economic status are confirmed only for IUGR. However, also prematurity component presents some socio-economic inequalities according to father's employment status.

These results may inform policies for reducing inequalities in LBW (in particular, in the IUGR components): in a context of universal access to health care, interventions providing emotional support should be offered in addition to usual care. Interventions should be focused on reducing psychological stress for women in pregnancy, in particular, for those who are not married and who have

low educational levels. Support benefits expectant mothers in terms of better psychological and physical well-being, but support may also motivate the mother to engage in positive health behaviors and to make lifestyle changes to improve her physical health.

Future researches need additional data: besides more in depth information on socio-economic status (income measure and type of occupation), data on health variables, on psychosocial factors (such as levels of stress, social support, and depression) or on behavioural aspects (such as smoking habits) would allow an in depths analysis of the mediating factors through which socio-economic status acts.

References

- Agresti A. 1996. *An Introduction to Categorical Data Analysis*. John Wiley and Sons, Inc.
- Arntzen A., Samuelsen S.O., Magnus P., Bakketeig L.S. 1994. *Birth weight related to social indicators in Norway*, The European Journal of Public Health, 4(2), 92-97.
- Huxley R.R., Shiell A.W., Law C.M. 2000. *The role of size at birth and postnatal catch-up growth in determining systolic blood pressure: a systematic review of the literature*, Journal of Hypertension, 18, 815-831.
- Kerckhoff A.C., Raudenbush S.W., Glennie E. 2001. *Education, cognitive skill, and labor force outcomes*, Sociology of Education, 74, 1-24.
- Koupilová I., Vager D., Leon D.A., Pikhart H., Prikazsky V., Holcik J. 1998. *Social variation in size at birth and preterm delivery in the Czech Republic and Sweden. 1989-1991*, Paediatric and Perinatal Epidemiology, 12, 7-24.
- Kramer M.S., Séguin L., Lydon J., Goulet L. 2000. *Socio-economic disparities in pregnancy outcome: Why do the poor fare so poorly?*, Paediatric and Perinatal Epidemiology, 14, 194-210.
- Istat 2006. *L'indagine campionaria sulle nascite: obiettivi, metodologia e organizzazione*, Collana "Metodi e Norme", n. 28, Istat, Roma.
- Nobile C.G.A., Raffaele G., Altomare C., Pavia M. 2007. *Influence of maternal and social factors as predictors of low birth weight in Italy*, BMC Public Health, 7, 192-201.
- Olsen P., Laara E., Rantakallio P., Jarvelin M.R., Sarpola A., Hartikainen A.L. 1995. *Epidemiology of preterm delivery in two birth cohorts with an interval of 20 years*, American Journal of Epidemiology, 142(11), 1184-93.
- Pallas A.M. 2000. *The effects of schooling on individual lives*. In: Hallinan M.T. (ed.), Handbook of Sociology and Education, pp. 499-525, Kluwer/Plenum, New York.
- Peacock J.L., Bland J.M., Anderson H.R. 1995. *Preterm delivery: effects of socioeconomic factors, psychological stress, smoking, alcohol, and caffeine*, British Medical Journal, 311, 531-536.
- Pickering R.M., Deeks J.J. 1991. *Risks of Delivery during the 20th to the 36th Week of Gestation*, International Journal of Epidemiology, 20, 456-466.

- Raum E., Arabin B., Schlauda M., Waltera U., Schwartz F.W. 2001. *The impact of maternal education on intrauterine growth: a comparison of former West and East Germany*, International Journal of Epidemiology, 30, 81-87.
- Rodriguez C., Regidor E., Gutierrez-Fisac J.L. 1995. *Low birth weight in Spain associated with sociodemographic factors*, Journal of Epidemiology and Community Health, 49, 38-42.
- Strauss R.S. 2000. *Adult functional outcome of those born small for gestational age: twenty-six-year follow-up of the 1970 British birth cohort*, JAMA, the Journal of the American Medical Association, 283, 625-632.
- Tuntiseranee P., Olsen J., Chongsuvivatwong G.V., Limbutara S. 1999. *Socioeconomic and work related determinants of pregnancy outcome in southern Thailand*, Journal of Epidemiology and Community Health, 53, 624-629.

SUMMARY

This paper studies the disparities in low birth weight in Italy according to socio-economic status, using data updated to 2005. In particular, the two components of low birth weight, prematurity and intrauterine growth retardation, are considered separately. Results of multinomial models show an effect of maternal socio-economic characteristics only for growth retarded births: women who are unmarried and who have low educational level have the highest risks to have restricted foetal growth, even controlling for physiological aspects. The prematurity component presents, instead, some inequalities according to father's employment status. This paternal factor influences also intrauterine growth retardation.

SOGLIE INFERIORI E SUPERIORI DI TALUNI INDICI DI CONNESSIONE

Giovanni Portoso, Antonio Lucadamo

1. Introduzione

Oggetto della presente nota è l'individuazione di soglie di massimizzazione di due noti indici, atti a misurare il legame associativo tra due variabili categoriche non ordinali, congiuntamente distribuite in tavole di contingenza quadrate e rettangolari.

In particolare si sono analizzati i massimi legati al numeratore dell'indice del Mortara, fondato, come è noto sugli scarti, in valore assoluto, tra le frequenze empiriche e teoriche, ancorate quest'ultime a condizioni d'indipendenza stocastica ed al *Chi* quadrato di K. Pearson, basato invece sui quadrati degli scarti.

La ricerca dei massimi, vincolati a condizioni variamente configurate, è oltremodo importante per definire i fattori di normalizzazione degli indici in questione.

Si sono riesaminate dette condizioni, individuando estremi superiori, che, assunti come fattori di normalizzazione, contribuiscono ad aumentare il valore degli indici e quindi l'intensità del legame associativo tra le due variabili categoriche considerate.

Più che interessante anche ai fini applicativi, ci pare la rideterminazione del normalizzante del numeratore dell'indice del Mortara, rettificato attraverso la considerazione della varianza di uno dei vettori dei totali marginali.

Si sono precisati anche estremi inferiori, abbinati a condizioni di dipendenza minimale in modo da valutare il distacco da quelli superiori e definire pertanto il quantum di oscillazione dei massimi ancorati alla tavola di contingenza osservata.

2. Massimo del numeratore dell'indice del Mortara : Tavole quadrate

L'indice di contingenza media assoluto del Mortara per tavole quadrate $s \times s$, ${}_q I_1$, come è noto, è espresso dalla formula :

$${}_q I_1 = \frac{\sum_i \sum_h |c_{ih}|}{(2N)} \quad i = 1, 2, \dots, s; \quad h = 1, 2, \dots, s \quad (1)$$

in cui $c_{ih} = n_{ih} - n_{ih}^*$ e $n_{ih}^* = n_{i0} n_{0h} / N$

N indica la dimensione del collettivo, i e h sono rispettivamente gli indici di riga e di colonna, c_{ih} le contingenze, n_{ih} le frequenze osservate riportate nella cella del rigo i -esimo e della colonna h -esima ed n_{ih}^* le frequenze teoriche, ottenute dal prodotto dei totali marginali, $n_{i0} n_{0h}$, rapportato al totale generale N e desunte sulla base dell'assunzione della condizione d'indipendenza stocastica.

Va sottolineato che il denominatore, $2N$, costituisce un maggiorante del numeratore, del tutto virtuale, che non può essere mai raggiunto; il che si traduce in un livellamento verso il basso dell'indice ${}_qI_1$.

Esso può essere corretto con il fattore di riduzione : $(1-1/s)$, che emerge allorché si consideri l'ipotesi di una distribuzione massimante della tavola di contingenza legata alla massima eterogeneità delle variabili categoriche, in cui cioè sussistono solo gli n_{ii} sulla diagonale principale ed essi sono pari a $n_{i0} = n_{0i} = N/s$. Si dimostra facilmente infatti che, nel caso di perfetta uguaglianza tra le frequenze diagonalizzate ed i totali marginali di riga e di colonna e considerando il numeratore di ${}_qI_1$, si ottiene : $2N(1-1/s)$.

Detta condizione, che, in realtà, è duplice in quanto basata sia sulla uguaglianza dei totali omologhi dei vettori marginali che sulla loro invarianza pari a N/s , è difficilmente riscontrabile in pratica per cui anche il normalizzante, rettificato come sopra, si configura come un estremante di difficile realizzazione.

Si è indirizzata pertanto la ricerca verso la determinazione di un fattore di normalizzazione, che prescindesse dall'uguaglianza dei totali marginali.

Quindi imponendo la condizione per cui $n_{ii} = n_{i0} = n_{0i} \gg N/s$, si è trovato che l'estremante precedentemente considerato, $2N(1-1/s)$, viene ulteriormente diminuito di un addendo negativo pari a $2s \sigma_M^2 / N$, in cui σ_M^2 sta ad indicare la varianza più alta tra i due vettori dei totali marginali.

Pertanto l'indice di contingenza media assoluto del Mortara, di cui alla (1), andrebbe corretto c. s. :

$${}_qI_3 = \sum_i \sum_h |c_{ih}| / [2N(1-1/s) - 2s \sigma_M^2 / N] \quad (2)$$

L'importanza della rettifica sta nel fatto che la normalizzazione dell'indice del Mortara viene a legarsi, sia pure in parte, alla variabilità più elevata di uno dei vettori dei totali marginali con il risultato di elevare il valore dell'indice relativizzato ed aumentare in tal modo la valutazione del legame associativo.

Il denominatore di ${}_qI_3$ rappresenta quindi una soglia superiore di facile calcolo, che viene ad incorporare una componente del tutto nuova quale è la σ_M^2 , che

introduce nella formula la considerazione di alcuni totali marginali della distribuzione empirica .

Rimarrebbe fuori dal calcolo l'altra componente cioè quella legata ai totali marginali, che hanno varianza più ridotta.

Essa può rientrare facendo ricorso ad una distribuzione massimante, che utilizzi entrambi i vettori dei totali marginali attraverso una procedura di massimizzazione, basata sulla diagonalizzazione o cograduazione (Salvemini 1939) delle frequenze d'incrocio, riportata in precedenti lavori (Portoso 2007/1, 2007/2, 2008/1) e che può essere utilizzata per il calcolo del denominatore dell'indice del Mortara. In questo modo si ancora il calcolo del normalizzante del numeratore dell'indice suddetto alla massima associazione compatibile con i vettori dei totali marginali.

Può essere interessante chiedersi quanto può abbassarsi il valore del normalizzante soprattutto al fine di quantificarne il campo di oscillazione ed indirettamente valutare anche la distorsione apportata all'indice in oggetto dall'uso dei normalizzanti presentati finora.

A tal uopo è necessario ricorrere ad un'ipotesi di dipendenza minimale tra due variabili categoriche, che si realizza allorquando sussiste la massima dissomiglianza tra i vettori dei totali marginali (Portoso 2008/2).

Si dimostra – scegliendo tra due schemi di calcolo – che la soglia inferiore del normalizzante dell'indice del Mortara per tavole quadrate, ${}_qS_i$, risulta essere :

$${}_qS_i = 4 (s-1) (1-1/s) \quad (3)$$

Essa è del tutto indipendente da N e dipende solo dal valore di s ; pertanto, a parità di condizioni, il range - qualunque denominatore dell'indice si consideri : quello di cui alla (1) o alla (2) - sarà tanto più esteso quanto più alto è N .

A scopo esemplificativo la Tab. 1 riporta una tavola di contingenza ipotetica in cui figurano nella riga marginale gli indici del Mortara calcolati con diversi normalizzanti e la varianza più alta, calcolata sui totali di colonna.

Tabella 1 – *Tavola di contingenza (valori ipotetici) 3x3 ed indici relativi.*

X Y	y₁	y₂	y₃	Totali
x₁	3	1	4	8
x₂	4	4	11	19
x₃	8	27	38	73
Totali	15	32	53	100
qI₁=0,075	qI₂=0,113	qI₃=0,177	qI₄=0,221	σ_M²=806,89

Nella Tab.1 ${}_qI_1$ sta ad indicare il valore dell'indice del Mortara calcolato secondo la (1); ${}_qI_2$ indica quello avente come normalizzante : $2 N (I-1/s)$; ${}_qI_3$ fornisce il valore dell'indice calcolato secondo la (2) ; ${}_qI_4$ riporta il valore dell'indice ottenuto, seguendo la procedura di massimizzazione, basata sulla diagonalizzazione o cograduazione delle frequenze di cella.

Con riferimento al rigo marginale di Tab. 1 , va evidenziata la progressione in senso positivo che subisce l'indice allorché si passa dal normalizzante, di cui alla (1) , alla procedura di massimizzazione delle frequenze di celle. Il valore dell'indice normalizzato, nel caso in esame, diventa quasi triplo, passando da 0,075 a 0,221.

Risulta quindi che l'ultima procedura esalta sensibilmente il valore dell'indice e, conseguentemente, la valutazione del legame associativo soprattutto perché considera entrambi i vettori dei totali marginali e rappresenta il percorso ottimale ai fini della normalizzazione. Però essa richiede la ricerca della distribuzione massimante, che non è sempre di facile ed immediata acquisizione.

Si può osservare però che anche la normalizzazione eseguita secondo la (2) , che considera solo uno dei vettori dei totali marginali - quello a più alta variabilità - eleva il valore dell'indice del Mortara e può rappresentare una comoda "scorciatoia" per una più opportuna misurazione dell'intensità del legame associativo.

3. Massimo del numeratore dell'indice del Mortara : Tavole rettangolari

Bonferroni (1940, pag. 295) ha dimostrato che, per tavole rettangolari, $s \times t$, il massimo valore che può acquisire il numeratore dell'indice del Mortara, è il seguente :

$${}_rS_2 = 2 N (I-1/v) \quad \text{in cui } v = \min (s,t) \quad (4)$$

Pertanto l'indice, ${}_rI_2$, si configura c.s. :

$${}_rI_2 = \sum_i \sum_h |c_{ih}| / [2 N (I-1/v)] \quad (5)$$

Anche il suddetto normalizzante – analogamente a quanto rilevato per le tavole quadrate - costituisce una limitazione superiore, difficilmente riscontrabile nella realtà e pertanto del tutto virtuale.

Si è trovato, attraverso simulazioni, che l'estremo superiore, ${}_rS_3$. risulta essere :

$${}_rS_3 = 2 N (I-1/u) - 2 u \sigma_M^2 / N \quad (6)$$

in cui u è il numero delle righe o delle colonne, associato alla varianza più alta : σ_M^2 .

Pertanto anche l'indice rI_2 è suscettibile di rettifica attraverso la considerazione della varianza più elevata dei totali marginali e l'assunzione della (6) come normalizzante in modo da ottenere :

$$rI_3 = \frac{\sum_i \sum_h |c_{ih}|}{[2N(1-1/u) - 2u\sigma_M^2/N]} \quad (7)$$

rI_3 contribuisce a modificare verso l'alto l'indice del Mortara e ad accentuare, in senso positivo, la misurazione e la valutazione del legame associativo tra le due mutabili.

Nel caso di minima dipendenza tra due variabili categoriche (Portoso, 2008/1) si dimostra che la soglia inferiore, rS_i , per $s > t$,

$$rS_i = 4(s-1)(1-1/t) \quad (8)$$

E' ovvio che, per $s < t$, basta scambiare s con t nella (8).

I rilievi e le considerazioni fatte per le tavole quadrate circa il range di oscillazione possono essere ricalcate per le rettangolari. Anche per quest'ultime va ribadito che il più idoneo denominatore di normalizzazione è quello che si ottiene, usando i criteri di diagonalizzazione e cograduazione delle frequenze d'incrocio in assonanza ai totali dei vettori marginali.

Il caso, sia pure ipotetico, riportato nella Tab. 2, può valere a scopo esemplificativo.

Si precisa che l'indice del Mortara, rI_2 , è stato calcolato usando come normalizzante : $2N(1-1/v)$; rI_3 utilizzando : $2N(1-1/u) - 2u\sigma_M^2/N$; rI_4 è stato ottenuto, generando la distribuzione massimante, seguendo le procedure suaccennate.

Tabella 2 – Tavola di contingenza (valori ipotetici) 2x4 ed indici relativi.

X \ Y	y₁	y₂	y₃	y₄	Totali
x₁	10	10	0	0	20
x₂	15	20	40	55	130
Totali	25	30	40	55	150
	rI₂=0,338	rI₃=0,731	rI₄=0,760		σ_M²=3.025

E' di tutta evidenza l'incremento che viene a subire l'indice allorché si modifica il normalizzante; dal valore di 0,338 si passa a 0,731 con un salto

d'intensità del legame associativo più che notevole mentre il miglioramento che si ottiene passando da ${}_rI_3$ a ${}_rI_4$ - almeno per questo caso - sembra essere di poco conto.

Anche per le tavole rettangolari risulta confermato che la normalizzazione più idonea, che tiene conto della variabilità espressa da entrambi i vettori dei totali marginali e pertanto più ancorata alla distribuzione empirica, è quella ottenuta attraverso la generazione della distribuzione massimante.

Comunque va sottolineato che, a seconda del tipo di normalizzazione usata, può cambiare notevolmente il valore dell'indice e quindi la valutazione della connessione tra le due variabili.

4. Massimo dell'indice *Chi* quadrato : Soglie e range di oscillazione

Per tavole quadrate, la soglia superiore, comunemente accettata per la normalizzazione del *Chi* quadro, ${}_qS_{\chi^2}$, si deduce dalla condizione : $n_{ii} = n_{i0} = n_{0i} >> N/s$; però anch'essa è difficilmente riscontrabile nella realtà :

$${}_qS_{\chi^2} = N (s-1) \quad (9)$$

La soglia inferiore, ${}_qS_{\chi^2}$ è stata determinata, come per l'indice del Mortara, ipotizzando la massima dissomiglianza tra i vettori dei totali marginali :

$${}_qS_{\chi^2} = N (s-1)^2 / (N-s+1) \quad (10)$$

Il range di oscillazione, ${}_qW_{\chi^2}$, è molto ampio e copre quasi l'intera soglia superiore, essendo l'ultimo rapporto di poco inferiore a 1 :

$${}_qW_{\chi^2} = N (s-1) - N (s-1)^2 / (N-s+1) = N (s-1) (N-2s+2) / (N-s+1) \quad (11)$$

Quanto alle tavole rettangolari, occorre solo riconsiderare le suddette relazioni alla luce della diversità di s da t .

La soglia superiore, ${}_rS_{\chi^2}$ e la soglia inferiore ${}_rS_{\chi^2}$, per $s > t$, sono rispettivamente :

$${}_rS_{\chi^2} = N [\min(s,t)-1] \quad (12)$$

$${}_rS_{\chi^2} = N (s-1) (t-1) / (N-s+1) \quad (13)$$

Il range di oscillazione, ${}_rW_{\chi^2}$, risulta essere, per $s > t$:

$${}_rW_\chi = N(t-1) - N(s-1) \frac{(t-1)}{(N-s+1)} = N(t-1)(N-2s+2)/(N-s+1) \quad (14)$$

E' ovvio che, per $s < t$, basta scambiare t con s nelle formule suddette.

Nella Tab. 3 con χ_n^2 si è indicato il *Chi* quadro normalizzato usando la (12) ed invece con ${}_o\chi_n^2$ quello normalizzato, ottenuto ricorrendo alla distribuzione massimante, generata utilizzando la procedura di diagonalizzazione o cograduazione, che ha prodotto un chi quadro massimo, ${}_oM(\chi^2)$, pari a 126; i risultati evidenziano un aumento dell'indice normalizzato da 0,213 a 0,338, pari a circa il 59%.

Tabella 3 – *Tavola di contingenza (valori ipotetici) 2x4 ed indici relativi.*

X \ Y	y ₁	y ₂	y ₃	y ₄	Totali
x ₁	10	10	0	0	20
x ₂	20	30	50	80	180
Totali	30	40	50	80	200
	$\chi^2=42,6$	$\chi_n^2=0,213$		${}_oM(\chi^2)=126$	${}_o\chi_n^2=0,338$

5. Considerazioni conclusive

Ci pare di poter delineare, con riguardo all'indice del Mortara, che sono disponibili almeno tre livelli di normalizzazione: il primo, legato a condizioni teoriche molto restrittive, difficilmente ravvisabili nella pratica; il secondo che tiene conto anche della varianza più alta dei totali dei vettori marginali; il terzo che considera entrambi i vettori della tavola di contingenza nell'ambito della procedura di generazione della distribuzione massimante.

Per il *Chi* quadro invece la scelta si pone soltanto tra il primo ed il terzo livello.

I risultati comunque consentono di rilevare un innalzamento degli indici analizzati man mano che si passa dalla prima alla terza opzione e quindi ad una reinterpretazione verso l'alto del legame associativo.

I range di oscillazione tra soglie superiori ed inferiori sono tutti alquanto ampi e lasciano intuire che quanto più le tavole di contingenza reali si avvicinano a condizioni di dipendenza minimali, cioè a condizioni di massima dissomiglianza tra i vettori dei totali marginali, tanto più marcate risultano le differenze tra gli indici abbinati a tipi diversi di normalizzanti e tanto più falsata può risultare la valutazione del legame associativo tra le due variabili categoriche esaminate.

Riferimenti bibliografici

- Bonferroni C.E. (1940) *Elementi di statistica generale*, Litografia Gilli, Torino.
- Portoso G. (2007/1) Indici e Distribuzioni Massimanti in Tavole di Contingenza quadrate con marginali prefissati, *Rivista Italiana di Economia, Demografia e Statistica*, Vol. LXI, n. 3-4, 391-400.
- Portoso G. (2007/2) Una proposta di correzione dell'Indice di Cramer per Tavole di Contingenza quadrate, *Quaderno n. 25 del Dipartimento Semeq*, Novara, 1-15.
- Portoso G. (2008/1) Rilievi ed Osservazioni sulla massimizzazione del *Chi* quadrato, *Rivista Italiana di Economia, Demografia e Statistica*, Vol. LXII, n. 2-4, 417-425.
- Portoso G. (2008/2) Sul range di oscillazione del massimo *Chi* quadrato, *Quaderno n. 25 del Dipartimento Semeq*, Novara, 1-17.
- Salvemini T. (1939) Sugli indici di omofilia, *Atti della I Riunione della Società Italiana di Statistica*, Pisa.

SUMMARY

Lower and upper thresholds of some indices of association

We analyzed the upper thresholds, which are used as normalization factors for both the Mortara index and the Chi square Pearson. Since they are linked to restrictive conditions, such higher thresholds are hardly verified in practice.

For the Mortara index, we identified thresholds, which consider also the highest variability, expressed by one of the vectors of marginal totals of the observed contingency table. We verified that the most qualified normalization is the one obtained by including both of the marginal vectors in the computation of diagonalization or cograduation, which have been examined in previous works.

Considering these normalization factors, we also determined the lower thresholds linked to minimal conditions of dependence and consequently the range of oscillation between them and the upper thresholds.

For the sake of example, we have considered some hypothetical cases, which indicate the differences between the values of the indices, obtained by considering different types of normalization factors.

It is obvious that the evaluation of the associative link between the two examined categorical variables changes depending on the adopted normalization factors.

Giovanni PORTOSO, Associato di Statistica, Dipartimento SEMEQ – Università del Piemonte Orientale “A. Avogadro” portoso@eco.unipmn.it

Antonio LUCADAMO, Assegnista di Ricerca di Statistica, Dipartimento SEMEQ – Università del Piemonte Orientale “A. Avogadro” alucadam@unina.it

IMMIGRATI STRANIERI IN PIEMONTE E FECONDITÀ UNA PRIMA ANALISI DESCRITTIVA*

Mauro Reginato, Tiziana Barugola

1. Premessa, dati, metodi

Da più di trent'anni la fecondità dell'Europa è sotto la soglia di rimpiazzo della popolazione, ed è generalmente riconosciuto che rimarrà debole ancora per lungo tempo, continuando ad alimentare l'invecchiamento demografico.

L'Italia rispecchia perfettamente la situazione europea. Fino alla metà degli anni Settanta, il tasso di fecondità totale (TFT) garantiva il ricambio generazionale, con valori che tenevano la soglia critica di 2,12 figli per donna; il 1975 è stato l'ultimo anno con un tasso adeguato al mantenimento della stazionarietà, ma nel breve periodo di cinque anni l'Italia, paese mediterraneo, è scesa al livello dei principali paesi nordici (1,6) ed ha proseguito in costante decrescita fino al 1995, anno dopo il quale è iniziata la timida ripresa che ha portato agli attuali valori di circa 1,3 (o poco più). Una lenta risalita coincidente con tre fattori concomitanti riguardanti la popolazione straniera: la crescita sempre più rapida degli ingressi in Italia, la serie di regolarizzazioni della presenza irregolare, il forte incremento della proporzione di nati provenienti da madri straniere¹.

Analogamente, il Piemonte è perfettamente inserito nel quadro italiano rispecchiando i medesimi andamenti nell'invecchiamento, nella fecondità, nella presenza straniera e nelle nascite da essa prodotte. Le poco meno di 3.000 nascite di inizio decennio dovute alle madri straniere (su un totale di circa 36.000) sono attualmente raddoppiate (quasi 6.200) ed incidono per il 16 per cento sul totale delle nascite regionali (38.565 nel 2007). In questo ambito territoriale tutte le nascite avvenute nel corso dell'anno di calendario (indipendentemente dal luogo di residenza della madre) sono documentate tramite i certificati di assistenza al parto (Cedap) che risultano, pertanto, una fonte primaria, privilegiata ed esaustiva per l'analisi degli eventi nascita².

* Il lavoro è risultato della collaborazione dei due autori. Sono da attribuirsi a M. Reginato i paragrafi 1 e 2; a T. Barugola il paragrafo 3".

¹ In Italia, dalla metà degli anni '90 le nascite dovute a residenti stranieri sono triplicate, da meno del 4 per cento sono arrivate all'attuale quasi 12 per cento.

² Il certificato di assistenza al parto, istituito con il decreto ministeriale n. 349 del 16 luglio 2001, è strutturato in sezioni che riportano le informazioni epidemiologiche ed anche altre informazioni utili per la ricerca demografica e sociale.

L'analisi delle informazioni riportate nei Cedap è alla base del lavoro che segue e che verte sulla fecondità delle immigrate straniere in Piemonte. Con esso si vuole dare una visione d'insieme del comportamento riproduttivo, tramite una analisi di tipo statico che non escluderà, tuttavia, anche l'osservazione dei principali caratteri socio-demografici, così come appaiono nelle schede di registrazione delle nascite.

Definita la fonte, si è provveduto a depurarla da quei casi che, per residenza e per età, non rispondevano alla finalità della ricerca: la fecondità delle immigrate straniere residenti in Piemonte. Tale operazione ha ridotto i Cedap effettivamente utilizzabili da 36.908 a 35.737 (96,8 per cento, tav.1)³ e su questi sono state condotte le prime analisi che hanno tracciato il profilo sociodemografico delle madri e dei padri, nonché altre più mirate verso particolari aspetti delle nascite (ad esempio, genitori molto giovani, molto anziani, con elevato scarto di età, di nazionalità diverse, ecc.)⁴.

Tavola 1 – Schema generale dei Cedap in Piemonte nel 2006 e principali nazionalità.

Madri	Totale	Italiane	Straniere	Romania	Marocco	Albania
Totale generale	36908	29445	7463	1986	1657	898
Schede escluse	1171	815	356	136	35	41
Schede valide	35737	28630	7107	1850	1622	857
(% sul totale generale)	(96,8)	(97,2)	(95,2)	(93,2)	(97,9)	(95,4)

Successivamente si è passati all'aspetto della fecondità, la cui analisi è stata condotta con il tradizionale metodo dei tassi specifici. Va ricordato che questi possono essere intesi come una media ponderata dei tassi delle italiane e delle straniere, i quali assumono come pesi la numerosità media delle donne nelle singole età. In tal modo, tenendo conto della scomposizione contemporanea del numero dei nati e del numero delle donne a seconda della combinazione nazionalità (italiana, straniera) ed età, si può pervenire alla determinazione del contributo portato dalle donne italiane e dalle donne straniere al tasso di fecondità complessivo⁵.

³ Sono state escluse dall'analisi le schede che non riportavano alcuna indicazione di residenza, quelle riferite a donne residenti fuori regione, quelle con età non compresa tra gli estremi comunemente ritenuti fertili (15-49 anni). Per coerenza, analoghi controlli sono stati effettuati sui dati dei padri.

⁴ In questa comunicazione non sono riportati i risultati dettagliati delle analisi effettuate, in quanto non confacenti con l'oggetto della comunicazione stessa.

⁵ In forma molto sintetica si è proceduto nel modo seguente: per l'anno generico di calendario "t" definiamo con "x" l'età delle madri, con ${}_tN_x$, ${}_tN_x^i$ e ${}_tN_x^s$ i nati a ciascuna età e quelli da madre italiana e da madre straniera nella stessa età, con ${}_t\bar{F}_x$, ${}_t\bar{F}_x^i$, ${}_t\bar{F}_x^s$ il numero medio di donne per età e con ${}_t f_x$, ${}_t f_x^i$, ${}_t f_x^s$ il tasso specifico di fecondità per età. Con alcuni

2. Fecondità: i tassi delle italiane e delle immigrate straniere

Nel 2006, gli stranieri residenti in Piemonte erano circa 252.000 (1'8 per cento del totale italiano, Istat), equilibrati nel rapporto tra i sessi (femmine 50,9 per cento) e con una distribuzione per nazionalità molto variegata. Quasi altrettanto numerose erano le nazionalità delle donne divenute madri nel corso dell'anno⁶, l'apporto delle quali alle nascite ha raggiunto il 20 per cento del totale della Regione.

La prima caratteristica che emerge dall'analisi dei Cedap riguarda l'età: per le italiane oltre il 50 per cento dei parti avviene tra i 25 ed i 34 anni, tra le straniere si osserva la stessa proporzione tra i 20 ed i 29 anni. La maternità adolescenziale (madri con età tra i 15 ed i 19 anni) è quattro volte più alta nelle straniere, tra le quali sono circa il 2,7 per cento le giovani (in particolare rumene) che hanno dato alla luce un figlio a fronte dello 0,7 per cento delle italiane.

La recente e rapida trasformazione dei modelli sociali e familiari, nonché la struttura per età della popolazione sono notoriamente all'origine dell'addensarsi delle nascite nelle età più avanzate delle madri; ciò porta il Piemonte ad avere un calendario tardivo della fecondità, espresso da un'età media alla maternità tra le più elevate (31,8 anni) e da un livello di fecondità tra i più bassi d'Italia (1,23)⁷. Un calendario del tutto differente da quello più precoce delle straniere, per le quali non si potrebbe escludere una probabile sovrastima della fecondità del momento. Infatti, le considerazioni legate ai differenti livelli di fecondità non possono ignorare la presenza perturbatrice di elementi quali l'età della donna al momento dell'immigrazione e il desiderio di recuperare gli eventuali "vuoti" provocati dalla separazione della coppia con delle maternità dagli accelerati intervalli intergenesici.

segue nota a pagina precedente: semplici passaggi si ottiene la seguente relazione tra i tassi di fecondità: $t_{f_x} = r \cdot t_{f_x}^i + (1-r) \cdot t_{f_x}^s$ nella quale t_{f_x} risulta la somma dei tassi specifici delle italiane e delle straniere, ponderati con la proporzione di donne italiane (r) e straniere ($1-r$) sul totale delle donne all'età x .

⁶ In Piemonte, nel 2006 erano presenti stranieri di 172 nazionalità e le nascite hanno fatto riferimento a madri di 125 paesi diversi. In Torino, dove si concentra la metà dei residenti stranieri della Regione, le nazionalità erano 147 e 87, rispettivamente.

⁷ L'età al parto in Piemonte è salita in modo parallelo e quasi sovrapposto a quella dell'Italia; attualmente è superiore ai 31 anni per le primipare italiane ed ai 27 per le straniere.

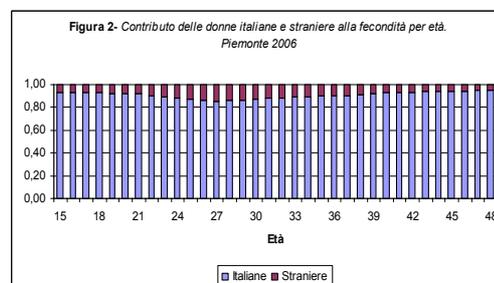
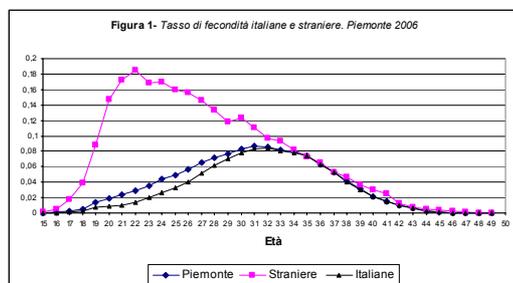


Tavola 2 – Tassi di fecondità per età e contributi specifici. Piemonte 2006.

Età	Tassi di fecondità			Contributi		Età	Tassi di fecondità			Contributi	
	Italiane	Straniere	Totale	Italiane	Straniere		Italiane	Straniere	Totale	Italiane	Straniere
15	0,00012	0,00174	0,00023	0,93460	0,06540	33	0,08095	0,09340	0,08233	0,88949	0,11051
16	0,00062	0,00539	0,00092	0,93561	0,06439	34	0,07837	0,08119	0,07867	0,89347	0,10653
17	0,00174	0,01711	0,00274	0,93536	0,06464	35	0,07463	0,07332	0,07450	0,89780	0,10220
18	0,00269	0,03868	0,00523	0,92938	0,07062	36	0,06232	0,06578	0,06266	0,90207	0,09793
19	0,00698	0,08864	0,01324	0,92334	0,07666	37	0,05228	0,05272	0,05232	0,90639	0,09361
20	0,00880	0,14713	0,01936	0,92367	0,07633	38	0,04044	0,04705	0,04103	0,91051	0,08947
21	0,01041	0,17267	0,02347	0,91952	0,08048	39	0,03060	0,03692	0,03110	0,92134	0,07866
22	0,01324	0,18460	0,02924	0,90663	0,09337	40	0,02108	0,02956	0,02167	0,93019	0,06981
23	0,01963	0,16804	0,03496	0,89672	0,10328	41	0,01492	0,02497	0,01558	0,93374	0,06626
24	0,02695	0,16937	0,04379	0,88176	0,11824	42	0,01004	0,01277	0,01021	0,93575	0,06425
25	0,03232	0,15992	0,04906	0,86880	0,13120	43	0,00604	0,00798	0,00616	0,93661	0,06339
26	0,04010	0,15647	0,05631	0,86064	0,13936	44	0,00272	0,00462	0,00283	0,94073	0,05927
27	0,05173	0,14553	0,06513	0,85708	0,14292	45	0,00135	0,00324	0,00146	0,94256	0,05744
28	0,06133	0,13344	0,07140	0,86030	0,13970	46	0,00037	0,00228	0,00048	0,94412	0,05588
29	0,07054	0,11805	0,07698	0,86456	0,13544	47	0,00014	0,00064	0,00016	0,94905	0,05095
30	0,07750	0,12314	0,08343	0,87005	0,12995	48	0,00007	0	0,00007	0,95326	0,04674
31	0,08339	0,11117	0,08676	0,87887	0,12113	49	0	0	0	0	0
32	0,08387	0,09693	0,08536	0,88605	0,11395	TFT	1,06827	2,57446	1,22884	0,91142	0,08858

La figura 1 mostra il calendario della fecondità, più alto ed anticipato per le straniere e più basso e ritardato per le italiane; la tavola 2 riporta i valori dei tassi specifici di fecondità per età per i due gruppi di madri e la quota relativa (o contributo) da assegnare a ciascuno di essi per giungere al tasso complessivo per età. La scomposizione dei tassi, visibile nella stessa tavola e nella figura 2, mette in evidenza come sia ancora molto basso il concorso delle immigrate straniere alla fecondità osservata in Piemonte; solamente tra i 25 ed i 28 anni il contributo delle straniere arriva al 15 per cento, nelle altre età è notevolmente inferiore. Complessivamente, la fecondità totale nel 2006 si è manifestata con una media di

1,23 figli per donna; l'apporto delle italiane è stato basso (1,07) ma ha contribuito per nove decimi (0,911), a differenza di quello alto delle straniere (2,57) che, però, ha pesato debolmente (0,089) (tav.2).

3. Fecondità: l'influenza di alcuni caratteri sociali.

Il declino del tasso di fecondità appare quale esito di un processo di cambiamento dovuto a delle scelte indotte dall'evoluzione di modelli culturali, sottoposto allo stesso tempo, anche a costrizioni di tipo economico. Un processo che coinvolge in modo speciale gli stranieri per i quali la migrazione rappresenta un'interruzione del percorso di vita. In questo contesto diviene sempre più importante l'osservazione di quei fattori, strettamente collegati alla maternità (età, stato civile, livello di istruzione, occupazione) che, desumibili anche dai Cedap, possono essere interpretati separatamente per ciascun genitore, oppure congiuntamente per entrambi. In più, i Cedap consentono di determinare il numero di figli di ciascuna donna divenuta madre nel 2006; questa informazione permette di determinare la loro quota nell'ambito delle famiglie ed approfondire le connessioni con i fattori sociali sopraccennati⁸.

Oltre all'età, della quale si è già detto, in questo cambiamento gioca un ruolo importante anche la diversa interpretazione della funzione sociale del matrimonio e della coabitazione come forma accettata di unione, e la considerazione della maternità realizzata al di fuori del matrimonio come un fatto del tutto naturale. In Piemonte la quota delle madri nubili si presenta superiore a quella del resto della nazione con quasi due donne su dieci al momento del parto e non vi sono differenze tra le italiane e le straniere (20,7 e 18,7 per cento)⁹.

Nei riguardi dell'istruzione, l'analisi condotta sui Cedap ha messo in evidenza che in Piemonte il 50 per cento delle immigrate divenute madri ha una scolarità bassa, contro appena il 27 per cento delle italiane e, per converso, solo il 7 per cento delle straniere risulta laureata a fronte del 16 delle italiane. In tema di occupazione, il già basso tasso di occupazione delle straniere in Italia (50,7 per cento) rispetto ai principali paesi europei (ISTAT, 2008) è reso ancora più evidente a livello regionale dal divario del tasso delle madri italiane e delle madri straniere

⁸ Per le italiane: il 55,9 ha un figlio, il 35,4 due e l'8,9 per cento tre e più figli; per le straniere: il 53,6 ha un figlio, il 31,6 due e il 14,7 tre e più figli.

⁹ Assodato che le coniugate hanno un numero di figli di ordine superiore a quello delle nubili, il rischio collegato all'aumento della proporzione di madri nubili è quello di un'ulteriore depressione del tasso di fecondità totale. Nei Cedap del 2006, si osserva che il circa il 74 per cento delle nubili italiane e il 61 delle straniere ha un solo figlio, rispetto al 51 ed al 42 per cento delle coniugate.

(79 e 28 per cento) e dal tipo di attività esercitata¹⁰. Ciò spinge a chiedersi quale sia la vera quota di donne straniere non occupate rispetto a quante dichiarano di non esserlo; quante di quel 61 per cento che si dicono “casalinga” lo sono veramente; quanto l’istruzione garantisca alla donna la partecipazione nel mondo del lavoro anche in presenza di figli, a seconda del suo essere italiana o straniera.

Nei riguardi dell’ultimo interrogativo, un recente studio (Rosina, Saraceno, 2008) ha messo in luce il maggior rischio di abbandono dell’attività lavorativa da parte della donna se nella coppia prevale il titolo di studio del marito. Con i Cedap è anche possibile esplorare le relazioni tra occupazione e maternità, utilizzando la regressione logistica quale metodo di analisi. In particolare, per tutte le donne, sono state create le variabili dicotomiche “occupata” e “coniugata”; sono stati ricodificati il titolo di studio (per le madri e per i padri), il settore occupazionale (solo per i padri), il numero di figli, l’età e in ultimo, a livello di coppia, si sono determinate le variabili “parità” e “sup-padre”¹¹.

Nella tavola 3 sono riportati, distinti per italiane e straniere, gli odds ratio delle stime di massima verosimiglianza dei parametri di tre diversi modelli di regressione, ottenuti considerando come variabile dipendente “occupata”. Il primo modello considera solo le caratteristiche sociali della madre; in questo si osserva come il rischio di essere occupata sia praticamente identico tra italiane e straniere in riferimento all’età e come questo, invece, diminuisce al crescere del numero dei figli (per le italiane al terzo figlio si riduce del 75 per cento, per le straniere di oltre il 50). Per le italiane appare evidente l’effetto “protettivo” fornito dall’elevato livello di istruzione nei confronti della presenza sul mercato del lavoro¹², che non si manifesta, invece, nei confronti delle straniere. Se si considera il modello successivo, nel quale sono state inserite le caratteristiche di entrambi i genitori, si nota la perdita di significatività di alcuni fattori riferiti alla madre (coniugata e

¹⁰ Le madri straniere fanno ancora parte di quella che si definisce “prima generazione” di immigrati e, come tale, ancora strettamente legata a schemi culturali dei loro paesi, riflessi nei modelli insediativi e differenti in base alla provenienza. In tal senso si possono spiegare i bassi tassi di occupazione delle straniere.

¹¹ Valori di riferimento: occupata 1; coniugata 1; livello di istruzione 0 se scuola dell’obbligo, 1 se scuola professionale, 2 se diploma di scuola superiore, 3 se laurea; settore occupazionale 1 se industria, 2 se commercio/pubblici servizi/alberghi, 0 per tutti i settori rimanenti; parità 1 per analoghi livelli di istruzione tra padre e madre, sup-padre 1 quando il livello di istruzione del padre è superiore a quello della madre. Per la variabile età sono prima stati calcolati i decili e poi si è specificato un effetto quadratico. Inoltre, sono state ricercate le interazioni tra i diversi fattori non tutte presenti nella tavola che presenta i risultati dell’analisi condotta con il metodo stepwise.

¹² Ancor più evidente se nel modello non si inserisce l’effetto dell’interazione con il numero di figli: in questo caso l’odds delle laureate sale al 6,24.

l'interazione tra figli e istruzione) mentre tra i padri si leggono effetti significativi e quasi simili nei due gruppi per istruzione e settore di occupazione¹³. Il terzo modello vede inserite le due variabili (parità e sup-padre) che intendono cogliere l'effetto congiunto del livello di istruzione nella coppia. Nel caso di parità di livello di istruzione, per le italiane si osserva un rischio di essere occupata basso (16%), il doppio nel caso che il marito abbia un titolo di studio più elevato (30%); per le straniere i rischi si presentano superiori (36% e 56%). Nel caso delle sole madri italiane i risultati concordano con quanto emerge a livello nazionale, forse ancor interpretabili nell'ottica della New Home Economics di Becker (1991), come conseguenza dell'intesa da parte dei coniugi a massimizzare una comune utilità familiare, che si rivela, però, rischiosa per la donna in una società che vede aumentare il numero di divorzi. Allo stesso tempo, per le straniere la spiegazione dei rischi, anche se di poco superiori, può essere ricercata ancora nelle difficoltà del percorso migratorio intrapreso.

Tavola 3 – Italiane e straniere occupate e caratteri sociali di entrambi i genitori in base ai Cedap 2006 (gli odds ratio delle regressioni logistiche).

Caratteri sociali	Modello 1		Modello 2		Modello 3	
	Italiane	Straniere	Italiane	Straniere	Italiane	Straniere
Madre						
Età (dec)	1,5805***	1,6605***	1,5093***	1,7177***	1,5131***	1,7190***
Età (dec)^2	0,9683	0,9733***	0,9733***	0,9753***	0,9732***	0,9751***
Coniugata	1,2328**	0,8652	1,1019	0,8182	1,1075**	0,8119
Due figli	0,5196**	0,7595**	0,5167***	0,8077	0,5161	0,8043
Tre figli	0,2512**	0,4763***	0,2442***	0,8077**	0,2427***	0,5396**
Liv. Istruzione 1	2,3701***	2,7437***	2,0987***	2,0412***	2,3154***	2,5581***
Liv. Istruzione 2	2,4741***	1,3316**	2,4259***	1,1241**	2,8716***	1,5083**
Liv. Istruzione 3	4,5269***	0,9892	4,1267***	0,7544	5,3165***	1,1516
n. figli x istruzione	1,0638**	1,0509	1,0437=	1,0787	1,0439=	1,0781
Coniugata x n. figli	0,8266***	0,7414**	0,8704*	0,7417**	0,8686**	0,7464**
Padre						
Età (dec)			1,0479	0,9590	1,0487	0,9606
Età (dec)^2			0,9922**	0,9902**	0,9921***	0,9899**
Liv. Istruzione1			1,1152=	1,3689**		
Liv. Istruzione 2			1,2568***	1,3993***		
Liv. Istruzione 3			1,2666***	1,4211**		
Set. Occupazionale 1			1,1672***	0,9885	1,1665***	0,9861**
Set. Occupazionale2			1,3470***	1,2822**	1,3546***	1,3053
Caratteristiche della Coppia vs istruzione						
Parità					1,1609**	1,3632**
Superiore padre					1,2914***	1,5682***
Log likelihood	-10319,57	-3150,49	-99177,69	-2644,17	-9182,96	-2646,85

Nota : = p<0,10; * p<0,05 ; ** p<0,01 ; ***p<0,001

N.B : per motivi di spazio non è stata inserita la categoria di riferimento delle diverse variabili.

¹³ Si notano odds ratio più elevati tra i padri stranieri in riferimento all'istruzione. Per entrambi i gruppi gli odds ratio sono più elevati se i padri sono occupati nel commercio e/o pubblici servizi.

Riferimenti bibliografici

- ISTAT, 2009. *Gli stranieri nel mercato del lavoro (versione provvisoria)*. Argomenti n.38. Istituto Nazionale di Statistica, Roma.
- Becker G.S.1991. *A Treatise on the Family* (Enlarged edition). 2nd ed. Cambridge, University Press, Harvard.
- Caritas/Migrantes, 2007. *Dossier statistico immigrazione. XVII Rapporto*, Caritas, Roma.
- Keyfitz, N., 1985. *Applied Mathematical Demography*. New York, Berlin Heidelberg, Tokio, Springer Verlag.
- Rosina A., Saraceno C., 2008. *Interferenze asimmetriche. Uno studio della discontinuità lavorativa femminile*. Economia e Lavoro 2/2008

SUMMARY

The immigrant settlements have increased and now, when we study the population dynamics in Italy, we must count their contributions on future demographic change. We analyze the age specific fertility rate in Piemonte using the database of childbirths in 2006. This analysis was performed on three different cases: Italian mothers, foreign mothers and both together.

We decomposed the rates to compute the different contributions by the two groups of mothers (Italians and foreigners mothers) to specific and total fertility rates. We have also began the analysis of the socio-demographic characteristics.

Mauro REGINATO, Dipartimento di Statistica e Matematica Applicata, Università di Torino.

Tiziana BARUGOLA, Dipartimento di Statistica e Matematica Applicata, Università di Torino.

LA PERFORMANCE DEI MODELLI DI PREVISIONE DELLA VOLATILITÀ

Giuseppe Ricciardo Lamonica

1. Introduzione

Nell'ambito dei mercati finanziari la volatilità storica, come è noto, indica il grado con il quale le quotazioni delle attività finanziarie tendono a fluttuare nel tempo. Come tale, essa è un concetto strettamente legato a quello di rischio di investimento.

Per tale motivo, nell'ambito delle allocazioni delle risorse, l'analisi della volatilità ed in particolare del suo andamento futuro rappresenta un elemento di investigazione di fondamentale importanza. Al riguardo si rinvia, tra gli altri, ai lavori di: Yu (2002), Poon e Granger (2003), Andersen, Bollerslev, Christoffersen e Diebold (2005), Moore e Wang (2007), Corrado e Truong (2007).

Conseguentemente, se da un lato sono stati sviluppati tanti modelli di previsione da rendere impossibile una loro elencazione, dall'altro lato è ancora controverso quale sia il migliore.

Prendendo in considerazione il mercato finanziario italiano, che ad oggi ha ricevuto ancora poca attenzione, l'obiettivo di questo lavoro è quello di valutare la capacità di descrivere e prevedere la volatilità di quei modelli che più comunemente vengono utilizzati nella pratica. Così facendo si ritiene di fornire un utile contributo alla discussione.

Pertanto, nel paragrafo seguente vengono illustrati nelle loro caratteristiche salienti i modelli considerati, mentre in quello successivo si espongono i risultati della indagine condotta. Seguono infine alcune proposte conclusive.

2. Volatilità: misura e modelli di previsione

La volatilità storica (d'ora in poi solo volatilità) intesa come variabilità delle quotazioni di un titolo o di un portafoglio finanziario, può essere quantificata in diversi modi. Tuttavia, la varianza o lo scarto quadratico medio delle quotazioni (x_t) osservate in un certo periodo di tempo è l'indice che viene generalmente utilizzato nella pratica:

$$\sigma_T^2 = \frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})^2 \text{ per } \bar{x} = \frac{1}{T} \sum_{t=1}^T x_t \quad (1)$$

Come è stato detto, quando si valuta la opportunità di un investimento prevedere la sua volatilità è una operazione di fondamentale importanza. Al riguardo sono stati sviluppati in letteratura diversi modelli. In questo lavoro, limitando l'attenzione a quelli maggiormente utilizzati nella pratica, si considerano quelli di seguito elencati. Al riguardo sia T l'ultimo periodo osservato e $T+1$ fino a $T+s$ il periodo di previsione:

- *Il modello di random walk (RW)*; In accordo con l'ipotesi di efficienza debole dei mercati finanziari la previsione della volatilità di una attività finanziaria avviene nel seguente modo:

$$\sigma_{T+j}^2 = \sigma_{T+j-1}^2 \quad \text{per } j=1, \dots, s \quad (2)$$

- *Il modello della media aritmetica storica (MAS)*; Se la volatilità è più o meno costante nel tempo allora la previsione è data dalla media aritmetica delle volatilità osservate nel passato:

$$\sigma_{T+j}^2 = \frac{1}{T+j-1} \sum_{t=1}^{T+j-1} \sigma_t^2 \quad \text{per } j=1, \dots, s \quad (3)$$

- *Il modello della media mobile (MM)*; Rappresenta una variante del modello precedente e parte dal presupposto che, ai fini della previsione, i dati più remoti non hanno alcuna influenza e pertanto possono essere esclusi. Quindi, indicato con k il numero delle ultime osservazioni considerate:

$$\sigma_{T+j}^2 = \frac{1}{k} \sum_{t=T+j-k}^{T+j-1} \sigma_t^2 \quad \text{per } j=1, \dots, s \quad (4)$$

Poiché la scelta di k è del tutto soggettiva, nella analisi condotta in questo lavoro, sono stati considerati i seguenti livelli: 4, 12, 25, 50 e 100.

- *Il modello lisciamiento esponenziale (EXP-SMO)*; La logica del modello in questione è del tutto analoga alla precedente. Infatti, ai fini della previsione, si considerano tutte le osservazioni disponibili dando però ad ognuna di esse un peso decrescente man mano che si considerano quelle più remote. Rinviando per i dettagli alla letteratura, come è noto, le previsioni avvengono nel seguente modo:

$$\sigma_{T+j}^2 = \alpha \sum_{t=1}^{T+j-1} (1-\alpha)^t \sigma_{T+j-t}^2 \quad \text{per } j=1, \dots, s \quad (5)$$

- *Il modello media mobile esponenziale (EXP-MM)*; costituisce una combinazione degli ultimi due modelli considerati. Infatti, la previsione si ottiene considerando soltanto le ultime k osservazioni e dando ad ognuna di esse un peso decrescente secondo il metodo del lisciamiento esponenziale. Pertanto:

$$\sigma_{T+j}^2 = \frac{1}{\Gamma} \sum_{t=1}^k \alpha^t \sigma_{T+j-t}^2 \quad \text{per } \Gamma = \sum_{t=1}^k \alpha^t \text{ e } j=1, \dots, s \quad (6)$$

Anche in questo caso essendo la scelta di k soggettiva sono stati considerati gli stessi livelli del modello MM. Inoltre, seguendo la letteratura ed in particolare Ederington e Guan (2004), si è posto $\alpha=0.94$.

- *Il processo ARCH*; Rappresenta il primo processo stocastico che tiene conto del fatto che la volatilità può non essere costante nel tempo. Il processo è formalizzato nel seguente modo dove r_t per $t=1, \dots, T$ è la serie osservata dei rendimenti di un titolo finanziario:

$$r_t = \mu_t + a_t \quad (7)$$

mentre μ_t a seconda dell'evidenza empirica può essere nullo, coincidere con una costante oppure con un processo ARMA di piccolo ordine. Inoltre $a_t = \sigma_t \varepsilon_t$ è una variabile residuale che ha media nulla e varianza condizionata eteroschedastica specificata nel seguente modo:

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_p a_{t-p}^2 \quad (8)$$

Infine ε_t sono variabili indipendenti e distribuite secondo una normale standardizzata. Infine, al fine di assicurare che la varianza sia non negativa, $\alpha_0 > 0$ e $\alpha_i \geq 0$ per $i=1, \dots, p$. Alla luce di questa formalizzazione la volatilità viene prevista proiettando in avanti l'equazione della varianza nel seguente modo:

$$\hat{\sigma}_{T+j}^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \sigma_{T+j-i}^2 \quad \text{dove } \sigma_{T+j-i}^2 = a_{T+j-i}^2 \text{ se } j-i \leq 0 \text{ e } j=1, \dots, s \quad (9)$$

- *Il processo GARCH*; Costituisce una generalizzazione del modello ARCH ed in particolare, ha la caratteristica di essere più parsimonioso nella scelta dell'ordine dell'equazione della varianza. Inoltre, come è noto, prende in considerazione il fenomeno del volatility clustering (anche persistenza) e cioè l'esistenza di periodi in cui la volatilità tende a rimanere bassa e periodi in cui al contrario essa tende a restare elevata. I coefficienti β_j della equazione della varianza catturano questo effetto. formalmente:

$$r_t = \mu_t + a_t; \quad (10)$$

$$a_t = \sigma_t \varepsilon_t; \quad (11)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i a_{t-i}^2 + \sum_{k=1}^q \beta_k \sigma_{t-k}^2; \quad (12)$$

Con $\alpha_0 > 0$; $\alpha_i \geq 0$ per $i=1, \dots, p$; $\beta_k \geq 0$ per $k=1, \dots, q$; $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1$.

Rinviando per i dettagli a Baillie e Bollerslev (1992), la volatilità viene prevista, come è noto, in modo analogo al processo ARCH.

- *Il processo GARCH in MEAN (GARCH-M)*; E' una variante del processo GARCH in quando considera esplicitamente la circostanza che i rendimenti di una attività finanziaria possono dipendere direttamente dalla loro volatilità. Il parametro c della seguente equazione (13), chiamato in letteratura premio per il rischio, considera il trade-off tra il rendimento e la sua volatilità. Formalmente questo

processo è analogo al modello GARCH. L'unica differenza è, come è stato detto, nella equazione del rendimento che presenta la seguente forma:

$$r_t = \mu_t + c\sigma_t^2 + a_t \quad (13)$$

- *Il modello EXPONENTIAL GARCH (EXP-GARCH)*; Una caratteristica che a volte si osserva nelle analisi empiriche è che ai movimenti al ribasso dei rendimenti sono associati ad una volatilità maggiore rispetto a quella che si osserva nei movimenti al rialzo. Ciò comporta che l'effetto esercitato dalla variabile residuale a_t sulla varianza di r_t sia asimmetrico (effetto leverage). Al fine di considerare esplicitamente questa evidenza sono stati proposti diversi modelli. Fra questi, quello che ha riscosso un certo successo è il processo EXP-GARCH formalizzato nel seguente modo:

$$r_t = \mu_t + a_t; \quad (14)$$

$$a_t = \sigma_t \varepsilon_t; \quad (15)$$

$$\ln \sigma_t^2 = \alpha_0 + \beta_1 \ln \sigma_{t-1}^2 + \alpha_1 \left(\left| \frac{a_{t-1}}{\sigma_{t-1}} \right| - E \left(\left| \frac{a_{t-1}}{\sigma_{t-1}} \right| \right) \right) + \gamma \frac{a_{t-1}}{\sigma_{t-1}} \quad (16)$$

Rinviando per i dettagli alla letteratura, come è noto, per il modello in questione non esiste una formula analitica che consente di prevedere la volatilità futura per un orizzonte di $j > 1$. In questo lavoro, seguendo le impostazioni di Alexander (2001) e Marcucci (2005) le previsioni sono state effettuate nel seguente modo:

$$\sigma_{T+1}^2 = \exp(\alpha_0) (\sigma_T^2)^{\beta_1} \exp \left[\alpha_1 \left(\left| \frac{a_T}{\sigma_T} \right| - E \left(\left| \frac{a_T}{\sigma_T} \right| \right) \right) + \gamma \frac{a_T}{\sigma_T} \right] \quad (17)$$

mentre per $j=2, \dots, s$:

$$\sigma_{T+j}^2 = \exp(\alpha_0) (\hat{\sigma}_{T+j-1}^2)^{\beta_1} \quad (18)$$

3. La volatilità nel mercato finanziario italiano

Tenendo conto degli obiettivi del lavoro sono stati considerati i rendimenti giornalieri composti continui dell'indice MIBTEL della borsa valori di Milano dal 04/01/2000 al 19/09/2008. Premesso che tutte le analisi sono state condotte con il software statistico SAS, l'attenzione si è quindi concentrata sulla volatilità settimanale (σ_k^2) la quale è stata determinata mediante la varianza dei rendimenti giornalieri di ogni singola settimana del periodo analizzato. In particolare, indicato con r_{tk} il rendimento del giorno t -esimo della generica k -esima settimana e con n_k il numero di giorni di contrattazione di quella settimana:

$$\sigma_k^2 = \frac{1}{n_k} \sum_{t=1}^{n_k} (r_{tk} - \bar{r}_k)^2; \quad \bar{r}_k = \frac{1}{n_k} \sum_{t=1}^{n_k} r_{tk} \quad (19)$$

In generale quindi con i dati a disposizione è possibile calcolare 455 volatilità settimanali.

Tuttavia poiché i modelli considerati, ad eccezione del RW, MM, EXP-MM e del MAS prevedono la stima di alcuni parametri, il periodo di analisi è stato suddiviso in due sottoperiodi. Il primo che va dal 04/01/2000 al 30/9/2007 (periodo di valutazione) è stato utilizzato per la stima dei parametri della equazione dei rendimenti, della varianza condizionata dei vari modelli e della costante di liscio del modello EXP-SMO. Il secondo sottoperiodo che va dal 1/10/2007 al 19/09/2008 (periodo di previsione), costituito da 50 settimane, rappresenta invece l'orizzonte temporale di previsione.

In particolare, l'equazione dei rendimenti giornalieri dell'indice MIBTEL per i processi ARCH, GARCH, EXP-GARCH e GARCH-M, è stata individuata analizzando la funzione di autocorrelazione totale e parziale e sulla base dei criteri di Akaike e Schwarz. Da questa indagine risulta che la variabile in considerazione è bene approssimata da un processo autoregressivo del primo ordine.

Invece, l'ordine della equazione della varianza condizionata è stata individuata, attraverso il test dei moltiplicatori di Lagrange. I risultati di queste analisi che non vengono riportati per non appesantire l'esposizione sono disponibili su richiesta.

Dopo avere stimato i parametri dei vari modelli e condotte tutte le verifiche necessarie sui residui, sono state previste le volatilità giornaliere per tutto il periodo previsionale. Quindi quelle settimanali sono state ottenute nel seguente modo:

$$\hat{\sigma}_k^2 = \sum_{i=1}^{nk} \hat{\sigma}_{i,k}^2 \quad \text{per } k=1, \dots, 50 \quad (20)$$

dove $\hat{\sigma}_{i,k}^2$ indica la volatilità prevista per i-esimo giorno della k-esima settimana.

In seguito, al fine di valutare l'accuratezza delle previsioni, le volatilità settimanali previste sono state confrontate con le corrispondenti volatilità osservate (σ_k^2) mediante i seguenti indicatori:

$$- \left(\frac{1}{50} \sum_{k=1}^{50} (\sigma_k^2 - \hat{\sigma}_k^2)^2 \right)^{0.5} \quad \text{ERRORE QUADRATICO MEDIO (RMSE);} \quad (21)$$

$$- \frac{1}{50} \sum_{k=1}^{50} |\sigma_k^2 - \hat{\sigma}_k^2| \quad \text{ERRORE MEDIO ASSOLUTO (MAE);} \quad (22)$$

$$- \frac{1}{50} \sum_{k=1}^{50} \frac{|\sigma_k^2 - \hat{\sigma}_k^2|}{\hat{\sigma}_k^2} \quad \text{ERRORE MEDIO ASSOLUTO RELATIVO (MAPE).} \quad (23)$$

La Tabella 1 riporta i risultati di questa analisi. Come è possibile notare, secondo l'indice RMSE i modelli più accurati sono quelli a media mobile esponenziale. Al contrario invece tutti i modelli stocastici delle famiglie ARCH e GARCH risultano essere i meno precisi.

Questi ultimi, come mostrano i risultati, sono meno accurati del 46,6% rispetto al modello EXP-MM(25) che, secondo l'indice RMSE è il migliore.

Di seguito troviamo invece i modelli a media mobile i quali, a loro volta sono mediamente meno accurati 36% rispetto al modello EXP-MM(25) mentre risultano in media più precisi del 16,7% rispetto agli ultimi tre modelli (ARCH, EXP-GARCH e GARCH-M).

Considerando invece i risultati dell'indice MAE si ottiene una situazione leggermente diversa. Infatti, i modelli più accurati sono quelli a media mobile per $k=50$, $k=25$ ed il modello della media aritmetica storica. Invece i meno accurati sono, ancora una volta, i processi stocastici delle famiglie ARCH e GARCH i quali, rispetto precedente tre modelli, sono mediamente meno precisi del 36,7%.

Infine, analizzando dai valori dell'indice MAPE è possibile notare che l'unica differenza degna di attenzione rispetto ai risultati precedenti è la posizione medio alta occupata nella classifica dalla famiglia dei processi ARCH e GARCH.

Una analisi congiunta dei risultati ottenuti con questi tre indici di adattamento evidenzia che i modelli migliori sono quelli a media mobile ed a media mobile esponenziale. Al contrario invece quelli peggiori sono i modelli stocastici.

Quale ulteriore indice di valutazione è stato considerato anche il seguente modello di regressione:

$$\hat{\sigma}_k^2 = \alpha + \beta\sigma_k^2 + \varepsilon_k \quad \text{per } k=1, \dots, 50 \quad (24)$$

Infatti, se la volatilità prevista coincide con quella storica, le stime dell'intercetta e del coefficiente angolare non dovrebbero essere significativamente diverse, rispettivamente, da zero ed uno (ipotesi nulla). Quindi, il test F di Fisher consente di verificare se la precedente ipotesi è vera.

I risultati di questa indagine sono esposti nella seconda parte della Tabella 1 e concordano con quelli ottenuti precedentemente.

Infatti, analizzando la significatività della statistica F di Fisher è possibile notare per tutti i modelli a media mobile, quelli a media mobile esponenziale e per il modello della media storica è possibile accettare l'ipotesi nulla.

Al contrario invece per i restanti modelli la statistica in questione assume valori, che senza ombra di dubbio, comportano l'accettazione dell'ipotesi alternativa.

4. Conclusioni

Prevedere la volatilità di una attività o di un portafoglio finanziario è un ambito di ricerca della letteratura finanziaria molto importante. Al riguardo, sono stati sviluppati tanti i modelli di previsione.

In questo lavoro ne sono stati considerati 17 con l'obiettivo di valutare la loro abilità nel descrivere e prevedere la volatilità storica settimanale dell'indice MIBTEL della borsa valori di Milano.

La verifica è stata condotta confrontando, mediante quattro indici di valutazione, la volatilità prevista dai vari modelli con quella storica osservata.

I risultati della indagine condotta sono abbastanza chiari e, concordando con quelli ottenuti da altre ricerche (Taylor J. W. 2004), evidenziano con un buon grado di certezza che i modelli a media mobile e quelli esponenziali a media mobile producono le migliori previsioni.

Al contrario invece i modelli appartenenti alla famiglia dei processi stocastici ARCH e GARCH che sono formalmente più complessi, più popolari e considerano esplicitamente tante evidenze empiriche (fatti stilizzati) che si riscontrano nell'analisi dei dati finanziari, sono risultati i meno accurati.

Riferimenti bibliografici

- Alexander C. 2001. *Market models*. John Wiley & Sons, England.
- Andersen T. G., Bollerslev T., Christoffersen P.F. e Diebold X. F. 2005. *Volatility forecasting*. NBER working paper n. w11188.
- Baillie R. e Bollerslev T. 1992. *Prediction in dynamic models with time-dependent conditional variances*. Journal of Econometrics, vol. 52, pp 91-113.
- Corrado C. e Truong C. 2007. *Forecasting stock index volatility: comparing implied volatility and the intraday high-low price range*. The Journal of Financial Research, vol. XXX, pp 201-215.
- Ederington L. H. e Guan H. 2005. *Forecasting volatility*. Journal of Futures Markets, vol. 25, pp 465-480.
- Lopez J. A. 2001. *Evaluating the predictive accuracy of volatility models*. Journal of forecasting, vol. 20, pp 87-109.
- Marcucci J. 2005. *Forecasting stock market volatility with regime-switching GARCH models*. Studies in Nonlinear Dynamics and Econometrics, vol.9, pp 1-53.
- More T. e Wang P. 2007. *Volatility in stock returns for new EU member states: markov regime switching model*. International review of financial analysis, vol. 16, pp 282-292.
- Poon S. H. e Granger C. W. J. 2003. *Forecasting volatility market: a review*. Journal of economic literature, vol. XLI, pp 478-489.
- Taylor W. J. 2004. *Volatility forecasting with smooth transition exponential smoothing*. International Journal of forecasting, vol. 20, pp 273-286.
- Yu J. 2002. *Forecasting volatility in the New Zealand stock market*. Applied Financial Economics, vol. 12, pp193-202.

Appendice 1**Tabella 1** – Risultati degli indici statistici di valutazione e del test F di Fisher per la verifica dell'ipotesi nulla.

Modello	Indici statistici			Test F di Fisher			
	RMSE	MAE	MAPE	Media quadrati numeratore	Media quadrati Denominat.	Stat. F di Fisher	Signif.
EXP-MM(25)	0,000170	0,000136	1,2455	1,8417E-7	7,0139E-8	2,63	0,0828
EXP-MM(12)	0,000183	0,000138	1,1043	1,9587E-7	6,9727E-8	2,83	0,0692
EXP-MM(100)	0,000261	0,000119	0,9837	4,4688E-8	6,9332E-8	0,64	0,5294
EXP-MM(50)	0,000263	0,000125	1,1816	5,2243E-8	7,0137E-8	0,74	0,4803
MM(12)	0,000264	0,000121	1,0789	6,4704E-8	6,9879E-8	0,93	0,4037
MM(25)	0,000265	0,000118	1,0056	6,8817E-8	7,0131E-8	0,98	0,3822
MM(50)	0,000266	0,000113	0,7666	1,1183E-7	6,8848E-8	1,62	0,2077
MAS	0,000267	0,000117	0,8906	1,1358E-7	6,9295E-8	1,64	0,2048
MM(4)	0,000268	0,000126	1,0029	1,2740E-7	6,9446E-8	1,83	0,1707
EXP-MM(4)	0,000273	0,000139	1,1303	2,0018E-7	6,9265E-8	2,89	0,0653
MM(100)	0,000274	0,000120	0,6470	2,7164E-7	6,6997E-8	4,05	0,0236
EXP-SMO	0,000280	0,000131	0,8963	2,8405E-7	6,9950E-8	4,06	0,0235
RW	0,000296	0,000151	1,2048	5,1813E-7	6,9808E-8	7,42	0,0016
GARCH	0,000317	0,000182	0,9843	8,3948E-7	6,9915E-8	12,01	<0,0001
GARCH-M	0,000318	0,000183	0,9904	8,4656E-8	6,9891E-8	12,11	<0,0001
ARCH	0,000318	0,000184	0,9966	8,4487E-7	7,0029E-8	12,06	<0,0001
EXP- GARCH	0,000318	0,000184	0,9981	8,4424E-7	7,0142E-8	12,04	<0,0001

SUMMARY**The performance of the forecasting volatility models**

This study evaluates the performance of seventeen alternatives models for predicting the weak volatility in the Italian stock market.

The competing models contain both simple models such as random walk and historical mean and Arch, Garch, Exp-Garch and Garch-m type complex models.

Five different measures are used to evaluate the forecasting accuracy. The main result is that among the time series models considered the moving average and the exponentially moving average yields better forecasts.

Giuseppe RICCIARDO LAMONICA, Ricercatore di statistica economica presso la Facoltà di Economia dell'Università Politecnica delle Marche.

EVIDENZE POCO NOTE DELL'INCREMENTO DELLA POPOLAZIONE STRANIERA NEI COMUNI ITALIANI*

Angela Silvestrini, Alessandro Valentini

1. Introduzione

Nel corso del 2008 la popolazione residente sul territorio italiano ha superato, per la prima volta nella storia, la soglia dei 60 milioni di abitanti (Istat, 2009). La crescita, di circa 426.000 unità rispetto all'anno precedente (+0,7%), è interamente ascrivibile alle migrazioni dall'estero, una circostanza che si ripete ormai da alcuni anni. Infatti, l'immigrazione dall'estero rappresenta, insieme all'invecchiamento, la componente che determina le più importanti trasformazioni socio-demografiche che il nostro Paese sta sperimentando ed è per tale motivo che sia la statistica ufficiale, che molti studiosi e ricercatori, stanno dedicando numerosi studi e approfondimenti su dette tematiche (si vedano per esempio Golini, 2006; Istat, 2008a).

Ciò nonostante, scarsa o nulla attenzione viene riservata alle mere conseguenze finanziarie, giuridiche e amministrative sui singoli Comuni che derivano dalla variazione nel numero dei residenti indotto dalle immigrazioni. Una delle rare eccezioni è rappresentata dal Rapporto Caritas/Migrantes (Caritas, 2008). L'allargamento della prospettiva di analisi anche a questa dimensione può far emergere alcune evidenze poco note da mettere in relazione all'incremento della popolazione straniera nel nostro Paese e aprire la strada ad una visione nuova del fenomeno migratorio. Questo è l'obiettivo che ci proponiamo nel presente lavoro analizzando le ricadute "finanziarie" sui singoli Comuni nel caso in cui fosse venuto a mancare l'apporto della componente straniera all'incremento della popolazione residente dei Comuni italiani a partire dai dati del Censimento della popolazione dal 2001 ad oggi.

2. Principali provvedimenti legislativi legati all'ammontare della popolazione

Molte leggi del nostro Paese fanno riferimento alla popolazione legale, determinata sulla base del Censimento generale della popolazione, o alla

* Pur essendo frutto di un lavoro comune, i paragrafi 4 e 5 di questo rapporto sono stati scritti da Angela Silvestrini e i paragrafi 1, 2 e 3 da Alessandro Valentini.

popolazione residente, calcolata dall'Istat al termine di ciascun anno intercensuario, per la determinazione di una serie di provvedimenti, in particolare in materia economica ed elettorale, e dei loro effetti giuridici e/o amministrativi nei confronti dei Comuni e di altri Enti locali.

Alla popolazione residente determinata dal censimento si fa menzione, innanzitutto, nella Costituzione agli articoli 56 e 57, ai fini della distribuzione dei seggi della Camera dei deputati nelle diverse circoscrizioni elettorali, nonché della distribuzione dei seggi su base regionale, per il Senato.

A livello locale, il dato di popolazione legale influisce sul sistema di elezione del Sindaco e dei consiglieri comunali, nonché sulle indennità dei Sindaci, degli assessori, sul numero degli assessori e sulla qualifica del segretario comunale. Sempre alla popolazione legale si fa riferimento nella legge elettorale relativa ai consigli provinciali (L. 122/1951).

Il numero dei residenti in ciascun comune è importante fra gli altri, ai fini dell'apertura di nuove farmacie (L. 475/1968), nella classificazione delle imprese commerciali di vendita (D.Lgs. 114/1998), e, soprattutto, nel calcolo dei trasferimenti economici dallo Stato ai Comuni, alle Province e alle Regioni.

3. Il metodo di lavoro utilizzato

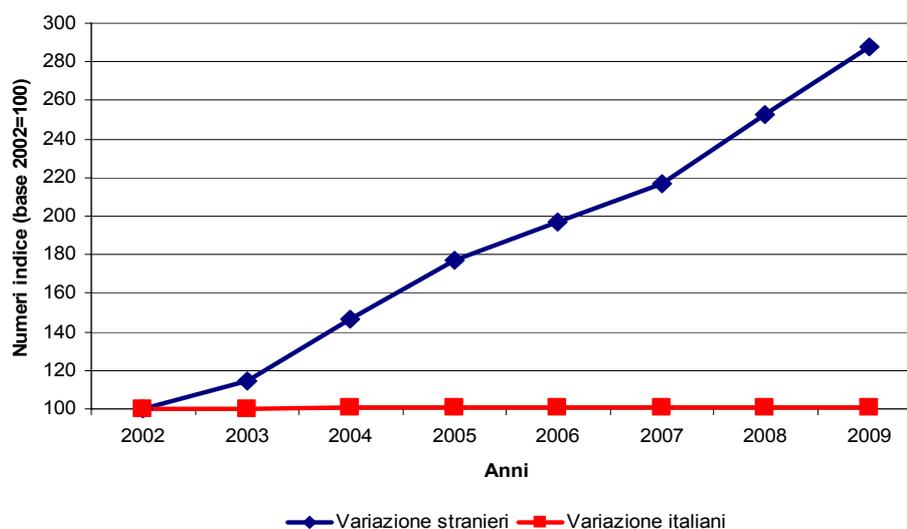
Nel corso degli ultimi anni la popolazione residente nel nostro Paese ha avuto un rapido incremento, passando da quasi 57 milioni al 21 ottobre 2001 (censimento) a più di 60 milioni al 31 dicembre 2008. Tale incremento è da ascrivere in maniera quasi esclusiva alla crescita della componente straniera, che sale da 1,3 a 3,9 milioni di unità tra il Censimento 2001 e il 2008, con una variazione di incidenza rispetto alla popolazione totale dal 2,3 al 6,5 per cento, come si può osservare dalla Figura 1. Senza questo incremento la popolazione residente complessiva nel 2008 non avrebbe superato i 58 milioni di unità¹. Ancora, saranno gli stranieri negli anni futuri a consentire la crescita, o quanto meno ad evitare il declino, della popolazione totale (Istat, 2008b).

Tuttavia, la variazione non è uniforme sul territorio nazionale, tanto che a livello comunale si verificano tendenze di segno opposto di cui è necessario tenere conto. In particolare, l'incremento nel numero di residenti tende a localizzarsi nel centro-nord e specificamente nel nord-est per effetto di una maggiore

¹ Analizzando i dati, potrebbe apparire un incremento anche della sola componente italiana della popolazione residente. In realtà, questo è dovuto alla pratica dei recuperi post-censuari, cioè all'inserimento nei modelli di calcolo della popolazione di persone residenti ma sfuggite al censimento.

concentrazione dei flussi migratori dall'estero. Questo implica, in molti casi, ripercussioni di tipo amministrativo e finanziario, spesse volte trascurate nei dibattiti relativi alle conseguenze dell'immigrazione nel nostro Paese.

Figura 1 – *Variazione della popolazione residente in Italia dal 1° Gennaio 2002 al 1° Gennaio 2009.*



Per fornire un quadro esemplificativo di tali conseguenze nei contesti micro-territoriali abbiamo confrontato per ogni Comune, attraverso una simulazione², il dato di popolazione attuale con quello che sarebbe risultato nel caso di invarianza della popolazione residente straniera tra il Censimento 2001 e il 2008.

Per realizzare la simulazione abbiamo predisposto un Database con i principali dati di nostro interesse per ciascuno degli 8.101 Comuni italiani³: la popolazione censita al 2001 italiana, straniera e complessiva e quella residente in anagrafe al 1°

² La simulazione consiste nel confrontare, per ogni Comune, la popolazione iscritta in anagrafe al 1° Gennaio 2008 con quella ipotetica che sarebbe risultata dalla "sterilizzazione" nel numero complessivo di residenti stranieri tra il Censimento 2001 e il 31 Dicembre 2007.

³ Le variazioni territoriali tra il 2001 e il 2007 sono del tutto marginali e trascurabili. Si segnala soltanto che Bollate (MI) ha perso parte della popolazione a favore di Baranzate (MI) e che è stato istituito il comune di San Siro come fusione tra Santa Maria Rezzonico e Sant'Abbondio (CO).

gennaio 2008 italiana, straniera e complessiva, ulteriormente suddivisa in tre grandi classi di età (0-4 anni; 5-64 anni; 65 anni e oltre). Abbiamo successivamente estrapolato dal database due livelli di popolazione:

Popolazione anagrafica - Popolazione complessiva iscritta in anagrafe al 1° gennaio 2008.

Popolazione simulata - Popolazione straniera al Censimento 2001 + Popolazione italiana iscritta in anagrafe al 1° gennaio 2008.

Nei due casi abbiamo determinato le soglie demografiche di pertinenza di ciascun Comune utilizzando le classi dimensionali richiamate dalle varie normative vigenti. In particolare ci siamo soffermati sul numero di passaggi di fascia tra il caso reale (popolazione anagrafica) e quello al netto dell'incremento degli stranieri (popolazione simulata) e abbiamo stimato alcune conseguenze amministrative e finanziarie legate a tali (ipotetici) passaggi di fascia.

4. Alcune esemplificazioni: conseguenze amministrative e finanziarie sui Comuni nel passaggio dalla popolazione anagrafica a quella simulata

Il conteggio relativo all'ammontare della popolazione residente in ciascun Comune non viene effettuato soltanto per finalità di tipo statistico. A ciascuna variazione temporale nell'ammontare sono infatti associate - sulla base delle disposizioni normative vigenti - precise e talvolta rilevanti conseguenze di tipo finanziario e amministrativo. Per esempio, per quanto concerne il calcolo dei trasferimenti economici dallo Stato ai Comuni si segnalano due tipologie di contributi ordinari per i Comuni con meno di 5.000 abitanti, in base all'art.1 c. 703 L. 296/06 e successive modifiche (art 2 c.10 L. 244/07 e art. 2 C. 40 L. 203/08). Il primo contributo (per un ammontare complessivo di 45 milioni di Euro nel 2008) è riservato ai Comuni la cui percentuale di ultrasessantacinquenni è superiore rispetto ad una certa soglia (più del 25% della popolazione complessiva). Il secondo contributo (per un ammontare complessivo di 81 milioni di Euro nel 2008) viene invece attribuito ai Comuni nei quali la percentuale di bambini al di sotto di 5 anni sia superiore rispetto al 4,5% del totale (in passato tale percentuale era del 5%).

I 45 milioni di Euro per gli anziani sono stati assegnati nel 2008 a 2.005 Comuni, distribuiti prevalentemente nel Centro-Nord (Tavola 1). Il contributo medio si assesta attorno ai 22.500 Euro per Comune, ma in alcune realtà esso si aggira attorno ai 100.000 Euro, come nel caso di Campione d'Italia (CO), Scansano (GR), Licciana Nardi (MS). Si trovano però nel Sud i due Comuni dove il contributo risulta più elevato in assoluto: San Fele (PZ) con 110.000 Euro e Castiglione di Sicilia (CT) con 130.000 Euro.

Nel caso in cui non si fosse verificato alcun incremento degli stranieri tra il

2002 e il 2008, come nella nostra simulazione, ulteriori 372 Comuni avrebbero potuto beneficiare del contributo, causando la redistribuzione del totale.

In questa specifica circostanza, infatti, l'effetto di traino esercitato dagli stranieri sulla popolazione complessiva risulta penalizzante: 61 realtà locali perdono il contributo perché viene superata la soglia dei 5.000 abitanti, altri 311 Comuni, che rimangono sotto la soglia, lo perdono perché gli stranieri, mediamente più giovani degli italiani, fanno sì che la quota di anziani rispetto alla popolazione complessiva scenda al di sotto del 25%. L'effetto paradossale della disposizione legislativa è che laddove ci sono molti anziani, e di conseguenza molti stranieri addetti ai servizi di cura, la popolazione "ringiovanisce" e perde il diritto al contributo! Si noti, per curiosità, che in una unica circostanza l'incremento degli stranieri fa sì che un Comune che in base ai dati del Censimento 2001 non avrebbe avuto diritto al contributo, lo abbia acquisito. È il caso di Airole (in provincia di Imperia), dove circa il 10% dei 460 abitanti è composto da stranieri ultrasessantacinquenni.

Gli 81 milioni di Euro di contributo destinati ai Comuni con meno di 5.000 abitanti, che hanno una percentuale di bambini da 0 a 4 anni superiore al 4,5%, nel 2008 sono stati allocati tra 2.103 Comuni (mediamente 38.500 Euro per Comune), la metà dei quali si trova nel Nord-Ovest. In diversi casi come quelli di Brembilla e di Pontirolo Nuovo (in provincia di Bergamo), il contributo supera i 100 mila Euro. Il Comune cui va la somma più elevata è però quello di Valdidentro (in Provincia di Sondrio), che riceve 330 mila Euro per i suoi 4.029 residenti (di cui 210 unità al di sotto dei 5 anni), quasi 1.600 Euro per bambino.

In questo caso l'incremento della popolazione straniera, ed in particolare dei bambini stranieri, tra il 2002 e il 2008 fa sì che 347 Comuni in più riescano a rientrare nei parametri di legge. È infatti la giovane struttura per età della popolazione immigrata (con un rilevante numero di figli) che fa salire la popolazione in età 0-4 anni al di sopra del 4,5%.

Paradossalmente però, questo effetto strutturale è parzialmente frenato dal fatto che l'incremento della popolazione straniera fa salire anche quella complessiva oltre la soglia dei 5.000 abitanti (oltre i quali non è più possibile ricevere i contributi) in ben 61 casi.

All'ammontare della popolazione residente sono legate molte altre disposizioni giuridiche. Tra queste, le principali riguardano le soglie per l'insediamento di nuove farmacie (ve ne possono essere una ogni 5.000 abitanti nei Comuni fino a 12.500 residenti, una ogni 4.000 nei Comuni più popolosi), la necessità o meno di rispettare il patto di stabilità (limite dei 5.000 abitanti), la tipologia di sistema elettorale per le elezioni amministrative comunali (sistema a turno unico con vittoria da parte della coalizione che riceve la maggioranza relativa nei Comuni al di sotto dei 15.000 abitanti; maggioranza assoluta con eventuale doppio turno nei

Comuni più grandi). Alla dimensione demografica sono legati anche il numero massimo di Assessori e Consiglieri comunali, la remunerazione del Segretario comunale e del Sindaco e, conseguentemente, delle altre cariche politico-amministrative, nonché i gettoni di presenza alle sedute del Consiglio comunale.

Tavola 1 – Distribuzione del contributo agli anziani per regione.

Regione	Anagrafe 01.01.2008			Variazioni non tenendo conto dell'incremento degli stranieri		
	Comuni al di sotto di 5.000 ab		>5.000 ab	Al di sotto di 5.000 ab		>5.000 ab
	Anziani >25%	Anziani <=25%		Anziani >25%	Anziani <=25%	
Piemonte	531	542	133	94	-91	-3
Valle d'Aosta	10	63	1	3	-3	0
Lombardia	176	922	448	86	-60	-26
Trentino Alto Adige	15	293	31	12	-10	-2
Veneto	40	277	264	15	-11	-4
Friuli	45	114	60	12	-10	-2
Liguria	133	50	52	1	-1	0
Emilia Romagna	85	71	185	18	-13	-5
Toscana	81	54	152	23	-21	-2
Umbria	34	26	32	16	-13	-3
Marche	86	92	68	24	-23	-1
Lazio	96	158	124	19	-16	-3
Abruzzo	149	102	54	18	-17	-1
Molise	81	43	12	5	-4	-1
Campania	105	229	217	4	-2	-2
Puglia	25	61	172	1	-1	0
Basilicata	47	51	33	3	-3	0
Calabria	85	239	85	6	-2	-4
Sicilia	85	112	193	9	-8	-1
Sardegna	96	216	65	3	-2	-1
Totale	2.005	3.715	2.381	372	-311	-61

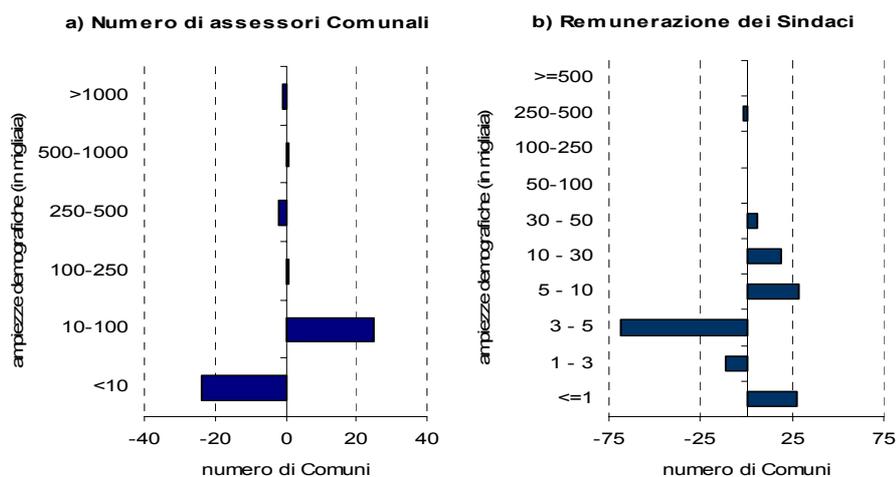
Consideriamo, a titolo esemplificativo, le soglie demografiche relative al numero di assessori e all'indennità di funzione del Sindaco. Il numero massimo di assessori è pari a 4 nei Comuni fino a 10 mila abitanti, a 6 in quelli da 10 mila a 100 mila abitanti, a 10 tra 100 e 250 mila abitanti, a 12 tra 250 e 500 mila abitanti, a 14 tra 500 mila e 1 milione di abitanti, a 16 oltre 1 milione di abitanti. Per effetto dell'incremento nel numero di residenti tra il 2002 e il 2008 ben 48 Comuni (situati prevalentemente del Centro-Nord) hanno superato la soglia demografica di pertinenza al Censimento 2001 (Figura 2a). Questo significa, tra l'altro, l'incremento nel numero di assessori comunali per un totale di 103 unità.

Consideriamo ora l'indennità mensile di funzione del Sindaco. Secondo la legge essa è di 1.291 Euro nelle realtà al di sotto di 1.000 abitanti, di 1.446 in

quelle tra 1.000 e 3.000, di 2.169 tra 3 e 5 mila, di 2.789 tra 5 e 10 mila, di 3.099 tra 10 e 30 mila, di 3.460 tra 30 e 50 mila, di 4.132 tra 50 e 100 mila, di 5.010 tra 100 e 250 mila, di 5.784 tra 250 e 500 mila, di 7.798 nei Comuni con più di 500 mila abitanti.

L'incremento nel numero di stranieri negli ultimi anni fa sì che 115 Comuni passino ad una classe demografica superiore rispetto a quella in cui si trovavano al Censimento 2001 per l'attribuzione dell'indennità mensile di funzione del sindaco (Figura 2b). Questo comporta un incremento della remunerazione aggregata per tutti i Sindaci di circa 1,5 milioni di Euro complessivi annui. L'elemento paradossale che emerge da queste due esemplificazioni riguarda il fatto che la popolazione straniera, priva del diritto di voto, contribuisce in maniera molto significativa all'incremento del numero e della remunerazione degli amministratori locali.

Figura 2 – Ampiezze demografiche rilevanti per: a) la nomina del numero di assessori comunali; b) la remunerazione dei sindaci. Variazione nel numero di Comuni tra il Censimento 2001 e il 2007.



5. Conclusioni

Se venisse effettuato oggi il censimento della popolazione, le ripercussioni in termini di ricaduta sul sistema politico-amministrativo e finanziario dei Comuni sarebbero quelle analizzate.

In realtà, il costante incremento demografico della maggior parte dei comuni amplificherà questi effetti da qui al 2011. Al momento della effettiva rilevazione

censuaria, in particolare, saranno i Comuni del Nord e del Centro che potranno trarre i maggiori benefici da tale redistribuzione demografica.

L'analisi effettuata mostra come spesso il dato di popolazione sia utilizzato dalla normativa vigente senza considerare gli effetti e le contraddizioni prodotte o come i cambiamenti demografici sia scarsamente considerati. Il rapporto tra numerosità della rappresentanza politica e dei rappresentati è particolarmente eloquente: comuni in cui diminuisce o nel migliore dei casi rimane stabile il corpo elettorale (popolazione italiana), aumentano il numero dei loro rappresentanti e dei loro amministratori, mentre coloro che producono tali variazioni (popolazione straniera) rimangono del tutto esclusi dal sistema politico-elettorale, dato che solo per i cittadini comunitari residenti è possibile partecipare alle elezioni amministrative comunali, qualora ne facciano richiesta.

Riferimenti bibliografici

Caritas, 2006. *Immigrazione: dossier statistico 2008. XVIII rapporto sull'immigrazione*, edizioni Idos

Golini, 2006 (a cura di). *L'immigrazione straniera: indicatori e misure di integrazione*, Il Mulino, Bologna

Istat, 2007. *Rapporto annuale. La situazione del Paese nel 2006*, pp. 311-382.

Istat, 2008a. *La presenza straniera in Italia: l'accertamento e l'analisi*. Atti del Convegno di Roma 15-16 Dicembre 2005.

Istat, 2008b. *Previsioni demografiche. 1° Gennaio 2007 – 1° Gennaio 2051*. Nota informativa del 19 Giugno 2008

Istat, 2009. *Rapporto annuale. La situazione del Paese nel 2008*, pp. 241-296.

Istat, 2009, *Popolazione. Bilancio demografico Nazionale*, Comunicato Stampa del 23 Giugno 2009.

SUMMARY

In the latest years the total population resident in Italy has grown up due to immigration of foreigners: from less than 57 millions at 2001 Census, to more than 60 millions at the end of 2008. Many laws and rules are referred to resident population of each municipality, and bring different effects at administrative and economic level. This work aims to illustrate some of the changes introduced from the increase of immigrants, in particular in the Centre and North of the Country.

Angela SILVESTRINI, ricercatore Istat

Alessandro VALENTINI, ricercatore Istat

INDUSTRIA E TERRITORIO TRA ASPETTATIVE LOCALI E ESIGENZE GLOBALI. PRIME EVIDENZE DI UN CASO STUDIO

Valeria Stragapede, Annamaria Stramaglia*

1. Introduzione

Il c.d. distretto del salotto, per anni indicato nella letteratura distrettuale come esempio virtuoso, rappresenta attualmente un caso di crisi strutturale che coinvolge l'intera economia locale con aspetti degenerativi sotto il profilo sociale oltre che produttivo. Partendo da questa considerazione in questo lavoro si presentano alcuni risultati preliminari di un'indagine sul campo condotta con l'idea-obiettivo di dimostrare come la dinamica economica virtuosa che i distretti comportano, possa svilupparsi in maniera distorta fino a trasformarsi in una spirale di sfruttamento, in una corsa all'abbattimento dei costi di produzione creando vincoli alla crescita anche civile e sociale. E' lecito infatti domandarsi se il successo a livello internazionale che i salotti "made in Puglia" hanno registrato, sia stato il meritato risultato di un settore in grado di innovare, porre sul mercato prodotti di elevata qualità e adottare una vincente strategia di marketing, oppure l'esito di una strategia competitiva basata sullo sfruttamento di alcuni operatori a vantaggio di altri, che fa somigliare in modo preoccupante le dinamiche produttive locali a quelle, tanto criticate, adottate in Cina e negli altri paesi a crescita forzata.

E' del tutto evidente che nell'attuale quadro globale la possibilità di sopravvivenza di un settore, che finora è ricorso in misura limitata ad innovazioni e ha basato il proprio successo sulla riduzione dei costi dei fattori mobili, è strettamente dipendente da economie che conseguono i medesimi risultati con costi del tutto incomparabili e per questo attraggono i processi di delocalizzazione che negli ultimi anni molte imprese del distretto hanno deciso di attuare. Il pericolo di spiazzamento da parte delle aree geoeconomiche emergenti è ancor più concreto se si considera che tutt'oggi i tagli dei costi di produzione sono indicati come principale via per la soluzione della profonda crisi che attanaglia il settore.

* Il lavoro è il risultato di una ricerca comune. Tuttavia, V. Stragapede ha curato in modo particolare il par. 2, A. Stramaglia i par. 3 e 4, Introduzione e Conclusioni sono comuni.

2. Le peculiarità del settore

L'attenzione verso il "Triangolo del salotto" è notevolmente cresciuta a partire dagli anni '80, trainata dal successo internazionale delle imprese leader, tanto che si è giunti al riconoscimento istituzionale e alla sua definizione in letteratura di "isola felice nel Mezzogiorno" (Morrison, 2006)¹. Il comparto presenta le caratteristiche tipiche del modello di "economia locale a specializzazione flessibile" (Priore e Sabel, 1987), non è una semplice interconnessione tra una impresa dominante ed altre contigue, ma un insieme di imprese in un contesto di conoscenze e pratiche organizzative tanto sociali quanto produttive, elemento peculiare ed ulteriore che distingue il distretto industriale dal cluster produttivo.

Come per altre aree distrettuali il "triangolo del mobile imbottito" si è affermato grazie alla presenza di fattori favorevoli quali la possibilità tecnica di frazionare il processo produttivo in fasi logicamente separabili e imprenditorialmente autonome - pur legate da un patrimonio condiviso di saperi locali - associata alla disponibilità dal lato dell'offerta a soddisfare i repentini incrementi della domanda assicurando flessibilità qualitativa e quantitativa (Sacchi, 2006). Se tali sono state le premesse su cui è cresciuto il distretto del salotto, è giusto domandarsi quali siano le ragioni per cui si registra oggi una conclamata e profonda crisi. Finora le dinamiche di questo comparto produttivo sono state studiate guardando in modo preferenziale alle strategie di mercato e ai successi a livello internazionale delle poche imprese che si trovano al vertice della struttura piramidale che connota il distretto. Con la presente indagine al contrario, si è inteso porre attenzione alle imprese che sono alla base delle piramidi produttive, nella convinzione che questa analisi bottom-up potesse offrire una chiave interpretativa della crisi attuale.

L'idea di partenza è che sulle micro imprese subfornitrici, spesso non tenute nella dovuta considerazione, sia stato costruito il successo a livello internazionale del distretto, con uno sfruttamento degli anelli più deboli della filiera tanto da provocarne l'asfissia. In letteratura si concorda nell'individuare i punti di forza del comparto nella ibridazione di conoscenze locali con tecniche e principi di organizzazione industriale postfordisti, nel forte controllo dei tempi di ciclo, nella standardizzazione delle fasi, nell'attivazioni di fitte reti di subfornitura e nella metodologia just-in time; scelte peraltro agevolate da una esacerbata competizione tra i subfornitori per via di scarse barriere all'entrata e bassi costi iniziali (Altieri, Belussi, Oteri, 1998). Tuttavia, sebbene il modello distrettuale trae la propria forza dalla divisione del lavoro, dalla cultura locale e dal patrimonio di competenze che sono sedimentate sul territorio, esso si distingue dalla tradizionale impresa di produzione di matrice fordista e ne è in qualche modo un'evoluzione, per la sua dinamica auto-organizzativa. La grande impresa di produzione introduceva le innovazioni con modalità gerarchiche: i centri di

¹ Regione Basilicata, l.r. n.1 del 23.01.2001.

produzione della conoscenza la trasferivano ad esecutori materiali privi della capacità di interazione e retroazione che ne qualificassero l'attuazione; i distretti, invece, hanno successo quando rivelano capacità evolutiva e sviluppano un ciclo di trasformazione del sapere locale, generando varietà, selezionando le conoscenze, diffondendo le innovazioni e replicando le soluzioni migliori.

Il successo dei distretti risiede nella capacità di trasformarsi in modo coerente con gli stimoli provenienti dal mercato, come ogni sistema basato sulle conoscenze è mosso da processi di cambiamento negli assetti organizzativi e di marketing, nelle tecnologie e nelle competenze necessarie al loro utilizzo nonché nei legami delle imprese con l'ambiente locale (Corò, Micelli, 2007). La strategia produttiva adottata nel distretto del salotto, sottolinea la forte attenzione verso obiettivi di abbassamento dei costi e massima accelerazione dei processi produttivi chiaramente perseguiti a spese delle micro imprese subfornitrici, che evidentemente, *supportano* la piramide produttiva e con tutta probabilità *sopportano* i contraccolpi derivanti dalla esasperata competitività che i processi di globalizzazione comportano.

A riprova di tale idea si riscontra negli ultimi anni un significativo processo di delocalizzazione all'estero favorito da bassi costi della manodopera e cambio favorevole. La lavorazione dei divani infatti, vede una ripartizione dei processi produttivi in fasi distinte: il design e lo studio del modello del salotto, produzione dei fusti in legno (intelaiature), produzione dell'imbottito per i cuscini e delle spugne per la seduta, taglio e cucito della pelle o di qualsiasi altro rivestimento ed, infine, assemblaggio ed imballaggio. In conformità ad un meccanismo che solitamente si riscontra nell'evoluzione naturale delle attività manifatturiere in ambito distrettuale, anche nel distretto del salotto con gli anni le fasi ad alto contenuto di manodopera sono state esternalizzate. Tale processo ha riguardato soprattutto le fasi di realizzazione dei fusti e dell'imbottito e spesso anche le attività di taglio e cucito dei rivestimenti; mentre sono rimaste appannaggio delle imprese al vertice della piramide distrettuale, o meglio gelosamente custodite per limitare il flusso di informazioni a più alto valore aggiunto, le funzioni legate alle strategie commerciali, amministrative, vendita e approvvigionamento e, design e scelta di nuovi modelli. Per sua stessa natura pertanto, il distretto si presta ad una scelta di delocalizzazione delle attività a più scarso valore aggiunto che in passato hanno interessato le imprese subornatrici del territorio circostante alle imprese leader, ma che negli ultimi anni si stanno indirizzando verso i Paesi dell'estremo oriente o dell'est europeo dove, seppure ponderati con una minore produttività, i costi del lavoro sono notevolmente più bassi anche grazie al tasso di cambio favorevole che riduce i costi degli investimenti all'estero e al contempo aumenta la profittabilità di importazioni di componenti e fasi di lavorazione. E' fondato, quindi, il rischio che tali politiche di delocalizzazione spiazzino tutto l'indotto costituito dai subornatori specializzati, con conseguenti forti tensioni sociali e aumento della disoccupazione in un'area già sofferente da questo punto di vista.

3. Il comparto del salotto nella provincia di Bari: alcune evidenze

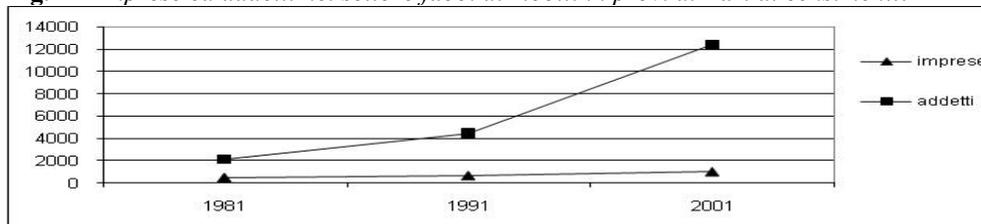
Il sistema economico regionale presenta un forte squilibrio causato dal fatto che la quota più rilevante della produzione manifatturiera (c.ca 40%) deriva dai due comparti, il tessile calzaturiero e quello del salotto imbottito, maggiormente in crisi negli ultimi anni sotto la pressione della concorrenza internazionale.

Il comparto del salotto imbottito pugliese, dopo una fase di grande sviluppo durata c.ca due decenni, sta vivendo, negli ultimi anni, una grave crisi le cui matrici sono essenzialmente due, l'ormai nota concorrenza internazionale, in particolare cinese, più competitiva di prezzo, ed il cambio dollaro – euro non favorevole alle esportazioni che ha fatto perdere rilevanti quote di mercato negli Stati Uniti.

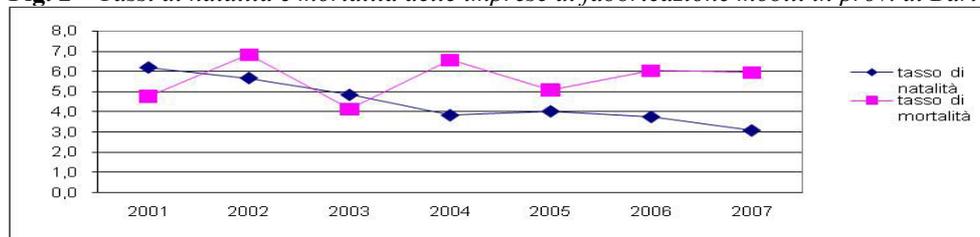
Dei due differenti andamenti il primo, positivo, di lungo periodo, ed il secondo manifestatosi da alcuni anni e che solo negli ultimissimi sta coinvolgendo anche le imprese leader del settore, possono essere assunti ad indicatori rispettivamente i dati censuari e quelli relativi alla demografia di impresa, informazioni che in prima battuta forniscono una, seppur grossolana, dimensione del fenomeno.

Il settore della fabbricazioni di mobili nella provincia di Bari, nell'ambito del quale la produzione di poltrone e divani rappresenta la quota più rilevante, ha subito una crescita estremamente significativa tra il 1981 ed il 2001 sia in termini di imprese attive che, soprattutto di addetti. In particolare la sua importanza risiedeva proprio in quest'ultimo aspetto, si tratta di un settore che nell'arco di venti anni ha incrementato il numero di addetti di oltre 10.000 unità. Dall'esame dei tassi di natalità e mortalità delle imprese afferenti al settore fabbricazione di mobili relativi agli ultimi anni, emerge la negativa performance che il settore sta sperimentando: le nuove iniziative imprenditoriali risultano in costante calo a fronte di un tasso di mortalità tendenzialmente in crescita. Ciò appare espressione in primo luogo della percezione, da parte di potenziali imprenditori locali, di minori opportunità offerte dall'operare nel settore, in secondo luogo l'andamento del tasso di mortalità delle imprese, in netta controtendenza rispetto a quanto emerge fermandosi all'ultima fonte censuaria disponibile, evidenzia la presenza delle difficoltà che il settore sta recentemente sperimentando.

Fig. 1 – *Imprese ed addetti nel settore fabb. di mobili in prov. di Bari ai censimenti.*



Fonte: Istat

Fig. 2 – Tassi di natalità e mortalità delle imprese di fabbricazione mobili in prov. di Bari

Fonte: elaborazioni su dati Movimpresa

4. Primi risultati del caso studio

L'indagine –rilevazione diretta mediante questionario strutturato - ha coinvolto complessivamente 125 imprese operanti nella provincia di Bari nell'ambito del comparto del salotto, si tratta sia di imprese specializzate per fasi, che forniscono solo i fusti in legno, le sole imbottiture o specializzate nella fase di assemblaggio di componenti, che di salottifici in senso stretto.

La quota più rilevante delle imprese rilevate (86,4%) ha iniziato l'attività dopo il 1996, in particolare risulta rilevante la quota di imprese che ha avviato l'attività solo da pochi anni, in quanto sintomo di una forma di ristrutturazione dell'organizzazione territoriale del comparto in risposta alla difficoltà di gestire imprese di dimensioni più consistenti in una fase di crisi del settore. Tali considerazioni sono suggerite dall'esame congiunto dell'età delle imprese rilevate e dell'attività svolta dall'imprenditore prima di avviare l'impresa. In particolare nel 48% dei casi le imprese rilevate rappresentano degli spin-off aziendali, ovvero frutto dell'iniziativa imprenditoriale di un ex dipendente di un'azienda operante nel medesimo settore. In linea generale gli spin-off aziendali possono derivare dall'iniziativa del lavoratore che decide di porsi in concorrenza con il suo ex datore di lavoro, combinando le conoscenze e l'esperienza acquisite con le abilità innate che lo orientano positivamente verso l'attività imprenditoriale, ovvero indotti dalle stesse imprese madri che incentivano dirigenti e/o lavoratori specializzati a dar vita ad una impresa indipendente (o parzialmente indipendente) in particolare in momenti di forte crescita così da evitare, ad esempio, il rischio connesso ad eccessiva burocratizzazione e gerarchizzazione dell'impresa o piuttosto in periodi di crisi, quale risposta ad esigenze di ristrutturazione. Nel caso in esame buona parte degli spin-off sono stati indotti, incentivati con la garanzia di commesse future e, talvolta, sostenuti finanziariamente nella fase di avvio.

Tali imprese paiono espressione di una precisa strategia difensiva esercitata dall'impresa madre in periodi di domanda instabile e talvolta proposti come alternativa, per il lavoratore/potenziale neo imprenditore, alla partecipazione a forme di ristrutturazione aziendale. Si tratta di strategie che rispondono all'esigenza di scaricare i costi, tra i quali si ricorda il lavoro rappresenta la quota più rilevante, sulle imprese subfornitrici.

Nel caso della ricerca in oggetto le informazioni raccolte direttamente presso le imprese delineano un quadro che suggerisce l'applicazione non tanto, o non solo, di strategie di competitività, ma quanto piuttosto il ricorso a tutti gli strumenti possibili di contenimento dei costi in un periodo di grave difficoltà. Ciò trova ulteriore riscontro quantitativo nella distribuzione delle imprese per numero di dipendenti. In linea generale prevale piccola e piccolissima dimensione, in maniera particolarmente significativa per le imprese sorte dopo il 1996, circostanza che può essere espressione di una vocazione artigiana del settore ma anche sintomatica della fragilità stessa del settore con buona parte delle imprese che mantengono una dimensione minima al fine di limitare i costi in un periodo si domanda instabile, ricorrendo al lavoro sommerso per far fronte a commissioni per le quali le risorse interne sono insufficienti.

Altro tratto caratteristico delle imprese in oggetto è l'analisi delle funzioni svolte e l'ambito nel quale le stesse vengono svolte. In primo luogo si rileva che in numerosi casi ci sono funzioni aziendali che non vengono gestite a livello aziendale, in particolare si tratta della R&S, dell'acquisto di materie prime e di macchinari, del trasporto e logistica. Tale circostanza è per certi versi sintomatica della tipologia prevalente di imprese, c/terziste specializzate su una specifica fase del ciclo produttivo, talvolta per una sola impresa committente, che le rende unità produttive "semplici" per le quali risulta imprescindibile il legame con l'impresa committente, preposta allo svolgimento di tutte le funzioni, e pertanto molto più esposte a subire le conseguenze di andamenti incostanti della domanda.

Inoltre si rileva una generalizzata tendenza a svolgere gran parte delle funzioni aziendali all'interno dei confini stessa dell'azienda, è infatti scarsamente rilevante la tendenza ad avvalersi delle specializzazioni presenti nell'ambito distrettuale/provinciale. Allo stesso modo, in ragione presumibilmente della opzione effettuata di esaminare la base del comparto, non si registrano imprese che svolgono le principali funzioni in ambito internazionale, ovvero imprese che hanno ad esempio delocalizzato all'estero specifiche fasi. Eccezione è rappresentata dalla funzione di trasporto, internalizzata solo in caso di imprese c/terziste per un unico committente locale. Nella maggioranza dei casi il trasporto è affidato all'esterno a vettori operanti spesso nell'ambito territoriale del cliente.

5. Conclusioni

Sebbene la ricerca in oggetto sia ancora in corso, soprattutto per via dell'evoluzione che il settore sta vivendo, i risultati preliminari riportati in questo lavoro suggeriscono alcune riflessioni di interesse poiché forniscono un'immagine della base della piramide produttiva precedente alla capillare diffusione degli effetti della crisi. In particolare è emerso che l'ex distretto sta proseguendo lungo una linea di intensificazione dei ritmi produttivi e di disgregazione delle grandi imprese, mediante la creazione forzata di piccole nuove unità locali in grado di conservare i vantaggi connessi alla propria dimensione, tuttavia, numerosi segnali indicano che tali vantaggi potrebbero non essere

sufficienti per la sopravvivenza del distretto in un contesto di crescente competizione globale ed in una fase congiunturale sfavorevole quale quella corrente, anche a causa della tendenza a decentrare non il processo di produzione ma le fasi che lo compongono.

Condividendo l'assunto per cui la crescente globalizzazione comporta la rivalutazione della dimensione locale e regionale dei vantaggi competitivi, è necessario che le realtà distrettuali recuperino la propria flessibilità e capacità adattiva cogliendo le sfide della competizione globale. Il peso della tradizione del distretto di tipo manifatturiero, rende più difficile nel nostro Paese riconoscere come elemento di vitalità del modello un ripensamento in termini organizzativi, logistici e tecnologici. Come conferma la letteratura, anche nei settori più tradizionali è possibile creare delle filiere che sviluppino congiuntamente innovazioni complementari, dei "consorzi tecnologici" che possano rappresentare le forme più evolute di nuovi distretti produttivi (Baumol 2002) e numerosi sono i contributi scientifici in tema di capacità innovativa delle imprese inserite in realtà distrettuali (cfr. Baptista, 2001, Morrison, 2005). Un ulteriore sostegno a questa forma di rinnovamento verrebbe anche dalla Commissione Europea che nel Settimo Programma Quadro ha previsto forme di finanziamenti per imprese e istituzioni che effettuino investimenti congiunti su progetti di innovazione. E' necessario prendere atto che nelle economie avanzate il lavoro non è più subordinazione ad una norma tecnica, ma è capacità di auto-organizzazione e di apprendimento, disponibilità al miglioramento continuo e alla ricerca di soluzioni innovative. Queste sono caratteristiche legate alle attività a maggior valore aggiunto: progettazione, design, ricerca e sviluppo tecnologico, comunicazione e marketing, distribuzione e finanza sulle quali sempre più si basa il vantaggio competitivo.

Il distretto del salotto, per superare questa difficile crisi e crescere, dovrà rispondere a queste nuove esigenze (Corò, Micelli 2007). Al contrario, nella realtà in esame, l'impresa leader non guida più il processo di proiezione internazionale del distretto, viceversa sembra favorirne la de-distrettualizzazione introducendo elementi di rivalità tra i fattori su cui costruisce il proprio successo e, quindi creando barriere basate non sulla conoscenza e sulle attività a maggior valore aggiunto che si guarda bene dal porre nel giro della cooperazione che contraddistingue il distretto, ma sulla competizione tra gli stakeholders più deboli facendo così venire meno il principio di base della distrettualizzazione produttiva, cioè quello della condivisione del rischio e della cooperazione interna, così come della competitività del "sistema" rispetto all'esterno. Si configura così, una sorta di strategia predatore-preda applicata alla scala territoriale ed un modello di specializzazione produttiva, molto prossimo al modello fordista.

Riferimenti bibliografici

- Altieri G., Belussi F., Oteri C. (1998), *Il distretto industriale dell'imbottito localizzato nell'area Matera, Altamura, Santeramo*, in IRES Nazionale "Adapt Learning Region".
Baptista R. (2001), *Geographical Clusters and Innovation Diffusion*, in "Technological

Forecasting and Social Change”, vol.66.

Baumol W. (2002), *The Free Market Innovation Machine*, Princeton University Press.

Corò G., Micelli S.(2007), *The industrial districts as local innovation systems*, in “Review of Economic Conditions in Italy, vol.1.

Morrison, A. (2006), *Le determinanti della performance innovativa delle imprese distrettuali: capitale relazionale vs absorptive capacity*. In: Scienze Regionali – The Italian Journal of Regional Science, n.1.

Priore M. J., Sabel C.F. (1987), *Le due vie dello sviluppo industriale: produzione di massa e produzione flessibile*, ISEDI Torino.

Sacchi S. (2006), *Distretti, cluster, o ..? Esperimenti di analisi per le politiche di sviluppo locale*, in “Rivista dell’Agenzia Umbria Ricerche”, Anno 3, n.5-6.

SUMMARY

The aim of this paper is to study the links between a high level of global competition and the connected crisis of productions and traditional model of industrial organization. Empirical analysis has been carried out on a sample of firms operating in the Murge sofa district named in the past as case of success of the “made in Italy”.

Annamaria STAMAGLIA – Ricamatore di Economia Politica – Università di Foggia

Valeria STRAGAPEDE – Dottoranda di ricerca – Università di Palermo

INFORMAZIONI GENERALI

1. La collaborazione alla Rivista è aperta soltanto ai Soci della Società Italiana di Economia Demografia e Statistica e a coloro che vengono invitati a collaborare dalla Direzione.
2. I lavori per la pubblicazione vanno presentati alla Direzione della Rivista, adottando le regole indicate di seguito, con l'indicazione dei rispettivi titoli accademici ed eventuali qualifiche.
3. I lavori inviati verranno sottoposti al vaglio di un Comitato di lettura. Quelli non accettati per la pubblicazione non si restituiscono.
4. *Gli autori degli articoli e delle note pubblicate rinunciano, in favore della Rivista, alla proprietà letteraria.* La riproduzione integrale o parziale di articoli e note pubblicate nella Rivista è consentita alla condizione che risulti citata la fonte. Gli autori ricevono, gratuitamente, 15 estratti. Le spese per tirature supplementari o per stampa anticipata degli estratti sono a totale carico degli Autori.
5. La Rivista è distribuita ai Soci in regola col pagamento della quota di associazione.
6. La Rivista pubblica recensioni o segnalazioni delle opere ricevute in omaggio o in cambio.

Le quote di adesione alla S.I.E.D.S. sono le seguenti:

Soci individuali: Euro 90

Soci collettivi: Euro 150

Soci vitalizi, secondo accordi

Il versamento della quota sociali può essere effettuato tramite:

* c/c postale n. 42223008 intestato a Società Italiana di Economia Demografia e Statistica – Roma;

* bonifico bancario sul c/c n. **000400852787** presso Unicredit Banca di Roma AG. 153 Città Universitaria, Piazzale Aldo Moro, 5 – 00184 Roma;

CODICE IBAN: IT97 D 03002 03371 000400852787

INFORMAZIONI PER GLI AUTORI REGOLE PER LA COMPOSIZIONE DEI TESTI

Foglio formato A4. Spazio riservato al testo rettangolo di 13x18 cm. Impostazione margini: Superiore 5,8 cm; Inferiore 5,8 cm; Sinistro 4 cm; Destro 4 cm.

Il **titolo** deve essere scritto in **Times New Roman 12 punti**

Il **testo** deve essere scritto in **Times New Roman 11 punti**

Interlinea singola o 0,5 cm o 12 punti

Rientro prima riga (**capoverso**) **0,5 cm**

La numerazione delle pagine deve essere fatta fuori del rettangolo 13x18.

Testo, note, riferimenti bibliografici e appendici in Times New Roman 10 punti

Interlinea delle note e dei riferimenti bibliografici 10 punti.

PRIMA PAGINA

Nella prima pagina va indicato il titolo del lavoro (massimo 3 righe) lasciando una linea bianca, in carattere Times New Roman 12 punti, tutto maiuscolo, in grassetto centrato.

Il Nome e Cognome dell'autore va scritto in carattere Times New Roman 11 punti (come il testo), sotto il titolo lasciando una linea bianca, maiuscolo e centrato.

Il testo deve iniziare a cm 5 dalla prima riga del rettangolo di cm 13x18 e deve essere **giustificato**.

PARAGRAFI

I paragrafi devono essere numerati progressivamente con numeri arabi seguiti dal punto e dal titolo. Sia il numero, sia il titolo devono essere in grassetto e allineati a sinistra. Esempio:

1. Introduzione

2. Metodi statistici

I sottoparagrafi devono essere numerati progressivamente all'interno del paragrafo, sempre in numeri arabi e il titolo in corsivo. Analogamente al paragrafo vanno allineati a sinistra. Esempio:

3.1 Descrizioni dati

3.2 Strategie di modellazione

Sia i paragrafi, che i sottoparagrafi devono essere preceduti da due linee bianche e seguiti da una linea bianca; dopo il paragrafo (sotto-paragrafo) si inizia con un rientro (capoverso). Per le ulteriori numerazioni si procede in modo analogo.

FORMULE

Le formule devono essere numerate progressivamente con numeri arabi tra parentesi. Il numero deve essere allineato a destra. Il richiamo avviene nel testo tramite il numero della formula. Esempio: "...come espresso nella (7)..."

NOTE

Le note al testo devono essere numerate consecutivamente, ridotte al minimo, e riportate a piè di pagina. Il numero della nota nel testo va scritto in esponente con carattere più piccolo. Esempio: "...come affermato¹..."

TABELLE

Le tabelle devono essere numerate consecutivamente con numeri arabi. Le tabelle devono essere contenute nel formato 13x18 cm, usando anche caratteri più piccoli. L'intestazione inizia con **Tabella n** (in grassetto) allineata a sinistra e il titolo è separato da un trattino e scritto in corsivo. Esempio:

Tabella 2 – *Stranieri residenti in Italia nei censimenti del 1991 e del 2001.*

FIGURE E GRAFICI

Le figure devono essere numerate consecutivamente con numeri arabi. Le figure e i grafici devono essere stampati in **bianco nero e non a colori**. L'intestazione inizia con **Figura n** (in grassetto) allineata a sinistra e il titolo è separato da un trattino e scritto in corsivo. Esempio:

Figura 1 – *Funzione integrata di rischio per maschi e femmine.*

RIFERIMENTI BIBLIOGRAFICI

Le citazioni bibliografiche nel testo si effettuano con il nome dell'autore o degli autori e l'anno. Esempio: Cicchitelli, Herzel e Montanari (1992), oppure (Cicchitelli, Herzel, Montanari, 1992).

Dopo le conclusioni seguono i Riferimenti bibliografici (in grassetto) e allineato a sinistra. I riferimenti devono essere in ordine alfabetico come segue: il cognome precede il nome puntato e l'anno di pubblicazione. Il titolo del lavoro deve essere scritto in corsivo. Esempio:

Riferimenti bibliografici

Barr S.R. e Tuner J.S. 1990. *Qualità Issues and Evidence Statistical File*. In: Lieping G.E. e Uppuluri V.R.R. (eds), *Data Quality Control. Theory and Pragmatics*, pp 245-313, Marcel Dekker, New York.

Cicchitelli G., Herzel A. e Montanari G.E. 1992. *Il campionamento statistico*. Il Mulino, Bologna.

Trivellato U. 1990. *Modelli di comportamento e problemi di misura nelle scienze sociali: alcune riflessioni*, Atti della XXXV Riunione Scientifica della S.I.S., vol. 1, pp 11-31, Cedam, Padova.

APPENDICI

Dopo i riferimenti bibliografici seguono le appendici (lasciare soltanto due righe bianche dai riferimenti bibliografici) che devono essere numerate con numeri arabi. L'intestazione inizia con Appendice n (in grassetto) allineata a sinistra e il titolo è separato da un trattino e scritto in corsivo. Esempio:

Appendice 1 – *Eventuale titolo*

Appendice 2 – *Eventuale titolo*

SUMMARY

In fondo al testo segue il riassunto in inglese (summary) che deve essere contenuto in 1 pagina del formato 13x18 cm. Il titolo, in carattere 11 punti, deve essere scritto minuscolo, in grassetto e centrato.

TITOLI ACCADEMICI E ALTRO

Indicare in fondo al testo i titoli accademici o altro al momento della presentazione del testo.

Consigli utili:

- per evitare inutili sprechi nell'impaginazione definitiva del volume della rivista non inserire pagine o spazi bianchi;
- usare il carattere **grassetto neretto** con parsimonia;
- evitare il più possibile le maiuscole, nelle sigle di enti o altro usare il carattere MAIUSCOLETTO;
- per le parole straniere inserite nel testo usare il carattere corsivo;
- usare le virgolette soltanto quando si riportano brani originali;
- per gli elenchi puntati, i sottoelenchi ecc. non inserire il rientro del capoverso (0,5 cm), usare eventualmente i diversi tipi di elenchi numerati messi a disposizione dal programma;
- nella riquadratura delle tabelle inserire semplicemente quei bordi orizzontali ($\frac{1}{4}$ pt) che servono per una migliore lettura della tabella ed evitare i bordi verticali e troppo marcati;
- inserire tabelle, grafici e immagini dopo il punto di fine capoverso;
- spostare in fondo al testo tabelle, grafici e immagini che non rientrano nel formato e necessitano quindi di una riduzione fotografica, segnalando l'inserimento a tutta pagina nel testo.

INVIARE IL TESTO IN LINGUAGGIO WORD PER WINDOWS ALL'INDIRIZZO E-MAIL DI VOLTA IN VOLTA INDICATO

Esempio: lasciare una linea bianca

**STUDI EMPIRICI SULLE PICCOLE
E MEDIE IMPRESE ITALIANE**

Mario Rossi

Prima riga del testo (a cm 5 del formato 13x18)

SOCIETÀ E RIVISTA ADERENTI AL SISTEMA ISDS
ISSN ASSEGNATO: 0035-6832

Direttore Responsabile: Prof. ENRICO DEL COLLE

Iscrizione della Rivista al Tribunale di Roma del 5 dicembre 1950 N. 1864



Associazione all'Unione Stampa Periodica Italiana

TRIMESTRALE

Spedizione in abbonamento postale art. 2 comma 20/c legge 662/96 - Filiale di Roma

La copertina è stata realizzata da Pardini, Apostoli, Maggi p.a.m.@tin.it - Roma