

Volume LXVI N. 1
Gennaio-Marzo 2012

Rivista Italiana di Economia Demografia e Statistica

1



SOCIETÀ ITALIANA DI
ECONOMIA
DEMOGRAFIA
E STATISTICA

Organo della
Società italiana
di Economia Demografica
e Statistica

Poste Italiane S.p.A.
Spedizione in abbonamento postale -70% DCB Roma

SIEDS
SOCIETÀ ITALIANA
DI ECONOMIA DEMOGRAFIA E STATISTICA

CONSIGLIO DI PRESIDENZA

Presidente Onorario: LUIGI DI COMITE

Presidente: GIOVANNI MARIA GIORGI

Vice Presidenti: GIAN CARLO BLANGIARDO, ENRICO DEL COLLE,
OLGA MARZOVILLA

Segretario Generale: CLAUDIO CECCARELLI

Consiglieri: GIOVANNI CARIANI, FRANCESCO CHELLI, ANGELO DELL'ATTI,
PIERPAOLO D'URSO, MARGHERITA GEROLIMETTO,
DOMENICA FIORDISTELLA IEZZI, VENERA TOMASELLI, ROBERTO ZELLI

Segretario Amministrativo: FABIO FIORINI

Revisori dei conti: MATTEO MAZZIOTTA, ALESSANDRO POLLI, DOMENICO SUMMO

Revisori dei conti supplenti: STEFANIA GIRONE, GIUSEPPE NOTARSTEFANO

SEGRETERIA GENERALE:

Via delle Cave di Pietralata, 13 – 00157 ROMA

TEL. e FAX 06-43589008, E-mail: sieds@tin.it; rivista.sieds@gmail.com

ORGANO DELLA SOCIETÀ:

RIVISTA ITALIANA DI ECONOMIA DEMOGRAFIA E STATISTICA



Volume pubblicato con il contributo dell'Istituto Nazionale di Statistica
e l'Istituto Banco di Napoli



VOLUME FUORI COMMERCIO – DISTRIBUITO GRATUITAMENTE AI SOCI

VOLUME LXVI – N. 1



GENNAIO-MARZO 2012

RIVISTA ITALIANA DI ECONOMIA DEMOGRAFIA E STATISTICA

COMITATO SCIENTIFICO

Prof. LUIGI DI COMITE, Prof. GIOVANNI MARIA GIORGI,
Prof. ALBERTO QUADRIO CURZIO, Prof. CLAUDIO QUINTANO,
Prof.ssa SILVANA SCHIFINI D'ANDREA, Prof. GIOVANNI SOMOGYI

COMITATO DI DIREZIONE

Dott. CLAUDIO CECCARELLI, Prof. GIAN CARLO BLANGIARDO, Prof. ENRICO DEL COLLE,
Prof. PIERPAOLO D'URSO, Prof.ssa OLGA MARZOVILLA, Prof. ROBERTO ZELLI

DIRETTORE

Dott. CLAUDIO CECCARELLI

REDAZIONE

Dott. RAFFAELE FERRARA, Dott.ssa MARGHERITA GEROLIMETTO,
Dott.ssa CHIARA GIGLIARANO, Dott.ssa STEFANIA GIRONE,
Dott. ALESSIO GUANDALINI, Dott. MATTEO MAZZIOTTA



Direzione, Redazione e Amministrazione
Via delle Cave di Pietralata, 13
00157 ROMA
TEL. e FAX 06-43589008
E-mail: sieds@tin.it; rivista.sieds@gmail.com

IN QUESTO NUMERO

Questo volume accoglie una selezione, effettuata in sede di referaggio, delle comunicazioni dei Soci presentate in occasione della XLVIII Riunione Scientifica della Società Italiana di Economia, Demografia e Statistica tenutasi a Roma dal 26-28 maggio 2011, sul tema “150 Anni di Statistica per lo Sviluppo del Territorio: 1861-2011”. La Riunione Scientifica è stata organizzata dall’Istituto Nazionale di Statistica e dalla “Sapienza” Università di Roma.

Un sentito e doveroso ringraziamento va all’opera dei referee per l’importanza e l’accuratezza del lavoro svolto.

Claudio Ceccarelli

INDICE

| | |
|------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Anna Maria Altavilla, Francesca Galizia, Angelo Mazza <i>Indicatori di carico demografico ed invecchiamento della popolazione.....</i> | 7 |
| Anna Maria Altavilla, Angelo Mazza, Antonio Punzo <i>On the upward bias of the dissimilarity index</i> | 15 |
| Lucrezia Andria, Roberta Pace, Alessandro Polli <i>Dalla stagnazione alla crescita sostenuta: un’evoluzione lunga oltre un secolo.....</i> | 21 |
| Francesco Antonio Anselmi <i>Governare il turismo tra globale e locale agli inizi del XXI secolo.....</i> | 29 |
| Fabrizio Antolini <i>Pil, benessere economico e distribuzione del reddito: il caso dell’Italia.....</i> | 37 |
| Vincenzo Asero, Rosario D’Agata, Venera Tomaselli <i>Mercato turistico e organizzazione del territorio: i distretti turistici in Sicilia.....</i> | 45 |

| | |
|-------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Nidia Batic | |
| <i>Fattori generazionali e di genere nel rapporto con il fumo</i> | 53 |
| Gian Carlo Blangiardo, Simona Maria Mirabelli | |
| <i>L'orientamento del mondo cattolico su demografia e famiglia. Un indice per confrontare realtà e problematiche nel panorama europeo</i> | 61 |
| Cinzia Castagnaro, Antonella Guarneri, Claudia Iaccarino | |
| <i>Matrimoni con almeno uno sposo straniero: evoluzione del fenomeno e nuove tendenze. Il ruolo del contesto normativo</i> | 69 |
| Claudio Ceccarelli, Antonio R. Discenza, Francesca Fiori, Carlo Lucarelli | |
| <i>Weighting issues in LFS longitudinal data</i> | 77 |
| Claudio Ceccarelli, Giovanni Maria Giorgi, Alessio Guandalini | |
| <i>Stimatori di ponderazione vincolata nel caso di totali noti provenienti dalla stessa indagine campionaria</i> | 85 |
| Cinzia Conti, Domenico Gabrielli, Salvatore Strozza | |
| <i>Dati amministrativi per le statistiche ufficiali sulle migrazioni</i> | 93 |
| Andrea Cutillo, Daniela Lo Castro, Isabella Siciliani | |
| <i>Popolazione a rischio di povertà e intervento pubblico: confronti tra paesi europei</i> | 101 |
| Rosario D'Agata, Venera Tomaselli | |
| <i>Misurare l'integrazione: una proposta di indicatore composito</i> | 109 |
| Giuseppe De Bartolo, Manuela Stranges | |
| <i>Sull'inchiesta del CISP del 1938 su alcune popolazioni albanofone della Calabria</i> | 117 |
| Stefania Della Queva, Daniela Bonardo, Alessio Menonna, Flavio Verrecchia | |
| <i>Tra autoimpiego e micro-imprenditorialità: tratti dell'immigrazione in Lombardia</i> | 125 |
| Pasquale De Muro, Matteo Mazziotta, Adriano Pareto, Valentina Talucci | |
| <i>Un indice di sviluppo umano delle regioni dell'unione europea</i> | 133 |

| | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Grazia Di Bella, Daniela Bonardo, Lorena Galiè, Valentina Talucci <i>L'utilizzo degli archivi amministrativi per il monitoraggio delle politiche del lavoro: il ruolo delle imprese</i> | 141 |
| Agostino Di Ciaccio, Giovanni Maria Giorgi <i>Una nuova procedura di imputazione di dati mancanti basata sugli alberi di decisione</i> | 149 |
| Maria Grazia Didonna, Gianfranco Servodio <i>Scenari occupazionali di un'Italia che cambia</i> | 157 |
| Marco Giansoldati, Claudio Pizzi <i>Fragmentation of production: a cointegration analysis</i> | 165 |
| Marika La Rosa, Daniela Corso, Paola M. Chiodini, Francesca Cuppone, Giancarlo Manzi, Bianca M. Martelli, Flavio Verrecchia <i>Particolari forme di economia del crimine: il business della pedofilia on line</i> | 173 |
| Fabio Lipizzi <i>La distribuzione geografica della popolazione italiana nei censimenti dal dopoguerra ai giorni nostri</i> | 181 |
| Silvia Lombardi, Paola Naddeo <i>Il ruolo delle microimprese nei processi di entrata e uscita dal mercato del lavoro</i> | 189 |
| Eleonora Mussino, Giuseppe Gabrielli, Anna Paterno, Salvatore Strozza, Laura Terzera <i>Il comportamento riproduttivo e le strategie migratorie degli stranieri presenti in Lombardia</i> | 197 |
| Marilena Stigliano <i>In tema di paesi a sviluppo minimo</i> | 205 |
| Agata Zirilli, Angela Alibrandi, Massimiliano Giacalone <i>La soddisfazione degli studenti universitari in merito agli esami sostenuti: un'analisi statistica su due atenei siciliani</i> | 213 |

INDICATORI DI CARICO DEMOGRAFICO ED INVECCHIAMENTO DELLA POPOLAZIONE¹

Anna Maria Altavilla, Francesca Galizia, Angelo Mazza

1. Introduzione

Da alcuni decenni, in numerosi paesi, l'invecchiamento della popolazione suscita preoccupazione sulla sostenibilità dei sistemi pensionistici (OECD 1998). Se fino alla fine del XIX secolo l'assistenza agli anziani si realizzava attraverso trasferimenti volontari all'interno stesso delle famiglie o delle comunità, nelle società contemporanee la maggior parte dell'assistenza agli anziani è erogata attraverso strutture pubbliche.

Nei paesi in cui tali misure sono adottate, compresa l'Italia, queste hanno certamente contribuito a ridurre i differenziali di reddito tra la popolazione attiva e quella non più tale; tuttavia, il progressivo invecchiamento della popolazione, insieme al metodo di finanziamento delle pensioni, generalmente fondato su trasferimenti trasversali, tipici dei sistemi a ripartizione, ha reso il sistema sempre più oneroso al crescere dell'importanza numerica degli anziani.

In Italia, al problema della sostenibilità economica e sociale della spesa pensionistica a livello nazionale si sovrappone quello sugli effetti della progressiva redistribuzione delle risorse in ambito regionale derivante dall'attuazione del federalismo fiscale, dato che il processo di invecchiamento non si è realizzato e, presumibilmente, non si realizzerà in maniera omogenea in tutte le regioni del territorio italiano.

In questo lavoro, attraverso opportuni indicatori, si vuole misurare l'attuale livello d'invecchiamento delle regioni italiane e verificare, attraverso proiezioni sui trend demografici, la diversa evoluzione del loro profilo demografico.

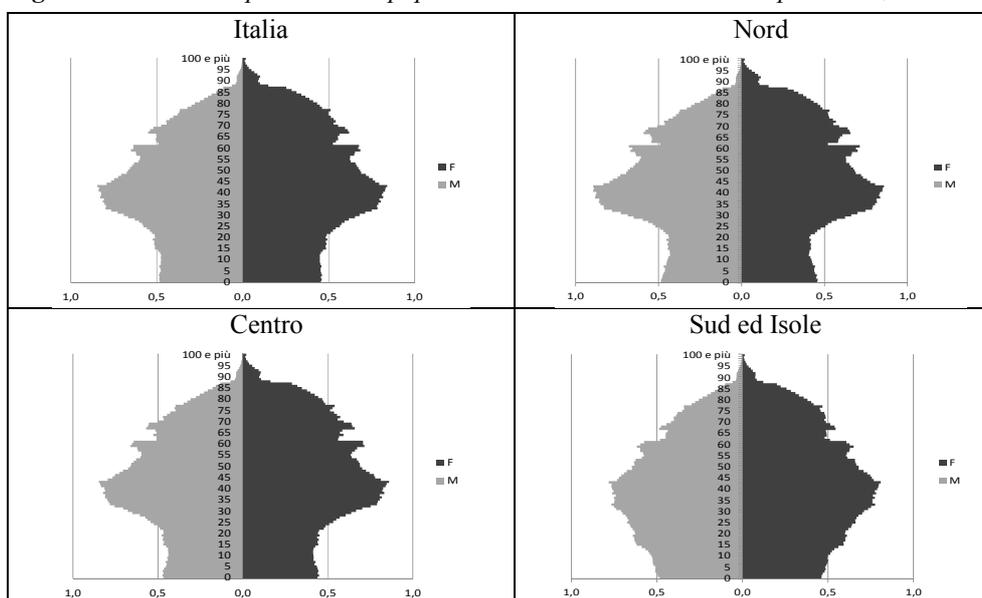
¹ Il lavoro è stato svolto dai tre Autori in stretta collaborazione: tuttavia, per quel che concerne la stesura dello stesso, ad Anna Maria Altavilla va attribuita l'introduzione, a Francesca Galizia il punto 2 e le conclusioni, ad Angelo Mazza i punti 3 e 4.

2. L'invecchiamento demografico in 150 anni di storia: l'Italia divisa

L'Italia ha visto cambiare profondamente la propria popolazione nell'arco degli ultimi 150 anni di storia. Un cambiamento demografico che ha interessato non solo la consistenza numerica della popolazione, ma anche la sua struttura per età.

Al tempo dell'unificazione fecondità e mortalità erano ancora ai livelli di antico regime: le donne avevano in media cinque figli nell'arco della propria vita feconda, l'aspettativa di vita superava di poco i 30 anni, gli over 65 erano meno del 5%.

Figura 1 – Piramide per età della popolazione dell'Italia e delle sue ripartizioni, 2008.



Fonte: Istat, 2008.

La transizione demografica ha completamente stravolto gli equilibri del passato ed una fecondità in continuo declino associata a un progressivo incremento della speranza di vita, attualmente tra le più alte in Europa, hanno condotto ad una struttura per età che può considerarsi invecchiata. Questo processo ha interessato in un primo momento le regioni del Nord ovest² per poi espandersi progressivamente

² Nel corso degli ultimi decenni del XIX secolo e della prima metà del XX secolo, alcune regioni settentrionali dell'Italia (Liguria, Piemonte, Valle d'Aosta) hanno sperimentato per prime il declino della fecondità rispetto alle altre regioni italiane. Per queste regioni territorialmente contigue alla Francia, si è da più parti ipotizzato una transizione della fecondità di "tipo" francese anziché meridionale.

nel resto del Paese. Nell'ultimo mezzo secolo anche le trasformazioni sociali ed economiche hanno influito sulla sua evoluzione demografica, soprattutto sulla fecondità che, raggiunto il suo minimo storico nel 1995, ha ripreso a salire visibilmente nelle regioni del Nord economicamente avvantaggiate e dove l'occupazione delle donne è più elevata. Nel Sud Italia all'opposto si è assistito ad un declino tardivo, ma più marcato, della fecondità che attualmente stenta ancora a stabilizzarsi. Il declino demografico del Sud, tuttavia, è legato non solo alla riduzione della fecondità, ma anche ad altri fattori come l'emigrazione verso il Nord o l'estero di molti giovani qualificati e la minore incidenza dell'immigrazione dall'estero. Emerge, pertanto, che a rendere più complesso il processo di invecchiamento della popolazione italiana contribuiscono in modo significativo oltre alla velocità, alla intensità e alla durata del fenomeno, anche le forti differenze interregionali che si rilevano all'interno del nostro Paese e che vedono nel Meridione l'area sfavorita.

Secondo le ultime previsioni Istat nei prossimi decenni si assisterà nel Mezzogiorno ad un notevole incremento della popolazione anziana sia in termini assoluti che relativi rispetto all'intero contesto italiano. Ciò fa emergere delle incertezze sul futuro della sostenibilità economica e sociale di questo impellente invecchiamento demografico. Il Sud è, infatti, il tallone d'Achille dell'Italia avendo un sistema sanitario e di welfare pubblico più carenti e meno efficienti che rischiano di essere ulteriormente indeboliti da un federalismo poco attento alle reali esigenze del paese.

3. Indici di carico sociale e divari regionali

Prendendo spunto da un lavoro di Bongaart J. (2004), sono stati costruiti alcuni indicatori per lo studio del fenomeno invecchiamento:

1. Indice di dipendenza degli anziani (Old Dependency Ratio)

$$ODR = \frac{\text{pop (65+)}}{\text{pop (15-64)}} \quad (1)$$

2. Tasso di pensionamento (Pensioner Ratio)

$$PR = \frac{\text{pensionati}}{\text{pop (65+)}} \quad (2)$$

3. Tasso di occupazione (Employment Ratio)

$$ER = \frac{\text{occupati}}{\text{pop (15-64)}} \quad (3)$$

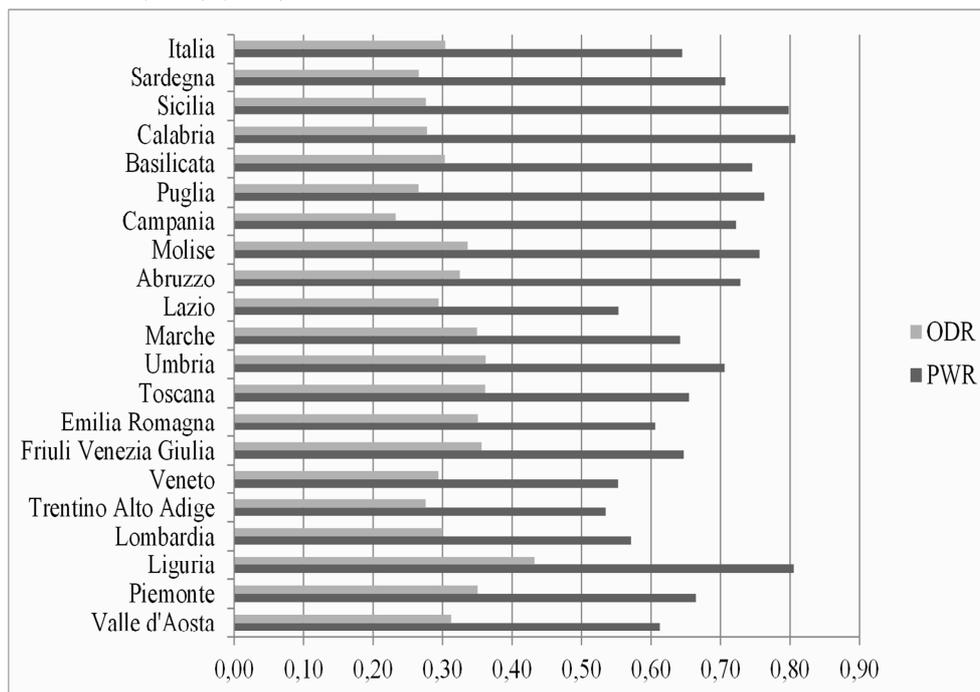
4. Pensionati per lavoratore (Pensioner/worker Ratio)

$$PWR = ODR * PR / ER \quad (4)$$

Attraverso questi indicatori è stato possibile evidenziare i divari regionali registrati con riferimento all'anno 2008. In particolare la fig. 2, dove sono riportati l'indice

ODR ed il rapporto PWR, evidenzia che il numero di pensionati per lavoratore è particolarmente alto (circa 0,8) in tre regioni Italiane: Liguria, Calabria e Sicilia.

Figura 2 – *Indice di dipendenza anziani (ODR) e indice di pensionati per lavoratori (PWR) (2008).*



Fonte: Istat, 2008.

Se però in Liguria anche l'ODR è piuttosto alto, lo stesso non può dirsi per le altre due regioni del Sud dove, invece, è la bassa occupazione a determinare gli alti valori del PWR. Dalla figura 3 si evince, infatti, che i tassi di occupazione³ sono inferiori al 50% in 4 regioni meridionali: Sicilia, Calabria, Puglia e Campania. Per queste ultime due regioni, inoltre, il tasso di pensionamento si presenta piuttosto elevato.⁴ Emerge, pertanto, un quadro preoccupante per il Sud caratterizzato da bassi tassi di occupazione, un welfare inefficiente e una popolazione che, sebbene

³ In queste regioni i tassi di occupazione sono particolarmente bassi per le donne.

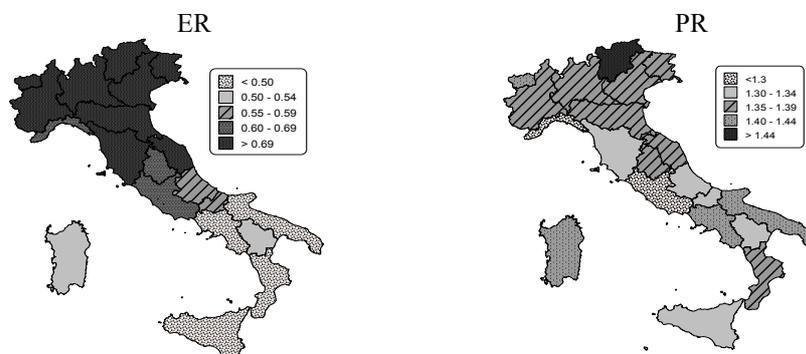
⁴ Il numero dei pensionati è generalmente più alto della popolazione con 65 anni e più. Ciò dipende dalle leggi sul pensionamento in vigore. In Italia l'età media della pensione per le donne è solo di 58,2 anni, 4 anni più bassa dei 62 anni ufficiali. Gli uomini si ritirano dal lavoro ad un'età relativamente giovane, che si attesta sui 59,4 anni, circa 6 anni prima dei 65 anni ufficiali.

al momento attuale si presenta più giovane rispetto a quella del Nord, è destinata ad invecchiare più velocemente. Virtuose appaiono, invece, le regioni settentrionali dove il divario tra ODR e PWR è più contenuto e dove l'occupazione è superiore o prossima al 70%. Un caso a parte nel Centro Italia è rappresentato dalla regione Lazio che, sebbene presenti un tasso di occupazione di poco inferiore rispetto a quello osservato per il Nord Italia, presenta con la Liguria il tasso di pensionamento più basso del paese.

4. Previsioni sull'invecchiamento

Come è già stato osservato nel secondo paragrafo, la popolazione italiana attraversa un processo di invecchiamento. Man mano che svanisce l'effetto transitorio delle generazioni del baby boom, il richiudersi della finestra demografica ed il progressivo invecchiamento della popolazione lasciano prevedere nel medio e lungo termine che la popolazione in età attiva incontri una fase di consistente riduzione ed, al suo interno, un processo di invecchiamento. Né il pur sempre più rilevante peso dei cittadini stranieri sembra in grado, se non parzialmente, di contrastare tali tendenze.

Figura 3 – Tasso di occupazione (ER) e di pensionamento (PR) nelle regioni italiane, 2008.



Fonte: Istat, 2008.

Le previsioni demografiche 1° gennaio 2007-2051 (Istat, 2009) prefigurano una situazione per la quale i potenzialmente attivi si ridurrebbero a 37,4 milioni, entro il 2031, e a 33,4 milioni entro il 2051, facendo registrare un saldo negativo finale di 5,6 milioni di individui rispetto al 2007 (scenario centrale). Ciò comporterebbe per l'Italia un raddoppio del valore dell'indice di dipendenza degli anziani, che passerebbe dal livello 0,3 registrato nel 2010 al valore 0,6 nel 2051. Ma tali

dinamiche, come si vedrà nel seguito, non si manifesterebbero in maniera uniforme nelle diverse regioni.

Nel paragrafo precedente è stato evidenziato come il Meridione sia attualmente caratterizzato da valori più favorevoli degli indici di dipendenza degli anziani, malgrado poi i tassi di occupazione (ER) siano talmente bassi da spingere il numero di pensionati per lavoratore (PWR) delle regioni meridionali ben al di sopra della media nazionale. Ma per il periodo 2010-2051 per le regioni meridionali si prefigurano incrementi dell'ODR ben maggiori della media, al punto da giungere, nel 2051, ad una inversione della posizione relativa delle diverse regioni.

Figura 4 – *Indice di dipendenza anziani nelle regioni italiane 2010 e 2051.*

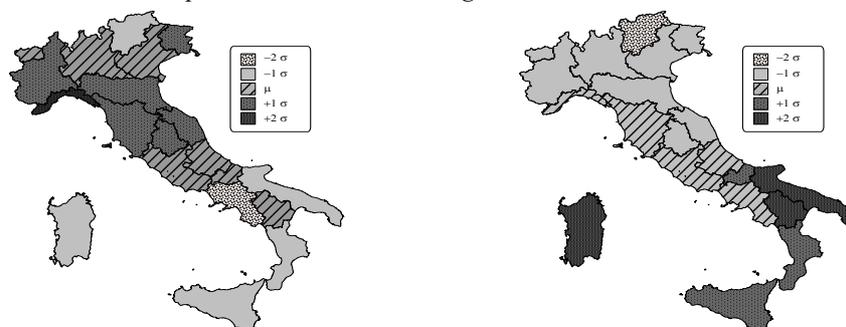
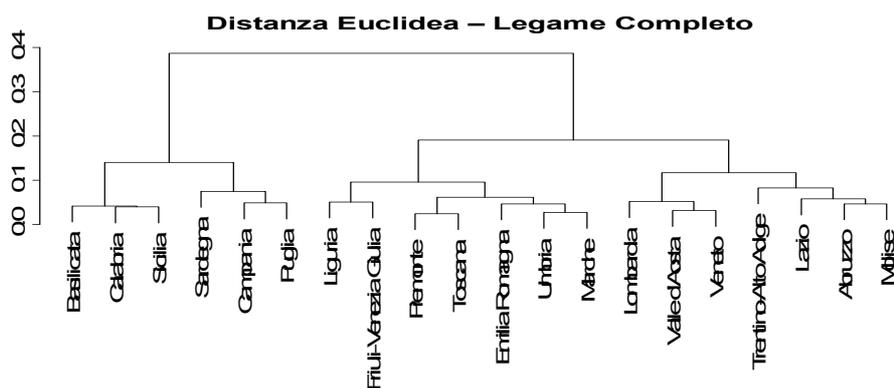


Figura 5 – *Dendrogramma della variazione percentuale complessiva dell'indice ODR nel periodo 2010-2051 (distanza euclidea – legame completo).*



Fonte: Nostre elaborazioni su dati Istat, Previsioni demografiche 1° gennaio 2007-2051, scenario centrale, 2008.

Come evidenzia la fig.4, gran parte delle regioni che attualmente presentano i valori più elevati dell'indice di dipendenza, quali Liguria, Umbria, Toscana e Friuli, nelle previsioni per il 2051 occuperebbero le ultime posizioni, mentre regioni come Sardegna, Calabria, Puglia e Sicilia, che attualmente hanno i valori più bassi, nelle previsioni per il 2051 si collocherebbero nei primi posti. Il dendrogramma della figura 5, ottenuto utilizzando un metodo di classificazione gerarchica, mostra delle aggregazioni delle regioni effettuate sulla base di andamenti simili delle variazioni percentuali complessive dell'indice ODR su base decennale (2010-2020, 2020-2030, 2030-2040 e 2040-2050). È interessante notare come, ad una distanza di 0.15, vengono individuati tre cluster. Un primo cluster è composto dalle regioni Campania, Puglia, Basilicata, Calabria, Sicilia e Sardegna. Si tratta delle regioni del Meridione d'Italia per le quali sono previsti gli aumenti più consistenti dell'indice di dipendenza, con incrementi medi complessivi intorno al 150%. Il secondo cluster è composto da regioni per le quali si prevede in media un raddoppio dell'indice, e comprende Piemonte, Liguria, Friuli-Venezia Giulia, Emilia Romagna, Toscana, Umbria e Marche. Infine, il terzo cluster comprende le regioni nelle quali l'indice complessivamente registra aumenti appena superiori al 60%; vi afferiscono Valle d'Aosta, Lombardia, Trentino-Alto Adige, Veneto, Lazio, Abruzzo, Molise.

5. Conclusioni

A 150 anni dall'Unificazione d'Italia notevoli divari economici e demografici continuano a dividere il nostro paese e, secondo le previsioni Istat, il divario è destinato a persistere nei prossimi anni anche con riferimento alla struttura per età della popolazione italiana. Dall'analisi effettuata in queste pagine emerge chiaramente che ad essere maggiormente interessate dal fenomeno invecchiamento saranno le regioni meridionali e le isole. È, pertanto, indispensabile adottare delle politiche efficienti che possano attenuare gli effetti negativi di un processo che inevitabilmente coinvolgerà quest'area del paese. Come proposto da Bongaarts, nei prossimi anni sarà indispensabile intervenire sugli indicatori analizzati: il tasso di dipendenza (ODR), il tasso di occupazione (ER) ed il tasso di pensionamento (PR). Per ridurre l'ODR sarebbe opportuno incrementare i tassi di fecondità attraverso politiche di sostegno alla famiglia, in particolare politiche indirette. Ma si potrebbe intervenire anche facendo ricorso all'immigrazione dall'estero, purché questa sia sostenibile dal contesto sociale ed economico. Sul tasso ER potrebbe essere utile intervenire adottando politiche che possano favorire l'entrata nel mondo del lavoro delle donne e delle giovani generazioni, e contrastandone l'ingresso tardivo. Infine, un provvedimento dovrebbe concernere la riduzione del PR attraverso l'incremento

dell'età al pensionamento. Scegliere una politica adeguata diventa una priorità per garantire sostenibilità all'invecchiamento nel nostro paese.

Riferimenti bibliografici

- ALTAVILLA A.M. (2001). Sistemi di pensionamento in una popolazione stabile. *Annali della Facoltà di Economia*. Università di Catania.
- BILLARI F., DALLA ZUANNA G. 2008. *La rivoluzione nella culla. Il declino che non c'è*. Milano: Università Bocconi.
- BONGAARTS J. (2004) Population Aging and the Rising Cost of Public Pensions, *Population and Development Review* 30(1).
- GALIZIA F. (2009). *La fecondità in Spagna e Italia: Catalogna e Piemonte/Liguria a confronto*, Tesi di dottorato, Università degli Studi di Bari.
- ISTAT (2009). *Previsioni demografiche 1 gennaio 2007-2051*. <http://demo.istat.it/uniprev>.
- OECD (1998). *Maintaining Prosperity in an Ageing Society*. Paris: OECD Publications.
- SALVINI S., DE ROSE A. (2011). *Rapporto sulla popolazione. L'Italia a 150 anni dall'Unità*. Bologna: il Mulino.

SUMMARY

As in most areas of the developed world, the declines observed in fertility and mortality in Italy will inevitably speed up the aging process. The social burden of the elder population is increasing, and this because the elderly people's support in Italy - pension, health care and social services - is still provided by the government. This study examines variations in the current levels of old-age dependency among the Italian Regions, as well, presents projections in demographic trends. The results reveal that the Southern Regions currently have better dependency ratios, though they have also higher pension-to-worker ratios (due to lower employment ratios). According to current population projections, mainly due to both lower fertility rates and lower foreign immigrations, dependency ratios of the Southern Regions are expected to catch-up the ones of the Northern Regions by 2030, and definitely surpass them by 2051. The current trend looks unsustainable in the long-period, so policy-options directed to increase fertility and immigration appear both quite urgent.

Anna Maria ALTAVILLA, Università di Catania, altavil@unict.it
Angelo MAZZA, Università di Catania, a.mazza@unict.it
Francesca GALIZIA, Università di Bari, francescagal@hotmail.it

ON THE UPWARD BIAS OF THE DISSIMILARITY INDEX

Anna Maria Altavilla, Angelo Mazza, Antonio Punzo

1. Introduction

The dissimilarity index of Duncan and Duncan (1955), commonly denoted by D , is widely used in a broad range of contexts to assess the overall extent of segregation in the distribution of two socio-demographic groups (*e.g.* male and female, blacks and whites, high and low skills workers, etc.) in two or more units (such as neighborhoods, workplaces, schools, etc.).

Generally, the observed settlement pattern is the resultant of a mix of behavior-based forces; thus it should be seen as one of the many possible outcomes of a stochastic – rather than deterministic – allocation. Usually researchers are interested in understanding the “systematic” characteristics of the allocation process, apart from random fluctuations that may affect a single observed pattern. In this view, the observed dissimilarity \hat{D} is merely an estimator of a true but unknown level of dissimilarity in the population D . So, it should be clear why this randomness also holds even if the index is computed on a full-count census data.

A problem with the use of this index is that \hat{D} appears to be an upward biased estimator of D . Accordingly, different correction approaches have been proposed in literature (see, *e.g.*, Allen *et al.*, 2009, and Altavilla, Mazza and Punzo, 2010 for two examples of bootstrap-based bias correction). So far, such bias has been computed analytically only for small numbers of units and small unit sizes; in the other cases, to analyze the bias of \hat{D} , simulation approaches have been pursued (see, *e.g.*, Carrington & Troske, 1997 and Allen *et al.*, 2009). In this paper we will demonstrate analytically that \hat{D} actually is an upward biased estimator of D .

The paper is organized as follows. In section 2, inferential framework and notation are given and in section 3 an analytic proof of the non-negativity of the bias of \hat{D} is provided. Finally, in section 4, conclusions are drawn.

2. Inferential framework and notation

Consider an area subdivided into k subareas (or units), denoted by $j = 1, \dots, k$, being populated by n individuals having a dichotomous characteristic indexed by $c = 0, 1$. Examples of common dichotomous characteristics are black or white ethnicity, male or female gender, and so on. The number of individuals with status c is denoted by n^c , $c = 0, 1$, with $n = n^0 + n^1$. There will be n_j^c individuals in

unit j having status c , with $n^c = \sum_{j=1}^k n_j^c$, $c = 0, 1$.

The observed settlement – characterized by the two sets denoted by n_1^0, \dots, n_k^0 and n_1^1, \dots, n_k^1 – is, however, only one of the possible realizations of an underlying *allocation process* P . If it is plausible to assume that individuals allocate themselves independently and that unit sizes are not fixed, then the process will be governed by the conditional probabilities

$$p_j^c = P(\text{unit of membership} = j | c), \quad j = 1, \dots, k, \quad c = 0, 1. \quad (1)$$

that an individual will belong to the unit j , given his/her status c

Social scientists are usually interested in making inferences on a particular function of these probabilities; this function, commonly called “segregation index”, should express the degree of segregation that characterize the process P . Before to introduce any kind of segregation index, it is important to define the concept of systematic segregation, occurring when there is at least one subarea in which individuals belonging to the two groups have a different probability to allocate themselves; in mathematical terms this means that:

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

Among the many segregation indexes existing in literature (see, e.g., Duncan and Duncan (1955), White (1986), and Massey and Denton (1988)), the most popular one is without doubt the segregation index of Duncan and Duncan (1955), usually denoted by D , characterized by the formula:

$$D = \frac{1}{2} \sum_{j=1}^k |p_j^1 - p_j^0| \quad (2)$$

Obviously, the index in (2) takes values on the compact interval $[0,1]$ and it increases as systematic segregation grows. Furthermore, it is straightforward to note that the case $D = 0$ (absence of systematic segregation) is achievable if, and only if

$$p_j^1 = p_j^0 \quad \forall j.$$

Unfortunately, we can only observe the crude counterpart of D

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

where \hat{p}_j^c , proportion of individuals with status c in the unit j , $c = 0,1$, is the plug-in estimator of p_j^c . The word “unfortunately” is justified if one thinks that the observed settlement pattern is only one of the numerous possible patterns arising from \mathbf{P} , each of them with probability given by the product of two independent multinomial distributions (see Allen *et al.*, 2009), one for $c = 0$ and one for $c = 1$:

$$P(n_1^c, \dots, n_k^c; p_1^c, \dots, p_k^c, n^c) = \prod_{j=1}^k \prod_{c=0}^1 n^c! \frac{(p_j^c)^{n_j^c}}{n_j^c!}. \quad (4)$$

As already mentioned, this framework assumes that individuals allocate themselves independently and that the quantities n^c , $c = 0,1$, are fixed after sampling.

3. Non-negativity of the bias of \hat{D}

As \hat{D} is an estimator for D , it is possible to define its bias as

$$Bias(\hat{D}) = E(\hat{D}) - D. \quad (5)$$

The expectation in the right-hand side of (5), taken over the independent multinomial distributions with probabilities p_j^c , $j = 1, \dots, k$ and $c = 0,1$, can be explicitated as follows:

$$E(\hat{D}) = \frac{1}{2} \sum_{(n_1^0, \dots, n_k^0)} \sum_{(n_1^1, \dots, n_k^1)} \left[\left(\sum_{j=1}^k \left| \frac{n_j^1}{n^1} - \frac{n_j^0}{n^0} \right| \right) \prod_{j=1}^k \prod_{c=0}^1 n_j^c! \frac{(p_j^c)^{n_j^c}}{n_j^c!} \right] \quad (6)$$

where the first two summations run across all possible patterns n_1^c, \dots, n_k^c satisfying the constraint $\sum_j n_j^c = n^c$, $c = 0, 1$.

As it has been shown through simulations (see Carrington and Troske, 1997, Allen *et al.*, 2009), $Bias(\hat{D})$ tends to be positive, that is, \hat{D} tends to overestimate the true but unknown value D . The common justification of this is related to the impact of the so-called “random” allocation that produces a positive bias that can be even considerably high.

In the following, we will prove that the non-negativity of (5) can be easily shown reasoning unit by unit.

Lemma 1: Bias(\hat{D}) assumes only non-negative values.

Proof. The allocation mechanism \mathbf{P} , in each unit j , $j = 1, \dots, k$, is governed by a binomial distribution with parameters n^c and p_j^c , denoted with $\text{Bin}(n^c, p_j^c)$; this is the j -th marginal distribution of the multinomial adopted in (4) (see Johnson, Kotz and Balakrishnan, 1997, p. 32).

Now, let $\hat{d}_j = \hat{p}_j^1 - \hat{p}_j^0$ be the plug-in estimator of the true but unknown difference d_j . Since we know that $n_j^c \sim \text{Bin}(n^c, p_j^c)$, it is easy to show that \hat{d}_j is an unbiased estimator of d_j ; in fact we have:

$$E(\hat{d}_j) = \frac{E(n_j^1)}{n^1} - \frac{E(n_j^0)}{n^0} = \frac{n^1 p_j^1}{n^1} - \frac{n^0 p_j^0}{n^0} = d_j. \quad (7)$$

Thus, using the result in (7), equation (5) can be expressed as follows:

$$Bias(\hat{D}) = \frac{1}{2} \sum_{j=1}^k E(|\hat{d}_j|) - \frac{1}{2} \sum_{j=1}^k |d_j| = \frac{1}{2} \sum_{j=1}^k E(|\hat{d}_j|) - \frac{1}{2} \sum_{j=1}^k |E(\hat{d}_j)| =$$

$$= \frac{1}{2} \sum_{j=1}^k \left[E\left(\left|\hat{d}_j\right|\right) - \left|E\left(\hat{d}_j\right)\right| \right] \quad (8)$$

The quantity in square brackets can be considered as the contribute to the bias given by the j -th unit. Being the absolute value a convex function, from the well-known Jensen's inequality (Jensen, 1906), we immediately have that

$$E\left(\left|\hat{d}_j\right|\right) \geq \left|E\left(\hat{d}_j\right)\right|.$$

Thus, the quantity in square brackets of (8) is non-negative and the thesis is confirmed.

4. Conclusions

In literature it is documented, but only through simulations, that the (observed) index of dissimilarity of Duncan & Duncan (1955) is an upward biased estimator of the true population counterpart and that the size of this bias is related to several quantities like the minority proportion, unit sizes and systematic segregation (see, e.g., Carrington & Troske, 1997 and Allen *et al.*, 2009).

In this paper, following the multinomial framework of Allen *et al.*, (2009), we have shown analytically the non-negativity of the bias of the dissimilarity index. This result might be used, in future works, in order to provide mathematical confirmation of the relation between the bias and the above mentioned quantities and to investigate on possible analytical bias corrections.

References

- ALLEN, R., BURGESS, S. and WINDMEIJER, F. (2009). *More Reliable Inference for Segregation Indices*. The Centre for Market and Public Organization, University of Bristol.
- ALTAVILLA, A. M., MAZZA, A. and PUNZO, A. (2010). Sull'impiego di un indice di dissimilarità nello studio della disposizione di popolazioni straniere su un territorio urbano. *Rivista Italiana di Economia Demografia e Statistica*, Vol. LXIV, No 1-2, pp.7-14.
- CARRINGTON, W. J. and TROSKE, K. R. (1997). On measuring segregation in samples with small units. *Journal of Business & Economic Statistics*, 15(4), pp. 402-409.
- DUNCAN, O. D. and DUNCAN, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, 20(2), pp. 210-217.

- JENSEN, J. L. W. V. (1906). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica*, 30(1), pp. 175-193.
- JOHNSON, N. L., KOTZ, S. and BALAKRISHNAN, N. (1997). *Discrete Multivariate Distributions*. New York: John Wiley & Sons.
- MASSEY, D. S. and DENTON, N. a. (1988). The dimensions of residential segregation. *Social forces*, 67(2), pp. 281-315.
- WHITE, M. J. (1986). Segregation and diversity measures in population distribution. *Population index*, 52(2), pp.198-221.

SUMMARY

In literature it is common to think about segregation as the outcome of a stochastic allocation process, characterized by the preferences and the constraints of both the individuals and the allocation units. In this perspective, several authors have already documented, using simulations, that the observed value of the Duncan & Duncan dissimilarity index is an upward biased estimator of the true population counterpart. In this paper, following a multinomial framework, we provide a mathematical proof of the non-negativity of this bias.

Anna Maria ALTAVILLA, Professor, University of Catania, Department “Impresa, Culture e Società”, altavil@unict.it

Angelo MAZZA, Assistant Professor, University of Catania, Department “Impresa, Culture e Società”, a.mazza@unict.it

Antonio PUNZO, Assistant Professor, University of Catania, Department “Impresa, Culture e Società”, antonio.punzo@unict.it

DALLA STAGNAZIONE ALLA CRESCITA SOSTENUTA: UN'EVOLUZIONE LUNGA OLTRE UN SECOLO

Lucrezia Andria, Roberta Pace, Alessandro Polli

1. Introduzione

A proposito delle relazioni dinamiche tra crescita e transizione demografica, Galor e Moav sviluppano una teoria evolucionistica della crescita economica allo scopo di “inserire l’intricata evoluzione della popolazione, della tecnologia e dell’output nella lunga transizione da un regime di stagnazione malthusiana ad un altro di crescita economica sostenuta” (Galor e Moav, 2001).

Obiettivo dello studio è quello di arrivare a definire un modello¹ che descriva le relazioni tra trend demografico, distribuzione del reddito e crescita in un’economia a due settori, di cui uno tradizionale ed uno *technology-intensive*, con agenti caratterizzati da livelli differenziati di capitale umano.

Nel presente lavoro ci soffermeremo, dunque, sull’osservazione empirica del caso italiano, in cui il passaggio dalla stagnazione ad un periodo di crescita sostenuta – così come accaduto in tutti i paesi sviluppati – può essere illustrato attraverso i processi di transizione demografica.

2. Framework teorico

Sino agli anni '80 del XX secolo la teoria economica ha ritenuto che la disuguaglianza esplicasse effetti positivi sul tasso di crescita dell’economia. Per Kaldor (1956) gli agenti economici caratterizzati da elevati livelli di reddito hanno una propensione marginale al risparmio più alta. Fenomeni di polarizzazione nella distribuzione dei redditi determinerebbero pertanto un aumento del risparmio e dell’investimento e, in definitiva, effetti positivi in termini di crescita economica.

¹ Si rimanda ad un lavoro successivo in cui si ipotizza di adottare un modello a generazioni sovrapposte, che descriva le relazioni esistenti tra andamenti demografici, distribuzione del reddito e crescita. In cui si ipotizza di adottare un modello a generazioni sovrapposte, che descriva le relazioni esistenti tra andamenti demografici, distribuzione del reddito e crescita economica.

L'evidenza storica, individuando gli eventi chiave che separano i due regimi – accelerazione del processo tecnico e transizione demografica – rende le tradizionali teorie della crescita per lo più inconsistenti. Infatti, alla luce del ruolo centrale rivestito dall'interazione tra crescita della popolazione ed innovazione, le nuove teorie rendono endogeni i due fenomeni.

Nelle formulazioni di Galor e Weil (1999, 2000) il legame positivo sussistente tra aumento della popolazione e numero di innovazioni ha gradualmente innalzato il tasso di progresso tecnico, inducendo maggiori investimenti in capitale umano, transizione demografica e crescita sostenuta.

In due contributi di Galor e Moav (2000, 2001) si sottolinea che gli esseri umani si confrontano con il trade off biologico tra quantità e qualità della discendenza. La preferenza per la 'quantità' dei figli ha un effetto positivo sui tassi di fertilità ma ne influenza negativamente la 'qualità', il benessere ed i futuri tassi di fertilità. Di conseguenza, i differenziali di fertilità tra gruppi socioeconomici evolvono non-monotonicamente nel tempo, in accordo con quanto evidenziato da alcuni demografi (Livi Bacci, 1997).

Nel regime malthusiano i tassi di fertilità risultano sensibili a temporanei aumenti nel reddito pro capite, ma manifestano una sostanziale stabilità intertemporale. Tuttavia, se la distribuzione del reddito riflette differenze nei tratti genetici e culturali (De la Croix, 2001), la pressione sui tassi di fertilità ed il successo evolutivo degli individui che investono maggiormente nella 'qualità' della discendenza influenzano il mix genetico della popolazione.

Ogni tratto ereditario che influenza la preferenza per la 'qualità' accresce, attraverso la selezione naturale, l'investimento aggregato in capitale umano; di conseguenza aumentano la produttività del lavoro, la speranza di vita e mutano stabilmente la struttura delle preferenze individuali in direzione della 'qualità' della discendenza. L'accelerazione del progresso tecnico, quindi, risulterebbe sempre associata alla transizione demografica e ad un'attenuazione della disuguaglianza, in accordo con alcuni risultati parziali ottenuti già alla fine degli anni Novanta da Dahan e Tsiddon (1998).

2.1. *Oltre Malthus*

Il concetto di 'trappola malthusiana'² sviluppato da Leibenstein (1954, 1957) e Nelson (1956) evoca l'equilibrio stabile ad un livello prossimo a quello di

² Secondo il regime malthusiano, vi sono mutamenti del reddito e della popolazione legati sia all'ambiente fisico sia a fasi irregolari nel progresso tecnologico tali da determinare, nel lungo periodo, una crescita economica vicina allo zero (stagnazione). La trappola malthusiana, dunque, si realizza quando la crescita della popolazione impedisce incrementi nel tenore di vita

sussistenza. Secondo Nelson, evitare la trappola è possibile ma richiede una serie di condizioni, quali: modernizzazione della struttura sociale (crescita del risparmio e dell'imprenditorialità, promozione della produzione e limitazione delle dimensioni familiari), aumento del tasso della popolazione attiva e migliore produttività (Verón, 1989).

Altri autori concordano con questa visione affermando che non ci sia un esempio di crescita generalizzata della popolazione accompagnata da un calo costante del prodotto pro capite. Nelson, analizzava il caso dell'Africa, dove le risorse naturali non vengono utilizzate e la crescita della popolazione è utile per l'aumento dell'offerta di lavoro (Kuznets, 1965). Una larga parte della letteratura introduce come variabile chiave il capitale umano in vista del rapporto intergenerazionale, in quanto potrebbe spiegare le dinamiche demografiche ed economiche più recenti. Il miglioramento della speranza di vita induce i genitori a sostituire la qualità per la quantità e trascina quindi l'economia in un percorso di crescita sostenuta (Ehrlich e Lui, 1991).

Il tentativo di analizzare l'evoluzione complessa della popolazione, della tecnologia e della produzione nel passaggio dalla stagnazione alla crescita, è oggetto della teoria della crescita unificata sviluppata da Galor e Moav (2000), cui si è già fatto riferimento. A questo riguardo l'interazione malthusiana tra il livello tecnologico, le dimensioni e la composizione della popolazione ha un effetto positivo sul progresso tecnologico e permette di delineare l'importanza del capitale umano nel processo di produzione (Galor, 2005).

All'interno di questa teoria, la valorizzazione del ruolo del capitale umano e la transizione demografica (con il declino della fertilità e della mortalità) risultano aver contribuito in maniera significativa alla crescita economica sostenuta. In tal senso, quindi, l'impatto della crescita demografica sulla produttività sembra non avere più un'influenza negativa.

3. Verso la Transizione Demografica e la crescita sostenuta

La crescita della popolazione, com'è noto, va di pari passo con i cambiamenti rilevanti nei settori agricolo ed industriale e, normalmente, porta ad un aumento del tenore di vita e del PIL pro capite (Kuznets, 1986).

Data la mancanza di dati storici, risulta estremamente complesso descrivere lo sviluppo demografico ed economico della popolazione italiana prima dell'unificazione del Paese. Tuttavia, grazie agli studi storici è possibile identificare gli eventi chiave che hanno segnato il Paese durante la rivoluzione industriale e la transizione demografica.

Tra questi è possibile individuare (Malanima, 2002):

1) introduzione di mais, riso e patate, che ha permesso un incremento della produttività del terreno; 2) crescita della disponibilità di beni alimentari, a seguito di una maggiore produttività agricola; 3) passaggio dal calore all'energia meccanica; 4) scoperta dell'elettricità ed introduzione di fertilizzanti in agricoltura.

Tali eventi, succedutisi in più di due secoli, hanno determinato sostanziali cambiamenti nel contesto economico.

Il periodo napoleonico rappresenta una fase importante di riforme (socio-sanitarie ed amministrative) che hanno determinato la fine delle malattie infettive in forma epidemica e reso altre malattie meno frequenti ed aggressive. In questo periodo la mortalità comincia ad essere percepita come un fenomeno selettivo, dove fattori socioeconomici e di contesto assumono un ruolo fondamentale nella possibilità di sopravvivenza (Bengtsson et al., 2004).

Nonostante tali miglioramenti, questo periodo è ancora antecedente alla prima transizione demografica, che ha avuto luogo in Italia solo alla fine del XIX secolo.

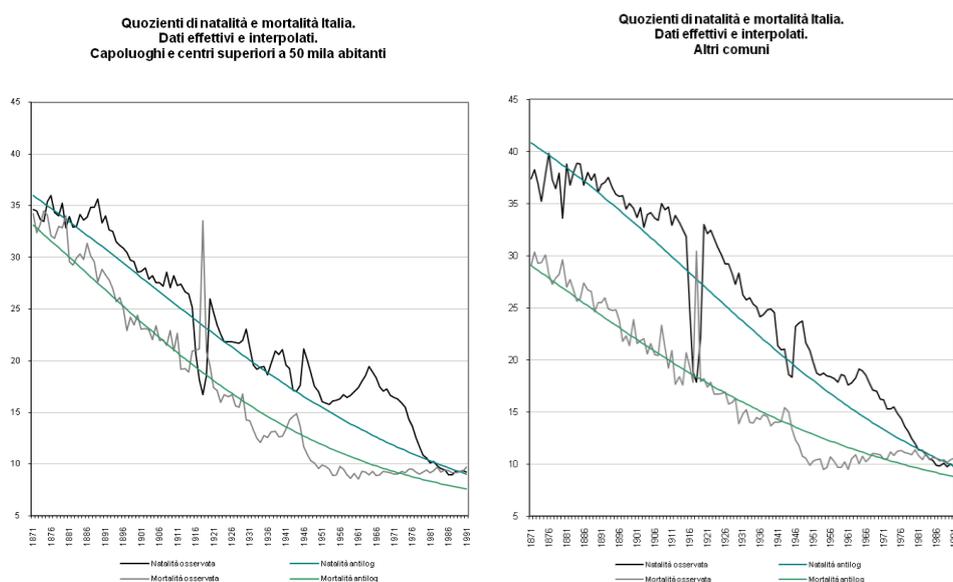
Bisogna, tuttavia, partire dalle grandi trasformazioni avvenute durante il XIX ed il XX secolo – in particolar modo l'urbanizzazione, la fine delle società contadine e la progressiva industrializzazione – per poter affermare che la transizione demografica ha giocato un ruolo importante per ciascuno di questi cambiamenti epocali. In termini generali, durante tale processo si osserva il passaggio da un regime in cui la popolazione è di tipo stabile con incremento antieconomico (alta natalità ed alta mortalità con un reddito procapite invariato) ad un regime in cui la popolazione è in fase di stabilizzazione (bassa natalità e bassa fecondità) con un incremento di tipo economico (Natale, 2002).

Il processo di transizione demografica rientra in un generalizzato processo di cambiamento della società che, nello stesso periodo, sperimenta il passaggio da una società rurale ad una urbana ed industrializzata, con conseguenti e profondi cambiamenti in tutte le sfere del vivere sociale. Il minore o maggior grado di urbanizzazione, nonché il maggior o minor livello di industrializzazione contribuiscono a scandire il calendario del declino dei tassi di mortalità e di fecondità della popolazione, nonché la durata e la portata della transizione demografica.

L'evidenza empirica della storia demografica italiana conferma quanto detto. Nelle figure 1 e 2 sono rappresentati i processi transazionali italiani nei capoluoghi e città con più di 50mila abitanti e nei comuni fino a 20mila abitanti che, in teoria, identificano rispettivamente un ambito presumibilmente urbano ed uno pressoché rurale. Emerge in modo chiaro la profonda differenza nei percorsi transazionali tra città e campagna. Non solo i livelli di partenza dei due quozienti sono profondamente diversi – soprattutto in termini di natalità, che risulta più alta nei comuni rurali – ma anche la portata dell'intero processo è diversa: identificando un periodo di crescita della popolazione molto più esteso e rilevante in ambito rurale.

E' possibile affermare, quindi, – seguendo anche l'impostazione di Galor – che questo processo si verifica con maggiore lentezza e gradualità man mano che ci si sposta dalle aree urbane a quelle rurali.

Figure 1 e 2 – Tassi di natalità e di mortalità in Italia (1871-1991).



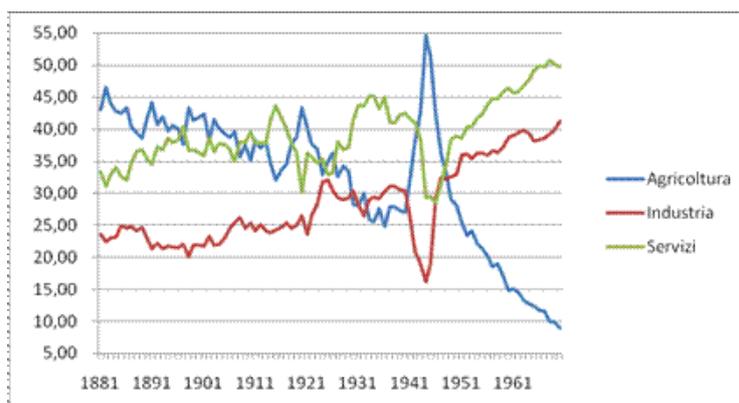
Fonte: Elaborazioni proprie.

A livello economico numerosi studiosi collocano l'inizio di una positiva performance italiana verso la fine del XIX secolo, ed in particolare a partire dal 1880, anno in cui il PIL comincia ad aumentare in maniera stabile. La crescita italiana registrata in questo periodo (sino ai primi decenni del XX secolo) dipese essenzialmente dall'aumento della dotazione di risorse naturali, di capitali e da un generale aumento dell'efficienza del sistema produttivo (Malanima, 2006).

Come già osservato, il graduale processo di industrializzazione, che ha investito il paese tra la fine del XIX ed il XX secolo, ha certamente contribuito ad un abbandono dei comportamenti tipici di una società tradizionale, in cui il settore agricolo rappresenta il principale comparto economico, per lasciare spazio a valori e costumi di una società moderna. Di fatto, durante le fasi principali del processo di transizione demografica, in Italia si assiste ad una graduale diminuzione della parte di valore aggiunto rappresentato dall'agricoltura e ad un contestuale aumento nel peso del comparto industriale. Entrambe le tendenze sono interrotte solo in

corrispondenza degli eventi bellici (ed in particolare della seconda guerra mondiale) per la temporanea interruzione che interessò la produzione (Figura 3).

Figura 3 – Valore aggiunto. Anni 1881-1961.



Fonte: Elaborazioni proprie.

Un'ulteriore conferma empirica della particolare e positiva fase di miglioramento delle condizioni economiche e sociali della popolazione italiana, tra la fine del XIX e l'inizio del XX secolo, si ha osservando i dati relativi all'indice di sviluppo umano ed a tutte le sue componenti. In particolare, i dati relativi al *Tasso di alfabetizzazione degli adulti* ed al *Tasso di scolarità* (Tabella 1) confermano l'importanza crescente che assume l'investimento in capitale umano per la storia sociale ed economica del paese.

La teoria suggerisce che un'aspettativa di vita più lunga incoraggi gli investimenti educativi, in quanto un orizzonte temporale più ampio aumenta il valore del rendimento degli investimenti. *Ceteris paribus*, un investimento che renda una certa quantità ogni anno è più importante se il flusso di pagamenti dura più a lungo.

Un'importante implicazione è che i miglioramenti nell'aspettativa di vita dovrebbero determinare un incremento degli investimenti in capitale umano che, a loro volta, creerebbero un *push-up* per la crescita economica.

Quindi, le scelte relative all'investimento in capitale umano sono anche funzione dei rendimenti attesi che si manifestano come maggiori redditi per le generazioni future (Birdsall, 1999). A parità di condizioni, quando il rendimento atteso degli investimenti in capitale umano aumenta, aumenterà il loro flusso.

Nel periodo temporale considerato, la mortalità subisce una forte contrazione (Tabella 1) e la speranza di vita alla nascita aumenta sensibilmente nel giro di poche generazioni. Allo stesso tempo la società risulta profondamente cambiata e

ne è conferma l'andamento del tasso di alfabetizzazione degli adulti che, dall'essere – alla fine del 1800 – appannaggio di una ristretta élite (circa il 44% della popolazione), si diffonde sino a raggiungere nel 1938 circa l'80% della popolazione, confermando l'ipotesi della crescente preferenza della 'qualità' della discendenza.

Tabella 1 – *Indicatori demografici ed economici in Italia (1871-1991).*

| Anni | Speranza di vita | Tasso di alfab. degli adulti | Tasso di scolarità | PIL/pro capite | Indice di Sviluppo Umano |
|------|------------------|------------------------------|--------------------|----------------|--------------------------|
| 1871 | 33,1 | 32,1 | 19,4 | 1,473 | 0,288 |
| 1891 | 39,3 | 43,9 | 23,7 | 1,615 | 0,358 |
| 1911 | 44,1 | 61,4 | 30,0 | 2,407 | 0,453 |
| 1928 | 52,8 | 75,6 | 34,1 | 2,949 | 0,549 |
| 1938 | 58,1 | 80,2 | 34,3 | 3,243 | 0,594 |
| 1951 | 65,0 | 85,8 | 34,0 | 3,658 | 0,651 |
| 1961 | 69,7 | 90,5 | 43,4 | 6,236 | 0,728 |
| 1971 | 72,3 | 93,9 | 59,0 | 9,598 | 0,791 |
| 1981 | 74,0 | 96,5 | 63,4 | 13,145 | 0,828 |
| 1991 | 77,0 | 97,6 | 65,0 | 16,112 | 0,861 |

Fonte: Conte, Della Torre, Vasta, 2007.

Riferimenti bibliografici

- BENGTSSON T. et al. 2004. *Life under pressure: mortality and living standards in Europe and Asia, 1700-1900*. Cambridge (Mass), USA.
- BIRDSALL N. 1999. Education: The People's Asset. *Centre on Social and Economic Dynamics, Working Paper n.5*.
- CONTE L. et al. 2007. The Human Development Index in Historical Perspective: Italy from Political Unification to the Present Day. *Quaderni del Dipartimento di Economia Politica*, Università degli Studi di Siena, gennaio, n. 491.
- DAHAN M., TSIDDON D. 1998. Demographic Transition, Income Distribution, and Economic Growth. *Journal of Economic Growth*, No. 3, pp. 29-52.
- DE LA CROIX D. 2001. Growth Dynamics and Education Spending: The Role of Inherited Tastes and Abilities", *European Economic Review*, No. 45, pp. 1415-1438.
- EHRLICH I., LUI F.T. 1991. *Intergenerational Trade, Longevity and Economic Growth*, Journal of Political Economy, No. 99, pp. 1029-1059.
- GALOR O. 2005. *From Stagnation to Growth: Unified Growth Theory*, Handbook of Economic Growth.
- GALOR O., MOAV O. 2000. Natural Selection and the Origin of Economic Growth. *Providence RI, Brown University, Working Paper N. 18-00*,.
- GALOR O., MOAV O. 2001. Evolution and Growth. *European Economic Review*, No. 45, pp. 718-729.

- GALOR O., WEIL D.N. 1999. From Malthusian Stagnation to Modern Growth. *American Economic Review*, No. 89, pp. 150-154.
- KALDOR N. 1956. Alternative theories of distribution. *Review of Economic Studies*, No.23 (2), pp. 83-100.
- KUZNETS S. 1955. Toward a Theory of Economic Growth. In LEKACHMAN R. (Ed.), *National Policy for Economic Welfare at Home and Abroad*, Garden City, NY: Doubleday.
- KUZNETS S. 1986. *Modern Economic Growth: Rate, Structure, and Spread*, Yale University Press, New Haven.
- LEIBENSTEIN H. 1954. *A Theory of Economic Demographic Development*, Princeton University Press.
- LEIBENSTEIN H. 1957. *Economic Backwardness and Economic Growth*. New York: John Wiley and Sons.
- LIVI BACCI M. 1997. *A Concise History of World Population*. Second edition, Oxford: Blackwell.
- MALANIMA P. 2002. La ricchezza e la povertà dell'Italia. Le risorse naturali. In CIOCCA P., TONIOLO G. (Ed.) *Storia economica d'Italia*, Vol. 3, Laterza.
- MALANIMA P. 2006. Alle origini della crescita in Italia 1820-1913. *Rivista di Storia Economica*, XXII.
- NATALE M. 2002. *Economia e popolazione. Alcuni aspetti sulle interrelazioni tra sviluppo demografico ed economico*, Milano: Franco Angeli.
- NELSON R.R. 1956. A Theory of Low Level Equilibrium Trap in Underdeveloped Economies. *American Economic Review*, 46.
- VERON J. 1989. *Eléments du débat population développement*, Les Dossiers du CEPED, No. 9, Paris.

SUMMARY

Starting from an overview of the literature on the dynamic relationships between economic growth and demographic transition, we will attempt to analyze the complex evolution of the population, technology and production, in the transition from a system of Malthusian stagnation to a regime of sustained economic growth. According to some theories, the enhancement of the role of human capital and demographic transition (with the decline of fertility and mortality) appear to have contributed significantly to economic growth. The empirical evidence will provide an interpretation of the dual dynamics of population growth and economic increase in Italy, distinguishing urban from rural areas.

Lucrezia ANDRIA, Ricercatore, CNAF, Parigi, lucreziaandria@hotmail.com
Roberta PACE, Professore associato di Demografia, Università di Bari "Aldo Moro",
r.pace@scienzepolitiche.uniba.it
Alessandro POLLI, Ricercatore in Statistica Economica, Università di Roma "La Sapienza", a.polli@dte.uniroma1.it

GOVERNARE IL TURISMO TRA GLOBALE E LOCALE AGLI INIZI DEL XXI SECOLO

Francesco Antonio Anselmi

1. Introduzione

Il tema di fondo di questo contributo è quello di contemperare due aspetti del turismo, globalizzazione ed identità locale, quali processi egualmente fondamentali nell'ambito della società agli inizi del XXI secolo. La globalizzazione, da un lato sembra tutto unire e tutto appiattare, ma dall'altro, sviluppa una infinita molteplicità di possibili percorsi di turismo individuale o per gruppi differenziati. Il senso del locale si riproduce e sconfigge dal di dentro la globalizzazione. Il turista della globalizzazione pone un rapporto diretto con i luoghi visitati, con il territorio e la sua memoria storica, la sua cultura e il suo ambiente, manifestando pertanto una domanda di autenticità, unicità, spontaneità di prodotti a volte non nati per il turismo e non ritrovabili in altra parte del mondo (aspetto locale). Si rileva una complessa domanda turistica che è anche multidimensionale e un turista molto pretenzioso con le sue strategie della scelta "imprevedibile" e "imponderabile".

In tale contesto si vuole evidenziare in questo contributo che l'offerta turistica si configura come integrazione di funzioni diverse e il prodotto turistico viene richiesto quando è in grado di coniugare due termini fondamentali: le caratteristiche proprie del territorio e le abilità imprenditoriali locali, nell'ambito di una sostenibilità turistica, economica, sociale ed ambientale e i sistemi locali di offerta turistica (SLOT) sono la strategia vincente per lo sviluppo di una destinazione, che sappia offrire anche servizi, strutture ed infrastrutture che consentano al turista un'esperienza positivamente unica.

2. Aspetti locali e globali del turismo

I cambiamenti in atto nel mercato turistico, acceleratisi negli ultimi anni, la progressiva globalizzazione dei mercati e l'evoluzione della domanda verso esigenze sempre più personalizzate ed articolate, hanno imposto agli operatori

turistici e ai soggetti pubblici territoriali lo sviluppo di nuove competenze e di capacità che costituiscono altrettanti elementi fondamentali per potere ottenere il proprio vantaggio competitivo.

In un contesto a così forte intensità competitiva, occorre far leva sui fattori di differenziazione delle singole realtà territoriali attraverso la mobilitazione di tutti i soggetti, di natura pubblica e privata, che, a vario titolo, influiscono sull'offerta turistica dell'area. Se in passato era possibile essere competitivi anche nonostante un'offerta dispersa e individualizzata, puntando su un vantaggio posizionale conferito dalle risorse, in un'ottica di "rendita di posizione" determinata dalle stesse, nel nuovo contesto in base a quanto osservato in precedenza, per ottenere successo risulta essenziale ricorrere ad un'aggregazione del prodotto in una logica di integrazione a livello di destinazione (Antonioli, 1999).

Lo sviluppo del settore turistico, soprattutto nell'arco dell'ultimo decennio, ha ampiamente dimostrato la necessità di un approccio sistemico al fenomeno, non solo per il suo carattere collettivo e la sua forte valenza territoriale, ma anche per i numerosi vantaggi che possono derivare da economie di scala e da esternalità positive che riguardano sia le singole aziende interessate che i territori e le comunità coinvolte, come conseguenza dei rilevanti impatti locali che caratterizzano questo settore.

Essendo il prodotto turistico particolarmente complesso, sia per l'eterogeneità dei servizi, sia per la responsabilità di ognuno di essi che è da imputare a soggetti diversi, al fine di sostenere la competitività di una destinazione, occorre ragionare non solo in una logica territoriale, ma anche adottare degli strumenti di gestione integrata di tali realtà che consentano di coordinare le decisioni, di valorizzare la complementarità tra risorse naturali e servizi e di stimolare l'instaurarsi di solide relazioni cooperative tra *stakeholders* pubblici e privati dell'area. La destinazione turistica secondo Antonioli, Mottironi, Viganò, richiede che la gestione della destinazione si focalizzi su quattro aspetti fondamentali (Antonioli, Mottironi, Viganò, 2005). Essi attengono:

- 1) al fatto che il successo di una destinazione è esclusivamente determinato dal grado di apprezzamento espresso da parte della domanda che è orientata al mercato;
- 2) alle principali fonti del vantaggio competitivo di una destinazione turistica che sono espresse dalle esigenze di differenziazione manifestate dal turista e di attenzione ai fattori *no price* (es. qualità d'area e tutela del consumatore) che vanno intese oggi quali principali fonti del vantaggio competitivo di una destinazione turistica;
- 3) alle esigenze di valorizzare le specificità locali e tutte le componenti da parte degli operatori della destinazione, al fine di ottimizzare l'uso delle risorse e gestire gli impatti derivanti dall'uso turistico del territorio;

- 4) al coinvolgimento del soggetto pubblico e degli enti che presiedono alla tutela e gestione dei siti del patrimonio naturale e culturale e comunque di chi è responsabile della legislazione e dei finanziamenti allo sviluppo.

In sintesi, il bene “turistico” considerato come elemento integrante e centrale di un sistema territoriale diviene tanto più efficiente, economicamente indipendente e capace di mantenere la propria memoria, quanto più è in grado di dotarsi di una organizzazione e gestione sistemica.

Il mercato turistico emergente reca in sé i presupposti per una competizione totale fra tutte le aree, tutte le risorse, tutte le modalità di pratica turistica. Esso si manifesta sia a livello macro-territoriale con la cosiddetta globalizzazione, sia a livello micro-territoriale, per il venir meno della separatezza tra gli spazi e tra i sottosistemi di relazione compresenti nella stessa regione: vita quotidiana ed esperienza straordinaria, produzione e consumo, lavoro e divertimento.

In un primo momento la globalizzazione si andava istituzionalizzando e autonomizzando intorno a codici propri che prescindevano da specifiche località, culture, tradizioni e sensibilità. La *comunità locale* in quanto tale, depositaria delle risorse culturali e ambientali di un determinato territorio, frutto di una specifica evoluzione, nell’ambito turistico è stato osservato l’emergere di un “vacanziero” che ha utilmente trovato spazio in un’immagine più complessa del territorio, più suscettibile di interpretazioni soggettive e di percorsi di accesso differenziati e continuamente rinnovati (A. Savelli, 2007).

Oggi il prodotto turistico richiede una nuova generazione territoriale basata su diverse modalità di offerta e rivolta a soddisfare interessi speciali che sono le attitudini. Infatti, sono in forte aumento le richieste di conoscenza dello spazio, la consapevolezza del rispetto dell’ambiente, il confronto delle culture, l’approccio con le diverse tradizioni, che fanno nascere nuovi turismi, come l’ecoturismo, il turismo culturale, naturalistico, enogastronomico, ecc.

A sua volta la globalizzazione, attraverso la connessione diretta e in un tempo reale in ogni luogo del globo con qualunque altro, richiede che i suoi nodi siano costituiti dai sistemi turistici territoriali locali o sistemi locali di offerta turistica (SLOT). I mutamenti strutturali della domanda turistica, caratterizzati da una segmentazione sempre più spinta con forti differenze tra un segmento e l’altro impongono un’offerta turistica gestita da una pluralità di imprese organizzate territorialmente in uno SLOT.

La Battilani (2001) ricorda che nel turismo globale “...*Non sono più le mete a fare quindi la differenza, ma le esperienze che si possono vivere: tutti possono raggiungere paesi esotici, anche i più lontani, ma non tutti possono pranzare in un determinato ristorante, alloggiare in un determinato albergo, assistere ad un particolare spettacolo, utilizzare servizi riservati etc. È quindi il modo in cui si visita e si vive una determinata località a segnare la differenza fra i vari ceti*

sociali: si tratta del cosiddetto turismo-teatro nel quale il viaggiatore è al centro di una rappresentazione teatrale e sceglie il personaggio da interpretare...”.

3. I sistemi locali di offerta turistica nella fase del turismo globale

I mutamenti strutturali della domanda turistica, caratterizzati da una segmentazione sempre più spinta, con sovrapposizioni tra turisti e residenti fruitori del tempo libero, impongono all’offerta turistica una gestione basata su una pluralità di imprese organizzate territorialmente. D’altro canto, la stessa nuova legge quadro (n. 135/2001) nell’ambito di un fortissimo decentramento funzionale associa un generale riconoscimento al turismo e alle risorse turistiche come fattori essenziali dello sviluppo territoriale locale.

La capacità competitiva di una destinazione turistica trova il proprio fulcro nella qualità e nell’efficienza dei sistemi territoriali e nel loro orientamento alla funzione ospitale. Un Sistema Turistico Locale è infatti un sistema territoriale, un network di destinazioni turistiche che hanno in comune prima di tutto il territorio. Ed è possibile che le realtà locali si aggregino anche a prescindere dai confini amministrativi provinciali e regionali, sulla base di una storia, di un vissuto, di un’identità e di una progettualità comune. Le cellule di base di un sistema turistico vengono quindi identificate in aree-destinazione, nel cui ambito gli enti locali, le imprese e l’intera comunità locale decidono di lavorare insieme per migliorare la qualità dei prodotti, per fare marketing, promozione e accoglienza, sviluppando la funzione ospitale del luogo.

Per Rispoli-Tamma lo SLOT configura «un insieme di attività e fattori di attrattiva che, situati nello spazio definito (sito, località, area) siano in grado di proporre un’offerta turistica articolata e integrata, ossia rappresentino un sistema per l’ospitalità turistica specifica e distintiva che valorizza le risorse e la cultura locali» (Rispoli M., Tamma M., 1995, pag. 41).

I caratteri portanti di uno SLOT si possono riassumere nei seguenti aspetti.

Sistema: indica insiemi di attività integrate fra loro e con il territorio coinvolgendo e coordinando il maggior numero di attori del territorio. Il sistema di ospitalità risiede in uno spazio definito (localismo) e assume caratteri peculiari che derivano dalla valorizzazione della cultura e dei fattori di attrattiva che lo caratterizzano.

Offerta Turistica: il sistema deve offrire una gamma di prodotti turistici, costituita da beni, servizi, fattori di attrattiva di natura sociale, culturale, economica, naturale, ecc., idonea a soddisfare la "domanda globale" di visita e soggiorno espressa da un turista.

L’assetto dello SLOT si fonda da un lato, sull’identificazione e convergenza dei diversi obiettivi verso uno scopo comune degli attori e sulle strategie di pianificazione che è lo sviluppo dell’offerta locale e dall’altro su un solido corredo

professionale e senso di responsabilità delle imprese (Manente, Furlan, Scaramuzzi, 1998).

Nello SLOT deve essere forte la convinzione tra gli attori di raggiungere delle performance per le imprese certamente non realizzabili dalle ordinarie regole del mercato (Palmer, 1997). Lo SLOT necessita di un soggetto strategico leader¹ che può essere una società (pubblica, privata o mista) o una coalizione tra i soggetti strategicamente diversi, che ha capacità di progettazione, di coordinamento per la realizzazione di iniziative di sviluppo del prodotto della località, secondo un principio di ottimizzazione del sistema.

Il ruolo dello SLOT è quello di favorire lo sviluppo del valore desiderabile dal cliente e delle condizioni che possono potenziare il dialogo tra proposta/ immagine della località e delle imprese in essa operanti, ed aspettative dei turisti (U. Morelli, 1995). L'obiettivo dello SLOT è nella ricerca delle fonti del vantaggio competitivo sostenibile per una *destination* rispetto alla concorrenza.

Le fonti del vantaggio competitivo sostenibile per una *destination* rispetto alla concorrenza dipendono: dalle risorse proprie della località; dalla capacità e competenza delle imprese locali nel promuovere un'offerta articolata, fondata sulla qualità dei servizi turistici e sulla valorizzazione delle risorse locali in una prospettiva non semplicemente temporanea ma di lungo termine.

Le risorse rappresentano il patrimonio strategico dello SLOT, invece le competenze indicano ciò che il management è in grado di fare.

Nella logica dello SLOT il problema non è più l'allocazione ottimale delle risorse esistenti, quanto la capacità di generare il massimo potenziale strategico con le risorse disponibili nel territorio (ambiente/milieu) che potrebbero essere anche molto scarse.² Nello SLOT, l'ambiente/milieu è una componente endogena che è chiamata a partecipare alla formazione di un'offerta territoriale complessa ed articolata. L'espressione "milieu" indica non un insieme di condizioni naturali e culturali date in un territorio, ma un complesso di valori condivisi, di potenzialità espresse da un territorio alle quali la rete locale di soggetti riconosce la funzione di innescare processi di sviluppo locale. Uno degli aspetti fondamentali del "milieu", nota Camagni (1991), è quello di ridurre l'incertezza, che caratterizza i processi innovativi, aumentando l'efficienza e riducendo i costi di produzione ("socializzazione del rischio").

¹ Le caratteristiche peculiari del soggetto leader devono essere: il carisma (la capacità di assumere un comportamento trasparente, ammirazione e fiducia) la motivazionale ispirazione (cioè la capacità di conferire significato e sfida nel lavoro), lo stimolo intellettuale (cioè stimolare la creatività).

² Tale concetto risulta molto importante nell'ambito dello sviluppo turistico in aree con risorse naturali o artistiche insufficienti (ad esempio Riccione, Rimini) dove tuttavia si è riusciti ad avviare fattori di attrattiva, creando domanda turistica notevole e duratura.

Il sistema delle risorse di uno SLOT deve essere valorizzato e sviluppato attraverso delle competenze in un processo dinamico di cambiamento morfogenetico (E. Rullani, 1994) caratterizzato da mutamenti a sbalzi ed evolutivi.

Le principali competenze delle imprese di un sistema locale di offerta turistica si suddividono in: *competenze specifiche* delle diverse categorie di impresa e *competenze di sistema* che richiamano il ruolo del soggetto leader.

Significative esperienze sono state osservate nel Nord-America e Australia. In alcuni Paesi come l'Inghilterra e la Spagna, lo SLOT si è realizzato con l'esistenza di enti strategicamente leader.

4. Conclusioni

Il turismo della fase globale è fenomeno trasversale che coinvolge una pluralità di attori sociali operanti sul territorio e attorno alle risorse locali, in grado di creare relazioni fondate su: 1) equilibri delicati tra visitatori, comunità locali e ambiente; 2) consapevolezza delle criticità ambientali e delle differenze culturali tra luoghi; 3) forte dipendenza delle attività turistiche dalla qualità delle risorse locali.

Lo sviluppo turistico deve garantire l'integrità complessiva dell'ecosistema nel quale si colloca e va comunque inserito in un disegno complessivo di sviluppo socio-economico locale, coerente con le scelte più generali dello sviluppo alla scala regionale, garantendo il pieno coinvolgimento degli stakeholder, in modo che ciascuno possa esprimere il proprio potenziale di sviluppo.

In conclusione, la globalizzazione, attraverso la connessione diretta e in un tempo reale in ogni luogo del globo con qualunque altro, richiede che i suoi nodi siano costituiti dai sistemi turistici territoriali locali. Queste considerazioni assumono rilevanza notevole quando si parla di offerta turistica nel nuovo scenario mondiale e nella specificità dei localismi.

Questo contributo ha focalizzato l'analisi sull'esigenza di un nuovo modo di fare turismo rispetto al passato. I bisogni dei turisti sono complessi e multidimensionali: si cerca di soddisfarli con la differenziazione, la personalizzazione e la specializzazione dell'offerta turistica che necessita dell'integrazione e della partecipazione di tutti gli attori coinvolti. In particolare, si vuole evidenziare che la competitività turistica duratura si gioca sulla qualità globale dell'offerta risultante dall'insieme delle attrattive turistiche dei servizi generali e delle attività turistiche imprenditoriali. Il turismo per sua natura richiede la messa in luce di azioni pubbliche (sfera pubblica) che influenzano una serie di attori e condizioni di diverso tipo e di azioni private (sfera privata o di mercato) che devono fornire un prodotto che deve divenire sistemico. I sistemi locali di offerta turistica (SLOT) sono la strategia vincente per lo sviluppo di una

destinazione turistica duratura, sia per il Mezzogiorno d'Italia che per il resto del paese che dispone di notevoli potenzialità turistiche. Le carenze di forme sistemiche di gestione del prodotto e del territorio non consentano di integrare concretamente prodotto e territorio in modo tale che il prodotto diventi globale ma che conservi una propria identità e tradizione.

In sintesi, il turismo globale può essere paragonato ad una sorta di imponente migrazione planetaria che funge da catalizzatore della dialettica tra globale e locale la cui *governance* si carica della forte opportunità e necessità di conservare e rafforzare l'identità locale nella sua connotazione più ampia (Trillo, 2003).

Riferimenti bibliografici

- AA.VV. 2008. *Trasformazioni territoriali e progettualità turistica*. Roma: Gangemi.
- ANSELMINI F. A. 2004. Il turismo nel sud d'Italia: punti di forza e opportunità strategiche. Il caso Sicilia. *Rivista Italiana di Economia, Demografia e Statistica*, Vol. LVIII, No. 3/4.
- ANTONIOLI CORIGLIANO M. 1999. I distretti turistici e le aggregazioni fra attori per lo sviluppo del prodotto-destinazione, In COLANTONI M. (Ed.) *Turismo: una tappa per la ricerca*, Bologna: Pàtron Editore.
- ANTONIOLI CORIGLIANO M., MOTTIRONI C., VIGANO G. 2005. Mobilità non motorizzata e risorse territoriali: le nuove forme di turismo. Lo stato dell'arte. In *Mobilità non motorizzata e risorse territoriali*, progetto MIUR, Technical Report No.7, Università degli Studi di Brescia, Dipartimento di Ingegneria Civile, Brescia.
- BATTILANI P. 2001. *Vacanze di pochi, vacanze di tutti. L'evoluzione del turismo europeo*, Bologna: Il Mulino.
- BOYER M. 2004. L'Italia nello sguardo dei viaggiatori turisti. In SAVELLI A. (Ed.) *Città, turismo e comunicazione globale*. Milano: Franco Angeli.
- CIANCIULLO A., REALACCI E. 2006. *Soft economy*. Milano: BUR.
- CIRCECHIA A. 2009. *Risorse culturali e turismo sostenibile. Elementi di pianificazione strategica*. Milano: Franco Angeli.
- FRANCH M. 2002. *Destination Management. Governare il turismo tra locale e globale*. Torino: G. Giappichelli.
- ICLEI 2003. *Tourism and Local Agenda 21. The Role of Local Authorities in Sustainable Tourism*, UNEP.
- LEMMI E. 2009. *Dallo spazio consumato ai luoghi ritrovati. Verso una geografia del turismo sostenibile*. Milano: Franco Angeli.
- MORELLI U. (1995). *Management delle imprese turistiche*. Milano: Etas Libri.
- PALMER A. 1997. Public-Private Sector Tourism Marketing: Exploratory Study of Antecedents of Effectiveness. *Working Paper, De Monfort University, Leicester Business School*.
- PENATI C. 2007. *Governare con il Territorio*. Roma: Formez-Ufficio Stampa ed Editoria.
- RULLANI E. 1994. L'impresa reale e le sue prospettive, *Economia e Politica Industriale*, No 82.

SAVELLI A. 2004. *Turismo, territorio, identità. Ricerche ed esperienze nell'area mediterranea*. Milano: Franco Angeli.

TRILLO O. C. 2003. *Territori del turismo. Tra utopia e atopia*. Firenze: Alinea.

SUMMARY

The basic theme of this paper is to reconcile two aspects of tourism, globalization and local identity, which equally fundamental processes in society in the early twenty-first century. Globalization, on the one hand, everything seems to connect and flatten everything, but on the other hand, develops an infinite range of possible paths for different groups or individual tourists. The sense of the local comes out and defeats from within globalization. Tourist of globalization poses a direct relationship with the places visited, with the area and its historical memory, its culture and its environment, showing a demand for authenticity, uniqueness and spontaneity of products sometimes not born for tourism and do not find anywhere else in the world (local aspect). In this context, the tourism product must become a product of systemic and local systems tourism (SLOT) is a winning strategy for the development of a destination that respects in the time a tourism that is not only attractive, but services, facilities and infrastructure that allow the tourist a positive and unforgettable experience.

PIL, BENESSERE ECONOMICO E DISTRIBUZIONE DEL REDDITO: IL CASO DELL'ITALIA

Fabrizio Antolini

1. Rappresentatività ed evoluzione del Pil

L'ampio dibattito che si sta sviluppando nel corso di questi anni circa l'inadeguatezza del Pil nel rappresentare anche performance sociali, in realtà era già presente nell'immediato dopoguerra (Jaszi G., 1951). Nonostante l'esigenza fosse già avvertita, sia dai policy maker, sia dagli analisti, lo schema di contabilità nazionale si limitava a rappresentare le operazioni condotte dagli operatori economici. Le prime tavole di contabilità nazionale nascono così "redditocentriche", anche perché il reddito prodotto in ciascun paese veniva utilizzato per ripartire le risorse finanziarie del piano Marshall¹.

Negli anni immediatamente successivi alla guerra, in un contesto storico di povertà diffusa l'incremento del reddito prodotto divenne il principale obiettivo dei policy maker; dal suo livello dipendeva il raggiungimento del consumo di sussistenza. Successivamente durante il periodo della crescita economica, il reddito nazionale finì con l'assumere anche il significato di *progresso*, venendo associato al suo maggior livello il passaggio da sistemi economici prevalentemente *rurali*, a sistemi economici più *complessi*². Il sistema dei conti economici nazionali trova però la sua prima ufficiale veste con lo Sna68³, il quale, pur rappresentando un sistema fondato su una teoria economica di riferimento e sul meccanismo dei saldi contabili, continuava a non rispondere a quelle esigenze informative precedentemente illustrate⁴. Difficoltà nell'adottare definizioni omogenee nei confronti di fenomeni complessi, scongiurarono però di allargare l'area di interesse

¹ Si veda per una descrizione sintetica Lequiller F. Blades D. (2006).

² Durante la guerre era la disponibilità di ferro a segnare la potenza ed il potenziale di progresso delle nazioni.

³ In verità l'origine dello Sna deve essere fatta risalire al 1947, mentre una prima versione si ebbe nel 1953. Per maggiori approfondimenti si veda <http://unstats.un.org/unsd/nationalaccount/hsna.asp>, ma è nel 1968 che si giunge ad una armonizzazione dei diversi aggregati.

⁴ Per maggiori approfondimenti si veda la "Conference on the Measurement of Economic and Social Performance, Novembre 1971, Princeton University.

della contabilità nazionale anche alle tematiche sociali. Nel corso del tempo molti paesi occidentali riescono a raggiungere un livello di reddito di “opulenza” e l’attenzione si sposta dal *reddito prodotto* a quello *distribuito*. A questo preciso riguardo lo Sna93 interviene articolando ulteriormente la distribuzione primaria del reddito e, soprattutto, introducendo quella terziaria (in natura). Si consente così di comprendere meglio il funzionamento dei sistemi di welfare. Alla fine degli anni Novanta però, con l’avvento della globalizzazione, i sistemi economici subiscono nuovi importanti mutamenti: le imprese iniziano a delocalizzare le proprie produzioni; la società della conoscenza – alla base della teoria della crescita endogena – diviene una realtà sempre più diffusa; la graduale finanziarizzazione dell’economia produce un incremento del numero di istituzioni finanziarie; vi è l’esigenza di ridefinire meglio il settore pubblico, anche per comprendere meglio l’attività delle società miste

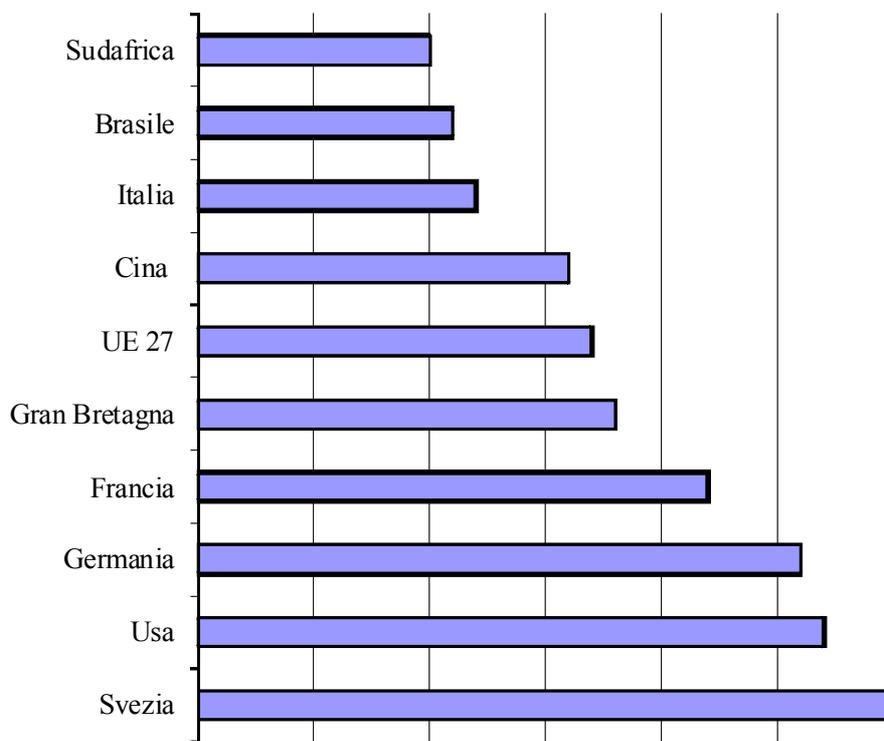
Come già avvenuto in passato, al mutamento istituzionale, corrisponde una nuova esigenza informativa, in parte recepita dallo *Sna2008*.

2. Sna 2008 e rapporto Stiglitz: l’evoluzione continua

Lo Sna2008, che entrerà in vigore a partire dal 2014, pur conservando la struttura del precedente sistema di contabilità nazionale, apporta alcune (limitate) modifiche allo Sna93⁵. Si introduce un nuovo trattamento per le spese per *R&D*, considerate *investimento* ma, più in generale, si presta maggiore attenzione alle *attività intangibili*; inoltre vi è una diversa considerazione delle imprese pubbliche/private con una parziale revisione del settore pubblico e, infine, una differente considerazione viene riservata al debito pensionistico nei sistemi senza costituzione di riserve e alle assicurazioni in seguito a catastrofi naturali (Caricchia A., 2009; Giovannini E., 2009)

In verità i cambiamenti intervenuti nelle classificazioni, possono avere un impatto molto differente sul Pil, a seconda del paese e della sua “organizzazione istituzionale”. Si consideri ad esempio la spesa pubblica per *R&D*, la cui riclassificazione in investimenti può generare la sensazione che i paesi che presentano valori monetari più alti, siano anche quelli che beneficeranno di maggiori incrementi di Pil (Figura 1).

⁵ Per avere una visione chiara dell’ampio dibattito che si è sviluppato sulla revisione della contabilità nazionale si veda <http://unstats.un.org/unsd/nationalaccount/snaUpdate.asp>

Figura 1 – Spesa Intervento Pubblico per la ricerca sul Pil.

Fonte: Oecd.

In realtà per comprendere quale sia il reale impatto sul Pil, si dovrebbero considerare molti aspetti, tra i quali ad esempio, la quota di spesa in R&D *non market*, oppure se la stessa apporta dei benefici economici⁶. Infine, la stessa definizione di spesa R&D dovrebbe essere coerente con quanto contenuto nello Sna, mentre alcuni paesi utilizzano la definizione contenuta nel Manuale di Frascati. Nonostante queste novità, la rappresentatività degli aggregati contabili, primo tra tutti il Pil, rimangono ancora saldamente ancorati ad una visione economica.

Durante la pubblicazione dello Sna 2008, accade però un fenomeno nuovo: si verifica, a seguito di scandali finanziari e comportamenti speculativi, una violenta crisi economica che però, in questo caso, è anche etica. Nasce così una nuova esigenza informativa volta a misurare anche i comportamenti umani che generano

⁶ Se la spesa è non market essa, contabilizzata dal lato dei costi, potrebbe già essere inclusa nel Pil.

l'incremento di ricchezza e, quindi, a valutare l'*eticità* dei sistemi economici e la felicità degli individui⁷.

Ma, sebbene come precedentemente illustrato, lo Sna 2008 non muti la struttura del precedente sistema, una ulteriore novità riguarda la migliore conoscenza ed utilizzazione della spesa delle famiglie, per meglio valutare l'attività di *welfare* (United Nations 2009, *par.1.75* e ss.).

Diversamente dallo Sna2008, uno dei principali obiettivi del "Rapporto Stiglitz" è stato di definire i limiti del Pil come indicatore della performance economica e del progresso sociale. Si voleva comprendere meglio il problema riguardante la sua misurazione, nonché riuscire ad individuare strumenti o fonti che potessero essere utilizzati per la descrizione del progresso sociale. Ad esempio "quando ci sono grandi cambiamenti nella disuguaglianza (più in generale, un cambiamento nella distribuzione del reddito), il prodotto interno lordo o qualsiasi altro aggregato calcolato pro-capite, può non fornire una valutazione precisa della situazione in cui molte persone si trovano. Se aumenta molto la differenza rispetto alla media di aumento del Pil pro capite, la maggior parte delle persone potrebbe trovarsi peggio, anche se il reddito medio sta registrando un incremento" (Rapporto Stiglitz, 2010).

Un ulteriore obiettivo del Rapporto nel comprendere il legame esistente tra Pil e benessere, era di individuare le linee guida per riuscire a sviluppare un sistema statistico che riuscisse ad integrare le misure delle attività di mercato con quelle incentrate sul benessere dei popoli e sulla sostenibilità. Ancora una volta però, nel riuscire a stabilire dei legami tra il Pil ed il benessere, sarebbe importante avere una definizione condivisa di benessere, ad esempio è importante stabilire se ci si debba riferire al benessere economico materiale, oppure al benessere sociale o, infine, al benessere individuale. La principale criticità consiste proprio nel riuscire ad individuare il punto di divergenza tra il benessere economico materiale e quello sociale, perché questo contribuirebbe a modificare il paradigma di riferimento per valutare il progresso della società. Così facendo però si assume una visione "organicista" della società con valori culturali che la alimentano e la soddisfano, a prescindere del valore monetario dei beni prodotti. Se ci si riferisce invece al benessere individuale, esso *equivale allo star-bene* e, quindi, la sua analisi sarebbe molto vicina a quella della *felicità*; rimane da stabilire se sia sufficiente procedere per semplice aggregazione dei benessere individuali per acquisire una dimensione sociale⁸.

⁷ L'affermazione non deve essere intesa nel senso che è a partire da questo momento che inizia il dibattito sul pil della felicità; l'analisi tra il Pil e la felicità inizia ad Istanbul nel 2004, riprendendo iniziative che in maniera non organica si erano sviluppate nei diversi paesi.

⁸ Si neherebbe in questo caso l'impostazione che interpreta la società come espressione organica.

Ancora una volta, prima di procedere a qualsiasi processo di misurazione è importante definire il collettivo di riferimento ma, per questo, è importante concettualizzare il fenomeno, utilizzando un sistema induttivo-deduttivo.

3. L'analisi della distribuzione del reddito in Italia

Analizzare i fenomeni economici e sociali è sempre difficile, particolarmente in Italia, dove l'esistenza di evidenti disparità territoriali, può produrre sintesi interpretative non pienamente rappresentative. Gli stessi principali indicatori macro-economici utilizzati per valutare le performance economiche nazionali, spesso presentano una elevata variabilità se ricondotti in ambito territoriale locale.

Tabella 1 – *Indice di Gini regionale.*

| Regioni | Indice di concentrazione di Gini |
|------------------------------|----------------------------------|
| Piemonte | 0,291 |
| Valle d'Aosta/Vallée d'Aoste | 0,310 |
| Lombardia | 0,295 |
| Liguria | 0,290 |
| Bolzano/Bozen | 0,298 |
| Trento | 0,280 |
| Veneto | 0,266 |
| Friuli-Venezia Giulia | 0,265 |
| Emilia-Romagna | 0,297 |
| Toscana | 0,283 |
| Umbria | 0,280 |
| Marche | 0,289 |
| Lazio | 0,324 |
| Abruzzo | 0,263 |
| Molise | 0,319 |
| Campania | 0,327 |
| Puglia | 0,310 |
| Basilicata | 0,289 |
| Calabria | 0,314 |
| Sicilia | 0,335 |
| Sardegna | 0,292 |
| Italia | 0,314 |

Fonte: Istat.

Fatta questa premessa, l'analisi della distribuzione del reddito, rappresenta un primo passo per stabilire il benessere economico di una società. Nel caso

dell'indice di concentrazione di Gini, oltre a presentare una propria variabilità territoriale (Tabella 1), si dimostra scarsamente sensibile alle variazioni distributive e, quindi, la sua sintesi interpretativa, non riesce a descrivere compiutamente il mutamento nella distribuzione del reddito. Un tentativo per comprendere meglio la distribuzione del reddito può essere fatto mettendo in *relazione il prodotto interno lordo con il reddito disponibile delle famiglie*. Com'è noto il settore istituzionale famiglie (S14) viene ulteriormente articolato nei conti economici per settore istituzionale, in *famiglie produttrici e famiglie consumatrici*.

La distinzione non è irrilevante; sarà sufficiente al riguardo considerare come nel vecchio Sec'70, all'interno del settore famiglie rinvenivamo le imprese al di sotto dei 20 addetti, mentre nel Sec95 troviamo le unità al di sotto dei 5 addetti. Ma in un sistema economico dove il 56 per cento delle unità economiche sono unipersonali e il 95 per cento sono al di sotto dei 10 addetti, considerare il reddito disponibile delle famiglie nel loro complesso, oppure delle sole *famiglie consumatrici*, non è equivalente almeno sotto il profilo distributivo.

Per questa ragione, ma anche per tener presente quanto raccomandato dal *Rapporto Stiglitz*⁹, la nostra attenzione sarà focalizzata nell'analizzare la relazione che intercorre tra il prodotto interno lordo dell'intero sistema economico, ed il reddito disponibile delle sole famiglie consumatrici.

L'esistenza di trend stocastici nelle variabili considerate richiede, prima di procedere all'identificazione del modello, di procedere ad una differenziazione delle variabili opportunamente trasformate in logaritmi. In questo modo riusciamo ad ottenere anche una misura dell'elasticità del reddito disponibile delle f.c. al variare del reddito prodotto all'interno sistema economico. Il modello considerato sarà quindi il seguente:

$$\Delta \log Red_dis_f.c. = \alpha + \beta \Delta \log P_{it} + \varepsilon \quad (1)$$

Per la stima dei coefficienti, la presenza di autocorrelazione seriale nei residui, consiglia di utilizzare come metodo di stima l'FGLS Cochrane-Orcutt (CO)¹⁰.

Il coefficiente stimato pari a 0,88 indica una parziale rigidità del reddito disponibile delle famiglie consumatrici rispetto al reddito prodotto nel sistema economico. Però prima di concludere che le famiglie italiane hanno peggiorato la distribuzione del proprio reddito, è forse opportuno estendere la nostra riflessione.

⁹ In particolare la Raccomandazione 2, che sollecita di rafforzare l'analisi sotto il profilo delle famiglie.

¹⁰ L' R^2 è pari a 0,57, mentre il D.W. è 2,17; anche gli altri test rifiutano l'ipotesi di correlazione dei residui. La stima non cambia utilizzando il metodo Prais-Winsten, diversamente, utilizza l'OLS, il valore del coefficiente *sale a* 1,06, ma la presenza di autocorrelazione nei residui, suggerisce una metodologia di stima differente.

Tabella 2 – *Elasticità del reddito disponibile delle famiglie consumatrici rispetto al Pil FGLS- Cochrane-Orcutt.*

| | <i>Coefficiente</i> | <i>Errore Std.</i> | <i>rapporto t</i> | <i>p-value</i> | |
|--------|---------------------|--------------------|-------------------|----------------|-----|
| Const | -0,00222548 | 0,00952932 | -0,2335 | 0,81872 | |
| ld_Pil | 0,876318 | 0,210544 | 4,1622 | 0,00096 | *** |

Note: Variabile dipendente: *ld_Red_disp_f_c*.

Per poter effettuare una valutazione esaustiva, occorre considerare anche il modo in cui è strutturato il nostro welfare, il quale, prevedendo anche l'erogazione di servizi nell'istruzione, nella sanità/assistenza oppure nei trasporti pubblici, produce una componente non monetaria di reddito. Considerando questa componente, il valore complessivo della redistribuzione sarebbe più alto rispetto al coefficiente stimato. Anche questa affermazione però può essere oggetto di ulteriori riflessioni dal momento che questi servizi, sono finanziati/erogati con una imposizione fiscale differenziata a seconda della fascia di reddito dei beneficiari (Antolini F., Petrei F., 2008). I dati di contabilità nazionale non colgono questo aspetto e considerano le famiglie consumatrici (finanziatori e beneficiari) come fossero tutte una grande famiglia. Infine, dobbiamo considerare che gli aggregati considerati, ovvero il Pil ed il reddito delle famiglie consumatrici, espressi a prezzi correnti, non danno la possibilità di esaminare gli effetti distributivi generati dall'inflazione.

Riferimenti bibliografici

- ANTOLINI F., PETREI F. 2008. L'Isee può essere considerato un indicatore di povertà economica? *Rivista Italiana di Economia, Demografia e Statistica*. Vol. LXII, No 2/4, pp. 269 – 276.
- RAPPORTO DELLA COMMISSIONE PER LA MISURAZIONE DELLA PERFORMANCE ECONOMICA E DEL PROGRESSO SOCIALE. 2010. *La misura sbagliata delle nostre vite: perché il PIL non basta più per valutare benessere e progresso sociale*. Roma: Edizione Etas.
- GIOVANNINI E. 2008. The Revision of the 1993 Sna: a missed opportunity? Presentato a: *Convegno Iscona*, Roma.
- CARICCHIA A. 2009. Intervento sulla Revisione dello Sna 2008, Presentato a: *Convegno Iscona*, Roma.
- GIORGI G. M. 2011. Corrado Gini. The man and the scientist. *Metron*, Vol.1.
- JASZI G. 1951. Statistical Sources and Methods in National Accounts Estimates and the problem of Reliability. *International Association for Research in Income and Wealth*. Serie III.

- LEQUILLER F., BLADES D. 2006. *Understanding National Account*, Oecd Manual, Paris.
- UNITED NATION. 2009. *System of National Accounts 2008*. New York.
- UNITED NATION. 2002. Use of National Accounts per policy analysis. Series F. No.81, New York.
- WOOLDRIDGE J. 2009. *Introductory Econometrics, Fourth Edition*.
- UNITED NATION. 2000. *Integrated Environmental and Economic Accounting, An operation Manual*. Series F., No.7.

SUMMARY

During the last years, in Italy as in other countries, GDP's volatility caused Government interventions aimed to support household income. Income distribution is a key element of social cohesion, however, as suggested by Report Stiglitz, it also provides relevant information on the economic *well-being*.

The aim of this paper is to analyze the dynamic of the income distribution in Italy respect to the National GDP over the last 20 years. The analysis is performed using as data set the National Accounts data.

MERCATO TURISTICO E ORGANIZZAZIONE DEL TERRITORIO: I DISTRETTI TURISTICI IN SICILIA¹

Vincenzo Asero, Rosario D'Agata, Venera Tomaselli

1. Turismo e territorio

Il turismo è considerato il settore su cui puntare per lo sviluppo socio economico di un territorio, grazie agli effetti che il fenomeno è in grado di determinare in termini di creazione di redditi e di occupazione, di stimolo agli investimenti e di capacità di attivare altri settori dell'economia. La natura eterogenea del prodotto turistico insieme al complesso di relazioni che si sviluppano tra imprese, turisti e residenti connotano il turismo come sistema. In rapporto alla dimensione di un territorio è possibile identificare sistemi di scala differenti, complessi ed aperti, ai quali danno il loro apporto sub-sistemi o altri sistemi, e settori produttivi diversi. In questo quadro vanno inserite le azioni di *destination management*, volte all'organizzazione ed alla gestione dell'offerta turistica (Asero *et al.*, 2010). Strumentali a queste azioni sono l'analisi della spesa turistica e dei comportamenti di spesa, da cui ricavare informazioni per la comprensione dei benefici economici per una destinazione e alla definizione delle strategie di mercato.

In Sicilia il concetto di turismo come sistema trova riscontro negli istituendi distretti turistici (DT), che costituiscono l'unità territoriale d'analisi del nostro studio, orientato all'identificazione di profili tipologici di presenze turistiche all'interno dei 16 DT territoriali in fase di riconoscimento: 1. *Selinunte il Belice e Sciacca Terme*; 2. *Sicilia Occidentale*; 3. *Iblei*; 4. *Il mare dell'Etna*; 5. *Siracusa e Val di Noto*; 6. *Golfo di Castellammare*; 7. *Palermo Costa Normanna*; 8. *Valle dei Templi*; 9. *Tirreno-Nebrodi*; 10. *Thyrrenium Tyndaris - Parco dei Miti*; 11. *Venere di Morgantina*; 12. *Taormina Etna*; 13. *Isole e Arcipelaghi di Sicilia*; 14. *Monti Sicani e Valle dei Platani*; 15. *Sicilia Centro Meridionale*; 16. *Cefalù e Parchi delle Madonie e di Himera*. Il nostro obiettivo è comprendere come si configura la domanda rispetto alle caratteristiche strutturali ed ai comportamenti di spesa al loro interno.

¹ Il contributo degli autori al presente lavoro è da considerarsi paritetico sotto ogni aspetto. I dati di riferimento provengono dall'indagine effettuata all'interno del progetto PRIN 2007-2009 "Mobilità del turismo regionale *incoming*. Aspetti socio-economici dei comportamenti e delle motivazioni".

2. I distretti turistici in Sicilia

Nel rapporto con il territorio il turismo evolve e si sviluppa con modalità proprie di un'entità sistemica, attraverso processi d'integrazione e di cooperazione tra le componenti della filiera turistica, incentrati su forme di progettualità endogena, capaci di determinare forme di specializzazione dell'offerta territoriale. Quanto più strette e funzionali sono le dinamiche organizzative interne ad una destinazione, tanto maggiori sono per una certa località le probabilità di successo in termini di mercato (Giannone, 2006).

La dimensione sistemica determinata dal binomio turismo e territorio appare trovare conferma in Italia nella Legge 135/2001, "Riforma della legislazione nazionale del turismo", che all'art. 5 definisce i "Sistemi Turistici Locali" (STL). I STL costituiscono un nuovo modello di riferimento per la riorganizzazione del settore turistico delle regioni italiane e per la definizione dell'offerta turistica territoriale. In questa direzione si è mossa anche la Sicilia con la L. R. 10/2005 che utilizza all'art. 6 la dizione "distretti turistici", enfatizzando la valenza del cosiddetto 'effetto distretto', determinato, come evidenziato dalla letteratura sui distretti, dalla concentrazione su uno spazio territoriale di imprese che operano lungo la stessa filiera produttiva. Sia per i STL sia per i DT, il modello aggregativo di riferimento segue un approccio *bottom-up*, poiché offre la possibilità agli attori locali, pubblici e privati, di organizzare la propria offerta turistica, attraverso l'autonoma definizione di ambiti geografici di riferimento che, accompagnati ad una comune idea progettuale, creano sistema all'interno di un dato territorio (Antonoli Corigliano, 1999). Sistemi e distretti turistici mirano allo sviluppo complessivo del territorio locale, incrementando l'efficienza del prodotto e aumentando la concorrenzialità di una destinazione (Candela e Figini, 2005).

Il lungo iter per riorganizzare il settore turistico in Sicilia, segnato da accesi dibattiti e numerose proposte, sembra ormai giunto a compimento. Attualmente, infatti, sono in attesa di definitiva valutazione e riconoscimento le proposte di 16 distretti "territoriali" e 11 distretti "tematici", definiti sulla base di quanto disposto dal Decreto n. 4/2010 che fissa criteri e modalità per il loro riconoscimento, prescrivendo alcuni requisiti, pena l'inammissibilità.

3. Studiare il mercato: il turista sul territorio

Ai fini dello studio della domanda turistica secondo la suddivisione in unità distrettuali, è stato utilizzato il *database* dell'indagine campionaria effettuata da lu-

glio ad ottobre del 2009 e da marzo a giugno del 2010, sui turisti in partenza dai porti ed aeroporti di Catania e Palermo, nell'ambito del progetto PRIN (*infra*)².

Proprio in funzione di questi obiettivi, dal campione iniziale di 3935 turisti, sono stati selezionati 3233 turisti auto-organizzati³, che hanno prevalentemente lasciato le loro tracce di spesa all'interno del territorio del loro soggiorno. Di questi sono stati considerati soltanto coloro che hanno risposto alle domande relative alla quantificazione delle singole voci di spesa, pari a 2488. L'ulteriore selezione ha riguardato i 2238 turisti auto-organizzati che hanno scelto come meta della loro vacanza una singola destinazione corrispondente ad un comune ricadente all'interno di un distretto territoriale. Questi sono turisti che hanno effettuato almeno un pernottamento presso strutture ubicate nel comune di destinazione.

Le variabili selezionate dal *database* riguardano informazioni socio demografiche sui turisti, su modalità, motivazioni e caratterizzazione della vacanza ed articolazione in singole voci della spesa sostenuta al netto della spesa relativa al viaggio⁴. Tutti i valori delle altre voci di spesa - mobilità interna, pranzi e ristoranti, visite a musei, enogastronomia ed artigianato, shopping, spettacoli ed eventi, alloggio, prodotti alimentari da cucinare ed altro - sono stati calcolati come valori *pro-capite* e *pro-die*.

Dopo aver eseguito un'analisi esplorativa sui dati, abbiamo utilizzato soltanto quelle variabili che, selezionate in base alla massima variabilità che le loro distribuzioni presentavano, risultavano più discriminanti rispetto all'obiettivo di ottenere una struttura di classificazione (cfr. Tabella 1).

Tabella 1 – Caratteri considerati ai fini della classificazione.

| | |
|------------------------------------------------------------|------------------------------------------------------------------|
| Motivazione principale del viaggio | Piacere, svago / Visita a parenti-amici |
| Notti trascorse in Sicilia | ≤ 7 / 8-14 / 15-21 / 22-30 / >30 |
| Caratterizzazione della vacanza | Esclusivamente balneare / In parte balneare / Per nulla balneare |
| Tipologia di alloggio | Hotel-Residence / Alloggio gratuito |
| Spesa pro-capite e pro-die al netto della spesa di viaggio | Valori continui |

A tal fine è stata eseguita l'analisi di cluster (CLA) secondo la procedura denominata *Two-Step clustering* (Zhang *et al.*, 1996), considerando le modalità dei caratteri categoriali e la variabile continua, da cui sono stati ricavati opportuni rag-

² Per la descrizione del disegno della ricerca e del piano di campionamento si rimanda ai contributi contenuti nel volume a cura di Parroco e Vaccina (2010), citato in bibliografia.

³ Sono turisti auto-organizzati coloro i quali non fanno ricorso a *tour operator* e/o agenzie di viaggio.

⁴ Gli effetti della spesa di viaggio, infatti, non ricadono sullo specifico territorio/distretto visitato perché questa è prevalentemente effettuata nel luogo di provenienza del turista.

gruppamenti classificatori delle unità medesime⁵. I risultati conseguiti dall'applicazione della CLA consentono di individuare 2 cluster (cfr. Figura 1).

Figura 1 – *Classificazione delle unità d'analisi.*

| Gruppo | 1 | 2 |
|----------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|
| Dimensione |  63,8% (1040) |  36,2% (591) |
| Funzioni | Qual è la motivazione principale del viaggio? Piacere, svago (92,3%) | Qual è la motivazione principale del viaggio? Visita a parenti-amici (97,3%) |
| Importanza della funzione  | Alloggio prima destinazione Hotel, residence (36,3%) | Alloggio prima destinazione Alloggio gratuito (72,8%) |
| | SpesaPRC_SpViag PRC_Prodie 84,25 | SpesaPRC_SpViag PRC_Prodie 35,24 |
| | Ha fatto una vacanza balneare? In parte balneare (50,1%) | Ha fatto una vacanza balneare? Per nulla balneare (39,1%) |
| | Notti trascorse Fino a 7 notti (62,2%) | Notti trascorse Fino a 7 notti (51,9%) |

La variabile che contribuisce in misura prevalente alla definizione della struttura di classificazione è *Motivazione principale del viaggio*, seguita da *Tipologia di alloggio*. Prima di valutare i risultati ottenuti, è stata controllata la validità della classificazione, mediante l'analisi discriminante applicata alle variabili originarie⁶, dai cui emerge che il 95,6% delle unità è stato classificato correttamente.

L'analisi della distribuzione delle presenze nei DT siciliani ha consentito di evidenziare due differenti tipologie di turisti, rispettivamente da noi denominati 'Solo turisti' (cluster 1) e 'Non solo turisti' (cluster 2). Sono stati considerati, così, sia i risultati della procedura di classificazione sia altri caratteri rilevati - genere, fascia d'età, provenienza, modalità di vacanza effettuata da solo o con famiglia oppure con amici - utilizzati in funzione di meri descrittori dei 2 cluster.

Escludendo il genere e la fascia d'età perché non significativamente associati con i 2 gruppi, nel cluster dei 'Solo turisti' prevalgono coloro che scelgono le destinazioni turistiche siciliane prevalentemente come luoghi di piacere e di svago (92,3%). Questi turisti, in particolare, sono presenti in misura maggiore all'interno

⁵ La misura di distanza prescelta per classificare le unità d'analisi è la funzione di verosimiglianza. Tale funzione fornisce la probabilità di appartenenza di ciascuna unità ad un dato gruppo sulla base della maggiore somiglianza. Dalla matrice di somiglianza tra unità, si individuano i fattori che contribuiscono alla classificazione e le relative misure.

⁶ Per la validazione della struttura di classificazione, l'analisi discriminante crea modelli predittivi di gruppi di appartenenza in funzione delle variabili osservate per ciascuna unità (Fabbris, 1997, p. 335).

di quei DT dove sono localizzate le più rinomate località dell'Isola. Scelgono strutture ricettive ufficiali - hotel/residence (36,3%) - ed anche strutture ricettive extra-alberghiere (casa o stanza in affitto, B&B, agriturismo, camping ed ostelli). Spendono in media 84,25 euro al giorno *pro-capite* e relativamente ad alcune voci che connotano la spesa turistica (alloggio, ristorazione e mobilità interna), questa risulta più elevata rispetto a quella dei 'Non solo turisti'. Trascorrono una vacanza in parte balneare (50,1%) e soggiornano da 1 a 7 notti (62,2%).

La presenza dei 'Non solo turisti' prevale in quei DT formati da comuni dove è minore la specializzazione turistica. Questa tipologia di turisti è caratterizzata da soggetti che scelgono la Sicilia come meta delle loro vacanze per visitare parenti e/o amici (97,3%). Godono gratuitamente di un alloggio (72,8%). Spendono mediamente 35,24 euro al giorno *pro-capite*. Le loro vacanze sono assolutamente non balneari (39,1%) o in parte balneari. Anche i 'Non solo turisti' trascorrono al massimo 7 notti nella destinazione prescelta (51,9%).

Dal punto di vista descrittivo, nel primo cluster è presente il maggior numero di stranieri (67,9%) rispetto al gruppo dei 'Non solo turisti' (32,7%). Fra i turisti che viaggiano da soli prevalgono coloro che appartengono al secondo cluster (61,7%). Fra coloro che trascorrono le loro vacanze in famiglia, ben il 66,5% si colloca nel cluster dei 'Solo turisti', mentre il secondo cluster ne annovera il 33,5. Fra coloro che, invece, effettuano le loro vacanze in compagnia di amici, l'83,5% si ritrova fra i 'Solo turisti', mentre soltanto il 16,5% sono 'Non solo turisti'.

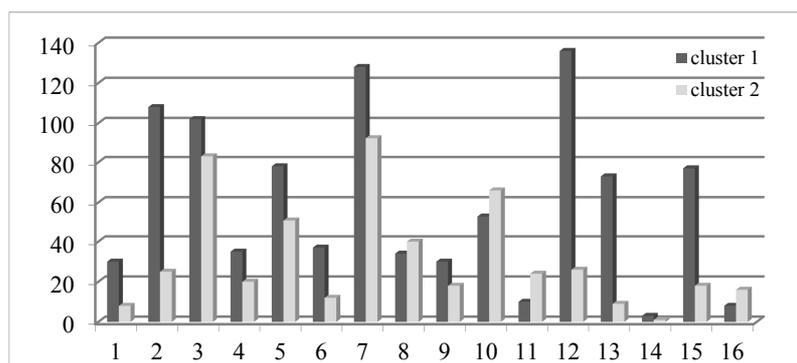
4. I distretti turistici tra presenze e comportamenti di spesa

La distribuzione del fenomeno secondo il criterio di suddivisione distrettuale rivela che il DT che ha registrato il maggior numero di presenze è il DT 7.*Palermo Costa Normanna* (cfr. Tabella 2). Se consideriamo la media delle notti trascorse, tuttavia, rileviamo che è il DT 16.*Cefalù e Parchi delle Madonie e di Himera* quello in cui la vacanza è mediamente più lunga, seguito dal distretto 8.*Valle dei Templi* e dal distretto 3.*Iblei*. Per quanto concerne la spesa *pro-capite*, notiamo come essa sia mediamente più elevata nel DT 13.*Isole e Arcipelaghi di Sicilia* e nel DT 12.*Taormina Etna*, all'interno dei quali ricadono alcune località turistiche tra le più note della Sicilia. Inoltre, nel DT 12.*Taormina Etna* si registra la spesa *pro-die* più elevata. In media, infatti, per ogni giorno trascorso sono stati spesi circa 140 euro, quasi il doppio della media della spesa *pro-die* calcolata sul totale dei DT (71,72 euro).

La distribuzione dei due cluster nei 16 DT è rappresentata nella Figura 2.

Tabella 2 – Presenze, notti trascorse e indicatori di spesa all'interno dei distretti.

| Distretto | Presenze | Notti trascorse | Spesa (€) | Media notti | Spesa pro cap. | Spesa pro die |
|-----------|----------|-----------------|-----------|-------------|----------------|---------------|
| 1 | 56 | 601 | 35147 | 10,73 | 627,63 | 58,48 |
| 2 | 186 | 1694 | 130392 | 9,11 | 701,03 | 76,97 |
| 3 | 207 | 2737 | 221749 | 13,22 | 1071,25 | 81,02 |
| 4 | 63 | 603 | 60169 | 9,57 | 955,06 | 99,78 |
| 5 | 178 | 1628 | 178784 | 9,15 | 1004,4 | 109,82 |
| 6 | 60 | 667 | 47385 | 11,12 | 789,75 | 71,04 |
| 7 | 413 | 3022 | 142083 | 7,32 | 344,03 | 47,02 |
| 8 | 100 | 1424 | 76471 | 14,24 | 764,71 | 53,7 |
| 9 | 65 | 752 | 38937 | 11,57 | 599,03 | 51,78 |
| 10 | 165 | 1597 | 109315 | 9,68 | 662,52 | 68,45 |
| 11 | 42 | 551 | 32235 | 13,12 | 767,5 | 58,5 |
| 12 | 199 | 1671 | 232857 | 8,40 | 1170,14 | 139,35 |
| 13 | 97 | 1077 | 114946 | 11,10 | 1185,01 | 106,73 |
| 14 | 7 | 69 | 1750 | 9,86 | 250 | 25,36 |
| 15 | 135 | 1461 | 101035 | 10,82 | 748,41 | 69,15 |
| 16 | 31 | 582 | 17720 | 18,77 | 571,61 | 30,45 |

Figura 2 – Distribuzione dei cluster all'interno dei distretti.

Prima di approfondire ulteriormente la composizione della spesa all'interno di ciascun DT, è utile appurare se esiste una relazione statisticamente significativa tra tipologia di turista e DT. A questo scopo, applichiamo il test χ^2 utilizzando la classificazione in cluster illustrata nel paragrafo precedente, da cui si ricava un valore pari a 154,86, con un livello di significatività molto elevato ($p\text{-value} < 0,00001$). Ciò induce a ritenere che esiste una relazione tra tipologia di turista e DT.

I 'Solo turisti', come sono stati definiti gli appartenenti al primo cluster, sono prevalenti nel DT 13. *Isole e Arcipelaghi di Sicilia*, dove rappresentano l'89% di tutti i turisti e nel DT 12. *Taormina Etna*, dove su 162 turisti classificati, 136 rientrano nel primo gruppo. In generale i 'Solo turisti' sono prevalenti in tutti i DT, con

l'eccezione dei DT 11. *Venere di Morgantina*, 10. *Thyrrenium Tyndaris - Parco dei Miti* ed 8. *Valle dei Templi*, dove prevalgono i 'Non solo turisti'.

L'ultima parte del lavoro focalizza l'attenzione sulla composizione della spesa osservata all'interno di ciascun distretto e considerando come elemento discriminante l'appartenenza ai cluster. Un preliminare test di Kolmogorov-Smirnov ha permesso di rifiutare l'ipotesi di normalità della distribuzione, impedendo, quindi, l'impiego del test dell'analisi della varianza. Il confronto tra i due cluster, pertanto, è operato utilizzando il test di Wilcoxon-Mann-Whitney (WMW) che considera la media dei ranghi (Siegel e Castellan, 1992, p. 180). La tabella 3 evidenzia le differenze per ciascuna variabile di spesa (*pro-capite*) all'interno di ciascun DT tra il rango medio del primo cluster ed il rango medio del secondo cluster.

Tabella 3 – Test di WMW: differenze tra le medie dei ranghi dei due cluster.

| Distretto | Mobilità | Ristoranti | Musei | Enogastronomia | Shopping | Alloggio | Alimentari | Spesa Tot. |
|-----------|----------|------------|---------|----------------|----------|----------|------------|------------|
| 1 | - | - | - | - | - | - | - | - |
| 2 | 28,27** | 27,66** | - | - | -21,92** | 33,36*** | - | 29,63** |
| 3 | - | - | 12,08* | 16,37* | - | 40,30*** | - | 18,13* |
| 4 | 12,85** | - | - | - | - | 15,79*** | - | 13,08** |
| 5 | 21,32** | 21,50** | 17,77** | 14,04* | - | 31,73*** | - | 17,61** |
| 6 | - | - | - | - | - | 18,54*** | - | - |
| 7 | 26,84** | 28,97** | 27,00** | 18,37* | - | 63,05*** | -15,9** | 28,56** |
| 8 | - | - | 11,64** | - | - | 15,51*** | -9,28** | - |
| 9 | - | - | - | - | - | - | - | - |
| 10 | 19,46** | 14,34* | 12,25** | - | - | 22,42*** | - | 20,02** |
| 11 | - | - | - | - | - | - | - | - |
| 12 | 30,72** | 22,82* | 22,31* | - | - | 57,13*** | - | 38,97*** |
| 13 | - | - | - | - | -22,47** | 27,52** | - | - |
| 14 | - | - | - | - | - | - | - | - |
| 15 | 28,96*** | - | - | - | - | 26,63*** | - | 24,13** |
| 16 | - | - | - | - | - | - | - | - |

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$. I valori statisticamente non significativi ($p > 0,05$) non sono stati riportati.

Come si può osservare, il gruppo dei 'Solo turisti' spende tendenzialmente di più dei 'Non solo turisti' nella maggior parte dei DT. In particolare, spende in misura maggiore rispetto al secondo gruppo per l'alloggio in quasi tutti i DT, con differenze elevate nel DT 7. *Palermo Costa Normanna* e nel 12. *Taormina Etna*. Il cluster dei 'Non solo turisti' mostra un rango medio più elevato del primo gruppo per la voce 'shopping' nei DT 2. *Sicilia Occidentale* e 13. *Isole e Arcipelaghi di Sicilia* e per la voce 'alimentari' nei DT 7. *Palermo Costa Normanna* e 8. *Valle dei Templi*. Se, infine, osserviamo le differenze nelle voci di spese quali 'mobilità' (ricordando che in questa voce non è compresa la spesa di viaggio), 'ristoranti e musei', tutte concettualmente legate al consumo turistico, la distinzione tra i due gruppi è confermata, come, allo stesso modo, la sua distribuzione all'interno dei DT.

5. Conclusioni

Il lavoro ha tentato di individuare connotazioni tipologiche di presenze turistiche nei DT siciliani attualmente in fase di riconoscimento definitivo.

L'analisi delle caratterizzazioni illustrate offre informazioni utili per gli operatori turistici ed i *decision maker* ai fini dell'implementazione delle azioni di *destination management*.

Riferimenti bibliografici

- ANTONIOLI CORIGLIANO M. 1999. I distretti turistici e le aggregazioni tra attori per lo sviluppo del prodotto-destinazione. In COLANTONI M. (Ed.) *Turismo: una tappa per la ricerca*. Bologna: Pàtron Editore, pp. 99-140.
- ASERO V., D'AGATA R., TOMASELLI V. 2010. Mercato turistico e territorio in Sicilia. In: PARROCO A. M., VACCINA F. (Ed.) *Mobilità ed altri comportamenti dei turisti: studi e ricerche a confronto*. Milano: McGraw-Hill, pp. 73-89.
- CANDELA, G. e FIGINI, P. 2005. *Economia dei sistemi turistici*, Milano: McGraw-Hill.
- FABBRIS L. 1997. *Statistica multivariata*. Milano: McGraw-Hill.
- GIANNONE M. 2007. La componente territoriale nei sistemi turistici locali. In: DALL'ARA G., MORANDI F. (Ed.) *I sistemi turistici locali. Normativa, progetti e opportunità*, Matelica: Halley Editrice, pp. 71-81.
- SIEGEL S., CASTELLAN N. J. 1992. *Statistica non parametrica*. Milano: McGraw-Hill.
- ZHANG, T., RAMAKRISHNON R., LIVNY. M. 1996. BIRCH: An efficient data clustering method for very large databases. In: *Proceedings of the ACM SIGMOD Conference on Management of Data*. Montreal, Canada: ACM.

SUMMARY

The Sicilian Regional law n. 10/2005 instituted the Tourist Districts (TD). A TD is characterized by the active presence of tourist firms, public institutions and a community of people in a geographical bounded area. This work analyses the distribution of tourists inside the TD in order to identify tourist profiles, using the data set of a survey carried out in Sicily. Employing a *Two-Step clustering* procedure we obtain two clusters that we defined 'Tourists only' and 'Not tourists-only', according to some specific features.

Vincenzo ASERO, Ricercatore in Economia Politica presso la Facoltà di Scienze Politiche dell'Università degli Studi di Catania, vasero@unict.it

Rosario D'AGATA, Ricercatore in Statistica Sociale presso la Facoltà di Scienze Politiche dell'Università degli Studi di Catania, rodagata@unict.it

Venera TOMASELLI, Professore Associato in Statistica Sociale presso la Facoltà di Scienze Politiche dell'Università degli Studi di Catania, tomavene@unict.it

FATTORI GENERAZIONALI E DI GENERE NEL RAPPORTO CON IL FUMO

Nidia Batic

1. Introduzione

Si stima che in Italia i fumatori siano 11,2 milioni, pari al 22% della popolazione con più di 15 anni (ISTAT, 2006). Secondo le stime dell'OSSFAD¹ (2010) in Italia il fumo di tabacco provoca tra i 70 e gli 83 mila morti all'anno (Martinella, 2009) ed è responsabile del 35% di tutti i tumori (Lega Italiana per la lotta contro i tumori): causa il 90% delle morti per tumore polmonare, il 25% delle morti per malattie cardiache, il 75% delle morti per malattie respiratorie. Tra quanti fumano, circa la metà morirà per ragioni direttamente collegate con questa abitudine e di questi la metà morirà prima di aver raggiunto l'età della pensione (tra 35 e 65 anni). A 70 anni l'80% dei non fumatori è vivo contro il 50% dei fumatori (Capuano, 2009).

L'abitudine al fumo da tabacco riguarda dunque più di un quinto della popolazione italiana, pertanto continua a rappresentare un argomento di interesse lo studio dei fattori di rischio che accomunano i fumatori, con particolare attenzione al confronto tra generazioni.

2. Obiettivi e metodo

L'indagine qui presentata si è posta l'obiettivo di verificare se il rapporto con il fumo e con i fattori che determinano la scelta di fumare o di non farlo più sono diversamente connotate a livello generazionale e di genere. In particolare si è indagato sui seguenti aspetti di tipo motivazionale e comportamentale:

- motivi per cui le persone fumano
- abitudine al fumo
- motivi per cui le persone riescono o non riescono a smettere di fumare
- abitudini al fumo della famiglia di origine e del gruppo amicale.

¹ Osservatorio fumo alcool e droga dell'Istituto Superiore di Sanità.

A tal fine si è voluto selezionare un campione ragionato² che delimitasse due generazioni: i giovani fino a 30 anni di età e gli adulti tra 45 e 70 anni. Per garantire la presenza di tutte le tipologie osservabili sono state definite le seguenti quote su dieci intervistati: almeno 4 giovani, almeno 1 ex fumatore, almeno 3 fumatori. Il campione è risultato così composto (cfr. Tabella 1):

Tabella 1 – *Composizione del campione per età e genere.*

| | Maschio | | Femmina | | Totale | |
|---------|---------|-------|---------|-------|--------|-------|
| | n | % | n | % | n | % |
| giovani | 201 | 64,3 | 259 | 71,5 | 461 | 68,2 |
| adulti | 112 | 35,7 | 103 | 28,5 | 215 | 31,8 |
| totale | 314 | 100,0 | 362 | 100,0 | 676 | 100,0 |

La somministrazione ha avuto luogo presso locali pubblici di comuni di grandi e piccole dimensioni della provincia di Udine nel periodo gennaio-marzo 2008.

3. Risultati principali

Il campione è risultato composto da 37,0% fumatori, 50,3% non fumatori e 12,7% ex fumatori, con una differenza significativa³ tra giovani e adulti⁴: i fumatori sono rappresentati in misura maggiore tra gli adulti (41,4% contro il 34,9% dei giovani). Per quanto riguarda il genere, si registra una differenza significativa⁵ tra i maschi del campione (fuma il 42,4%) e le femmine (fuma il 32,3%).

3.1. *Motivi per cui le persone fumano e loro abitudini*

Nell'immaginario collettivo le ragioni per cui si inizia a fumare sono principalmente legate ad un bisogno di emulazione dei coetanei, per non sentirsi esclusi dal gruppo dei pari o dalla necessità di sentirsi grandi mentre secondo i diretti interessati la molla che fa scattare il desiderio di cominciare a fumare è soprattutto la curiosità (55,1%); anche se comunque sono riconosciuti come importanti lo spirito di emulazione del gruppo (53,6%) o il bisogno di manifestare con la sigaretta in mano la conquista della condizione di adulti (45,8%).

L'abitudine al fumo può iniziare già in età infantile (uno ha iniziato a 6, uno a 8 e tre a 10 anni) ma il momento cruciale coincide con la frequenza della scuola

² Il campione ragionato non consente di trarre conclusioni di natura inferenziale, tuttavia fornisce interessanti spunti di riflessione per orientare azioni di informazione e prevenzione presso le fasce d'età maggiormente esposte al rischio di acquisire l'abitudine di fumare.

³ Le differenze segnalate nel testo sono significative dal punto di vista statistico.

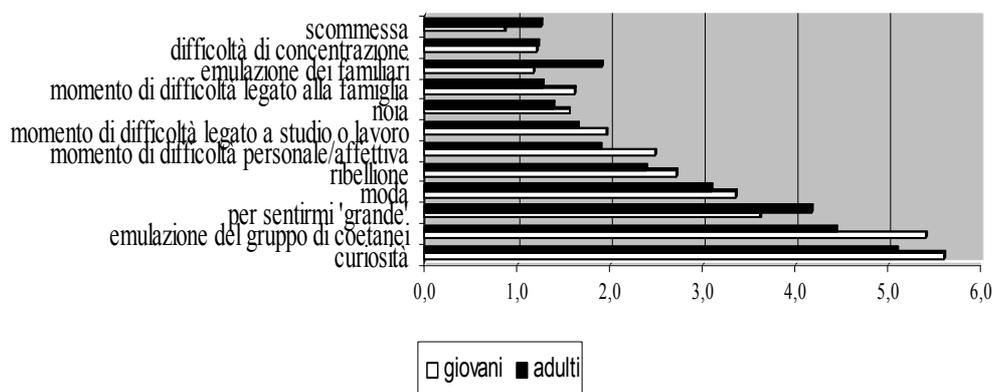
⁴ Test del chi quadro, con 2 gradi di libertà = 33,868, p<0,001.

⁵ Test del chi quadro, con 2 gradi di libertà = 16,361, p<0,001.

media superiore, infatti l'età media di inizio è stata di 16 anni e 2 mesi per i giovani e 17 anni e 7 mesi per gli adulti. Se si analizza il comportamento per genere si nota una sostanziale coincidenza media che si attesta attorno ai 16 anni e 9 mesi.

Le ragioni che hanno pesato nel determinare la scelta di iniziare a fumare sono state rilevate su 12 scale a 11 posizioni (cfr. Figura 1). Dalle dichiarazioni emerge che tra i giovani è stata più forte, rispetto agli adulti, la curiosità (media 5,6 su 5,1) e il bisogno di emulazione (media 5,4 su 4,4) mentre, sempre in termini relativi, ha pesato di più il bisogno di sentirsi grandi nella generazione degli adulti (4,2 su 3,6). Scarso è il peso che hanno avuto le scommesse per decidere di iniziare a fumare (media 1,0) o la necessità di migliorare la capacità di concentrazione (media 1,2) e

Figura 1 – *Ragioni determinanti nella scelta di iniziare a fumare per giovani e adulti (medie).*



non si apprezzano differenze significative di carattere generazionale. Si nota solo che, in termini relativi, l'emulazione dei familiari è stata più efficace tra gli adulti. L'analisi di genere non fa emergere differenze significative, anche se le femmine più dei maschi sono state sensibili alla curiosità, moda, ribellione e a momenti di difficoltà di varia origine.

Per valutare il peso che hanno avuto un insieme di fattori latenti sulla decisione di iniziare a fumare, sono state sottoposte ad analisi fattoriale⁶ le 12 scale di valutazione, unendo i fumatori con gli ex fumatori e differenziando quindi l'analisi per generazione e per genere.

⁶ È stata applicata l'analisi fattoriale sotto l'ipotesi che le modalità delle variabili qualitative ordinali siano rappresentabili su scala di rapporti. Sono stati estratti i fattori con auto valore maggiore di 1 ed è stata effettuata la rotazione ortogonale degli assi con il metodo Varimax. Sono state successivamente eliminate le scale con saturazione multifattoriale (Fabbris, 1997).

Dall'applicazione dell'analisi fattoriale a tutto il gruppo di fumatori ed ex fumatori (cfr. Tabella 2) sono stati estratti quattro fattori, i quali spiegano complessivamente il 59,6% della varianza osservata. Il primo fattore è riconducibile al vissuto psicologico della persona che decide di iniziare a fumare e può essere definito come *fattore di difficoltà*. Il secondo fattore, quello del *bisogno di emancipazione*, riunisce in sé delle motivazioni che il ragazzo associa al fumo in maniera simbolica e che, nella cultura giovanile, segnano il "grande passo" verso l'età adulta, riassumibili nelle espressioni "per sentirsi grandi" e per "emulazione del gruppo di coetanei". Il terzo fattore collega l'inizio dell'abitudine al fumo all'*emulazione* del comportamento dei familiari ma anche ad un inizio per scommessa. Il quarto fattore è correlato con *curiosità e noia*.

Tabella 2 – Analisi fattoriale riferita alle ragioni per cui si decide di iniziare a fumare.

| | Fattori | | | |
|-------------------------------------|--------------|--------------------------|--------------|------------------|
| | Difficoltà | Bisogno di emancipazione | Emulazione | Curiosità e noia |
| difficoltà personale/affettiva | 0,842 | -0,047 | -0,001 | 0,063 |
| difficoltà legate alla famiglia | 0,797 | 0,218 | 0,125 | -0,025 |
| difficoltà legate a studio o lavoro | 0,684 | -0,141 | 0,311 | -0,034 |
| per sentirmi 'grande' | 0,018 | 0,800 | 0,124 | -0,088 |
| emulazione del gruppo di coetanei | -0,154 | 0,662 | 0,025 | 0,154 |
| emulazione dei familiari | 0,145 | 0,111 | 0,727 | -0,033 |
| scommessa | 0,185 | 0,159 | 0,692 | 0,147 |
| curiosità | -0,045 | 0,215 | -0,048 | 0,792 |
| noia | 0,143 | -0,144 | 0,396 | 0,612 |

Se si considerano solo i giovani emergono quattro fattori che spiegano il 60,5% della varianza osservata e si presentano con una struttura uguale a quella del gruppo generale, con la sola aggiunta della ribellione tra le scale correlate con il fattore bisogno di emancipazione. Per quanto riguarda gli adulti i tre fattori estratti spiegano il 53,1% della varianza totale. Anche in questo caso il primo è associato alle *difficoltà* mentre con il secondo sono correlati elementi di *tedio ed emulazione* dei familiari. Il terzo fattore è quello che riconduce le motivazioni addotte ad un unico bisogno, quello di *emancipazione e di accettazione* dal gruppo dei coetanei, con una certa sensibilità anche alle mode del momento.

Se si ripete l'analisi fattoriale per genere, si conferma come primo fattore quello delle *difficoltà*, sia per i maschi che per le femmine. Per i maschi⁷ il secondo fattore è associato alla moda e al sentirsi grande, quindi un fattore che sintetizza il bisogno di compiere talune azioni per conquistare l'*emancipazione* in un certo gruppo sociale, mentre il terzo fattore è quello della *curiosità* associata alla ribellione alle

⁷ Sono stati estratti tre fattori che spiegano il 51,7% della varianza osservata.

regole e alla noia. Il secondo fattore per le femmine⁸ è quello già individuato per la popolazione adulta come *tedio ed emulazione* mentre il terzo è associato al bisogno di *emancipazione*.

Per quanto riguarda il fumo da tabacco, il 70,3% dei fumatori consuma sigarette leggere e il 30,9% fuma anche sigarette forti. Pochi sono i fumatori di pipa (4,8%) o di sigari (3,2%). A livello generazionale la distribuzione dei consumi non fa emergere comportamenti statisticamente diversi se non sulla voce “altro” che viene segnalata solo dal 12,5% dei giovani. A livello di genere sono le donne le più grandi consumatrici di sigarette leggere (83,6%) mentre spetta agli uomini il primato nel consumo di sigarette forti (41,4% contro il 19,0% delle donne). Il consumo mediano di sigarette o altri generi è di 10 al giorno e quello medio di 14, con un 33,5% che arriva ad un consumo massimo a 5 ed un 21,6% di fumatori forti che supera le 20 sigarette al giorno. A livello generazionale si apprezza una differenza nel consumo di fumo, laddove il 75,6% dei giovani fuma al massimo 10 sigarette al giorno mentre tra gli adulti il 54% fuma 20 e più sigarette (o altro).

3.2. *Motivi per cui le persone riescono o non riescono a smettere di fumare*

Il 33,0% di quanti hanno contratto l'abitudine al fumo non ha mai provato a smettere di fumare e non c'è differenza tra generazioni⁹. Subentra invece una differenza di genere nella distribuzione dei tentativi di smettere di fumare¹⁰ in quanto il 37,7% delle donne non ci ha mai provato contro il 29,1% degli uomini; il 51,9% delle donne ha fatto 1 o 2 tentativi contro il 46,2% degli uomini; il 10,4% delle donne ha fatto da 3 tentativi in su contro il 24,7% degli uomini.

Ma quali sono stati gli esiti di tali tentativi? Nel complesso il 25,6% delle persone che fumavano hanno smesso di farlo e il tasso di successo dei tentativi di smettere si attesta attorno al 38%. Con riferimento al genere le quote di successo non portano differenze significative: il 26,9% dei maschi e il 24,0% delle femmine che hanno fatto almeno un tentativo hanno smesso di fumare.

All'analisi delle ragioni per cui il 67,0% dei fumatori ha provato a smettere di fumare ha risposto appena il 38,2%, mettendo comunque in evidenza una sostanziale differenza generazionale nell'approccio al problema. Il 71,8% dei giovani rispondenti lo ha fatto perché convinti che il fumo faccia male, contro il 36,2% degli adulti. Il secondo motivo è il venir meno della necessità di fumare da parte del 51,3% dei giovani contro il 25,5% degli adulti. In terza posizione si registra una scelta di rispetto per chi sta vicino, comunque più sentita dai giovani (28,2%), mentre gli adulti in misura maggiore sono stati indotti da motivi di salute

⁸ Sono stati estratti tre fattori che spiegano il 53,7 della varianza osservata.

⁹ Test del chi quadro, con 4 gradi di libertà = 3,890, $p < 0,42$.

¹⁰ Test del chi quadro, con 4 gradi di libertà = 13,08, $p < 0,011$.

(23,4%). I giovani sono più sensibili a questioni di costi (23,1% contro il 12,8%). Pochi tentano di smettere per scommessa o perché costretti a farlo (3,5% sul totale). Le femmine più dei maschi sono convinte che faccia male (56,8% contro 49,0%) e hanno tentato di smettere per rispetto nei confronti di chi sta loro vicino (32,4% contro il 18,4% dei maschi). Anche la scomparsa del bisogno di fumare è più diffusa tra le donne (48,6%) che tra gli uomini (28,6%).

Ma perché tanti, pur provandoci, non riescono a smettere? Cosa accade se non fumano, quali reazioni si innescano, quali comportamenti si manifestano?

Tabella 3 – *Come reagiscono giovani e adulti se non possono fumare.*

| | Giovani | | Adulti | | Totale | |
|-----------------------------|---------|------|--------|------|--------|------|
| | n | % | n | % | n | % |
| divento nervoso | 61 | 38,9 | 36 | 40,0 | 97 | 39,3 |
| non saprei dire | 51 | 32,5 | 27 | 30,0 | 78 | 31,6 |
| mangio | 51 | 32,5 | 23 | 25,6 | 74 | 30,0 |
| nulla | 26 | 16,6 | 10 | 11,1 | 36 | 14,6 |
| mi annoio | 13 | 8,3 | 8 | 8,9 | 21 | 8,5 |
| perdo sicurezza/divento più | 9 | 5,7 | 6 | 6,7 | 15 | 6,1 |
| mancanza di concentrazione | 7 | 4,5 | 6 | 6,7 | 13 | 5,3 |

Si può affermare che il fumo unisce le generazioni nelle reazioni alla sua mancanza (cfr. Tabella 3). Prevale il nervosismo, riconosciuto dal 39,3% dei fumatori mentre un 30,0% trova compensazione e consolazione nel cibo. Molto elevata è la quota di chi non sa rispondere (31,6%). A livello di genere non si manifestano reazioni sensibilmente diverse, fatta eccezione per il nervosismo che sembra accomunare un maggior numero di maschi (42,1%) rispetto alle femmine (36,0%) e il senso di insicurezza, che è relativamente più presente tra le femmine (10,5%) che tra i maschi (2,3%).

3.3. *Abitudini al fumo della famiglia di origine e del gruppo amicale.*

La presenza di fumatori in famiglia sembra favorire la “carriera” dei fumatori: dove non si fuma il 63% segue l’esempio di genitori e fratelli mentre se fuma almeno un parente prossimo la percentuale di non fumatori si riduce sensibilmente attorno al 40%. Infatti¹¹ emerge una differenza significativa nella distribuzione dei fumatori, non fumatori ed ex fumatori sulla base della presenza di persone che fumano nella sfera familiare e amicale, inducendo a concludere che l’esempio visto in casa rende normale l’abitudine al fumo anche da parte dei figli. A livello generazionale i giovani si rivelano più autonomi nella scelta di non fumare rispetto agli adulti che in misura maggiore hanno intrapreso il consumo del fumo, soprattutto in presenza di familiari che avevano questa abitudine.

¹¹ È stato applicato il Test del chi quadro, con 10 gradi di libertà = 45,6, p<0,001.

4. Conclusioni

Dall'analisi comparata tra giovani e adulti si può quindi concludere che la differenza più rilevante tra le due generazioni sia da ricercare proprio nella diffusione dell'abitudine al fumo, più elevata nella generazione adulta (63,9%) rispetto a quella dei giovani (43,3%) e nell'età media di inizio (i giovani hanno iniziato a fumare in media a 16 anni e 2 mesi contro i 17 e 7 mesi degli adulti), mentre i comportamenti di quanti hanno acquisito questa abitudine sono sostanzialmente assimilabili, seppure con taluni distinguo. Le ragioni latenti per cui si è iniziato a fumare sono legate anzitutto a momenti di difficoltà di varia origine; per i giovani pesano inoltre fattori di bisogno di appartenenza al gruppo dei pari e a quello di riferimento, seguiti dall'emulazione degli adulti e infine dalla noia. Anche per gli adulti accanto ai momenti di difficoltà troviamo il tedio e l'emulazione e il bisogno di emancipazione. Il 33,0% di quanti hanno contratto l'abitudine al fumo non ha mai provato a smettere di fumare e non c'è differenza tra generazioni. Non c'è differenza nemmeno nelle reazioni che giovani e adulti hanno quando si trovano a non poter fumare: è diffusa tra tutti una reazione di nervosismo e una ricerca di compensazione nel cibo. Tutti, seppure in misura diversa, sono stati influenzati dai comportamenti dei familiari e la decisione di non fumare in un ambiente di fumatori è sensibilmente meno diffusa rispetto ai casi in cui nelle famiglie non si fuma.

Sia tra i giovani che tra gli adulti è più elevata la diffusione del fumo tra i maschi (poco più del 40% in entrambe le generazioni), mentre si arriva a superare il 62% delle giovani non fumatrici a fronte di un 46% di donne adulte che non fumano. Donne e uomini hanno iniziato a fumare in media attorno ai 16 anni e 9 mesi. Dall'applicazione dell'analisi fattoriale alle ragioni di tale scelta, il primo fattore comune a tutti è quello dalle difficoltà, seguito, per i maschi dal bisogno di emancipazione e poi dalla curiosità mentre il secondo fattore per le femmine è quello indicato con tedio ed emulazione e al terzo il bisogno di emancipazione. Se sono privati del fumo, maschi e femmine reagiscono allo stesso modo. Le differenze di genere qui emerse sono riconducibili per lo più alla decisione di provare a smettere di fumare, connotata più al maschile, e alle reazioni all'assenza di fumo che rendono i maschi più nervosi e le donne, seppure in misura molto contenuta, meno sicure di sé.

Da quanto emerso si è indotti a concludere che i comportamenti e le reazioni dei fumatori non sono correlabili con fattori di tipo generazionale o di genere ma dipendono solo dal loro rapporto con il fumo.

Da questa esplorazione si possono ricavare indicazioni utili per orientare iniziative di informazione e sensibilizzazione presso i ragazzi fin dalle scuole secondarie di primo grado. Non si può intervenire sulle cause che hanno portato i

ragazzi ad iniziare a fumare ma si può insegnare a loro come trovare modalità alternative per affrontare le difficoltà (correlate con il primo fattore) e per dimostrare di essere diventati grandi (secondo fattore).

Riferimenti bibliografici

- CAPUANO E. 2009. *Il fumo ucciderà un fumatore su due*,
<http://www.ecplanet.com/node/2116>.
- FABBRIS L. 1997. *Statistica multivariata*, Milano McGraw-Hill.
- ISTAT 2006. *I fumatori in Italia. Statistiche in breve*.
- LEGA ITALIANA PER LA LOTTA CONTRO I TUMORI. *Perché non ti conviene fumare. Istruzioni per l'uso*. (consultato il 15.02.2011).
<http://space.cinet.it/cinetclub/emmegi/magazine/fumononconviene/fumonon.ht>
- MARTINELLA V. 2009. Fumo: 80 mila morti ogni anno in Italia.
http://corriere.it/salute/sportello_cancro.
- OSSFAD 2010. *Rapporto annuale sul fumo 2010*. <http://www.iss.it/fumo>.

SUMMARY

The aim of this research was to verify if there are different smoking behaviours between young and adult people and between males and females. It was done a research on a champion of 676 people. Data regarding the reasons of smoking beginning were submitted to factorial analysis. The main results regarding the generations are: there are more adult than young smokers but young people begins to smoke earlier (at about 16 years). The first factor that justified the beginning is related with personal difficulties. For young people the second factor is related with the need of membership to peer group and the third is related with adult emulation and boredom. The 33,0% never tried to stop. There are no generational differences between young and adult when cannot smoke: both become nervous and eat. The greatest difference is related with diffusion of smoking custom. The other behaviours are related with the status of smoker, regardless the age of people.

Regarding the gender, the differences are referable to the attempt of given up smoking (males has tried more time than females) and to reactions when cannot smoke, when males become more nervous than females and females became more insecure.

L'ORIENTAMENTO DEL MONDO CATTOLICO SU DEMOGRAFIA E FAMIGLIA. UN INDICE PER CONFRONTARE REALTÀ E PROBLEMATICHE NEL PANORAMA EUROPEO

Gian Carlo Blangiardo, Simona Maria Mirabelli

Premessa

Il mondo cattolico si è recentemente confrontato, ai massimi livelli delle gerarchie ecclesiali, per dibattere sull'evoluzione della popolazione e delle famiglie in Europa, e sugli aspetti di natura demografica che ad essa si ricollegano. In occasione dell'incontro svoltosi a Zagabria a fine 2010, nell'ambito del Consiglio delle Conferenze Episcopali Europee (CCEE), è stato reso disponibile un ampio materiale informativo contenente dati e valutazioni sulla realtà demografica di ognuno dei 33 paesi oggetto di analisi e di confronto. Tale materiale ha reso possibile la costruzione di una serie di indicatori in grado di esprimere, anche nel dettaglio delle singole dimensioni oggetto di interesse, una misura sintetica del livello di problematicità percepita dalle comunità ecclesiali europee. In particolare, il presente lavoro vuole mettere a punto – in via sperimentale e per un opportuno insieme di paesi – una procedura per esprimere un confronto oggettivo basato sulla percezione di maggiore o minore gravità che viene riconosciuta alle tendenze demografiche in atto, da parte di testimoni privilegiati che sono autorevole espressione del mondo e della cultura cattolica nel contesto europeo. I risultati così ottenuti formano poi oggetto di valutazione alla luce di un insieme di appropriati indicatori, identificati nell'ambito delle più recenti statistiche di fonte Eurostat.

1. Obiettivi

Il lavoro si propone di cogliere il grado di conoscenza e di consapevolezza della Chiesa Cattolica sui più importanti temi demografici, unitamente alla sua valutazione circa la problematicità delle dinamiche in atto e delle conseguenze, di ordine sociale, economico e familiare, che ne derivano. Il presente contributo intende, altresì, verificare il livello di correlazione tra l'indice di problematicità

percepita e la realtà demografica accertata sulla base di alcuni indicatori statistici di fonte ufficiale che consentono di monitorare le dinamiche di interesse¹.

2. Aspetti metodologici

A partire dalla rilevazione svolta in occasione di un recente incontro tenutosi a Zagabria si sono acquisiti i risultati di un questionario, rivolto alle Conferenze Episcopali Europee dei 33 paesi partecipanti, strutturato in 25 domande a risposta aperta e raggruppate attorno ai seguenti temi: Evoluzione demografica, Struttura per età, Natalità e Sostegno alla famiglia. L'insieme delle risposte è stato ricondotto a due dimensioni di analisi: la prima riguardante la conoscenza e la consapevolezza della situazione demografica del paese in esame (cfr. Prospetto 1); la seconda riferita alla problematicità percepita rispetto alle tematiche sottoposte (cfr. Prospetto 2). In corrispondenza di ognuna di esse si è attivata una procedura di valutazione articolata in tre fasi:

- 1) Definizione di un modello teorico ottimale inteso a identificare, in corrispondenza di ogni sotto tema oggetto di analisi, sia il livello di conoscenza e di consapevolezza circa i fenomeni demografici di interesse, sia il grado di preoccupazione riguardo alla realtà demografica delle realtà nazionali considerate.
- 2) Identificazione di un appropriato set di domande attraverso il quale accertare l'atteggiamento nei riguardi delle due dimensioni indagate.
- 3) Attribuzione, in corrispondenza di ciascuna domanda, di un insieme di risposte articolate secondo una scala a 5 modalità che risulta inversamente orientata, sia rispetto al grado di problematicità percepita, sia al livello di conoscenza e consapevolezza circa la realtà demografica osservata nel paese in esame.

Con tali premesse, preso di volta in volta in esame il questionario corrispondente a ogni paese, si è proceduto ad attribuire i relativi punteggi da 1 a 5 agli items afferenti a ciascun sotto tema indagato. A partire da tali valori si è quindi pervenuti alla definizione di un indice sintetico di problematicità percepita (complessivo e parziale) costruito attraverso la media dei punteggi relativi ai temi indicati nel successivo prospetto 2². Gli indici, così costruiti, variano tra il valore 1 (a cui

¹ Lo studio si inserisce nell'ambito di un progetto sperimentale finalizzato all'attivazione di un Osservatorio Demografico delle CCEE.

² Si è operato inizialmente sia attraverso una media aritmetica semplice, sia mediante una media ponderata dell'insieme dei punteggi di ogni paese adottando un sistema di ponderazione che attribuisce maggiore importanza alle variabili mediamente meno correlate con le altre. Per le finalità di cui sopra si è proceduto, nei rari casi di mancata informazione, ad assegnare alla variabile del paese in oggetto il valore medio aritmetico rilevato sull'insieme dei restanti nove paesi.

corrisponde la massima problematicità percepita) e il punteggio 5 (assenza totale di problematicità). In vista di una successiva estensione al complesso dei 33 paesi che hanno fornito i dati in ambito CCEE, la sperimentazione è stata svolta con riferimento ai seguenti dieci membri dell'Unione Europea: Austria, Bulgaria, Francia, Germania, Grecia, Italia, Lituania, Polonia, Slovacchia e Slovenia.

Prospetto 1 – Dimensione relativa alla conoscenza e alla consapevolezza.

| SOTTO TEMA | MODELLO TEORICO OTTIMALE | CRITERI DI VALUTAZIONE/ DOMANDE COGNITIVE | SCALA |
|-----------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Evoluzione demografica e migrazione</i> | La crescita della popolazione è moderatamente positiva; il saldo naturale è positivo; i flussi migratori sono contenuti. | Qual è l'andamento demografico del Paese? | 1. La realtà descritta è molto distante dal modello teorico di riferimento. 2. 3. 4. 5. La realtà descritta è molto vicina dal modello teorico di riferimento. |
| <i>Invecchiamento e struttura demografica. Prospettive</i> | Situazione in miglioramento rispetto alle condizioni demografiche attuali | In futuro quale sarà la situazione demografica in relazione alla crescita demografica, al processo di invecchiamento e al rapporto tra giovani e anziani? | 1. Peggiora rispetto a quella attuale. 2. 3. Immutata rispetto a quella attuale. 4. 5. Migliora rispetto a quella attuale. |
| <i>Cambiamenti demografici. Conseguenze a livello economico</i> | I cambiamenti demografici in atto <u>non</u> producono conseguenze significative a livello economico. | I cambiamenti demografici producono conseguenze significative a livello economico nel Paese? | 1. Sì, molte 2. 3. 4. 5. Nessuna |
| <i>Cambiamenti demografici. Conseguenze a livello sociale</i> | I cambiamenti demografici in atto <u>non</u> producono conseguenze significative a livello sociale. | I cambiamenti demografici producono conseguenze significative a livello sociale nel Paese? | 1. Sì, molte 2. 3. 4. 5. Nessuna |
| <i>Cambiamenti demografici. Conseguenze a livello familiare</i> | I cambiamenti demografici in atto non producono conseguenze significative a livello familiare. | I cambiamenti demografici producono conseguenze significative a livello familiare nel Paese? | 1. Sì, molte 2. 3. 4. 5. Nessuna |
| <i>Fecondità</i> | Il livello di conoscenza e di consapevolezza delle tematiche oggetto di interesse è ritenuto ottimo ed esauriente. | Completezza ed esaustività delle risposte. Presenza di dati/informazioni dettagliate. | 1. Inconsistente/molto scarso 2. Scarso 3. Sufficiente 4. Buono 5. Ottimo/esauriente |

Prospetto 2 – Dimensione relativa al livello di problematicità percepita.

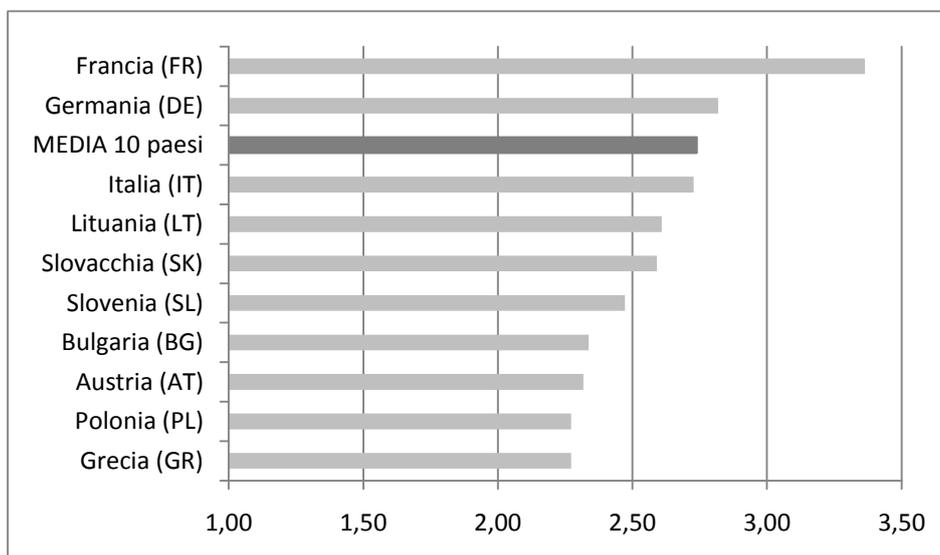
| SOTTO TEMA | MODELLO TEORICO OTTIMALE | CRITERI DI VALUTAZIONE/ DOMANDE COGNITIVE | SCALA |
|-----------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|
| <i>Livello di preoccupazione circa l'evoluzione demografica</i> | Non si manifesta preoccupazione. | Esistono patologie? Rappresentano una fonte di preoccupazione? | 1. Forte preoccupazione 2. 3. Lieve o moderata preoccupazione 4. 5. Nessuna preoccupazione |
| <i>Coscienza sociale</i> | Vi è un ampio e diffuso livello di coscienza sociale. | Esiste una coscienza sociale circa la situazione demografica del Paese? | 1. Livello inconsistente 2. Scarso 3. Sufficiente 4. Buono 5. Ottimo/esauriente |
| <i>Adozione</i> | Esistenza nel Paese di un programma di adozione. Gestione efficiente del sistema. | Esiste un programma di adozione? Qual è il giudizio sul suo funzionamento? | 1. Valutazione negativa. 2. 3. 4. 5. Valutazione positiva |
| <i>Il valore della genitorialità</i> | Il ruolo svolto da entrambi i genitori è riconosciuto come un valore. | E' considerato un valore il ruolo materno? E quello paterno? | 1. Valutazione negativa 2. 3. 4. 5. Valutazione positiva |
| <i>La specificità dell'intervento pubblico a favore delle famiglie numerose e della maternità plurima</i> | Sono previsti interventi a sostegno della famiglia e della maternità plurima. Le misure adottate sono indirizzate alla famiglia in quanto tale. | Esistono misure a sostegno della famiglia? Gli interventi sono specifici/selettivi? | 1. Valutazione negativa 2. 3. 4. 5. Valutazione positiva |
| <i>Il ruolo della Chiesa</i> | Azioni concrete. Ampiezza degli interventi implementati. | Qualità degli interventi realizzati dalla Chiesa a favore delle famiglie. | 1. Valutazione negativa 2. 3. 4. 5. Valutazione positiva |
| <i>Il supporto intrafamiliare: il ruolo dei nonni nella cura dei bambini (family running)</i> | Il ricorso al <i>family running</i> non rappresenta un indicatore di mancanza/insufficienza di strutture per l'infanzia o di altre anomalie. | L'eventuale ruolo dei nonni nel <i>family running</i> è una scelta di libertà e di valorizzazione della persona? | 1. No 2. 3. 4. 5. Sì, senz'altro |
| <i>Aspetti legislativi</i> | Presenza nel Paese di una normativa a sostegno delle famiglie. Incisività dell'azione. | Esiste una normativa <i>family-friendly</i> efficace? | 1. Per nulla 2. 3. 4. 5. Ampia e soddisfacente |
| <i>Giudizi di valore e livello di percezione rispetto al tema della maternità</i> | La maternità è valutata positivamente. Vi è attenzione e riguardo nei confronti delle madri <i>caregiver</i> (<i>maternal employment</i>). | La maternità e il ruolo di <i>caregiver</i> svolto dalle madri che valutazione riceve? | 1. Valutazione negativa 2. 3. 4. 5. Valutazione positiva |

| | | | |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------|
| <i>Presenza di organismi per la famiglia</i> | L'istituzione di un Ministero per la famiglia (o di altro organismo <i>ad hoc</i>) rappresenta un indicatore di attenzione nei confronti della famiglia. | Esistono istituzioni pubbliche delegate in modo specifico o prevalente alle problematiche della famiglia? | 1. In nessun caso 2. 3. 4. 5. In misura del tutto adeguata |
| <i>Legislazione finanziaria e sostegno economico. Specificità degli interventi a favore della famiglia in quanto tale e non come soggetto bisognoso di assistenza</i> | La presenza di una legislazione finanziaria che prevede misure a favore delle famiglie è indicatore di un reale e concreto sostegno ai nuclei familiari. | Esiste una legislazione finanziaria con forme di intervento specifiche a sostegno delle famiglie in quanto tali? | 1. No 2. 3. 4. 5. Sì |

3. Sintesi dei risultati

L'indice totale di problematicità percepita (cfr. Figura 1) segnala la leadership della Francia – sia nella versione rilevata attraverso la media aritmetica semplice dei punteggi relativi agli 11 temi trattati, sia in quella ove si tiene conto della ponderazione – cui segue la Germania con un punteggio di poco superiore alla media delle dieci realtà nazionali osservate.

Figura 1 – *Indice totale di problematicità percepita. Punteggi risultanti dalla media aritmetica semplice.*



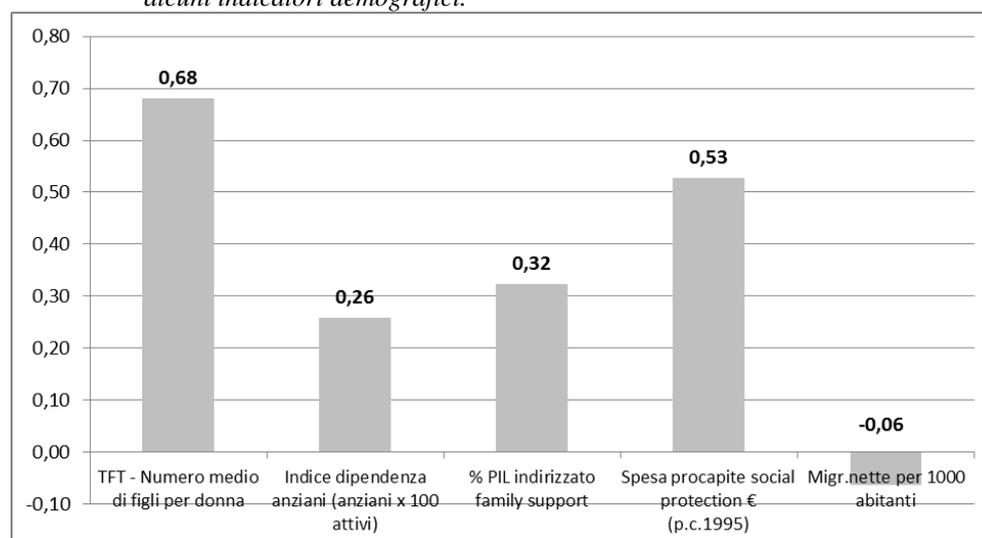
L'Italia detiene il terzo posto, mentre nelle posizioni di coda si collocano quattro paesi dell'Europa centro-orientale (Bulgaria, Austria, Polonia e Grecia).

Tuttavia, se andiamo a considerare la graduatoria relativa agli indici parziali riferiti ai singoli sotto temi oggetto di indagine (cfr. Tabella 1), emerge un quadro sensibilmente differenziato rispetto a quello testé delineato. Come si può osservare, la Francia si pone al vertice della classifica relativamente alle tematiche che riguardano le misure a sostegno dei nuclei familiari numerosi e della maternità plurima e che risultano adeguatamente supportate dalla legislazione finanziaria nazionale. Per quanto attiene il livello di problematicità generato dai cambiamenti demografici in atto, oltre alla Francia, anche la Grecia e la Bulgaria mostrano valori che esprimono un relativamente basso livello di preoccupazione; al contrario, per la Germania il corrispondente indice segnala il massimo grado di problematicità percepita. L'esistenza di una coscienza sociale circa la situazione demografica del paese risulta significativamente elevata in Germania e in Francia; viceversa, la Slovacchia e la Grecia sembrerebbero caratterizzarsi per un livello ritenuto insufficiente. In tema di adozione, quattro dei paesi considerati (Austria, Italia, Lituania e Polonia) si collocano al vertice della relativa graduatoria avendo espresso un giudizio positivo in merito alla gestione e al funzionamento del sistema adottivo in esame. Rispetto al riconoscimento del ruolo genitoriale, l'Italia spicca ancora al primo posto, mentre l'Austria e la Bulgaria, che segnalano una scarsa considerazione per tale valore, si collocano nelle ultime posizioni. La Lituania domina la classifica per l'ampiezza e la qualità degli interventi realizzati dalla Chiesa Cattolica a favore delle famiglie, avendo quest'ultima apportato il proprio contributo alle iniziative di carattere legislativo del Paese. Per quanto riguarda il ruolo dei nonni nelle attività di cura dei bambini, esso trova ampio riconoscimento sia in Austria che in Lituania. In tema di normativa a sostegno delle famiglie, domina invece la Slovacchia che segnala la presenza di una normativa *family-friendly* ritenuta soddisfacente. Infine l'esistenza (almeno sulla carta) di istituzioni pubbliche delegate in modo specifico alle problematiche della famiglia risulta adeguata sia in Italia che in Slovacchia e Slovenia.

Volendo poi verificare la rispondenza tra problematicità percepita e realtà oggettivamente descritta sulla base di alcuni indicatori di carattere socio-demografico, si è proceduto a calcolare il coefficiente di correlazione tra i corrispondenti valori. Dallo studio emerge chiaramente come ad una più elevata fecondità, misurabile attraverso il numero medio di figli per donna (TFT), corrisponda un minore grado di preoccupazione per la situazione demografica osservata; allo stesso modo, le maggiori risorse pubbliche destinate alla previdenza sociale e agli interventi *family support* risultano associati positivamente ad un minore livello di problematicità percepita (cfr. Figura 2).

Tabella 1 – Posizione in graduatoria per indici parziali di problematicità percepita.

| INDICE/PAESE | AT | BG | FR | DE | GR | IT | LT | PL | SK | SL |
|------------------------------------------------------------------------|-----|----|-----|-----|----|----|----|-----|-----|----|
| <i>Livello di preoccupazione circa l'evoluzione demografica</i> | 4° | 1° | 1° | 10° | 1° | 4° | 8° | 8° | 4° | 4° |
| <i>Coscienza sociale. Il livello</i> | 6° | 5° | 2° | 1° | 9° | 6° | 6° | 3° | 10° | 4° |
| <i>Adozione</i> | 1° | 6° | 6° | 9° | 6° | 1° | 1° | 1° | 5° | 9° |
| <i>Il valore della genitorialità</i> | 10° | 9° | 2° | 4° | 4° | 1° | 2° | 7° | 7° | 6° |
| <i>La specificità dell'intervento pubblico a favore delle famiglie</i> | 2° | 4° | 1° | 2° | 8° | 7° | 8° | 8° | 4° | 4° |
| <i>Il ruolo della Chiesa</i> | 9° | 5° | 8° | 3° | 2° | 9° | 1° | 5° | 3° | 5° |
| <i>Il supporto intrafamiliare. Il ruolo dei nonni</i> | 1° | 7° | 3° | 7° | 4° | 7° | 1° | 4° | 4° | 7° |
| <i>Aspetti legislativi</i> | 5° | 9° | 10° | 3° | 5° | 2° | 8° | 5° | 1° | 3° |
| <i>Giudizi di valore rispetto al tema della maternità</i> | 6° | 5° | 1° | 2° | 6° | 3° | 6° | 3° | 6° | 6° |
| <i>Presenza di organismi per la famiglia</i> | 5° | 7° | 4° | 5° | 9° | 1° | 7° | 9° | 1° | 1° |
| <i>Legislazione finanziaria e sostegno economico</i> | 6° | 3° | 1° | 2° | 8° | 4° | 5° | 10° | 8° | 7° |

Figura 2 – Grado di correlazione tra il livello di problematicità percepita e valore di alcuni indicatori demografici.

Riferimenti bibliografici

CONSIGLIO DELLE CONFERENZE EPISCOPALI EUROPEE (CCEE), Atti dell'Assemblea Plenaria dei Presidenti delle Conferenze Episcopali sul tema *Demografia e Famiglia*, Zagabria, 20 settembre-3 ottobre 2010.

EUROPEAN COMMISSION, Expert Group on Demographic Issues, Country profile, 2011.

SUMMARY

The Catholic world has recently been compared, at the highest levels of church hierarchy, to discuss the evolution of the population and households in Europe, and demographic aspects that are connected to it. At the meeting held in Zagreb at the end of September 2010 in the Council of European Episcopal Conferences (CCEE), was made available which contains a comprehensive information and assessments on the demographic reality of each of the 33 countries analyzed. This material has made possible the construction of a series of indicators that can express a summary measure of the level of concern felt by church communities in Europe. In particular, the work aims to develop, for a suitable set of countries, an objective comparison based on the perception of greater or lesser seriousness that is accorded to demographic trends, by the privileged witnesses who represent an authoritative expression of the world and Catholic culture in the European context. The results were then evaluated in light of a set of appropriate indicators, properly identified in the most recent statistics from Eurostat.

Gian Carlo BLANGIARDO, Professore Ordinario, Università degli studi di Milano-Bicocca, giancarlo.blangiardo@unimib.it

Simona Maria MIRABELLI, Assegnista di ricerca, Università degli studi di Milano-Bicocca, simona.mirabelli@unimib.it

MATRIMONI CON ALMENO UNO SPOSO STRANIERO: EVOLUZIONE DEL FENOMENO E NUOVE TENDENZE. IL RUOLO DEL CONTESTO NORMATIVO

Cinzia Castagnaro, Antonella Guarneri, Claudia Iaccarino

1. L'andamento dei matrimoni

Dal 1972, salvo sporadici picchi dovuti a eventi congiunturali (come ad esempio l'effetto attrattivo costituito dall'anno del Giubileo), i matrimoni celebrati in Italia continuano il loro *trend* decrescente. Tra il 2008 e il 2009 la diminuzione è stata però particolarmente evidente (-6,5%), conseguenza presumibilmente dell'effetto della congiuntura economica sfavorevole.

Un altro aspetto che sembrava ormai consolidato, a dispetto della tendenziale diminuzione delle nozze, era la continua crescita dei matrimoni con almeno uno sposo straniero (a parte una lieve flessione registrata nel 2007); nel 2009, invece, anche questa tipologia subisce una decisa contrazione che merita particolare attenzione anche alla luce delle recenti modifiche delle normative riguardanti i cittadini stranieri.

Tabella 1 — *Principali caratteristiche dei matrimoni celebrati in Italia. Anno 2009.*

| | Valori assoluti | % su totale matrimoni | % su totale matrimoni con almeno uno sposo straniero | Variazione percentuale rispetto all'anno precedente |
|------------------------------------------|-----------------|-----------------------|------------------------------------------------------|-----------------------------------------------------|
| Totale matrimoni | 230.613 | 100,0 | | -6,5 |
| Sposi entrambi italiani | 198.554 | 86,1 | | -5,3 |
| Matrimoni con almeno uno sposo straniero | | | 100,0 | -13,2 |
| Sposi entrambi stranieri | 10.702 | 4,6 | 33,4 | -13,5 |
| Matrimoni misti | 21.357 | 9,3 | 66,6 | -13,0 |
| <i>Sposo italiano, sposa straniera</i> | 16.559 | 7,2 | 51,6 | -9,2 |
| <i>Sposo straniero, sposa italiana</i> | 4.798 | 2,1 | 15,0 | -23,9 |

Lo studio focalizza dunque l'attenzione sui matrimoni con almeno uno sposo straniero, proponendosi di analizzare come sia cambiata negli ultimi anni la composizione per cittadinanza degli sposi e se si siano verificate trasformazioni

socio-demografiche evidenti. A questo scopo risulta di particolare interesse studiare le caratteristiche degli sposi per tipologia di coppia (coppie miste e coppie di sposi entrambi stranieri).

Tra i vari aspetti che influenzano il fenomeno troviamo:

- il contesto sociale e politico del paese di accoglienza che riveste un ruolo di primo piano nell'influenza - diretta o indiretta - dei comportamenti demografici;
- le peculiarità dei progetti migratori che caratterizzano le varie cittadinanze.

La tipologia di coppia per la quale si registra il numero più elevato di celebrazioni è quella in cui lo sposo è italiano e la sposa è straniera; tuttavia, come vedremo, se scendiamo nel dettaglio delle cittadinanze, emergono realtà tra loro molto differenziate.

2. Le fonti e la normativa

L'analisi utilizza i dati provenienti dalla Rilevazione dei matrimoni; si tratta di una rilevazione esaustiva, attiva in Istat dal 1926, che raccoglie le informazioni sui matrimoni celebrati in Italia e sulle caratteristiche socio-demografiche degli sposi, consentendo di calcolare le principali misure specifiche sulla nuzialità e le sue determinanti. La disponibilità dell'informazione sulla cittadinanza degli sposi risulta essenziale per interpretare la riduzione dei matrimoni che si è verificata negli ultimi anni.

In un contesto nel quale la scelta di sposarsi è sempre meno diffusa e, a differenza del passato, non costituisce più l'indiscussa tappa iniziale del complesso processo di formazione della famiglia¹, il matrimonio continua a giocare un ruolo attrattivo nei casi in cui assume anche una valenza "pratica" (ad esempio pensando ai tempi abbreviati di acquisizione della cittadinanza italiana quando questa viene richiesta per matrimonio e non per residenza). Altra cosa sono i matrimoni prettamente "di comodo" che sicuramente esistono ma sono difficilmente quantificabili. Fatte, quindi, le dovute premesse, le oscillazioni registrate nei dati, osservate in concomitanza con modificazioni intervenute in ambito legislativo, non possono e non devono essere totalmente attribuite al fenomeno dei matrimoni di convenienza.

La prima consistente contrazione delle nozze con almeno uno sposo straniero, avvenuta tra il 2006 e il 2007, ha riguardato soprattutto i matrimoni con almeno uno sposo rumeno (passati da 5.513 a 3.327) e può essere messa in connessione

¹ Emblematica è la costante crescita di nascite avvenute al di fuori del vincolo matrimoniale: dall'8,1% del 1995 al 19,6% del 2008. In particolare, se consideriamo i nati da coppie miste, oltre un bambino su tre nasce fuori dal matrimonio (il 36%).

con l'ingresso della Romania nell'Ue. Infatti, a partire dal 1° gennaio 2007 Romania e Bulgaria sono diventati Paesi membri dell'Unione Europea e, così come previsto dal Trattato istitutivo della Comunità Europea, i loro cittadini, godono del diritto di circolare e soggiornare liberamente nel territorio degli Stati membri².

Successivamente, tra 2008 e 2009, si avverte la seconda contrazione nei matrimoni con almeno uno straniero. Anche questa volta il contesto politico-normativo sembra giocare un ruolo di primo piano. Infatti, con l'art. 1 comma 15 della legge 15 luglio 2009, n. 94, viene imposto allo straniero che vuole contrarre matrimonio in Italia l'obbligo di esibire, oltre al tradizionale nulla osta (o certificato di capacità matrimoniale), anche "un documento attestante la regolarità del soggiorno nel territorio italiano". La verifica della regolarità del soggiorno³ deve essere effettuata sia al momento delle pubblicazioni sia al momento della celebrazione (cfr. circ. DCSD n. 19/2009).

Scopo della riforma è quello di impedire la celebrazione di matrimoni di comodo; tuttavia la normativa, riguardando l'individuo e non la tipologia di coppia, coinvolge anche i matrimoni con entrambi gli sposi stranieri. L'impossibilità di poter attestare la regolarità del soggiorno prevista dalla Legge 94 influenza presumibilmente le decisioni dei nubendi, inducendoli a rinunciare alla celebrazione del matrimonio in Italia.

3. Principali risultati

Focalizzando l'attenzione sul calo dei matrimoni con almeno uno straniero, come si evince dalla figura 1, la tipologia di coppia che ha subito la più evidente contrazione tra 2008 e 2009 è stata quella dei matrimoni misti in cui lo sposo è straniero (-23,9%), seguita da quella con sposi entrambi stranieri (-13,5%).

L'effetto della normativa sulla contrazione dei matrimoni con almeno uno straniero è, inoltre, particolarmente evidente se andiamo a considerare il mese di celebrazione (Figura 2). Mentre, infatti, fino al mese di luglio (mese dell'entrata in vigore della legge 94/2009) si osserva una sostanziale sovrapposizione tra la curva del 2009 e quella della media del periodo 2006-2008, successivamente si evidenzia una marcata differenza nei livelli.

² Per i soggiorni al di sotto dei tre mesi per i cittadini neocomunitari non è più prevista alcuna formalità per il soggiorno mentre per periodi superiori è necessario procedere all'iscrizione anagrafica presso il Comune di abituale dimora in luogo della precedente "carta di soggiorno per cittadini UE".

³ Nel caso in cui i futuri sposi stranieri non siano in grado di dimostrare la regolarità del proprio soggiorno, l'ufficiale di stato civile dovrà rilasciare un rifiuto scritto ai sensi dell'art. 7 d.P.R. n. 396/2000 con l'indicazione dei motivi del rifiuto; provvedere alla denuncia all'Autorità giudiziaria del reato previsto e punito dall'art. 10 bis del D.Lgs. n. 286/1998 (c.d. reato di clandestinità).

Figura 1 – *Andamento dei matrimoni per tipologia di coppia. Anni 1995-2009 (valori assoluti).*

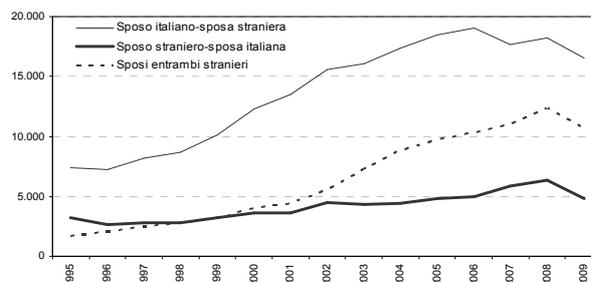
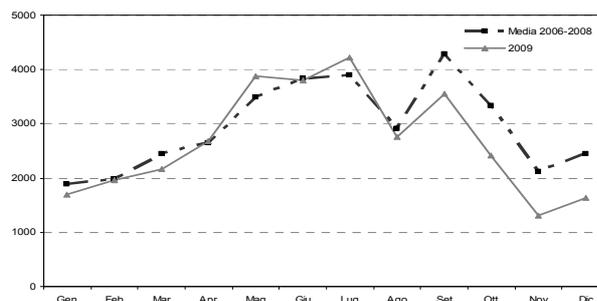


Figura 2 – *Matrimoni con almeno uno sposo straniero per mese di celebrazione. Media 2006-2008 e anno 2009 (valori assoluti).*



La figura 3 mostra la distribuzione dei matrimoni con almeno uno straniero per tipologia di coppia ed evidenzia come i comportamenti nuziali varino molto a seconda della cittadinanza degli sposi, essendo influenzati dal modello migratorio proprio dell'area di provenienza.

Considerando le cittadinanze di origine degli stranieri maggiormente coinvolti nel fenomeno oggetto di studio, la tipologia con sposo italiano e sposa straniera prevale nei matrimoni in cui le spose provengono dai paesi dell'est-europeo e dal Brasile; le cittadinanze in cui la distribuzione per tipologia di coppia è invece piuttosto bilanciata sono il Marocco e l'Albania. In questo panorama la Cina assume una posizione eccentrica, in quanto la quasi totalità dei matrimoni con almeno uno sposo straniero celebrati in Italia è costituita da matrimoni tra connazionali.

L'analisi dell'andamento dei matrimoni tra il 2008 e il 2009 per tipologia di coppia e cittadinanza consente di completare il quadro interpretativo del fenomeno - e del calo - osservato nell'ultimo anno.

Come già riscontrato nella figura 1, la tipologia di coppia che ha risentito meno del calo generalizzato dei matrimoni è quella con sposo italiano e sposa straniera; i matrimoni con spose rumene sono addirittura aumentati (Figura 4).

Figura 3 – *Matrimoni con almeno uno sposo straniero per tipologia di coppia e principali paesi di cittadinanza. Anno 2009.*

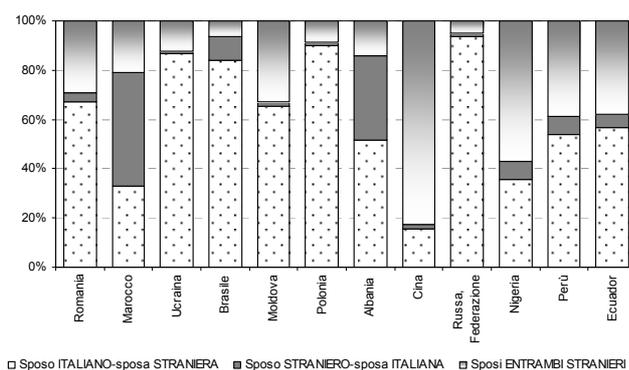
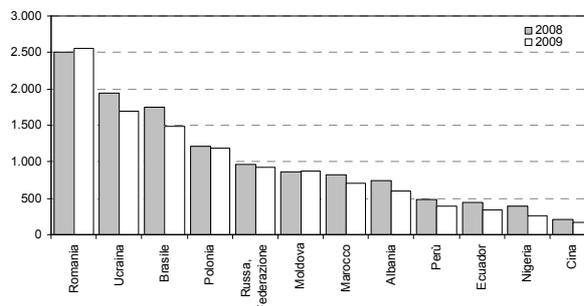


Figura 4 – *Matrimoni tra SPOSO ITALIANO e SPOSA STRANIERA per principali paesi di cittadinanza. Anni 2008 e 2009 (valori assoluti).*

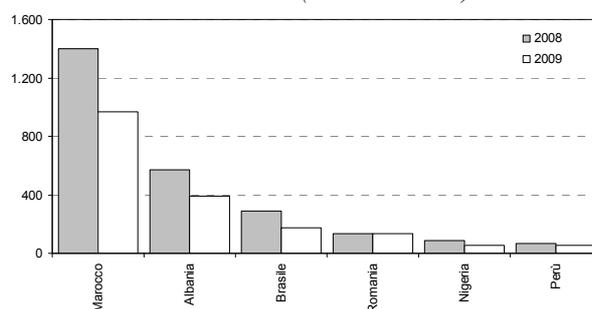


Se consideriamo invece la tipologia con sposo straniero e sposa italiana, vediamo come la riduzione dei matrimoni sia più consistente (Figura 5); i matrimoni con sposi marocchini e albanesi, i più rappresentativi del fenomeno, si sono abbattuti di oltre il 30%, mentre quelli con cittadini brasiliani, di oltre il 40%. Anche in questo caso fa eccezione, come già detto, la Romania, in quanto, probabilmente, l'ingresso nell'Ue ha agito da fattore "protettivo".

Se consideriamo, infine, i matrimoni tra cittadini entrambi stranieri - tipologia maggiormente esposta a recepire gli effetti di modificazioni normative riguardanti la popolazione straniera - e in particolare se selezioniamo solo i cittadini extra-Ue, vediamo come il calo sia generalizzato per tutte le principali cittadinanze, ad

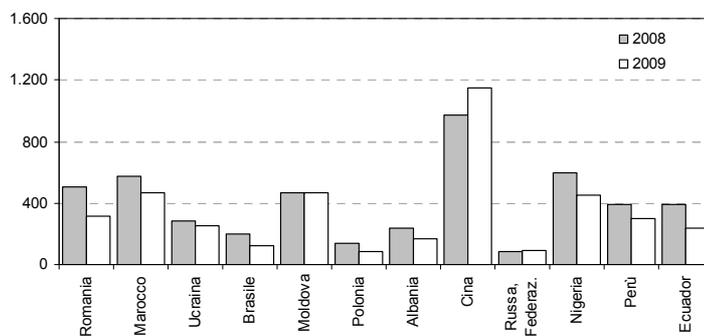
eccezione della cinese; quest'ultima, ricordiamo, è caratterizzata da una larghissima maggioranza di matrimoni tra connazionali e presenta una storia migratoria verso l'Italia ben consolidata e per lo più nell'alveo della regolarità (Figura 6).

Figura 5 – Matrimoni tra SPOSO STRANIERO e SPOSA ITALIANA per principali paesi di cittadinanza. Anni 2008 e 2009 (valori assoluti).



La recente normativa sembra colpire, infatti, in misura minore le cittadinanze che hanno più facilità ad inserirsi nel tessuto sociale e lavorativo e che hanno dunque maggiori opportunità di registrare una presenza regolare.

Figura 6 – Matrimoni tra SPOSI ENTRAMBI STRANIERI (extra UE) per principali paesi di cittadinanza*. Anni 2008 e 2009 (valori assoluti).



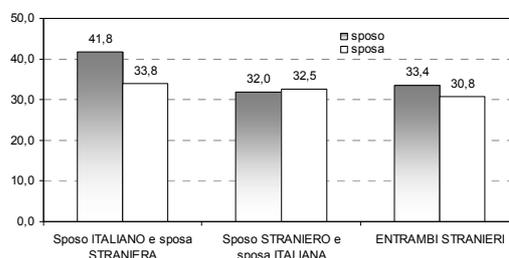
(*) Si fa riferimento alla cittadinanza della sposa

Come è dunque evidente, il comportamento nuziale, in seguito alle trasformazioni normative, si è modificato in maniera differente a seconda della tipologia di coppia considerata.

Pensando ai diversi percorsi di migrazione e di integrazione che i cittadini stranieri provenienti da diverse aree geografiche compiono nel nostro paese, è

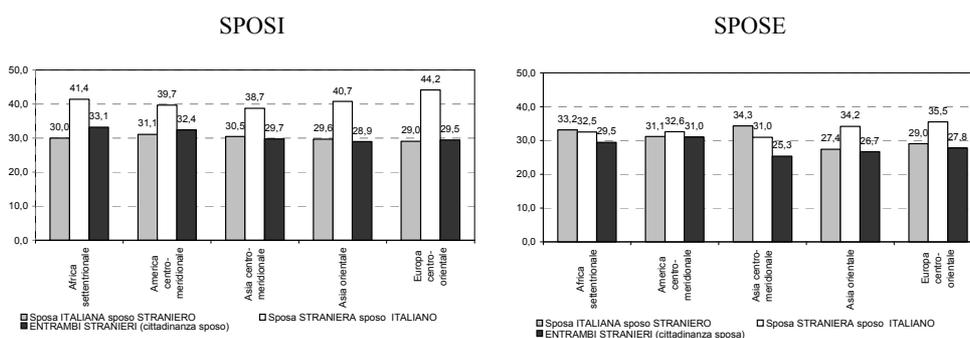
interessante inserire nell'analisi anche l'andamento di variabili puramente demografiche come l'età media al matrimonio. Come descritto nella figura 7, la differenza media di età tra coniugi è di circa 8 anni quando la sposa è straniera e lo sposo è italiano, mentre non raggiunge nemmeno 1 anno quando ad essere straniero è lo sposo; quando il matrimonio è tra sposi entrambi stranieri, l'age gap è di circa 2 anni e mezzo, ricalcando approssimativamente la differenza che si riscontra nelle età al matrimonio tra sposi entrambi italiani.

Figura 7 – Età media al matrimonio degli sposi nei matrimoni con almeno uno sposo straniero per tipologia di coppia. Anno 2009.



L'approfondimento per genere e cittadinanza mostra che l'età media al matrimonio dello sposo è sempre la più alta, per tutte le macro-aree di cittadinanza, nella tipologia sposo italiano-sposa straniera (Figura 8).

Figura 8 – Età media al matrimonio degli sposi e delle spose nei matrimoni con almeno uno sposo straniero per tipologia di coppia e principali macroaree di cittadinanza. Anno 2009.



Se consideriamo invece l'età media delle spose, la situazione per cittadinanza appare più eterogenea. Quando lo sposo proviene dai paesi dell'Africa settentrionale o dall'Asia centro-meridionale, la sposa italiana ha un'età media più alta rispetto alle altre tipologie; nelle restanti macroaree di provenienza, invece, è la sposa straniera che sposa un uomo italiano a presentare l'età relativamente più alta.

L'analisi per tipologia di coppia evidenzia diversi profili socio-demografici fortemente connotati dalle storie migratorie che contraddistinguono le varie cittadinanze; non va tuttavia trascurata l'indiscutibile importanza del contesto che, a sua volta, influenza e modifica i processi di integrazione dei cittadini stranieri.

Riferimenti bibliografici

- ASGI 2009. Il nuovo art. 116 del c.c. dopo l'entrata in vigore della legge n. 94/2009 e la condizione di regolarità del soggiorno del cittadino extracomunitario ai fini della capacità matrimoniale in Italia. *Segreteria organizzativa ASGI*, 10 agosto 2009.
- ISTAT 2011. *Il matrimonio in Italia. Anno 2009 e dati provvisori 2010*. Statistiche in breve, 18 maggio 2011.
- PALMIERI L. 2010. Il matrimonio dello straniero in Italia. Il delicato ruolo dell'ufficiale di stato civile dopo la modifica dell'art. 116 cod. civ. 30° *Convegno Nazionale Anusca*, Merano, Dicembre 2010.
- TERZERA L. 2006. Famiglia e fecondità. In: PATERNO A., STROZZA S. e TERZERA L. (Eds.), *Sospesi tra due rive: migrazioni e insediamenti di marocchini e albanesi*, Milano: Franco Angeli, pp. 105-129.

SUMMARY

In spite of the traditional growing trend of marriages with at least one foreign spouse, the year 2009 has seen a clear decrease. It is possible to presume the effect of the new legislation on public security introduced in 2009 (Law 94/2009). According to this law foreign people who wish to marry in Italy require a document that certifies their legal stay in the country. Especially created to discourage marriages of convenience, this law has drastically decreased the total amount of marriages with at least a foreign spouse.

The analysis by type of couple singled out different socio-demographic profiles of spouses deeply interrelated with the belonging to the various citizenships. As a matter of fact peculiar migratory projects and integration processes characterize the citizens coming from different geographical areas.

Cinzia CASTAGNARO, Ricercatore, Istat, cicastag@istat.it
Antonella GUARNERI, Ricercatore, Istat, guarneri@istat.it
Claudia IACCARINO, Collaboratore tecnico Enti di Ricerca, Istat, iaccarin@istat.it

WEIGHTING ISSUES IN LFS LONGITUDINAL DATA

Claudio Ceccarelli, Antonio R. Discenza, Francesca Fiori, Carlo Lucarelli

1. Introduction

The Labour Force Survey is a sample survey of households, which provides quarterly official estimates on the Italian labour market. The survey is designed to produce cross-sectional data; it has, however, a rotating sample design with scheme 2-2-2 so that each sample household is interviewed for 4 quarters. Therefore, the records can be linked together to produce a rich source of longitudinal data at 3, 9, 12 and 15 months. A complete series of 12 months (two-quarters) longitudinal datasets has been produced linking existing LFS data from 2004 to 2010. Imputation of longitudinal inconsistencies has been implemented and longitudinal weights have been computed. These longitudinal datasets can be used to produce estimates of gross flows over time between different economic activity statuses (transition matrices). This paper aims at providing a synthetic overview of some of the main methodological issues related to the production of these estimates, with a specific focus on the weighting procedure.

2. Methodological issues in weighting longitudinal data

Producing linked records is relatively straightforward given that a unique identification code for each individual is used for all waves. However, in order to produce longitudinal datasets and transition matrices we must be fully aware of the following aspects (Ceccarelli et al. 2002):

- a) LFS is not a panel survey, thus individuals which move out of the selected households, or households which move out of the municipality, are not re-interviewed;
- b) household non-response may occur at subsequent waves due to non-contact, refusal, etc.;
- c) the longitudinal sample is representative only of a specific longitudinal reference population, not of the whole population;

d) weights should account for the longitudinal population, for the total non-response (usually not at random) and should ensure consistency/coherence with the LFS quarterly estimates.

Given the Italian LFS sample and survey design, the most relevant methodological problems addressed to compute weights for longitudinal data, are the following:

- definition of a suitable reference population for the longitudinal sample;
- longitudinal non-responses and eligibility;
- coherence between cross-sectional and longitudinal estimates.

2.1 Definition of the reference population

The longitudinal component (sub-sample) of the Italian LFS requires the specification of a suitable reference population (Discenza 2004). For the Italian LFS, longitudinal data concerns only individuals who are resident, both at the beginning and at the end of the period, in the same municipality (however, for sampling frames which use dwellings or addresses the longitudinal population could be defined in a different way).

Figure 1 – Classification of individuals from the initial sample and eligibility in case of a real panel and absence of longitudinal non-response.

| INITIAL QUARTER | | FINAL QUARTER | | | LONGITUDINAL DATA | | | |
|-------------------------------|----------------------------------|-----------------------------------------------------------------------------------------|------------------------------------------------|-----------------------------|---------------------|------------------------------------|-------------|------------------------------------|
| PRESENT IN THE INITIAL SAMPLE | REPRESENT THE INITIAL POPULATION | CLASSIFICATION OF INDIVIDUALS FROM INITIAL SAMPLE | | PRESENT IN THE FINAL SAMPLE | LINK | PRESENT IN THE LONGITUDINAL SAMPLE | ELIGIBILITY | REPRESENT THE REFERENCE POPULATION |
| YES | YES | INDIVIDUALS STILL RESIDENT IN THE SAME MUNICIPALITY | RESPONDENTS | YES | LINKED (or MATCHED) | YES | YES | YES |
| | | | INTERNAL MIGRATION (TO ANOTHER MUNICIPALITY) | YES | | | | |
| | | INDIVIDUALS NOT ANYMORE RESIDENT IN THE SAME MUNICIPALITY (EXIT THE INITIAL POPULATION) | INTERNATIONAL MIGRATION (TO ANOTHER COUNTRY) | YES | | | | |
| | | | DEATHS | YES | | | | |

Hence, for the Italian LFS, the reference population is defined as the population which is resident in the same municipality over the 12 months (or 3 months) period of thus net of deaths and internal or international migration (Figure 1).

In principle there are several possibilities to define the reference population for the longitudinal LFS sample, but the final choice depends on the sample design, on the survey design and on the availability of population totals for weighting. The choice is anyway imposed by the fact that the LFS longitudinal component is not a real panel, and this has a direct effect on the transition matrix (matrix of stocks and flows) that can be built.

a) The reference population is the initial population

Ideally, longitudinal data from LFS should represent the whole initial population. However, the initial population actually modifies during the period of observation being exposed to several demographic events, such as deaths and internal/international migration. Thus longitudinal data could represent the whole initial population only if the LFS would have been designed like a “proper” panel, in which all the individuals in the initial sample were “followed” for a new interview at a later stage. In this “theoretical” situation we would be able to obtain nearly unbiased flow estimates for all the initial population, but in practice, we are quite far from it.

b) The reference population is the longitudinal population

Usually, in the LFS, people moving out of the household are not “followed” for re-interview. Moreover, information on deaths of individuals which were in the initial sample is not available. Under this situation we should think whether it is still correct to use the initial population as the reference population. In fact, we only have longitudinal information about those individuals still resident in the same municipality at the end of the period.

If we weight the longitudinal sample to the initial population we are making a very strong assumption: *the behaviour of individuals who move out of the household (dwelling or address) from one wave to another is similar to those who do not move.*

We have, thus, at least two problems:

- it has known that – at least as far as Italy is concerned – these are actually two very distinct groups in terms of labour market behaviours. Therefore, mobile and non-mobile individuals are likely to have different probability of permanence in/transition from a given labour market. Let’s make an example: at the beginning of the period we have two unemployed persons with the same characteristics, resident in a region where most of the population is employed in agriculture; if one person continues to stay in the same region and gets a job at the end of the period, it will be most probably in the agriculture sector; if the other person move to another region (where most of the labour demand is in construction) and gets a job at the end of the period, it will be most probably in construction.

→ if we use the longitudinal microdata to produce flow estimates, there are no records of individuals moving to other regions/country and/or dying (whereas they do exist in the population).

Bearing these considerations in mind, it seems that the best choice is given by the “longitudinal population” defined as the population which is resident in the same municipality for the period of 12 months (or 3 months), net of deaths and internal or international migration (Figure 2).

The advantages of this choice are the following:

- the assumption of similarity between the behaviour of the individuals moving out of the households from one wave to another and of those who do not move is not required;
- estimates of stock and flows from longitudinal data will probably have a much smaller bias.

Figure 2. – *Classification of individuals from the initial sample and eligibility in the Italian LFS (hypothesis of absence of longitudinal non-response).*

| INITIAL QUARTER | | FINAL QUARTER | | | LONGITUDINAL DATA | | | |
|-------------------------------|----------------------------------|-----------------------------------------------------------------------------------------|------------------------------------------------|-----------------------------|---------------------------|------------------------------------|-------------|------------------------------------|
| PRESENT IN THE INITIAL SAMPLE | REPRESENT THE INITIAL POPULATION | CLASSIFICATION OF INDIVIDUALS FROM INITIAL SAMPLE | | PRESENT IN THE FINAL SAMPLE | LINK | PRESENT IN THE LONGITUDINAL SAMPLE | ELIGIBILITY | REPRESENT THE REFERENCE POPULATION |
| YES | YES | INDIVIDUALS STILL RESIDENT IN THE SAME MUNICIPALITY | RESPONDENTS | YES | LINKED (or MATCHED) | YES | YES | YES |
| | | | INTERNAL MIGRATION (TO ANOTHER MUNICIPALITY) | NO | | NO | NO | NO |
| | | INDIVIDUALS NOT ANYMORE RESIDENT IN THE SAME MUNICIPALITY (EXIT THE INITIAL POPULATION) | INTERNATIONAL MIGRATION (TO ANOTHER COUNTRY) | NO | NOT LINKED (or UNMATCHED) | NO | NO | NO |
| | | | DEATHS | NO | | NO | NO | NO |

The only disadvantage is that we obtain flow estimates for a sub-population (smaller than the initial population, depending on the country’s levels of mobility/migration).

c) reduced sample and longitudinal population

Another relevant issue related to the longitudinal component of the LFS concerns the impact of unit non-response in subsequent waves. We may observe:

- Municipality non-response: some (very small) municipalities are replaced in July at the beginning of a new annual survey cycle and some others

may, for different reasons, fail to provide the interviews in subsequent waves;

- Household non-response: all the members of the household do not fill in the questionnaire because they refuse to respond;
- Individual non-response: some members of the household do not fill in the questionnaire because they refuse to respond, or they cannot be contacted or they left the household to create a new household in the same municipality.

Unit non-response may reduce the size of the longitudinal component, thus increasing the variance of the estimates. It can also produce bias if non-respondents have significantly different labour features with respect to respondents.

Given the longitudinal population defined as above, and due to non-responses, all the individuals interviewed at the initial quarter can be classified according to the Figure 3.

Figure 3 – Classification of individuals from the initial sample and eligibility in the Italian LFS (in presence of longitudinal non-response).

| INITIAL QUARTER | | FINAL QUARTER | | | LONGITUDINAL DATA | | | |
|-------------------------------|----------------------------------|-----------------------------------------------------------------------------------------|------------------------------------------------|-----------------------------|---------------------------|------------------------------------|-------------|------------------------------------|
| PRESENT IN THE INITIAL SAMPLE | REPRESENT THE INITIAL POPULATION | CLASSIFICATION OF INDIVIDUALS FROM INITIAL SAMPLE | | PRESENT IN THE FINAL SAMPLE | LINK | PRESENT IN THE LONGITUDINAL SAMPLE | ELIGIBILITY | REPRESENT THE REFERENCE POPULATION |
| YES | YES | INDIVIDUALS STILL RESIDENT IN THE SAME MUNICIPALITY | RESPONDENTS | YES | LINKED (or MATCHED) | YES | YES | YES |
| | | | REFUSALS * | NO | | NO | YES | YES |
| | | | UNREACHABLE ** | NO | | NO | YES | YES |
| | | INDIVIDUALS NOT ANYMORE RESIDENT IN THE SAME MUNICIPALITY (EXIT THE INITIAL POPULATION) | INTERNAL MIGRATION (TO ANOTHER MUNICIPALITY) | NO | NOT LINKED (or UNMATCHED) | NO | NO | NO |
| | | | INTERNATIONAL MIGRATION (TO ANOTHER COUNTRY) | NO | | NO | NO | NO |
| | | | DEATHS | NO | | NO | NO | NO |
| | | | | | | | | |

First of all they can be classified into two groups:

- eligible: they represent part of the longitudinal population (because still living in the same municipality), and should be re-interviewed at the subsequent wave. However, some of them are non-respondents in the final quarter, so that they must be considered in a model for treatment of non-response;

→ non-eligible: they left the initial population during the observed period (deaths and migrations) thus they should be excluded from a model for treatment of non-response.

When using models/methods for treatment of longitudinal non-response we face another important issue: usually, we don't have enough information to distinguish eligible from non-eligible individuals.

2.2. Coherence between cross-sectional and longitudinal estimates

The last problem is that the longitudinal component produces both cross-sectional and longitudinal estimates referred to the longitudinal population. These cross-sectional estimates, obtained from the longitudinal data, have to be consistent (i.e. not higher than) with the "official" estimates provided by the cross-sectional samples (the full sample) at the beginning and at the end of the observed period.

Given that longitudinal estimates have higher variability than quarterly estimates, it is not possible to control their consistency completely. However, it is possible to reduce the risk of obtaining inconsistent results by introducing specific constraints in a weighting process.

3. Weighting procedure for longitudinal data

Generalized raking procedures (calibration) have been used to account for, and thus reduce, the bias due to wave non-response. They are also used to ensure consistency with quarterly estimates. Raking involves modifying the original weights in order to simultaneously satisfy several marginal constraints (or to control totals) while minimizing the distance between original and adjusted weights. Final longitudinal weights have been obtained after two calibration stages: the first accounting for bias due to municipality non-response and due to differences in rotation groups; the last adjusting for bias due to individual non-response and to make weighted sample totals conform to the longitudinal population.

All the individuals who were *linkable* at the beginning of the period (i.e. all the individuals of the two rotation groups which overlap at 12 month and resident in those municipalities which provided interviews for both waves) have been selected; they can be considered like a random sub-sample of the whole cross-sectional sample (Rizzo et al. 1996).

Let k_i denote cross-sectional weight computed for the whole cross-sectional sample at the beginning of the period. Then, for *linkable* individuals, base longitudinal weights k_i^* are obtained from k_i applying the following correction

$$k_i^* = \left(k_i / \sum_{i \in \text{Linkable}} k_i \right) {}_1P \quad (1)$$

where ${}_1P$ is the cross-sectional population at the beginning of the period.

In order to ensure consistency between longitudinal and cross-sectional “official” estimates, the first calibration procedure makes *linkable* individuals at the beginning of the period representing exactly the same cross-sectional population of the whole cross-sectional sample. Moreover, it makes them provide the same cross sectional “official” estimates for a number of relevant figures (cross-classification of sex, region, age group, labour activity status, education, etc.). Thus, from the base longitudinal weight k_i^* and for all the *linkable* individuals the initial weight $g_i^* = k_i^* \delta_i^*$ has been obtained as result of a problem of constrained minimum as follows:

$$\sum_{i \in \text{Linkable}} g_{i,rse}^* = \sum_{i=1}^{N_1} w_{i,rse} = {}_1P_{rse}, \quad \sum_{i \in \text{Linkable}} g_i^* X_i = \sum_{i=1}^{N_1} w_i X_i, \quad \sum_{i \in \text{Linkable}} g_i^* X_i = \sum_{i=1}^{N_1} w_i X_i \quad (2)$$

$$\min \left\{ \sum_{i \in \text{Linkable}} D(k_i^* \delta_i^*, k_i^*) \right\} \quad (3)$$

where w_i is the first quarter cross-sectional final weight, ${}_1P_{rse}$ are the known totals for the first quarter cross-sectional population for area r , sex s and age group e , N_1 is the first quarter cross-sectional sample dimension, X_i are auxiliary variables, D is a logarithmic distance function (Deville, Sarndal, 1992).

In the second stage, the final longitudinal weight $w_i^* = g_i^* \gamma_i^*$, only for *linked* individuals, have been obtained applying a different calibration step to make weighted longitudinal component totals conform to totals derived from the longitudinal population under the following constraints:

$$\sum_{i \in \text{linked}} w_{i,rse}^* = {}_lP_{rse}, \quad \min \left\{ \sum_{i \in \text{linked}} D(g_i^* \gamma_i^*, g_i^*) \right\} \quad (4)$$

where ${}_lP_{rse} = {}_1P_{rse} - m_{rse} - c_{rse}$ is the known longitudinal population.

4. Conclusions

The procedure described above has been applied (until now) to six yearly longitudinal datasets from the 1st quarter of 2004 to the 1st quarter 2010 and transition matrices have been produced (a procedure has been set up to produce these data on quarterly bases). It has been found that it reduces the negative effects

of non-response bias providing consistent estimates for almost all the relevant figures. Few inconsistencies have been found for not-significant estimates and for those variables more likely to be subject to classification errors.

References

- CECCARELLI C., DISCENZA A. R., ROSATI S., PAGGIARO A., TORELLI N. 2002. *Le matrici di transizione della rilevazione trimestrale sulle forze di lavoro*. Nota metodologica, Istat, Roma.
- DEVILLE J.C., SARNDAL C.E. 1992. Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, Vol. 87, pp. 376-382.
- DISCENZA A. R. 2004. Weighting Procedure for Longitudinal Data of the Italian Labour Force Survey. In *Atti Convegno SIS 2004*, Bari.
- RIZZO L., KALTON G., BRICK J.M. 1996. A Comparison of Some Weighting Adjustment Methods for Panel Nonresponse. *Survey Methodology*, Vol. 22, pp.43-53.
- TATE P.F., CLARKE P.S. 1999. Methodological issues in the production and analysis of longitudinal data from the Labour Force Survey. *GSS Methodology Series*, Vol. 17.

SUMMARY

The Italian Labour Force Survey belongs to the framework of the European Union Quarterly Labour Force Survey and satisfies completely all the EU Regulations. It has a 2-2-2 rotating sample, thus, given a quarterly sample, 50% of the households are re-interviewed again after 3 and 12 months; 25% of them after 9 and 15 months. By matching (and linking) the records associated to the same individual for the different quarters, it is possible to build longitudinal datasets and transition matrices. Longitudinal data allow a fair evaluation of the level of labour market mobility, in particular in a time of economical crisis when dynamics result markedly accentuated. This paper presents some issues related to the production of gross flows estimates consistent with the quarterly estimates already disseminated; a specific focus is devoted to the weighting procedure, which account both for the longitudinal population and for the total non-response at subsequent waves, which is usually not at random.

Claudio CECCARELLI, Senior Researcher, ISTAT, clceccar@istat.it
Antonio Rinaldo DISCENZA, Senior Researcher, ISTAT, discenza@istat.it
Francesca FIORI, Research Fellow, Centre for Population Change, University of St.Andrews (UK), ff20@st-andrews.ac.uk
Carlo LUCARELLI, Researcher, ISTAT, calucare@istat.it

STIMATORI DI PONDERAZIONE VINCOLATA NEL CASO DI TOTALI NOTI PROVENIENTI DALLA STESSA INDAGINE CAMPIONARIA

Claudio Ceccarelli, Giovanni Maria Giorgi, Alessio Guandalini

Premessa

Nel panorama della statistica ufficiale si sta sempre più diffondendo l'utilizzo di disegni campionari complessi sia nella definizione della strategia di selezione delle unità campionarie sia nella fase di determinazione degli stimatori.

In tal senso è sempre maggiore la necessità di un utilizzo più complesso anche dello stimatore di ponderazione vincolata (*PV*), rispetto alla versione originaria messa a punto da Deville e Särndal nel 1992, in termini di numero, dettaglio e tipologia dei totali noti anche da rilevazioni campionarie.

In genere l'informazione proveniente da fonte campionaria è introdotta per soddisfare l'esigenza di produrre stime coerenti con altre indagini che si riferiscono allo stesso periodo o con stime ottenute dalla stessa indagine in occasioni precedenti. L'utilizzo di tali informazioni pone il problema di come l'errore associato al vincolo campionario possa ripercuotersi sull'efficienza della stima della variabile di interesse provenienti da altre rilevazioni campionarie.

L'obiettivo di questo lavoro, a partire dalla formulazione della varianza dello stimatore *PV* in presenza di vincoli campionari da indagini indipendenti (Ceccarelli, Giorgi, Guandalini, 2010), è quella di valutare la *performace* dello stimatore *PV* che utilizza variabili stimate da occasioni precedenti della stessa indagine campionaria.

La valutazione viene effettuata sulle stime mensili della Rilevazione sulle Forze di Lavoro (*RFL*) che rappresentano un esempio di impiego di vincoli campionari longitudinali provenienti, cioè, da occasioni precedenti della stessa indagine.

Il lavoro si sviluppa nel seguente modo: nel paragrafo 2 sono riprese le proprietà che servono, nel paragrafo 3, a determinare l'espressione analitica della varianza dello stimatore di ponderazione vincolata che utilizza informazioni campionarie provenienti dalla stessa indagine. Nei paragrafi successivi è illustrata l'applicazione e sono analizzati i risultati ottenuti applicando l'espressione determinata alle stime mensili della *RFL*.

1. Lo stimatore di ponderazione vincolata (PV)

Nelle indagini su larga scala, su ciascuna unità del campione vengono rilevate, oltre alla variabile di interesse Y , una serie di variabili ausiliarie ($j = 1, \dots, J$) sintetizzate dal vettore $\mathbf{x} = (x_{1k}, \dots, x_{jk}, \dots, x_{Jk})'$, così per ciascuna unità campionaria si ha un vettore (y_k, \mathbf{x}_k) composto da $J+1$ osservazioni. Il vettore delle informazioni ausiliarie, e in particolar modo il vettore dei totali di queste, $\mathbf{X} = \sum_{k \in U} \mathbf{x}_k = (\sum_{k \in U} x_{1k}, \dots, \sum_{k \in U} x_{jk}, \dots, \sum_{k \in U} x_{Jk})$, può essere utilizzato per migliorare le stime della variabile di interesse.

Deville e Särndal in base all'idea che "...weights that perform well for the auxiliary variable also should perform well for study variable" (Deville e Särndal, 1992, p. 376) hanno messo a punto lo stimatore di ponderazione vincolata (PV), che nella sua definizione originaria, è in grado di sfruttare l'informazione dei totali noti delle variabili ausiliarie riferita alla popolazione oggetto d'indagine per produrre stime coerenti (vincolate) per individui e famiglie (Deville e Särndal, 1992).

Lo stimatore di ponderazione vincolata, $\hat{Y}_{PV} = \sum_{k \in S} y_k w_k$, assegna a tutti gli individui dello stesso nucleo familiare pesi finali w_k (Lemaître e Dufour, 1987, p. 199), che sono, in media, per una data funzione di distanza, il più vicino possibile ai pesi base d_k (uguali all'inverso della probabilità di inclusione del primo ordine, π_k) assegnati dallo stimatore di Horvitz-Thompson (HT) (Deville, Särndal, Sautory, 1993, p. 1013) e che rispettano un sistema di vincoli con il vettore dei totali noti \mathbf{X} .

Nel caso di campioni di grandi dimensioni lo stimatore PV coincide asintoticamente con lo stimatore di regressione generalizzato (GREG) in quanto le funzioni utilizzate generalmente per misurare la distanza tra i pesi w_k e d_k sono lineari in \mathbf{x} (Deville e Särndal, 1992, pp.377-380 e Singh e Mohl, 1996, pp. 108-114). Questo risultato è fondamentale soprattutto per studiare le varianze dello stimatore PV, che altrimenti non sarebbe possibile determinare analiticamente a causa della complessità che il calcolo della probabilità di inclusione del secondo ordine (π_{kl}) può assumere.

Infatti, seguendo Deville e Särndal (1992, pp. 379-380), la varianza asintotica dello stimatore PV coincide con quella dello stimatore GREG e una stima corretta di tale quantità è:

$$var(\hat{Y}_{PV}) = \sum_{k \in S} \sum_{l \neq k} \frac{\Delta_{kl}}{\pi_{kl}} \left(\gamma_k \frac{\hat{e}_k}{\pi_k} \right) \left(\gamma_l \frac{\hat{e}_l}{\pi_l} \right), \quad (1)$$

dove $\Delta_{kl} = \pi_{kl} - \pi_k \pi_l$ ed \hat{e} e γ sono rispettivamente i residui stimati del modello di regressione e i correttori dei pesi base.

2. Lo stimatore di ponderazione vincolata con vincoli campionari (\widehat{PV})

Lo stimatore PV definito nel paragrafo 2 garantisce, per definizione, la coerenza con una fonte esterna (amministrativa o censuaria) priva di errore campionario in quanto la stima è vincolata ad un vettore \mathbf{X} di totali noti con certezza. È possibile, però, vincolare la stima ad un vettore di totali noti desunti da un'altra indagine campionaria, $\widehat{\mathbf{X}}$ (Deville, 1999, p. 207).

Nel caso in cui il campione dell'indagine per cui si calcola la stima e quello dell'indagine campionaria da cui sono desunti i totali utilizzati come vincoli sono indipendenti tra loro, secondo Ballin, Falorsi e Russo (2000), una stima asintoticamente corretta della varianza dello stimatore \widehat{PV} è:

$$\begin{aligned} \text{var}(\widehat{Y}_{\widehat{PV}ind}) &\cong \text{var}(\widehat{Y}_{GREGind}) = \\ &= \sum_{k \in S} \sum_{l \neq k} \frac{\Delta_{kl}}{\pi_{kl}} \left(\gamma_{ks} \frac{\hat{e}_k}{\pi_k} \right) \left(\gamma_{ls} \frac{\hat{e}_l}{\pi_l} \right) + \sum_{j=1}^J f_j \left(B_j^2 \text{var}(\widehat{X}_j) + \sum_{j' \neq j} B_j B_{j'} \text{cov}(\widehat{X}_j, \widehat{X}_{j'}) \right) \end{aligned} \quad (2)$$

dove il primo termine rappresenta l'errore nella stima della variabile Y per l'indagine in questione ed è uguale alla varianza (1) ed il secondo l'errore aggiuntivo che si commette vincolando la stima dei totali delle variabili ausiliarie ad un'altra indagine affetta a sua volta da errore campionario¹.

Nel caso in cui in maniera sistematica l'indagine per cui si vogliono produrre le stime e quella da cui si desumono i vincoli hanno delle unità in comune, è, invece, necessario introdurre una componente di variabilità congiunta che, sotto l'ipotesi di costanza nel tempo dei coefficienti di regressione e di correttezza della stima dei totali noti usati come vincoli, è uguale a $2\widehat{Y}$ (Ceccarelli, Giorgi, Guandalini 2010, pp. 24-25). La stima della varianza di PV con vincoli campionari da indagini dipendenti è, dunque:

$$\begin{aligned} \text{var}(\widehat{Y}_{\widehat{PV}dip}) &\cong \text{var}(\widehat{Y}_{GREGdip}) = \\ &= \sum_{k \in S} \sum_{l \neq k} \frac{\Delta_{kl}}{\pi_{kl}} \left(\gamma_{ks} \frac{\hat{e}_k}{\pi_k} \right) \left(\gamma_{ls} \frac{\hat{e}_l}{\pi_l} \right) + \sum_{j=1}^J f_j \left(B_j^2 \text{var}(\widehat{X}_j) + \sum_{j' \neq j} B_j B_{j'} \text{cov}(\widehat{X}_j, \widehat{X}_{j'}) + 2\widehat{Y} \right) \end{aligned} \quad (3)$$

¹ Nella proposta originaria di Ballin, Falorsi e Russo (2000) la (2) si riferisce al caso di un campione casuale semplice e di indipendenza delle due indagini e il secondo termine non viene scontato per la proporzione di individui per cui agisce il vincolo j , f_j . Qui si considera, invece, l'estensione al caso di campione complesso a due stadi con stratificazione delle unità di primo livello. Si veda per una trattazione più ampia Ceccarelli, Giorgi, Guandalini (2010, 2011).

La (2) e la (3) si possono generalizzare e scrivere in forma matriciale come:

$$\text{var}(\hat{Y}_{\overline{pV}}) \cong \text{var}(\hat{Y}_{GREG}) + \mathbf{1}'(\mathbf{F}\mathbf{L})\mathbf{1} + 2\hat{Y}\mathbf{f}'\mathbf{I}[j], \quad (4)$$

dove $\mathbf{1}$ è un vettore di uno di dimensione j , \mathbf{F} e \mathbf{L} sono due matrici quadrate di dimensione J così definite:

$$\mathbf{F} = \begin{bmatrix} f_1 & 0 & \dots & 0 & \dots & 0 \\ 0 & f_2 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f_j & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & f_j \end{bmatrix}$$

dove gli f_j rappresentano la proporzione degli individui del campione su cui agisce il vincolo j -mo,

$$\mathbf{L} = \begin{bmatrix} B_1^2 \text{var}(X_1) & B_1 B_2 \text{cov}(X_1, X_2) & \dots & B_1 B_j \text{cov}(X_1, X_j) & \dots & B_1 B_j \text{cov}(X_1, X_j) \\ B_2 B_1 \text{cov}(X_2, X_1) & B_2^2 \text{var}(X_2) & \dots & B_2 B_j \text{cov}(X_2, X_j) & \dots & B_2 B_j \text{cov}(X_2, X_j) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ B_j B_1 \text{cov}(X_j, X_1) & B_j B_2 \text{cov}(X_j, X_2) & \dots & B_j^2 \text{var}(X_j) & \dots & B_j B_j \text{cov}(X_j, X_j) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ B_j B_1 \text{cov}(X_j, X_1) & B_j B_2 \text{cov}(X_j, X_2) & \dots & B_j B_j \text{cov}(X_j, X_j) & \dots & B_j^2 \text{var}(X_j) \end{bmatrix}$$

dove il generico B_j è il relativo coefficiente di regressione associato alla j -ma variabile ausiliaria campionaria viene moltiplicato per le varianze e covarianze delle variabili ausiliarie². Invece $\mathbf{f}' = (f_1, f_2, \dots, f_j, \dots, f_j)$ e $\mathbf{I}[j]$ è un vettore dove ciascun elemento assume valore 1 nel caso in cui il totale noto della variabile j è noto da fonte campionaria dipendente e 0 altrimenti.

L'elemento $\mathbf{1}'(\mathbf{F}\mathbf{L})\mathbf{1}$ della (4) è uguale a 0 se tutti i totali delle variabili ausiliarie provengono da fonte amministrativa e l'elemento $2\hat{Y}\mathbf{f}'\mathbf{I}[j]_{[dip,ind]}$ è diverso da 0 solo quando ci sono vincoli campionari da indagini dipendenti.

² Se il totale della variabile ausiliaria è noto da fonte amministrativa $\text{var}(X_j) = 0$ e $\text{cov}(X_j, X_{j'}) = \rho_{j,j'} \sqrt{\text{var}(X_j)\text{var}(X_{j'})} = 0$ con $j \neq j'$.

3. La Rilevazione sulle Forze Lavoro

La Rilevazione sulle Forze di Lavoro (*RFL*)³ è una rilevazione continua effettuata in tutte le settimane dell'anno e rappresenta la fonte ufficiale del mercato del lavoro in Italia.

La popolazione di interesse è costituita da tutti i componenti delle famiglie residenti in Italia, anche se temporaneamente all'estero, mentre sono esclusi i membri permanenti delle convivenze (ospizi, istituti religiosi, caserme, ecc.).

Il disegno campionario per ciascuna rilevazione trimestrale è di tipo complesso a due stadi (comuni-famiglie) con stratificazione delle unità di primo stadio a livello provinciale in base all'ampiezza demografica e prevede uno schema di rotazione del campione del tipo $(2_T, 2_T, 2_T)$. Per produrre stime mensili, inoltre, l'Istat, in accordo con Eurostat, ha predisposto anche un disegno di campionamento nel tempo, in modo da garantire la rappresentatività dei campioni mensili.

Le stime mensili pubblicate, ottenute attraverso il ricorso allo stimatore *PV*, sono relative agli aggregati di occupati, disoccupati e non forze lavoro a livello nazionale e distinti per sesso.

Per rendere i campioni mensili più stabili e più simili come struttura a quelli dei campioni trimestrali e, dunque, per rendere meno volatili le stime mensili l'Istat ha deciso di seguire una strategia di ponderazione che prevede 302 vincoli, di cui 206 noti da fonte amministrativa⁴ e 96 relative alle stime di occupati, disoccupati e non forze lavoro di tre mesi e di dodici mesi prima⁵.

Per il particolare disegno campionario della *RFL*, che prevede una sovrapposizione teorica del 50% del campione a un trimestre e ad un anno (quattro trimestri) il caso in questione si colloca nella situazione di indagini dipendenti.

4. Applicazione e Risultati

L'espressione (4) è stata testata sui dati del terzo mese del terzo trimestre 2009 (settembre 2009), relativi a un campione di 20.928 famiglie e 49.114 individui. I

³ Per un quadro completo sull'indagine si rimanda a Gazzelloni (2006).

⁴ Le variabili demografiche sono desunte dal bilancio demografico della popolazione residente, bilancio demografico della popolazione straniera residente, stima rapida della popolazione per sesso, età e provincia di residenza, censimento della popolazione 2001.

⁵ Stima degli occupati per regione e sesso, stima dei disoccupati per ripartizione e sesso, stima delle non forze lavoro per regione e sesso.

vincoli campionari, invece, riguardano le stime dei vari aggregati di tre e di dodici mesi prima (giugno 2009 e settembre 2008)⁶.

I valori delle stime e degli errori relativi percentuali (*CV%*) ottenuti con un sistema di pesi determinati con e senza vincoli campionari⁷ sono riportati in Tabella 1.

Tabella 1 – *Stima di occupati, disoccupati e non forze lavoro ed errori campionari relativi percentuali (CV%) con lo stimatore di ponderazione vincolata senza vincoli campionari e con vincoli campionari – Settembre 2009, Italia per sesso.*

| | SENZA VINCOLI CAMPIONARI | | CON VINCOLI CAMPIONARI | | |
|---------|--------------------------|------------|------------------------|------------|------------|
| | [1] STIMA | [2] CV% | [3] STIMA | [4] CV% | [5] CV% |
| ITALIA | | | | | |
| OCC | 22.786.251 | 0,440 | 22.886.373 | 0,331 | 0,333 |
| DIS | 2.021.889 | 2,678 | 2.031.044 | 2,26 | 2,262 |
| NFL | 34.982.481 | 0,275 | 34.873.584 | 0,216 | 0,221 |
| MASCHI | | | | | |
| OCC | 13.599.617 | 0,491 | 13.647.567 | 0,385 | 0,387 |
| DIS | 1.092.265 | 3,508 | 1.093.438 | 3,094 | 3,097 |
| NFL | 14.371.974 | 0,435 | 14.323.053 | 0,341 | 0,345 |
| FEMMINE | | | | | |
| OCC | 9.186.634 | 0,828 | 9.238.806 | 0,610 | 0,612 |
| DIS | 929.624 | 3,926 | 937.606 | 3,259 | 3,262 |
| NFL | 20.610.507 | 0,360 | 20.550.531 | 0,275 | 0,281 |

In tutti i casi si ha un aumento dell'ammontare di occupati e disoccupati, con una conseguente diminuzione delle non forze lavoro.

L'introduzione dei vincoli longitudinali consente di stabilizzare le stime garantendo la stessa condizione di tre mesi prima e un anno prima a tutti coloro che non hanno effettivamente cambiato la propria condizione nel tempo. Si ha, infatti, un miglioramento nella *performance* dello stimatore testimoniato dalla diminuzione dell'errore campionario relativo percentuale (Tabella 1, [2] vs. Tabella 1, [4]). Il guadagno in termini di efficienza è maggiore per i disoccupati, sia a livello Italia che distintamente per maschi e femmine, e raggiunge anche 7 decimi di punto percentuale (nel caso delle femmine). Questo guadagno di efficienza, dovuto ad una migliore specificazione del modello di regressione, grazie all'inserimento dei vincoli campionari longitudinali, è tale da compensare

⁶ La sovrapposizione effettiva del campione nell'applicazione è del 26,5% a tre mesi di distanza (settembre 2009 – luglio 2009) e del 28,7% a dodici mesi di distanza (settembre 2009 – settembre 2008).

⁷ Nel caso con vincoli campionari il *CV%* è stato calcolato non considerando la componente di errore esogena, cioè come se i vincoli fossero tutti noti da fonte amministrativa (Tabella 1, [4]), e con la componente di errore esogena (Tabella 1, [5]).

ampiamente l'aumento dell'errore dovuto alla natura campionaria dei vincoli, mai superiore allo 0.005% (Tabella 1, [4] vs. Tabella 1, [5]).

5. Conclusioni

L'espressione funzionale determinata per calcolare la varianza dello stimatore di ponderazione vincolata in presenza di informazioni ausiliarie campionarie consente di misurare l'errore aggiuntivo commesso nello stimare un parametro ricorrendo ad informazioni affette anch'esse da errore.

L'applicazione al caso delle stime mensili della *RFL* evidenzia come la differenza tra il valore e l'errore commesso per le stime di occupati, disoccupati e non forze lavoro, ottenute con lo stimatore con e senza vincoli campionari, sono piccole ma non trascurabili. La funzione svolta dai vincoli campionari è, infatti, principalmente quella di rendere consistenti le stime di un campione che ha una numerosità pari ad un terzo rispetto a quella dei campioni trimestrali, oltre che vincolarle a variabili per cui è dimostrato essere molto elevata la correlazione.

Le differenze dei valori ottenuti nel calcolo degli errori, considerando o no la componente esogena, è piuttosto piccola (l'errore relativo aumenta leggermente quando si considera la componente di errore esogena). Nel complesso, quindi, il ricorso a stimatori di ponderazione vincolata che utilizzano, tra l'insieme di vincoli, anche informazioni ausiliarie campionarie, migliora la qualità della stima e comporta anche un miglioramento rispetto al caso in cui il sistema non contempla questi vincoli, a patto che le variabili ausiliarie campionarie siano altamente correlate con la variabile di interesse.

Ringraziamenti

Il presente lavoro è stato realizzato nell'ambito del progetto PRIN-2007: "Disuguaglianza, povertà ed esclusione sociale: analisi e coerenza delle fonti informative, nuove misure e modelli interpretativi".

Riferimenti bibliografici

- BALLIN M., FALORSI P. D., RUSSO A. 2000. Condizioni di Coerenza e Metodi di Stima per le Indagini Campionarie sulle Imprese. *Rivista di Statistica Ufficiale*, No. 2, pp. 31-52.
- CECCARELLI C., GIORGI G. M., GUANDALINI A. 2010. Lo Stimatore di Ponderazione Vincolata in Presenza di Informazioni Ausiliarie Campionarie. *Rapporto tecnico* No.18,

- Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università degli Studi di Roma "La Sapienza".
- CECCARELLI C., GIORGI G. M., GUANDALINI A. 2011. Varianza dello Stimatore Calibrato in Presenza di Informazioni Ausiliarie Campionarie. *Rivista Italiana di Economia Demografia e Statistica*. Vol. LXVI, No. 1-2, pp. 53-60.
- DEVILLE J. C., SÄRNDAL C. E. 1992. Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*. Vol. 87, No. 418, pp. 376-382.
- DEVILLE J. C., SÄRNDAL C. E., SAUTORY O. 1993. Generalized Raking Procedures in Survey Sampling. *Journal of the American Statistical Association*. Vol. 88, No. 243, pp. 1013-1020.
- DEVILLE J. C. 1999. Simultaneous Calibration of Several Surveys. *Proceedings of statistic Canada Symposium 99. Combining Data from Different Sources*. Ottawa: Statistics Canada, pp. 207-212.
- GAZZELLONI S. 2006. *La Rilevazione sulle Forze di Lavoro: Contenuti, Metodologie, Organizzazione*. Istat: Metodi e norme, No. 32.
- LEMAÎTRE G., DUFOUR. J. (1987). An integrated method for Weighting Persons and Families. *Survey Methodology*. Vol. 13, No. 2, pp. 199-207.
- SINGH A. C., MOHL C. A. (1996). Understanding Calibration Estimators in Survey Sampling. *Survey Methodology*. Vol. 22 , No. 2, pp. 107-115.

SUMMARY

The Italian National Institute of Statistics (Istat) introduced in Labour Force Survey (LFS) several constraints from the same survey of three and twelve months before to give at monthly samples the same structures of the quarterly samples and reduce the cyclical and seasonal differences. The aim of this article is to define a variance expression of calibrated estimator with sample auxiliary variables in complex sampling designs, and to measure the impact of additional error introduced.

Claudio CECCARELLI, Primo Ricercatore, Istituto Nazionale di Statistica,
Dipartimento delle statistiche sociali ed ambientali, cceccar@istat.it
Giovanni Maria GIORGI, Ordinario di Statistica, Università degli Studi di Roma
"La Sapienza", Dipartimento di Scienze Statistiche, giovanni.giorgi@uniroma1.it
Alessio GUANDALINI, Dottorando di Statistica Metodologica, Università degli
Studi di Roma "La Sapienza", Dipartimento di Scienze Statistiche,
alessio.guandalini@uniroma1.it

DATI AMMINISTRATIVI PER LE STATISTICHE UFFICIALI SULLE MIGRAZIONI¹

Cinzia Conti, Domenico Gabrielli, Salvatore Strozza

1. Introduzione

La produzione delle statistiche in campo sociale sta andando incontro a mutamenti radicali legati anche alla maggiore disponibilità e fruibilità degli archivi amministrativi. Un impulso decisivo è stato dato dallo sviluppo dell'ICT (*Information and Communication Technology*): grazie alle nuove tecnologie per la gestione e il trattamento dei dati individuali si è infatti assistito a un ampliamento di basi di dati amministrativi. In particolare, gli archivi tradizionali hanno fatto registrare significativi miglioramenti nella completezza, nella tempestività e, seppure parzialmente, nella qualità delle informazioni raccolte. A questa accresciuta mole di dati individuali sono applicate tecniche già note, ma che stanno trovando nuove occasioni di implementazione e di affinamento: ci si riferisce in particolare all'impiego di tecniche di *record-linkage*. La più facile trasformazione di documenti amministrativi in dati statistici rende più agevole il loro sfruttamento attraverso le tecniche di *linkage*, sia trasversali, cioè fra archivi diversi in uno stesso periodo, sia longitudinali, applicate cioè sullo stesso archivio riferito a tempi diversi. Di conseguenza cambiano e si accrescono gli indicatori disponibili che possono essere oggi, sempre più spesso, anche di tipo "longitudinale".

L'archivio dei permessi di soggiorno resta attualmente l'esempio più rilevante nel campo demografico-sociale di archivio di dati amministrativi sfruttato a fini statistici, grazie alle elaborazioni eseguite dall'Istat a partire dai dati forniti dal Ministero dell'Interno (Istat, 1998; 2009). Sebbene esso faccia riferimento "solamente" ai cittadini non comunitari, resta una fonte di primaria importanza per lo studio dell'immigrazione per il tipo di informazioni raccolte a livello individuale ed elaborate ormai da quasi vent'anni.

¹ Cinzia Conti ha scritto il paragrafo 3, Domenico Gabrielli il paragrafo 1 e Salvatore Strozza il paragrafo 2. Le conclusioni sono state redatte congiuntamente dagli autori.

Da tempo l'Istat ha avviato sull'archivio dei permessi uno studio per la valorizzazione delle informazioni raccolte in chiave "longitudinale" attraverso l'applicazione di tecniche di *record-linkage*. In questo lavoro verranno presentati alcuni primi risultati per mettere in luce alcune peculiari potenzialità di analisi sottese all'utilizzo di dati che consentono di seguire nel tempo gli individui, cioè le unità statistiche oggetto della rilevazione. In particolare l'attenzione si è focalizzata sul fenomeno della mobilità interna degli stranieri non comunitari, argomento che è stato già approfondito in letteratura (Casacchia *et al.*, 1999; Pugliese, 2006; Casacchia *et al.*, 2010; De Santis, 2010;), ma solitamente attraverso l'utilizzo di dati aggregati e con indicatori tradizionali di tipo trasversale.

2. Il *record-linkage* dei permessi relativi alla coorte degli immigrati del 2007

A livello di sperimentazione, si è deciso di considerare gli archivi dei permessi di soggiorno in corso di validità all'inizio del 2008 e del 2010, relativi pertanto ad un intervallo breve (due anni) e successivo alle novità intervenute nel 2007. Infatti, dall'inizio di tale anno la Romania e la Bulgaria sono entrate a far parte dell'Unione europea (Ue), passata da 25 a 27 paesi, e l'11 aprile in Italia è entrato in vigore il Decreto legislativo n. 30, che ha recepito la Direttiva 2004/38/CE del Parlamento europeo relativa al diritto dei cittadini dell'Unione e dei loro familiari di circolare e soggiornare liberamente sul territorio degli Stati membri. Si tratta di novità la cui importanza appare evidente in molteplici contesti, che hanno prodotto però variazioni significative anche nell'ambito delle statistiche sulla presenza straniera in Italia. Con riguardo specifico alla rilevazione dei permessi di soggiorno dall'inizio del 2008 si è determinata una cesura rispetto al passato poiché è cambiato l'universo di riferimento che non comprende più i cittadini dei 27 paesi dell'Ue. Pertanto, le analisi che verranno proposte riguardano esclusivamente i cittadini non comunitari titolari di permesso di soggiorno. In particolare, si è concentrata l'attenzione sul collettivo di quelli che hanno ottenuto nel corso del 2007 il loro primo permesso di soggiorno (oltre 250.000 stranieri non comunitari), con una qualche forzatura assimilabile alla coorte degli immigrati di tale anno².

Il ricorso al *record-linkage* tra l'archivio di inizio 2008 e quello di inizio 2010 dovrebbe consentire di determinare quanti degli immigrati non comunitari della coorte del 2007 sono ancora regolarmente presenti in Italia a distanza di due anni e, per questi ultimi, se si sono verificate variazioni significative con riguardo ad alcune informazioni disponibili nella rilevazione amministrativa (in particolare,

² Sono state considerate solo le persone con permesso di soggiorno individuale e non i minori iscritti sul permesso di un adulto per i quali non è al momento possibile avere l'informazione relativa al codice fiscale.

stato civile, provincia di rilascio e motivo del permesso). Quella proposta è una sperimentazione ancora non completata che per il momento ha riguardato esclusivamente i cittadini non comunitari per i quali si disponeva del codice fiscale nell'archivio di inizio 2008 (204.821 casi, pari all'81% della coorte di immigrati del 2007) che si è cercato di individuare nell'archivio di inizio 2010 attraverso il ricorso ad una procedura standard di *record-linkage* deterministico, utilizzando come chiave di abbinamento il codice fiscale. Si sono abbinati 162.604 record, pari al 79,4% di quelli per i quali si disponeva del codice fiscale all'inizio del 2008. In altri termini, in media circa 4 ogni 5 extracomunitari arrivati nel 2007 sono risultati ancora regolarmente presenti in Italia all'inizio del 2010. Si tratta di un risultato tutto sommato in linea con quanto registrato in precedenti sperimentazioni (Carfagna *et al.*, 2010), anche se converrà approfondire l'analisi sul collettivo dei permessi che non si sono abbinati per individuare i casi, da ritenersi in vero poco numerosi, dovuti a errori nell'indicazione del codice fiscale o al suo mancato inserimento nell'archivio di inizio 2010 (mancati abbinamenti). Inoltre, andrà estesa l'applicazione agli extracomunitari per i quali non si dispone del codice fiscale ad inizio 2008. In questi casi bisognerà determinare una chiave di abbinamento sulla base delle informazioni invariabili nel tempo disponibili nei due archivi e valutare l'adozione di eventuali procedure alternative di collegamento come il *record-linkage* probabilistico.

Il confronto tra gli non comunitari ancora presenti ad inizio 2010 e quelli che dovrebbero aver lasciato il paese o quantomeno non hanno più un permesso valido consente di notare come tra questi ultimi risultano sovra-rappresentati gli uomini e i permessi concessi per motivi diversi da lavoro e famiglia. Si tratta perlopiù di permessi per studio che si può presumere siano scaduti, senza essere stati rinnovati nel periodo considerato. Leggere differenze si riscontrano anche nella graduatoria delle cittadinanze tra gli abbinati e i non abbinati: tra questi ultimi sono meno rappresentati gli europei e hanno, invece, rilevanza maggiore asiatici e americani.

Per il sottoinsieme degli immigrati nel 2007 ancora presenti alla data più recente appare interessante, tra gli altri aspetti, analizzare la mobilità territoriale attraverso l'esame della variazione della provincia di rilascio del permesso tra l'inizio del 2008 e del 2010, cioè nel biennio 2008-09.

3. La mobilità degli stranieri non comunitari in Italia

Gli stranieri contribuiscono in maniera consistente alla mobilità interna in Italia (Casacchia *et al.*, 2010; De Santis, 2010). Gli immigrati hanno infatti, come è noto, una propensione alla mobilità più elevata degli autoctoni (de Filippo, Strozza, 2011). Nel nostro Paese possono essere facilmente individuati i principali poli di attrazione delle

migrazioni interne di stranieri, non necessariamente coincidenti con le aree che richiamano maggiormente le migrazioni dall'estero. I principali studi sul tema realizzati fino ad oggi sono stati portati avanti utilizzando macrodati, relativi soprattutto ai cambiamenti di residenza, attraverso l'utilizzo di indicatori classici di mobilità.

L'utilizzo di micro-dati dei permessi di soggiorno riferiti ad archivi di anni differenti collegati attraverso *record-linkage* consente di superare questo tipo di approccio e di individuare i "reali" percorsi degli immigrati sul territorio, anche in connessione con le caratteristiche degli individui³. In primo luogo è possibile mettere in luce quali siano le province italiane che riescono a "trattenere" sul territorio gli immigrati e quali li attirino una volta che sono giunti nel nostro Paese. È noto, infatti che alcuni territori rappresentano solo una porta di accesso al territorio italiano senza poi divenire area di insediamento effettivo per gli stranieri.

Della coorte di persone entrate in Italia nel 2007 circa il 18% ha un permesso rinnovato in una provincia diversa da quella di primo rilascio⁴. Le province del Mezzogiorno hanno evidentemente capacità di trattenimento molto più basse rispetto al Nord: nel Nord-ovest dopo due anni ha cambiato provincia il 15% degli entrati in Italia durante il 2007; nel Mezzogiorno coloro che hanno lasciato la provincia di rilascio del primo permesso sono oltre il 28%.

La provincia che perde di più è Crotone che all'inizio del 2010 ha trattenuto poco più del 26% dei cittadini non comunitari che hanno ottenuto un permesso nella provincia durante il 2007. Tra le province con maggiore capacità di trattenimento degli stranieri che hanno fatto ingresso in Italia nel 2007 troviamo Genova, Parma, Trento, Roma, Torino e Firenze che perdono meno del 13% dei soggiornanti che hanno avuto un nuovo permesso nella provincia nel 2007.

Roma non solo trattiene gli immigrati provenienti dall'estero, ma attira anche stranieri non comunitari che si spostano all'interno del Paese: oltre il 46% della coorte di entrati in Italia nel 2007 si è trasferita a Roma nel biennio considerato, dopo essere entrata per la prima volta in Italia in un'altra provincia. Ancora più forte la capacità attrattiva di Prato sui flussi di migrazioni interne della coorte di ingressi del 2007. Oltre il 63% di coloro che hanno un permesso rinnovato a Prato tra il 2007 e il 2009 sono entrati in Italia registrando il primo permesso di soggiorno in una provincia diversa.

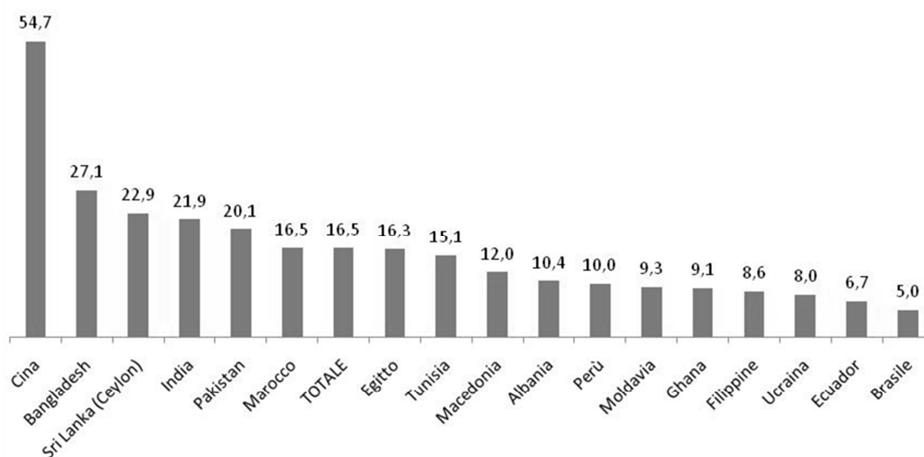
Molto differente la mobilità che ha interessato le diverse cittadinanze che compongono la coorte in osservazione. Tra quelle più rappresentate risulta

³ L'informazione utilizzata è relativa alla provincia in cui si è richiesto il permesso di soggiorno nel 2007 e quella di rinnovo registrata in archivio all'inizio del 2010. Non necessariamente coincide con la provincia di residenza.

⁴ Sono stati considerati nell'analisi solo coloro ai quali nel 2007 era stato concesso un permesso con scadenza precedente al 1° gennaio 2010, in modo da considerare solo coloro che hanno dovuto rinnovare il permesso nell'arco di tempo considerato.

particolarmente elevata la quota di cinesi spostatisi durante il biennio considerato: quasi il 55% della collettività ha rinnovato il permesso in una provincia diversa da quella di primo soggiorno. Oltre allo Sri Lanka (23% di trasferiti), particolarmente mobili risultano anche le collettività del subcontinente indiano: Bangladesh (27% di trasferiti), India (22%) e Pakistan (20%). Le collettività dell'Est Europa fanno registrare, invece, livelli di mobilità inferiori alla media (intorno al 18%): nel caso di Russia e Moldavia la quota di trasferiti è pari a circa il 9%, per l'Ucraina all'8%. A un livello intermedio si collocano le cittadinanze dell'Africa maghrebina e quelle della regione balcanica (Figura 1).

Figura 1 – Quota di cittadini non comunitari che hanno al 1° gennaio 2010 un permesso di soggiorno rinnovato in una provincia diversa da quella del primo rilascio nel 2007 per cittadinanza.



Fonte: elaborazioni Istat su dati del Ministero dell'Interno, dati provvisori.

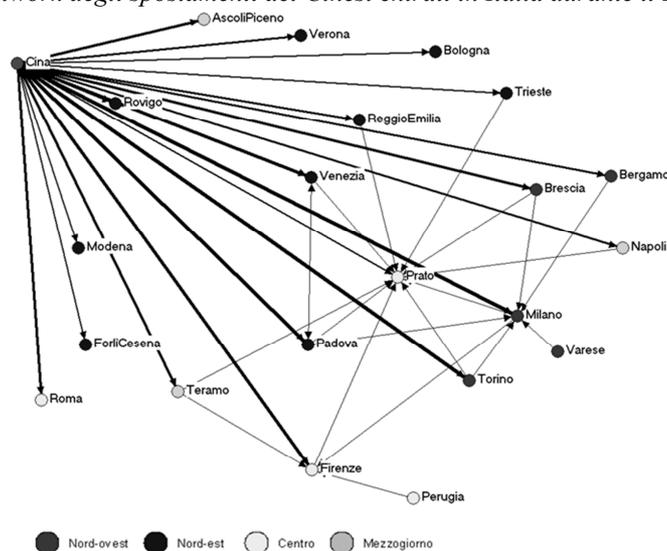
Le varie collettività hanno notoriamente caratteristiche demografiche e sociali differenti che possono influire sulle scelte di mobilità interna. È stata quindi realizzata un'analisi di regressione logistica per comprendere il legame della propensione alla mobilità (variabile risposta aver cambiato o meno la provincia di rilascio del permesso) con la cittadinanza e le principali caratteristiche individuali (sesso, età, cambiamento di stato civile tra il 2008 e il 2010, cambiamento di

motivo del permesso tra 2008 e il 2010)⁵ e migratorie (motivo del permesso, ripartizione territoriale di primo rilascio).

Anche a parità di altre caratteristiche come lo stato civile, la classe di età, il motivo del soggiorno e la ripartizione geografica di ingresso, le cittadinanze per le quali l'analisi descrittiva aveva messo in luce una maggiore mobilità sul territorio italiano - quelle del subcontinente indiano e soprattutto la Cina - confermano una più elevata propensione allo spostamento.

I modelli di regressione mettono inoltre in evidenza un legame tra i diversi cambiamenti di "status", coloro che hanno cambiato stato civile nel periodo considerato mostrano una maggiore propensione allo spostamento, così come coloro che hanno cambiato il motivo del permesso.

Figura 2 – Network degli spostamenti dei Cinesi entrati in Italia durante il 2007.



Fonte: elaborazioni Istat su dati del Ministero dell'Interno, dati provvisori.

Le diverse collettività danno luogo a modelli insediativi assai differenti sul territorio italiano con livelli di concentrazione peculiari. Il quadro che si coglie attraverso i dati trasversali di stock è il risultato della sedimentazione sul territorio sia di flussi provenienti dall'estero, sia delle dinamiche migratorie interne. La lettura dei dati elaborati con il *record-linkage*, ricostruendo i percorsi dei migranti

⁵ Per quanto riguarda la cittadinanza sono state considerate separatamente le prime otto cittadinanze per numero di primi rilasci nel 2007: Albania, Marocco, Moldavia, Ucraina, Cina, Bangladesh, India, Sri Lanka (Ceylon). Le rimanenti cittadinanze sono state raggruppate nella categoria residuale altro.

dal momento del loro ingresso in Italia consente di collegare la migrazione internazionale agli spostamenti interni. Nel caso della collettività più mobile all'interno della coorte degli immigrati nel 2007, quella cinese, i percorsi seguiti appaiono di grande interesse. Il grafo, nel quale lo spessore delle linee indica la rilevanza degli spostamenti e le frecce l'orientamento del flusso, mette in evidenza, come nelle attese, al centro del network degli spostamenti Prato, provincia che da sempre registra una quota notevole di presenza cinese (Figura 2).

Tuttavia, emerge che in realtà Prato è centrale rispetto ai flussi migratori interni e meno rilevante per quanto riguarda gli ingressi dall'estero della coorte del 2007. È interessante anche la posizione di Milano che accoglie un flusso rilevante dall'estero, ma anche dall'interno, specialmente dalle altre province del Nord-ovest. Numerose sono invece le province che hanno accolto nel 2007 rilevanti flussi dalla Cina e che nel biennio successivo hanno ridistribuito verso altre province i cittadini della coorte osservata. Di particolare interesse è, ad esempio, il caso di Napoli che è un importante polo di attrazione dei flussi internazionali di cittadini cinesi, ma che nei due anni seguenti al loro arrivo registra una perdita netta consistente con la provincia di Prato.

Il raffronto con i network territoriali di altre collettività, come il Bangladesh e l'India mettono, in evidenza dinamiche completamente differenti che spingono ancora una volta ad osservare separatamente le differenti collettività.

4. Conclusioni

L'utilizzo di tecniche di *record-linkage* applicate ad archivi ricchi di informazioni come quello dei permessi di soggiorno può rappresentare un passo fondamentale nello studio di fenomeni complessi e processuali come quello dell'integrazione della popolazione straniera in Italia.

Naturalmente il passo successivo, attualmente in corso di sviluppo, è il collegamento dell'archivio dei permessi con altri archivi per un ulteriore arricchimento delle informazioni e degli aspetti che è possibile indagare anche attraverso la costruzione di indicatori differenti da quelli classici di tipo trasversale.

Tuttavia anche per tematiche ormai approfondite in letteratura come quella della mobilità interna degli stranieri in Italia, l'ottica "longitudinale" consente modalità di osservazione differenti che gettano nuova luce su aspetti del fenomeno che fino ad oggi restavano in ombra.

Riferimenti bibliografici

- CASACCHIA O., NATALE L., STROZZA S. 1999. Migrazioni interne e migrazioni internazionali: il nuovo ruolo del Mezzogiorno nel sistema migratorio nazionale. In BONIFAZI C. (a cura di), *Mezzogiorno e migrazioni interne*, Monografie No.10, Roma: IRP-CNR, pp. 237-272.
- CASACCHIA O., CASSATA L., GIORGI P., REYNAUD C., STROZZA S. 2010. La mobilità interna in Italia: le principali caratteristiche dei cambiamenti di residenza, *Rivista Italiana di Economia Demografia e Statistica*, Vol. LXIV, No. 1-2, pp. 63-70.
- CONTI C., REYNAUD C. 2011. Mobilità della popolazione, immigrazione e presenza straniera. In DE ROSE A., SALVINI S. (a cura di) *Rapporto sulla popolazione. L'Italia a 150 anni dall'Unità*, Bologna: il Mulino, pp. 97-118.
- DE FILIPPO E., STROZZA S. 2011. Le migrazioni interne degli stranieri in Italia, *Sociologia del lavoro*, No. 121, pp. 168-195.
- DE SANTIS G. 2010. Mobilità a corto e lungo raggio e pendolarismo della popolazione italiana. In LIVI BACCI M. (a cura di) *Demografia del capitale umano*, Bologna: il Mulino, pp. 123-138.
- ISTAT 1998. *La presenza straniera in Italia negli anni '90*. Informazioni No. 61, Roma.
- ISTAT 2009. *Rapporto annuale. La situazione del Paese nel 2008*, Roma.
- PUGLIESE E. (2006). *L'Italia tra migrazioni internazionali e migrazioni interne*. Bologna: il Mulino.
- RICCIARDO LAMONICA G., ZAGAGLIA B. 2008. I flussi migratori delle regioni adriatiche: interazioni e determinanti. In MORETTI E. (a cura di) *Lungo le sponde dell'Adriatico*, Milano: FrancoAngeli, pp. 125-153.
- ROSSI F., STROZZA S. 2007. Mobilità della popolazione, immigrazione e presenza straniera. In GCD-SIS, *Rapporto sulla popolazione. L'Italia all'inizio del XXI secolo*, Bologna: il Mulino, pp. 111-137.

SUMMARY

Administrative data sources are already widely used for migration statistics, but it seems possible, by developing statistical harmonization and integration strategies, to make a better use of the collected data.

At present Istat has carried out a first attempt of longitudinal integration through the record linkage of residence permits databases. The principal results of this experience, are discussed in the paper. The focus is on internal mobility of No-EU citizens.

Cinzia CONTI, Ricercatore, Istat, ciconti@istat.it

Domenico GABRIELLI, Primo Ricercatore, Istat, gabriell@istat.it

Salvatore STROZZA, Ordinario di Demografia, Un. di Napoli Federico II, strozza@unina.it

POPOLAZIONE A RISCHIO DI POVERTÀ E INTERVENTO PUBBLICO: CONFRONTI TRA PAESI EUROPEI¹

Andrea Cutillo, Daniela Lo Castro, Isabella Siciliani

1. Introduzione

Questo lavoro rappresenta lo sviluppo di un precedente articolo degli autori (Cutillo, Lo Castro e Siciliani, 2009) in cui si analizzava l'impatto dei trasferimenti sociali sulla riduzione del rischio di povertà delle famiglie con minori in diversi Paesi europei tramite un modello gerarchico a effetti casuali basato sui microdati del sistema Eu-silc (*European Statistics on Income and Living Conditions*). Il presente lavoro amplia la prospettiva, analizzando l'effetto complessivo delle politiche redistributive del reddito dei diversi Paesi europei sulla riduzione dell'incidenza di rischio di povertà dell'intera popolazione determinatasi a seguito della distribuzione del reddito primario tra i fattori produttivi, attraverso quindi la considerazione congiunta dei trasferimenti sociali e dell'imposizione fiscale. Si continua ad adottare una prospettiva che combina aspetti micro e macro già usata in altri lavori sull'analisi della riduzione del rischio di povertà (Lohmann, 2008).

Oggetto di studio è la probabilità di uscita dalla condizione di rischio di povertà² a seguito del pagamento di tasse e contributi sociali allo Stato e, in direzione opposta, dell'integrazione del reddito con i trasferimenti sociali pubblici di tipo monetario. Sono presi in considerazione i dati di 25 paesi partecipanti ad Eu-silc (EU27 con l'eccezione di Germania, Francia e Slovenia, e la Norvegia).

Nel valutare l'effetto redistributivo, in questo lavoro si adottano delle semplificazioni rispetto alle definizioni adoperate in Contabilità Nazionale. Il reddito disponibile delle famiglie - dopo la redistribuzione - è dato dalla somma del reddito primario, i trasferimenti privati lordi e le componenti per intervento pubblico: in positivo, le pensioni dirette ed indirette e i trasferimenti sociali; in negativo, le imposte su redditi e patrimonio ed i contributi sociali.

¹ Sebbene il lavoro sia frutto delle riflessioni comuni degli autori, il paragrafo 1 è da attribuire a I. Siciliani, il paragrafo 2 a A. Cutillo e i paragrafi 3 e 4 a D. Lo Castro.

² La soglia di povertà è pari al 60 per cento della mediana del reddito equivalente sulla base della scala OCSE modificata.

Per reddito primario si intende invece quello che si determina per effetto dei meccanismi di mercato e cioè per la distribuzione del reddito tra i fattori produttivi, lavoro e capitale. Esso comprende quindi i redditi da lavoro dipendente lordi e il risultato lordo di gestione (cioè, un coacervo di redditi provenienti dalla capacità imprenditoriale e dal possesso dei mezzi produttivi). All'interno dei trasferimenti assistenziali si considerano i sussidi per disoccupazione, per invalidità, le indennità di malattia, per combattere l'esclusione sociale, i sussidi per l'istruzione e per l'abitazione, gli assegni familiari e comunque i trasferimenti monetari che mirano a combattere altre forme di esclusione sociale. A differenza di quanto accade in Contabilità nazionale, in questa nota le pensioni, dirette ed indirette, vengono considerate tra i redditi primari, quindi tra i redditi lordi, e non tra i trasferimenti sociali dello Stato. L'operazione trova giustificazione nel fatto che le pensioni erogate a seguito di attività lavorativa possono essere considerate una forma di reddito da lavoro differito nel tempo. In questa logica, tuttavia, dovrebbero essere escluse le pensioni non legate ad alcuna attività lavorativa, come ad esempio le pensioni sociali, che sono a tutti gli effetti trasferimenti assistenziali. Purtroppo la base informativa utilizzata non consente di distinguere tra i diversi tipi di pensione, ma è indubbio che le pensioni sociali rappresentano una parte marginale dei trasferimenti pensionistici (in Italia, ad esempio sono circa il 3 per cento).

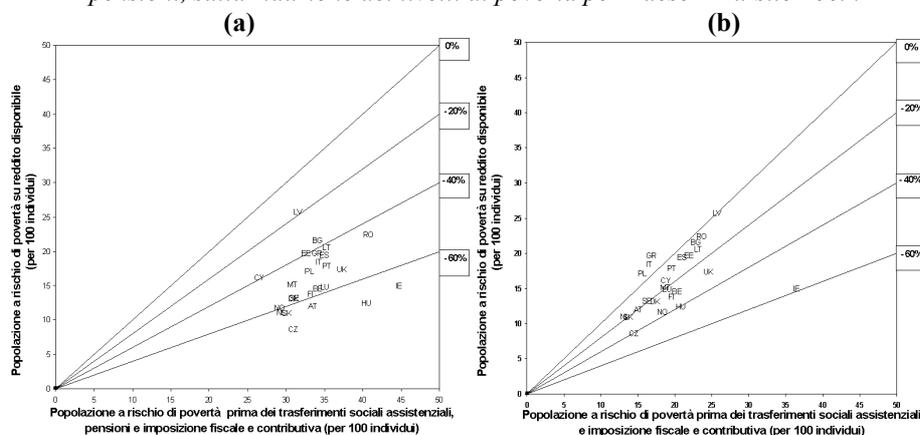
In Figura 1 (a) e (b) sono raffigurati gli effetti redistributivi del sistema di trasferimenti e tasse nei diversi Paesi in termini di uscita dal rischio di povertà, considerando le pensioni tra i trasferimenti sociali o all'interno del reddito primario rispettivamente. L'effetto delle pensioni è maggiore nel primo caso: in particolare per l'Italia si registra una riduzione del livello di povertà oltre il 40 per cento; quando invece si considera il solo effetto dei trasferimenti sociali assistenziali e dell'imposizione fiscale e contributiva nel nostro Paese l'incidenza di povertà è addirittura superiore (analogamente a Grecia e Polonia). Difatti, in Italia, il rapporto tra il reddito prima e dopo i trasferimenti assistenziali, l'imposizione fiscale e contributiva, è inferiore ad uno solo nel primo decile di reddito equivalente, mentre per il resto d'Europa questo accade anche per il secondo decile.

2. Il modello e le variabili

Il modello utilizzato in questo lavoro è un modello gerarchico con intercetta e pendenze casuali che permette di considerare la struttura gerarchica dei dati oggetto di studio, con le unità di primo livello raggruppate in unità di secondo livello, che quindi rappresentano il contesto (in questo caso, territoriale).

Il parametro di interesse che si vuole modellare è la probabilità di uscita dal rischio di povertà, ovvero $P_{ij} = Pr(Y_{ij}=1)$, dove Y_{ij} è pari a 1 se la famiglia i del Paese j , povera in base al reddito primario, non lo è più quando si aggiungono i trasferimenti assistenziali e si sottraggono le imposte e 0 altrimenti, ed è funzione sia di caratteristiche familiari che del Paese di appartenenza. Il collettivo di analisi è quindi costituito dalle sole famiglie esposte al rischio di povertà prima dell'intervento redistributivo.

Figura 1 – Effetto redistributivo del sistema di trasferimenti e tasse, con (a) e senza (b) pensioni, sulla riduzione dei livelli di povertà per Paese – Eu-silc 2009.



Il modello gerarchico di regressione logistica a due livelli, ad intercetta e pendenze casuali, può essere scritto nel seguente modo:

$$\text{Logit}(P_{ij}) = \beta_0 + u_{0j} + \beta_k K_{ij} + u_{wj} W_{ij} + \beta_r R_j \quad (1)$$

Per approfondimenti³ si rimanda a Gelman e Hill (2007) e Goldstein (1995).

Le variabili esplicative di primo livello sono: tipologia familiare (nelle modalità single, 2 adulti < 65 anni, 2 adulti di cui almeno uno 65+ anni, monogenitore, 2 adulti con 1, 2, 3+ figli a carico, altra tipologia); intensità lavorativa della famiglia (data dal rapporto tra numero di mesi in occupazione e numero di mesi destinabili all'occupazione); fonte di reddito principale (dipendente, autonomo, altro tipo);

³ Il vettore di variabili K_{ij} rappresenta le esplicative di primo livello (con $W \subseteq K$), il vettore R_j quelle di secondo; i parametri β sono i relativi coefficienti, e rappresentano gli effetti fissi del modello, costanti tra i Paesi; u_{0j} e u_{wj} sono la caratteristica distintiva dei modelli gerarchici, e rappresentano i residui di secondo livello (effetti casuali) associati all'intercetta e alle pendenze delle variabili esplicative; essendo liberi di variare tra i gruppi, rappresentano gli effetti dovuti al Paese di appartenenza.

numero di percettori; cittadinanza del principale percettore di reddito in famiglia, il *breadwinner*; livello di istruzione del *breadwinner* (basso, medio, alto); età del *breadwinner*; trasferimenti assistenziali ricevuti e tasse/imposte pagate. Per queste ultime due informazioni, che rappresentano le variabili di maggiore interesse per questo lavoro, sono anche considerati gli effetti di pendenza casuale. Inoltre, per renderli comparabili tra i diversi Paesi, il valore è espresso in percentuale delle diverse linee di povertà, assunte quali grandezze indicative dello standard di vita presente nel Paese.

Come variabili di secondo livello relative al contesto/Paese, sono state utilizzate: quota di spesa sociale su PIL; incidenza di famiglie a rischio di povertà e intensità di povertà prima della redistribuzione; quota di beneficiari di trasferimenti assistenziali; deviazione standard dei trasferimenti assistenziali.

Per valutare l'effetto delle diverse politiche il valore di $Logit(P_{ij})$ è stato stimato a livello di Paese e decomposto isolando le variabili oggetto di decisioni in termini **quantitativi** (trasferimenti (T), tasse (TA), spesa sociale (S)) e **distributivi** (deviazione standard dei trasferimenti (D), quota di famiglie con trasferimenti (P)):

$$Logit(P_{ij}) = F_j + u_{0j} + Q_j + E_j \quad (2)$$

$$\text{con } F_j = \beta_0 + \beta_x X_j + \beta_z Z_j + \beta_t T_m + \beta_{ta} TA_m + \beta_s S_m + \beta_d D_m + \beta_p P_m, \quad (3)$$

$$Q_j = \beta_t (T_j - T_m) + \beta_{ta} (TA_j - TA_m) + \beta_s (S_j - S_m) + u_{tj} T_j + u_{taj} TA_j, \quad (4)$$

$$\text{e } E_j = \beta_d (D_j - D_m) + \beta_p (P_j - P_m) \quad (5)$$

dove le variabili con pedice j sono espresse come media di Paese e quelle con pedice m sono espresse come media dei Paesi europei complessivamente considerati. \mathbf{X} ($\mathbf{X} \subset \mathbf{K}$) e \mathbf{Z} ($\mathbf{Z} \subset \mathbf{R}$) sono i vettori delle restanti variabili di primo e secondo livello. F_j indica quindi un effetto Base, cioè cosa accadrebbe senza considerare effetti casuali e se tutti i Paesi attuassero la medesima politica. Q_j e E_j indicano invece l'effetto aggiuntivo delle diverse politiche dei Paesi in termini quantitativi e distributivi. La somma di queste due quantità più l'effetto di intercetta casuale (u_{0j}) fornisce l'effetto Paese complessivo.

3. Applicazione e risultati

In Tabella 1 sono presenti i risultati del modello multilevel, in termini di effetti fissi. Le famiglie non composte da almeno 2 adulti con meno di 65 anni presentano maggiore probabilità di uscita dal rischio di povertà. Più è alta l'intensità lavorativa della famiglia e maggiore il numero di percettori in famiglia, maggiori sono le probabilità di uscita: queste famiglie, anche se esposte a rischio di povertà prima della redistribuzione, presentano generalmente condizioni meno critiche delle altre;

inoltre, è probabile che molti degli importi ricevuti siano funzione della propria condizione di occupati. Come atteso, le famiglie con reddito prevalentemente autonomo hanno una minore probabilità di uscita rispetto a quelle con fonte principale da lavoro dipendente: l'ipotesi è infatti che i lavoratori autonomi accedano più difficilmente ai trasferimenti per il rischio di essere sottoposti a controlli fiscali al momento della richiesta di sussidi legati a situazioni reddituali disagiate. I risultati del modello supportano anche l'idea che le famiglie con stranieri e quelle con basso titolo di studio del *breadwinner* abbiano una più bassa probabilità di uscita a causa di una probabile minore conoscenza delle opportunità offerte dallo stato sociale e delle procedure amministrative per accedere ai sussidi.

Tabella 1 – Modello multilevel sulla probabilità di uscita dalla povertà per i trasferimenti sociali assistenziali e il prelievo fiscale.

| Effetti Fissi | Coef. | Std.Error | Stat.Z | Pr(> Z) |
|-------------------------------------------------------------|----------|-----------|------------|--------------|
| Intercetta | -7.085 | 0.788 | -8.992 | < 2e-16 *** |
| Tipologia familiare (Rif=Due adulti con due figli a carico) | | | | |
| <i>Single</i> | 0.136 | 0.072 | 1.892 | 0.05851 * |
| <i>Due adulti < 65 anni</i> | -0.218 | 0.076 | -2.870 | 0.0041 ** |
| <i>Due adulti di cui almeno uno 65+ anni</i> | 0.729 | 0.088 | 8.298 | < 2e-16 *** |
| <i>Famiglia monogenitore</i> | 0.308 | 0.078 | 3.950 | 7.82E-05 *** |
| <i>Due adulti con un figlio a carico</i> | -0.161 | 0.081 | -1.982 | 0.04744 * |
| <i>Due adulti con tre o più figli a carico</i> | -0.229 | 0.084 | -2.713 | 6.66E-03 ** |
| <i>Altre famiglie</i> | -0.323 | 0.079 | -4.088 | 4.34E-05 *** |
| Intensità lavorativa (Rif=1) | | | | |
| <i>Intensità lavorativa = 0</i> | -0.482 | 0.063 | -7.623 | 2.49E-14 *** |
| <i>Intensità lavorativa in (0; 0.5]</i> | 0.088 | 0.074 | 1.180 | 0.23792 |
| <i>Intensità lavorativa in (0.5; 1)</i> | 0.282 | 0.064 | 4.422 | 9.76E-06 *** |
| Fonte principale di reddito (Rif=Lavoro dipendente) | | | | |
| <i>Lavoro autonomo</i> | -1.227 | 0.095 | -12.848 | < 2e-16 *** |
| <i>Altra fonte</i> | -1.583 | 0.058 | -27.231 | < 2e-16 *** |
| Numero di percettori | 0.717 | 0.030 | 23.728 | < 2e-16 *** |
| Cittadinanza straniera del <i>breadwinner</i> | -0.125 | 0.072 | -1.745 | 0.08099 * |
| Istruzione del <i>breadwinner</i> (Rif=Isced fino a 2) | | | | |
| <i>Isced 3 e 4</i> | 0.115 | 0.038 | 2.988 | 0.00281 ** |
| <i>Isced 5 e 6</i> | 0.147 | 0.060 | 2.432 | 0.01503 * |
| Età del <i>breadwinner</i> | 0.040 | 0.001 | 27.532 | < 2e-16 *** |
| Trasferimenti assistenziali | 0.057 | 0.003 | 18.020 | < 2e-16 *** |
| Tasse | 0.005 | 0.007 | 0.653 | 0.51385 |
| Quota spesa sociale su Pil | 0.024 | 0.032 | 0.753 | 0.45124 |
| Deviazione Standard trasferimenti | 0.009 | 0.028 | 0.324 | 0.74557 |
| Quota beneficiari trasferimenti | 0.008 | 0.007 | 1.294 | 0.19551 |
| Intensità di povertà prima redistribuzione | -0.145 | 0.06 | -2.424 | 0.01535 * |
| Incidenza famiglie a rischio povertà prima redistribuzione | 0.104 | 0.037 | 2.802 | 0.00508 ** |
| Effetti Casuali | | | | |
| | Varianza | Std.Dev. | Corr. | |
| Intercetta | 1.0314 | 1.0156 | | |
| Trasferimenti assistenziali | 0.0002 | 0.0148 | -0.7290 | |
| Tasse | 0.0009 | 0.0305 | -0.5270 | -0.1270 |
| Statistiche | | | | |
| | AIC | BIC | logVerosim | Devianza |
| | 25048 | 25312 | -12493 | 24986 |

Num osservazioni: 37661; Num. gruppi: 25; Significatività a livello: 0 **** 0.001 *** 0.01 ** 0.05 * 0.1 *

Mentre le tasse non sono significative, l'altra esplicativa di maggiore interesse, l'entità dei trasferimenti ricevuti, presenta, in termini statisticamente significativi, il logico effetto positivo atteso: torneremo successivamente sugli effetti casuali di Paese relativi a queste variabili.

Tra le variabili di secondo livello, solo l'intensità di povertà prima della redistribuzione presenta il segno negativo atteso in quanto maggiore è l'intensità e più alti dovranno essere i trasferimenti. L'incidenza di famiglie a rischio di povertà pre-redistribuzione ha invece un effetto significativamente positivo: probabilmente proprio la maggiore incidenza induce ad adottare misure politiche di lotta alla povertà più incisive che quindi ne favoriscono una maggiore uscita.

Le altre variabili di secondo livello sono quelle oggetto di decisioni politiche: dove maggiore è la quota di spesa sociale tanto più facile dovrebbe essere uscire dal rischio di povertà. La percentuale di famiglie che ricevono trasferimenti sociali può invece cogliere l'effetto di politiche diverse, con trasferimenti più o meno estesi sulla popolazione, mentre la deviazione standard dei trasferimenti ricevuti dalle famiglie a rischio di povertà prima della redistribuzione, può consentire di verificare quanto la variabilità nei trasferimenti erogati alle famiglie a rischio di povertà possa influire sull'uscita. Nessuna di queste risulta però significativa.

In Tabella 2 vengono riportate la probabilità stimata dal modello e gli effetti per ciascun Paese, decomposti nelle componenti Base, Quantità e Distribuzione individuate nelle formule (3), (4) e (5).

Complessivamente l'Italia è tra i paesi con la minore probabilità stimata di uscita dal rischio di povertà in seguito alla redistribuzione del reddito insieme a Lettonia e Bulgaria (colonna 2). L'effetto Paese è più basso nei Paesi mediterranei, in quelli baltici e in Romania e Bulgaria mentre è più elevato tra i paesi scandinavi, Irlanda, Olanda, Belgio, Regno Unito, Ungheria e Repubblica Ceca (colonna 4).

Si osserva un ruolo preponderante dell'effetto Quantità (Q_j) rispetto a quello di Distribuzione (E_j). I Paesi con effetto Quantità più elevato in termini positivi sono i Paesi Scandinavi, Ungheria e Olanda. All'estremo opposto si trovano invece Grecia, Lettonia, Cipro e Bulgaria.

Infine, gli u_{0j} incorporano le diverse politiche distributive in termini di famiglie beneficiarie e la diversa composizione delle esplicative all'interno dei Paesi. Essendo data questa composizione, i Paesi con intercetta casuale più negativa (Estonia, Italia, Lettonia, Polonia) dovrebbero variare qualcosa nei requisiti di accesso.

In Tabella 2 sono anche presentati gli effetti casuali u_{ij} e $u_{(aj)}$ sulla pendenza rispettivamente dei trasferimenti alle famiglie e dell'imposizione fiscale. Questi, essendo lo scarto di Paese rispetto all'effetto fisso, possono essere interpretati come una sorta di efficacia rispetto all'uscita dal rischio di povertà per ogni trasferimento e tassazione aggiuntivi. In particolare, se consideriamo la pendenza casuale u_{ij} , i

Paesi a maggiore efficacia sono generalmente quelli con un peggiore effetto Quantità, cioè l'Italia seguita da Grecia, Lituania e Estonia. I Paesi a minore efficacia sono invece alcuni tra quelli con effetto Quantità maggiormente positivo (Paesi Scandinavi, Olanda, Belgio): questo fatto è probabilmente dovuto ad un'utilità marginale generalmente decrescente per ogni trasferimento aggiuntivo.

Tabella 2 – Probabilità stimata dal modello, Effetti Base, Paese, Intercetta, Quantità e Distribuzione, Pendenze casuali di trasferimenti e tasse.

| PAESE | Probabilità stimata | Effetto Base | Effetto Paese | di cui | | | U _t | U _{ta} |
|-------|---------------------|--------------|---------------|----------------|------------------|-----------------------|----------------|-----------------|
| | | | | U ₀ | Effetto Quantità | Effetto Distribuzione | | |
| AT | 0.310 | -1.735 | 0.714 | 0.311 | 0.364 | 0.039 | -0.006 | 0.005 |
| BE | 0.381 | -2.084 | 1.405 | 0.895 * | 0.425 | 0.085 | -0.013 * | -0.013 * |
| BG | 0.102 | -0.641 | -2.485 | -1.083 * | -1.331 | -0.071 | 0.008 | 0.029 * |
| CY | 0.167 | -0.423 | -1.716 | -0.444 | -1.364 | 0.092 | 0.007 | 0.001 |
| CZ | 0.421 | -1.602 | 1.063 | 1.034 * | 0.114 | -0.085 | -0.003 | -0.015 |
| DK | 0.580 | -3.136 | 3.176 | 0.173 | 2.833 | 0.170 | -0.015 * | 0.018 * |
| EE | 0.159 | -0.824 | -2.341 | -1.358 * | -1.028 | 0.045 | 0.012 | 0.025 |
| ES | 0.173 | -1.221 | -1.034 | 0.084 | -0.918 | -0.200 | 0.004 * | -0.013 * |
| FI | 0.390 | -1.843 | 1.005 | 0.343 | 0.579 | 0.083 | -0.023 * | 0.051 * |
| GR | 0.166 | -1.101 | -1.662 | 0.375 * | -1.740 | -0.297 | 0.026 * | -0.088 * |
| HU | 0.446 | -1.305 | 1.078 | -0.111 | 1.108 | 0.081 | 0.001 | 0.005 |
| IE | 0.534 | -1.913 | 2.250 | 2.414 * | -0.391 | 0.227 | -0.028 * | -0.024 * |
| IT | 0.108 | -1.221 | -2.225 | -1.181 * | -1.001 | -0.043 | 0.026 * | 0.013 * |
| LT | 0.176 | -1.063 | -1.694 | -1.049 * | -0.734 | 0.089 | 0.009 | 0.015 |
| LU | 0.370 | -1.279 | 0.671 | -0.100 | 0.771 | -0.001 | -0.003 | 0.007 |
| LV | 0.064 | -0.673 | -3.067 | -1.741 * | -1.392 | 0.066 | 0.023 * | 0.011 |
| MT | 0.248 | -1.426 | -0.047 | 0.504 | -0.730 | 0.180 | -0.005 | -0.009 |
| NL | 0.511 | -2.664 | 2.237 | 1.427 * | 0.701 | 0.108 | -0.015 * | -0.026 * |
| NO | 0.497 | -2.036 | 1.678 | 0.241 | 1.287 | 0.150 | -0.004 | 0.004 |
| PL | 0.195 | -1.414 | -0.977 | -1.188 * | 0.343 | -0.132 | -0.001 | 0.045 * |
| PT | 0.199 | -0.996 | -0.991 | -0.176 | -0.758 | -0.057 | 0.004 | -0.010 |
| RO | 0.151 | -1.428 | -1.330 | -0.674 * | -0.721 | 0.065 | 0.010 * | 0.012 |
| SE | 0.388 | -2.109 | 0.966 | 0.335 | 0.552 | 0.079 | -0.015 * | 0.012 |
| SK | 0.270 | -1.402 | -0.295 | 0.112 | -0.446 | 0.040 | 0.006 | -0.014 |
| UK | 0.365 | -2.025 | 1.367 | 1.410 * | -0.142 | 0.099 | -0.012 * | -0.038 * |

Significatività a livello: 0.05 **

4. Conclusioni

Alla luce dei risultati evidenziati, è difficile fornire indicazioni su quali politiche i Paesi dovrebbero attuare per abbattere il rischio di povertà delle famiglie. Sicuramente in Italia gli interventi mirati sono scarsamente presenti e quindi scarsamente efficaci; altrove, invece, è maggiore il sostegno alle famiglie e l'intervento pubblico ha successo nella lotta alla povertà. L'effetto redistributivo di tassazione e trasferimenti assistenziali, nel nostro Paese, è limitato ma il peso delle

pensioni sul *welfare* rimane preponderante e rappresenta una differente forma di sostegno alle famiglie. Il modello utilizzato, inoltre, non considera la distanza iniziale dalla linea di povertà: se si trasferissero tutte le risorse alla coda più bassa delle famiglie a rischio di povertà, ad esempio, gli interventi non sarebbero efficaci rispetto all'uscita dalla condizione di povertà quanto piuttosto rispetto alla riduzione dell'intensità di povertà. Si ritiene utile approfondire questo aspetto in successivi sviluppi del lavoro.

Riferimenti bibliografici

- CUTILLO A., LO CASTRO D., SICILIANI I. (2009). Popolazione minorile a rischio di povertà: confronti tra Paesi europei, *Rivista Italiana di Economia, Demografia e Statistica*, Vol. LXIII, No. 3-4, pp. 63-70.
- GELMAN A., HILL J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
- GOLDSTEIN H. (1995). *Multilevel Statistical Models*. London: Edward Arnold.
- LOHMANN H. (2008). *Welfare states, labour market institutions and the working poor. A comparative analysis of 20 European countries*. Berlin: DIW.

SUMMARY

This paper analyses the impact of some European countries' income redistributive policies in terms of monetary social transfers and taxation on the reduction of the incidence of households poverty risk as it results after primary income distribution to the production factors. Using micro data collected in Eu-silc 2009 (*European Statistics on Income and Living Conditions*), we employ a two-level logistic regression model. It turns out an overall country effect, which could be splitted into two partitions: one due to different policies relative to quantitative aspects (Quantity), the other to distributive ones (Distribution). The Quantity effect is larger than the Distribution one. European countries with a greater positive effect on the exit from the poverty risk after transfers and taxes are the Nordic and the Western ones; the Mediterranean and Baltic ones have the opposite effect. Moreover, some countries with minor efficacy in the exit from poverty risk have a greater Quantity effect, maybe due to a decreasing marginal utility for each additional transfer.

Andrea CUTILLO, Ricercatore, ISTAT, cutillo@istat.it
Daniela LO CASTRO, Ricercatrice, ISTAT, locastro@istat.it
Isabella SICILIANI, Ricercatrice, ISTAT, sicilian@istat.it

MISURARE L'INTEGRAZIONE: UNA PROPOSTA DI INDICATORE COMPOSITO¹

Rosario D'Agata, Venera Tomaselli

1. Misurare un fenomeno complesso: il livello d'integrazione degli immigrati

Misurare un concetto tanto complesso come quello d'integrazione significa in primo luogo dover tener conto della sua elevata multidimensionalità, considerata in approcci teorici differenti (Cesareo e Blangiardo, 2009, p. 11). Mentre, da una parte, l'integrazione è definita sulla base di percezioni soggettive; dall'altra, non può prescindere dal ruolo che il contesto svolge nel determinare il luogo di permanenza in una specifica area territoriale (Strozza *et al.* 2008, p. 99).

Dal punto di vista soggettivo, la quantificazione della misura d'integrazione si rivela un'operazione complessa. Questo sia per le peculiarità biografiche che ciascun soggetto porta con sé sia perché un soggetto non può essere totalmente escluso da reti relazionali che condizionano e persino possono determinare la sua percezione a partire da capacità linguistiche, tratti di personalità (motivazioni, bisogni, idee personali), sentimenti, giudizi e valutazioni (Nunnally, 1978). Ai fini della costruzione del concetto, pertanto, operano indicatori di tipo soggettivo, finalizzati alla rilevazione di stati su proprietà individuali (Maggino e Mola, 2007, p. 4).

Secondo la dimensione oggettiva dell'integrazione, invece, assumono interesse alcuni aspetti della realtà contestuale, semanticamente connessi al concetto di integrazione. Tra questi: opportunità occupazionali, presenza di connazionali, forme strutturali della domanda sul mercato del lavoro, ecc.. La traduzione operativa della componente oggettiva del concetto include l'impiego di indicatori oggettivi, costruiti sulla base delle informazioni provenienti da archivi e fonti ufficiali. Essi, però, non soltanto possono risultare non esaustivi rispetto alle componenti del concetto nella sua complessità ma potrebbero pure non avere la stessa rilevanza nella determinazione del concetto stesso.

¹ Il contributo degli autori al presente lavoro è da considerarsi paritetico sotto ogni aspetto.

2. Combinare e pesare indicatori oggettivi e soggettivi: un approccio *multilevel*

Il nostro interesse di studio verte sulla costruzione di uno strumento sintetico che possa considerare in maniera congiunta e funzionale elementi di tipo soggettivo e forme oggettive che connotano il concetto d'integrazione al fine di ottenere misure quantitative opportunamente ponderate.

Nel processo di costruzione degli indicatori composti per aggregazione di indicatori elementari assume rilevanza la modalità di attribuzione dei pesi in funzione della esigenza di combinare in modo significativo le diverse e molteplici dimensioni. Questo obiettivo implica la scelta di un'opportuna procedura di aggregazione e di un adeguato modello di ponderazione degli indicatori elementari in base alla loro affidabilità e valore semantico (OECD-JRC, 2008).

I pesi hanno la proprietà di rendere espliciti gli obiettivi che sottendono la costruzione di un indicatore composto. Le metodologie disponibili sono tante e differenti sia per approccio paradigmatico sia per contenuti operativi. Per ognuna vale la considerazione del quadro teorico di riferimento.

Da un punto di vista metodologico, abbiamo scelto di aggregare le unità osservate a livello individuale in macro-unità territoriali, di ricavare valori ponderali mediante il ricorso all'approccio *multilevel* e, quindi, di sintetizzare indicatori elementari oggettivi già ponderati in un indicatore composto².

Per procedere alla stima del modello *multilevel*, i dati sono organizzati in una struttura gerarchica. In essa, gli immigrati costituiscono le unità individuali di primo livello, su cui sono osservate variabili esplicative riferite a stati soggettivi; le unità di secondo livello sono rappresentate dalle province sulla cui base sono aggregate specifiche informazioni - rilevate sulle unità di primo livello - utili ai fini dell'attribuzione del peso.

Il modello stimato è un modello:

$$Y_{ij} = \underbrace{\beta_0 + \beta_1 X_{ij} + \gamma_{01} Z_j + \gamma_{11} X_{ij} Z_j}_{\text{parte fissa}} + \underbrace{u_{0j} + u_{1j} X_{ij}}_{\text{parte casuale}} + \varepsilon_{ij} \quad (1)$$

in cui la covariata Z_j , rilevata su unità di secondo livello (le province), è la *j*-ma variabile che contribuisce a spiegare la variabile dipendente Y_{ij} (Snijders e Bosker, 1999). Così, il primo passo definito dalla (1) permette di valutare la varianza delle unità di secondo livello, attraverso la stima del modello a componenti di varianza in cui è inserito l'effetto della variabile esplicativa Z_j rilevata sulle unità macro. Il

² Sui modelli *multilevel* esiste una copiosa letteratura. In questo lavoro sono stati considerati come riferimenti bibliografici quelli orientati alla combinazione di elementi oggettivi e soggettivi (Rampichini e Schifini, 1998; Maggino, 2009).

modello permette il controllo dell'ipotesi sull'esistenza di due componenti di variabilità all'interno e fra le unità di secondo livello, conseguenza dell'annidamento gerarchico delle unità. In sintesi, la varianza totale $Var(Y_{ij})$ è scomposta nei due fattori:

$$Var(Y_{ij}) = Var(u_{0j} + \varepsilon_{ij}) = \tau + \sigma^2. \quad (2)$$

In cui:

τ è la quota di varianza fra le unità di secondo livello o di contesto (province)

σ^2 è la quota di varianza interna alle unità di primo livello o individuali (immigrati).

Il rapporto tra la varianza fra i gruppi τ o fra le unità di secondo livello (province) e la varianza totale $Var(Y_{ij})$ fornisce una misura del grado di omogeneità fra le osservazioni appartenenti al medesimo gruppo o provincia, ottenuta mediante il calcolo del coefficiente di correlazione intra-classe (ICC):

$$\rho = \frac{\tau}{\tau + \sigma^2} \quad (3)$$

Il coefficiente misura la proporzione di varianza totale spiegata dalla varianza fra le unità di secondo livello o di variabilità dovuta all'effetto del contesto: più è elevato il valore del coefficiente tanto più apprezzabile è l'effetto dovuto al raggruppamento e più importante è il ruolo della struttura gerarchica. Esso indica, pertanto, l'esistenza di differenze fra le unità di secondo livello (province). La riduzione del valore dell'ICC, a seguito dell'introduzione di una covariata o variabile rilevata sulle unità di secondo livello Z_j , individua la capacità della covariata di ridurre la variabilità dovuta alla strutturazione gerarchica. Risulta maggiore, pertanto, il peso attribuito a quella dimensione di contesto. La stima ed il valore della significatività del coefficiente permettono, quindi, la valutazione della significatività del livello gerarchicamente superiore.

Nella nostra applicazione, le variabili rilevate sulle unità individuali di primo livello sono operativizzate in indicatori soggettivi; mentre le variabili definite contestuali, ottenute per aggregazione di dati soggettivi sulla base di aree territoriali - assunte come unità di secondo livello - sono impiegate per ricavare i pesi (w_k). Da questa operazione deriva la costruzione di un indicatore composito per misurare il livello d'integrazione degli immigrati.

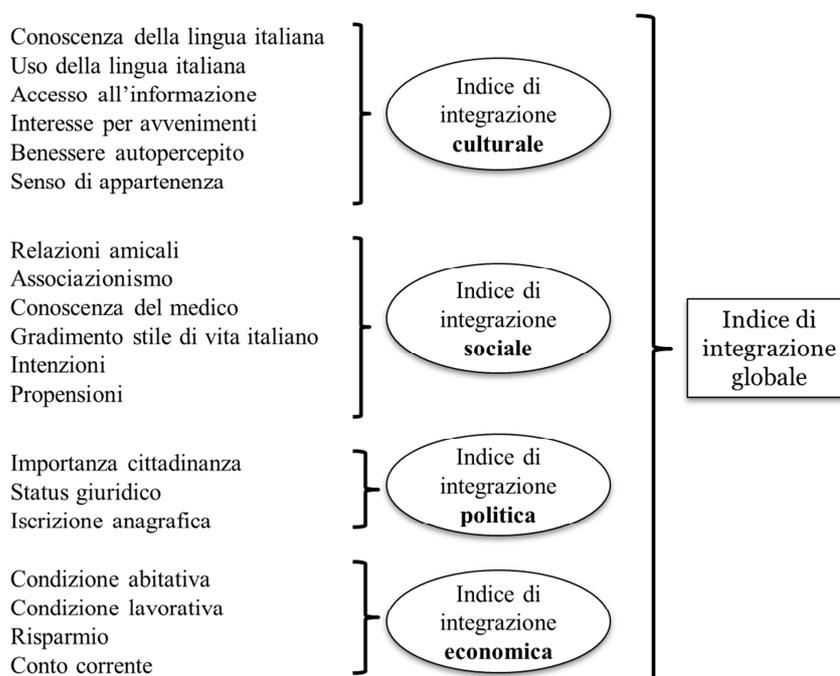
3. Gli indici di integrazione: ISMU e CNEL

Nel 2009 la fondazione ISMU pubblica i risultati di un'indagine empirica proponendo l'impiego di un indice soggettivo di integrazione globale (Cesareo e Blangiardo, 2009). Come si evince dallo schema riportato nella figura 1, gli autori costruiscono l'indice globale aggregando 4 indici di dimensione ottenuti, a loro

volta, per aggregazione di indicatori semplici (*ibid.* p. 38). Gli indicatori soggettivi impiegati tentano di ‘coprire’ le dimensioni relative all’integrazione culturale, sociale, politica ed economica³.

Tali dimensioni assumono un ruolo centrale nella costruzione dell’indicatore composito oggetto della proposta metodologica del presente lavoro, giacché, come vedremo, rappresentano il punto di partenza per la costruzione di un sistema di ponderazione da applicare ad indicatori oggettivi.

Figura 1 – Schema di costruzione dell’indicatore globale (ISMU).



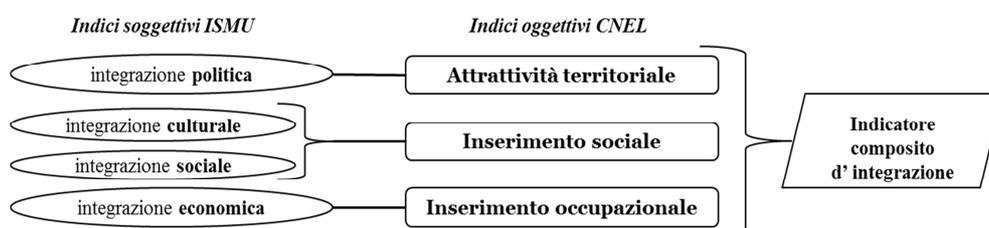
Individuati gli strumenti di misurazione della componente soggettiva dell’integrazione, per la scelta degli indicatori oggettivi, è apparso utile impiegare l’indice del *potenziale d’integrazione dei territori italiani* messo a punto dai ricercatori del CNEL nel settimo rapporto sull’integrazione (CNEL, 2010). Tale indice, aggrega tre dimensioni, ciascuna delle quali sintetizza 5 indicatori semplici. Le tre

³ Lo studio è stato condotto su 32 province del territorio italiano dove sono state raccolte 12.049 interviste ad immigrati. Per ulteriori approfondimenti relativi agli strumenti di misurazione utilizzati e alle procedure di campionamento adottate si rimanda al testo citato.

dimensioni considerate nel rapporto CNEL sono: attrattività territoriale, inserimento sociale e inserimento occupazionale.

Alla luce degli indici finora considerati e relativi alle due dimensioni (soggettiva ed oggettiva) che concorrono a definire il concetto generale di integrazione, nelle prossime pagine tenteremo di esporre la nostra proposta metodologica per la costruzione di un indicatore composito d'integrazione che comprenda entrambe le dimensioni. In particolare, l'obiettivo consiste nel trasformare ciascun indice soggettivo in peso da attribuire a ciascun indice oggettivo semanticamente affine come riportato schematicamente nella figura 2.

Figura 2 – Schema di costruzione dell'indicatore composito proposto.



A tal fine è apparso utile aggregare, attraverso il calcolo della aritmetica, le dimensioni di integrazione culturale e sociale in un unico indice di integrazione socio-culturale al fine di rendere possibile la ponderazione del corrispettivo indice oggettivo di inserimento sociale.

4. La costruzione dello schema di ponderazione

L'introduzione della componente soggettiva all'interno dell'indicatore composito avviene attraverso la trasformazione in pesi degli indici soggettivi. Questi ultimi, per definizione, sono però calcolati sugli individui mentre, gli indici CNEL, sono rilevati sulle province. Per risolvere il problema, dunque, è necessario cogliere l'importanza che ciascuna dimensione soggettiva ha a livello contestuale. In altri termini, maggiore è il contributo che a livello contestuale ciascuna dimensione (soggettiva) fornisce nel determinare l'integrazione degli individui, maggiore è il peso che quella dimensione ha nella costruzione di un indicatore composito. A questo scopo appare utile l'impiego dei modelli *multilevel*.

Dopo aver trasformato i valori degli indici attraverso la normalizzazione su campo di variazione (Nardo *et al.*, p. 47), si è proceduto alla costruzione di un mo-

dello *multilevel* con variabile dipendente l'indice globale di integrazione soggettivo calcolato sui 12049 casi. La specificazione del modello, quindi, ha previsto l'inserimento di due covariate di primo livello: il genere (con *baseline* = maschio) e l'anzianità migratoria, misurata in anni di permanenza in Italia e, alternativamente, l'inserimento dei 3 indici soggettivi di integrazione (politica, socio-culturale ed economica) come covariate di secondo livello, ottenute calcolando la media aritmetica di ciascun indice all'interno di ciascuna delle 32 province considerate.

Per valutare la capacità esplicativa di ciascuna covariata di secondo livello, si è scelto di considerare la riduzione del coefficiente di correlazione intraclasse (ICC). L'ipotesi è che ad una maggiore riduzione del coefficiente corrisponda una maggiore capacità esplicativa della covariata di secondo livello.

La tabella 1 mostra i coefficienti statisticamente significativi ($p\text{-value} < 0,001$). Dall'analisi dei coefficienti di correlazione intraclasse notiamo che il modello con l'integrazione politica come covariata di secondo livello riduce l'ICC di 0,042 punti in valore assoluto (71,17%) rispetto al modello con solo covariate di primo livello. La dimensione dell'integrazione politica appare, dunque, pesare maggiormente a livello contestuale sulla performance dell'indice globale. Segue la dimensione socio-culturale con una riduzione dell'ICC di 0,039 punti (65,77%) e, infine, la dimensione economica con una riduzione del coefficiente pari a 0,028 (46,42%).

Impiegando le informazioni ottenute dalle riduzioni dei coefficienti, il passo successivo ha previsto la riproporzionalizzazione delle suddette riduzioni in una scala 0–1, i cui valori sono contenuti nell'ultima riga della tabella 1.

Tabella 1 – Modelli multilevel impiegati per il calcolo dei pesi (w_k). Y_{ij} = indice globale di integrazione soggettiva.

| | β_{0j} | E. S. |
|--------------------------------|--------------|-------|--------------|-------|--------------|-------|--------------|--------|
| <i>Effetti fissi livello 1</i> | | | | | | | | |
| Intercetta | 34,120 | 0,696 | 26,65 | 0,983 | 29,014 | 0,814 | 27,635 | 1,3610 |
| Genere (Femmina) | 4,902 | 0,274 | 4,896 | 0,273 | 4,897 | 0,273 | 4,894 | 0,273 |
| Anzianità migrat. | 1,408 | 0,022 | 1,404 | 0,022 | 1,406 | 0,022 | 1,406 | 0,022 |
| <i>Effetti fissi livello 2</i> | | | | | | | | |
| Integraz. Politica | | | 0,122 | 0,015 | | | | |
| Integraz. Socio-cult. | | | | | 0,119 | 0,016 | | |
| Integraz. Economica | | | | | | | 0,109 | 0,021 |
| <i>Effetti casuali</i> | | | | | | | | |
| livello 2 u_{0j} | 13,075 | 3,451 | 3,615 | 1,077 | 4,3210 | 1,25 | 6,809 | 1,8790 |
| livello 1 e_{0j} | 210,392 | 2,733 | 210,387 | 2,733 | 210,381 | 2,133 | 210,392 | 2,7330 |
| ICC | 0,059 | | 0,017 | | 0,020 | | 0,031 | |
| w_k | | | 0,385 | | 0,358 | | 0,257 | |

Dopo aver calcolato i pesi, abbiamo costruito l'indicatore composito (IC) attraverso la media aritmetica dei 3 indici CNEL ponderati per i relativi pesi⁴.

La tabella 2 presenta il confronto tra la graduatoria delle province stilata attraverso l'indice CNEL e quella ottenuta attraverso l'IC proposto.

Tabella 2 – *Graduatoria dell'indice CNEL e graduatoria dell'IC con i pesi.*

| Provincia | Rango CNEL | Rango IC | Provincia | Rango CNEL | Rango IC |
|--------------|------------|----------|--------------------|------------|----------|
| Prato | 1 | 5 | Grosseto | 17 | 16 |
| Parma | 2 | 2 | Livorno | 18 | 17 |
| Vicenza | 3 | 1 | Lucca | 19 | 18 |
| Modena | 4 | 3 | Pistoia | 20 | 23 |
| Milano | 5 | 6 | Teramo | 21 | 19 |
| Treviso | 6 | 4 | Pisa | 22 | 25 |
| Forli-Cesena | 7 | 8 | Torino | 23 | 30 |
| Rimini | 8 | 15 | Palermo | 24 | 20 |
| Roma | 9 | 13 | Massa-Carrara | 25 | 24 |
| Padova | 10 | 10 | Campobasso-Isernia | 26 | 21 |
| Arezzo | 11 | 9 | Pescara | 27 | 31 |
| Trento | 12 | 7 | Catania | 28 | 26 |
| Ravenna | 13 | 14 | Chieti | 29 | 28 |
| Siena | 14 | 12 | Napoli | 30 | 32 |
| Ancona | 15 | 11 | Benevento | 31 | 27 |
| Firenze | 16 | 22 | Bari | 32 | 29 |

La classifica ottenuta attraverso l'IC vede al primo posto Vicenza che era terza nella classifica CNEL e Parma, prima per il CNEL, occupa la quinta posizione. Le differenze più rilevanti si osservano in relazione alle province di Rimini e Torino con una differenza di ben sette posizioni tra le due graduatorie. In entrambi i casi, infatti l'introduzione dei pesi 'peggiora' la performance delle due province; così Rimini passa dall'ottava alla quindicesima posizione e Torino dalla ventitreesima alla trentesima. Anche Firenze mostra una differenza di 6 posizioni mentre Pescara, Roma e Prato 'perdono' 4 posizioni. L'introduzione del sistema di ponderazione migliora la performance di Campobasso-Isernia e Trento che nel confronto con la graduatoria dell'IC 'acquistano' 5 posizioni. Allo stesso modo anche Ancona, Benevento e Palermo migliorano il loro grado di integrazione risalendo di 4 posizioni. Se l'indice CNEL, infine, vede in fondo alla classifica la provincia di Bari, la classifica dell'IC rileva che Napoli è la provincia con minore integrazione.

In conclusione, l'applicazione del modello *multilevel* ha permesso di attribuire pesi agli indicatori oggettivi, ottenendo un indicatore composito d'integrazione.

⁴ Per far ciò è stato necessario selezionare solo le 32 province in cui è stata svolta l'indagine dell'ISMU e trasformare i valori iniziali di ciascun indice di dimensione attraverso la normalizzazione su campo di variazione relativo alle sole province considerate.

Riferimenti bibliografici

- CESAREO V., BLANGIARDO G. (Ed.). 2009. *Indici d'integrazione. Un'indagine empirica sulla realtà migratoria italiana*. Milano: Franco Angeli.
- CNEL 2010. *Indici di integrazione degli immigrati in Italia. VII Rapporto*. Roma: CNEL
- MAGGINO F. e MOLA T. 2007. *Definizione e costruzione di indicatori soggettivi*. Firenze: Università degli Studi di Firenze.
- MAGGINO F., 2009. Methodologies to integrate subjective and objective information to build well-being indicators. *Proceedings of international conference "From GDP to well-being: economics on the road to sustainability"*. Ancona: Università Politecnica delle Marche, December 2009, pp. 1-27.
- NUNNALLY J. C., 1978. *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- OECD-JRC, 2008. *Handbook on Constructing Composite Indicators. Methodology and User Guide*. European Commission.
- RAMPICHINI C., SCHIFINI D'ANDREA S., 1997. A hierarchical ordinal probit model for the analysis of life satisfaction in Italy. *Social Indicators Research*, n. 44, pp. 41-69.
- SNIJDERS T. e BOSKER R.. 1999. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. London: Sage.
- STROZZA S., FORCELLATI L., FERRARA R. 2008. Il lavoro degli stranieri residenti nelle diverse Italie: differenze di genere e per cittadinanza. *Rivista Italiana di Economia Demografia e Statistica*, gennaio-marzo 2008, Vol. LXII, No. 1, pp. 99-128.

SUMMARY

In this paper we propose a composite indicator by aggregating objective indicators weighted by subjective indicators. Employing multilevel models, we use individual data as 1-level units nested into 2-level units in order to estimate contextual effects on the subjective measure of immigrants integration as dependent variable. Then, we compare the effects of three models, each one with a different 2-level covariate, evaluating the ICC decrease: the higher is the decrease, the higher is the contextual effect related to the 2-level covariate in the model. Doing so, we obtain different weights employed to construct a composite indicator of integration. In this aim, we use ISMU survey data, collected in 32 Italian provinces as subjective indicators to weight objective indicators of CNEL Report.

Rosario D'AGATA, Ricercatore in Statistica Sociale presso la Facoltà di Scienze Politiche dell'Università degli Studi di Catania, rodagata@unict.it

Venera TOMASELLI, Professore Associato in Statistica Sociale presso la Facoltà di Scienze Politiche dell'Università degli Studi di Catania, tomavene@unict.it

SULL'INCHIESTA DEL CISP DEL 1938 SU ALCUNE POPOLAZIONI ALBANOFONE DELLA CALABRIA

Giuseppe De Bartolo, Manuela Stranges

1. Introduzione

In Italia, il primo studio demografico sulle minoranze linguistiche fu realizzato da Corrado Gini, il quale nel 1928 fondò il CISP – Comitato Italiano per lo studio della popolazione, che aveva nel suo programma lo studio dei gruppi etnici isolati, con lo scopo di completare le conoscenze sull'evoluzione delle popolazioni umane. Per questo motivo organizzò nove spedizioni scientifiche per esaminare gruppi che presentavano varie gradazioni che andavano dalla più spiccata primitività ad una civiltà raffinata, man mano che si passava dai Seris, Dauada, Otomi, Zambo di Guerrero, Mixe, Cinantechi, Zapotечи, Coras, Huicholes, Beberi, Taraschi, Aztechi, Samaritani, Albanesi di Calabria, Caraimi di Polonia e Lituania, coloni liguri di Calasetta, coloni liguri di Carloforte (Gini, 1955).

Nell'ambito di questo progetto scientifico tra l'agosto e l'ottobre del 1938 vennero esaminati i gruppi allogeni di Calabria di Carfizzi, Caraffa e S. Nicola dell'Alto, tutti in provincia di Catanzaro, comuni scelti in virtù del fatto che queste popolazioni discendevano quasi unicamente da quella albanese, emigrata verso l'Italia meridionale e insulare nel XV-XVI secolo. Scopo delle ricerche giniane era quello di determinare il grado di persistenza dei caratteri degli immigrati rispetto alla popolazione di origine e di verificare le teorie dell'assimilazione delle stirpi che da tempo vivono in Italia. Per avere dei validi elementi di confronto, nell'anno seguente furono raccolti altri dati relativi alle famiglie albanesi di Kavaja, cittadina situata a sud-est di Tirana in Albania, in quanto proprio da questa cittadina e da altre limitrofe sembrava, infatti, che più intensa fosse stata la diaspora arbereshe verso l'Italia.

Mentre i risultati relativi all'inchiesta sulla popolazione albanese di Kavaja vennero pubblicati nel 1960 da Nicolichia su *Genus*, dell'inchiesta sulle popolazioni albanofone della Calabria si hanno solo risultati parziali, pubblicati da Nora Federici in un libretto presumibilmente del 1955¹. A distanza di oltre settanta

¹ In un elenco delle pubblicazioni prodotte dal CISP in relazione allo studio delle popolazioni isolate,

anni dalla raccolta dei dati sui tre paesi italo-albanesi, anche per onorare la memoria di Nora Federici, pubblichiamo in questo studio altri risultati di quell'inchiesta, ottenuti attraverso lo spoglio più puntuale dei 168 questionari compilati dall'equipe² di cui faceva parte la stessa Federici³.

E' bene precisare che, nella presentazione dei risultati della nostra ricerca, solo in casi particolari i dati saranno riferiti ai tre comuni distintamente, mentre per il resto faremo riferimento ai dati globali. In alcune circostanze, infatti, non si è ritenuto opportuno operare una suddivisione in quanto le caratteristiche e le tendenze erano più o meno omogenee e quindi un loro confronto non avrebbe fornito ulteriori elementi di diversificazione. I risultati così ottenuti sono stati confrontati in maniera sistematica con quelli dell'inchiesta sugli albanesi di Kavaja (903 questionari). Per ragioni di spazio verranno esposti solo i risultati salienti dello spoglio delle schede, illustrandoli in maniera descrittiva senza l'utilizzo di tabelle o figure. Ci si ripropone, chiaramente, la pubblicazione in altra sede dei risultati completi dello studio.

2. Strutture familiari

2.1. Caratteristiche generali delle famiglie

Le schede demografiche riportano per ciascuna famiglia notizie relative ai tratti demografici, antropometrici e medico-biologici dei singoli componenti, le loro condizioni economiche e lo stato delle abitazioni. In particolare riferiscono notizie su: 1. tribù o altro gruppo sociale; 2. abitazione della famiglia; 3. Capofamiglia; 4. ascendenti del capofamiglia e della moglie; 5. fratelli e sorelle del capofamiglia; 6. fratelli e sorelle della moglie del capofamiglia; 7. capofamiglia, mogli e figli viventi o premorti; 8. altri componenti la famiglia e altri conviventi abituali od

pubblicato su *Genus* (n. 1-2 nel 1989), compare anche un articolo di Nora Federici: "Caratteristiche demografiche differenziali di un gruppo allogeno di Albanesi di Kavaja e di alcune colonie allogene albanesi di Calabria (primi risultati)", *Genus*, XII, n. 1-4, 1956. In realtà tale articolo non apparve mai su *Genus*, ma i suoi contenuti vennero anticipati temporalmente per motivi accademici dell'autrice in un libretto della Tipografia Failli di Roma, senza data, ma verosimilmente del 1955.

² Furono compilate 49 schede riguardanti famiglie residenti nel Comune di San Nicola dell'Alto, 59 nel Comune di Caraffa e 60 nel Comune di Carfizzi. Nel suo articolo del 1960, Nicolicchia faceva erroneamente riferimento a 127 schede.

³ Della Commissione del Comitato, presieduta dallo stesso Gini, facevano parte anche Giovanni Cirillo e Silvio Orlandi. I dati su Kavaja furono raccolti da Cirillo che in quegli anni faceva il servizio militare in Albania. L'indagine albanese fu predisposta con lo scopo specifico di consentire il confronto con le popolazioni albanofone dei tre comuni calabresi, poiché il gruppo esaminato di Kavaja, a detta dei rilevatori, era rappresentativo della regione da cui erano originarie le colonie albanofone della Calabria.

occasionalmente; 9. osservazioni varie.

Come ci si attende per una comunità tradizionale come quella esaminata, nella maggior parte delle famiglie il capofamiglia era il marito (85,6% dei casi). I capifamiglia femmine (14,3%) erano per la maggior parte vedove (83,3%) e, in minor parte, nubili (12,5%) e separate o abbandonate dai rispettivi mariti (4,2%). Nell'inchiesta condotta a Kavaja non si coglievano sostanziali differenze con i paesi albanofoni: infatti, nell'89% dei casi il capofamiglia era maschio. Di questi il 94,68% era coniugato, il 2,5% celibe, lo 0,12% divorziato e il 2,7% vedovo (Nicolicchia, 1960, p.75). I questionari evidenziano una maggiore presenza di vedove rispetto ai vedovi, attribuibile al fatto che molti vedovi si erano risposati, mentre molte vedove, dopo la morte del marito, non avevano contratto un nuovo matrimonio.

In media le mogli erano più giovani dei mariti di 4 anni (precisamente l'età media dei mariti è risultata 48,6 anni mentre quella delle mogli era di 44,6 anni). La popolazione coniugata del gruppo Kavaja era più giovane dell'analoga popolazione di Calabria, con una maggiore distanza tra i coniugi⁴: l'età media è risultata, infatti, di 42,7 per i mariti e 34,5 anni per le mogli.

Tra le varie informazioni riguardanti il capofamiglia sono stati da noi scelte anche due variabili all'epoca "di moda", la razza e il sangue. In particolare le informazioni riguardavano la razza degli ascendenti del capofamiglia (genitori, nonni paterni e materni) che venivano trascritte nella IV parte della scheda demografica. Dei 168 capifamiglia, 151 erano classificati come appartenenti alla razza albanese (89,9%) mentre 17 (10,1%) erano di origine calabrese. Caraffa presentava un maggior numero di capifamiglia di razza calabrese. Il "sangue" veniva, invece, classificato come puro, se i capifamiglia erano esclusivamente albanesi o calabresi (cioè non si erano verificate mescolanze con altre popolazioni), o misto, se gli ascendenti avevano subito una fusione con altre popolazioni di razza diversa. Il 75,6% dei capifamiglia era di razza pura albanese o calabrese, mentre il 24,4% era di sangue misto. Il sangue misto era più presente a Carfizzi (46,7%) rispetto agli altri due comuni.

2.2. Numerosità delle famiglie

Nella parte VIII della scheda erano riportate informazioni riguardanti fratelli e sorelle del capofamiglia e della moglie. Da queste informazioni è possibile risalire alla numerosità delle famiglie di origine e confrontarla con quella di cui era a capo il capofamiglia stesso nel 1938. Dalla VI parte della scheda è, invece, possibile ottenere la numerosità della famiglia attuale del capofamiglia, inclusi i figli morti

⁴ Si tenga presente che, in quella inchiesta, essendo molti mariti poligami, si considerò solo l'età della prima moglie fra tutte le mogli conviventi.

successivamente alla nascita. Il numero complessivo delle persone che costituivano le 168 famiglie originarie dei titolari della scheda era pari a 1.405 persone, da cui si ottiene che la media dei componenti per famiglia era di 8,36.

Per quanto riguarda le mogli, poiché molte notizie al riguardo erano mancanti, è stato possibile risalire solo a 143 famiglie originarie con un numero di componenti di 1.197 persone e una media di 8,37, corrispondente a quella dei mariti. Considerando invece il capofamiglia, la moglie e figli nati vivi è possibile calcolare la media dei componenti di una famiglia alla data del 1938 (1.168 componenti, media 6,95), da cui si deduce che le famiglie del 1938 avevano una dimensione di circa 2 unità in meno rispetto alle famiglie formatesi alcuni anni prima.

3. Comportamenti demografici

3.1. *Periodo fecondo e fecondità*

Al 1938, le mogli viventi erano in totale 158, comprese vedove, separate e abbandonate. L'età alla prima mestruazione, dichiarata da 154 mogli, oscillava da 13 a 15 anni, in media 13,84 anni. Anche le figlie avevano avuto la prima mestruazione ad una età media pari a quella delle madri (13,87 anni) ma l'intervallo di età era stato più ampio (13-16 anni). Solo 66 donne avevano dichiarato l'età alla menopausa risultata essere in media di 50 anni.

Confrontando l'età media alla prima mestruazione e alla menopausa delle mogli italo-albanesi con quelle delle albanesi di Kavaja si riscontra una notevole differenza. Le mogli di Kavaja avevano dichiarato un'età media alla prima mestruazione di 12 anni con un campo di variazione da 10 a 15 anni e con una maggior frequenza intorno a 12 anni, mentre l'età media alla menopausa era risultata di 51 anni, con un intervallo di variazione da 42 a 56 anni e con una frequenza massima intorno a 53 anni. In media le mogli di Kavaja cominciarono il loro periodo fecondo 1,9 anni prima e lo terminavano quasi un anno dopo.

Se si escludono i celibi e le nubili, le famiglie italo-albanesi indagate risultarono 163. I figli avuti da queste famiglie, escludendo i nati morti e gli aborti, furono 843, di cui 419 femmine e 424 maschi, dunque un rapporto di mascolinità alla nascita pari a 101,2%, con forti differenze a livello territoriale: a Caraffa e S. Nicola dell'Alto i rapporti di mascolinità risultarono pari a 99%, per Carfizzi a 108%. Il numero medio di figli per famiglia è risultato di 5,17, con differenze significative nei tre centri abitati: Caraffa 5,53; Carfizzi 3,86; S. Nicola dell'Alto 6,27. In realtà, fra le famiglie esaminate 9 erano senza prole (tra cui anche quelle che avevano avuto solo aborti e nati morti) e, di conseguenza, il numero medio di figli per famiglia sale 5,47. Per la città di Kavaja la media di figli per famiglia era

risultata di 4,20 (4,62 tenendo conto, anche in questo caso, delle famiglie senza figli⁵), con un rapporto di mascolinità di 106 maschi per 100 femmine.

Per le italo albanesi l'età al primo matrimonio era stata pari a 22,4 anni, quella alla nascita del primo figlio a 23,6. Per le donne di Kavaja, tali età erano risultate, rispettivamente, di 20,6 e 22 anni. Da ciò si deduce che queste ultime disponevano di un più lungo periodo di convivenza matrimoniale rispetto alle discendenti immigrate in Calabria.

Nelle sue analisi sulle schede del 1938, la Federici prese in considerazione solo le informazioni riguardanti la nuzialità e la fecondità matrimoniale. Per la fecondità concluse che le donne italo-albanesi erano nettamente più prolifiche delle donne di Kavaja⁶, con valori analoghi a quelli delle donne calabresi del tempo (che erano tra le più feconde d'Italia⁷), attribuendo tale prolificità ad una più completa utilizzazione del periodo matrimoniale fecondo.

Queste ed altre considerazioni portarono la Federici ad affermare che le colonie allogene di Calabria potessero considerarsi assimilate alle popolazioni calabresi oltre che per i caratteri fisici anche per quelli demografici, ponendo l'accento sull'influenza dei fattori riconducibili all'ambiente di vita (ma non escludendo, comunque, l'effetto della selezione operata dall'immigrazione) in questo processo di assimilazione.

3.2. *Mortalità e cause di morte*

I mariti morti tra gli italo-albanesi erano 26 su 142, mentre le mogli decedute erano 17 su 158. Pertanto, i rapporti viventi/totale e deceduti/totale sono stati pari, rispettivamente, a 90,3% e 9,7% per le mogli e a 84,5% e 15,5% per i mariti. A Kavaja gli analoghi rapporti denotavano una supermortalità femminile, che non era in passato infrequente specialmente tra le popolazioni tradizionali (Tabutin, 1978): infatti, il rapporto deceduti/totale era pari a 17,6% per le femmine e a 10,5% per i maschi. L'età media alla morte dei 26 mariti deceduti è risultata di 56 anni (48,8 anni per i 95 capifamiglia deceduti nella città di Kavaja). La mortalità del gruppo italo-albanese

⁵ La percentuale delle donne che non aveva figli tra le italo-albanesi era del 5,47% mentre per le donne di Kavaja era ancora più alta, pari all'8,95%.

⁶ Considerando tutte le donne dell'indagine, infatti, calcolò un indice di produttività (rapporto tra il numero complessivo dei figli e il numero di donne feconde) di 4,85 e uno di prolificità (rapporto tra il numero complessivo dei figli e il numero di donne che hanno avuto almeno un figlio) di 5,10, mentre per le donne di Kavaja i valori corrispondenti erano 3,51 e 4,10. Considerando le donne con ciclo riproduttivo chiuso questi valori salivano a 5,72 per l'indice di produttività e 5,90 per l'indice di prolificità, mentre per le donne di Kavaja, rispettivamente, 3,12 e 3,73 (Federici, 1955, p.10).

⁷ Il tasso di fecondità totale in Calabria nel 1936 era di 4,18 figli per donna feconda e il tasso generico di natalità, attorno al 1938, era del 30‰ contro un valor medio nazionale del 23-24‰.

considerato è abbastanza bassa fino a 50 anni (solo il 26,9% del totale dei decessi dei mariti) e la maggior parte dei decessi in questo intervallo di età è da imputare a polmonite o lesioni violente. Dai 50 anni in poi, in particolare dopo i 70 anni, aumenta ovviamente il numero dei decessi, con polmonite, malattie di cuore e paralisi come cause più ricorrenti.

Le mogli morte erano, invece, 22, con un'età media alla morte molto più bassa di quella maschile (42 anni). La mortalità colpiva in misura maggiore le donne con meno di 40 anni (40,9% del totale dei decessi delle mogli), con una punta massima nella classe 30-40 (22,7%), per poi decrescere leggermente con l'età. Le cause di morte erano per la maggior parte malattie contagiose (tifo, tbc, malaria), mentre dai 40 anni si moriva, principalmente, di parto, malattie contagiose e, in misura ridotta, pleurite.

4. Mortalità ante-natale e mortalità infantile

Lo spoglio delle schede mette in evidenza che su 930 gravidanze, i nati vivi erano 843 (90,6%), i nati morti 23 (2,5%), mentre le gravidanze che avevano dato luogo ad un aborto erano 64 (6,9%). Il quoziente di natimortalità è stato, dunque, pari a 23,7‰, quello di abortività a 75,9‰, mentre il rapporto fra aborti e gravidanze è stato pari a 68,8‰. Mediamente, tra le donne italo-albanesi, vi sono stati 1,88 aborti in riferimento alle 34 donne che hanno avuto un aborto, valore che scende a 0,41 considerando tutte le donne che hanno avuto almeno una gravidanza. Tra le 1.098 donne di Kavaja prese in esame vi erano stati, invece, 212 aborti (in media 0,19 per donna che aveva avuto almeno una gravidanza). In relazione all'ordine del parto secondo cui si è verificato l'aborto, la frequenza maggiore è stata riscontrata nella seconda e nella terza gravidanza (16,7% ciascuna). Caraffa ha registrato la percentuale maggiore di aborti (12,4%), mentre a Carfizzi non ne sono stati osservati.

Come si accennava, degli 843 nati vivi, solo il 74% risultava essere in vita al 1938 (con un valore massimo del 79,6% a Carfizzi), mentre il 26% era deceduto. Dei 219 figli che risultavano deceduti al momento dell'indagine, il 53,4% era maschio mentre il 46,6% era femmina denotando, quindi, una supermortalità maschile. Con riferimento ai figli maschi il numero maggiore di decessi era avvenuto nei primi dieci anni di vita (84,6%). In particolare, ben il 41% dei decessi era avvenuto nel primo anno di vita. All'epoca molte cause di morte infantili erano sconosciute: fra quelle rilevate segnaliamo la bronchite e la paralisi infantile. Elevato era anche il numero dei decessi dal primo al decimo anno (43,6%), con cause di morte quali bronchite, gastroenterite e altre malattie infettive. Dal decimo anno in poi la mortalità maschile decresceva (15,4%) ed era da imputare a bronchite, malattie infettive, poliomielite, lesioni violente o a causa di guerra. Per le femmine nel primo anno di vita l'incidenza della mortalità era del 48%, con cause quali bronchite, gastroenterite, malattie

infettive e molte altre cause sconosciute, mentre il 33% delle cause erano mancanti. L'incidenza della mortalità decresceva tra il primo e il decimo anno (40,2%) e le cause di morte in questa fascia erano malattie infettive e ancora bronchite. Dal decimo anno in poi la percentuale di morti si riduceva notevolmente.

A Kavaja si riscontrò anche una supermortalità maschile (36,8% per i maschi e 33,87% per le femmine), ad eccezione del primo anno di vita. Per entrambi i sessi si registra, chiaramente, una diminuzione della mortalità all'aumentare dell'età.

5. Aspetti sociali ed economici

Per quanto riguarda il grado di istruzione dei capifamiglia e delle mogli, su 308 soggetti, il 48,1% era alfabeto ma altrettanto era analfabeta (il 3,8% dei dati erano mancanti). L'analfabetismo era predominante tra le femmine (66,7%) rispetto ai maschi (27,4%). Dalle schede è stato possibile rilevare dati anche sull'istruzione di 624 figli viventi: non considerando i bambini di età inferiore a 7 anni, il 50,8% era alfabeto, il 20,4% analfabeta, il 18,9% era composto da bambini non in età scolare e solo lo 0,2% aveva dichiarato un'istruzione superiore (il 9,7% erano dati mancanti).

Per quanto concerne la lingua, quella albanese era all'epoca ben conservata nei tre comuni: su 168 capifamiglia (maschi e femmine), il 93,5% parlava la lingua albanese e calabrese contemporaneamente, l'1,2% parlava esclusivamente la lingua calabrese (perché erano calabresi immigrati dai centri albanesi) e nessuno parlava solo l'albanese. Tutti gli indagati praticavano la religione cattolica.

Riguardo all'occupazione, il mestiere predominante tra gli uomini era quello di contadino (77,6%), mentre molto meno rappresentate erano le professioni di minatore (4,2%), agricoltore (4,2%) e calzolaio (2,8%). Il 9,1% era diviso fra commercianti, fattori, fornai, guardiani di buoi, mugnai, muratori, parrucchieri, pastori, sarti e venditori ambulanti, tutte attività non presenti tra quelle dei figli che, a loro volta, ne esercitavano altre (cameriere, cancelliere, fabbro, meccanico, medico). Tra i 307 figli maschi predominavano sempre i contadini (43,6%), a seguire minatori (3,3%), impiegati e agricoltori (2,6%), mentre il 4,2% svolgeva, come si diceva, piccole attività. Da rilevare che, sempre tra i figli, il 22,4% non svolgeva alcuna attività (ma in questi vi erano sia i bambini inferiori a 8 anni che gli invalidi), l'11,7% erano scolari (bambini tra 6 e 14 anni), mentre per l'11,2% non si avevano informazioni. Per quanto riguarda le figlie femmine, escludendo quelle che al momento non svolgevano alcuna attività (o perché inferiori a 8 anni di età o perché affette da malattie che non consentivano di lavorare), il 45,7% risultava casalinga, il 10,7% contadina, il 10,7% scolara, il 20,5% non esercitava alcuna attività e solo il 2% era insegnante.

6. Brevi conclusioni

Lo spoglio esaustivo delle schede dell'indagine del CISP realizzato in questo studio ha consentito di evidenziare particolari aspetti demografici, sociali ed economici relativi alle tre popolazioni italo albanesi di Calabria indagate. L'analisi, che integra i primi risultati pubblicati dalla Federici nel 1955 relativi solo ad aspetti quali nuzialità e fecondità, ha permesso di rilevare che i gruppi calabresi esaminati mostrano una divergenza significativa per molti tratti con gli albanesi di Kavaja da cui provengono. Tale risultato conferma il processo di assimilazione che il Gini stesso si era prefisso di indagare tramite queste indagini e che Nora Federici aveva confermato nelle sue conclusioni. L'obiettivo successivo, dopo questo primo studio, è quello di andare oltre alle analisi descrittive, utilizzando i dati originari per analisi di tipo esplicativo in grado di indagare la relazione tra le variabili studiate.

Riferimenti bibliografici

- NICOLICCHIA P.E. 1960. Le caratteristiche demografiche di un gruppo di albanesi di Kavaja. *Genus*. 1960, Vol. XVI, No.1-4, pp.72-115.
- FEDERICI N. 1955. *Caratteristiche differenziali di un gruppo di Albanesi di Kavaja e di alcune colonie allogene albanesi di Calabria (primi risultati)*, Roma: Tipografia Failli.
- CASANOVA E. 1940. Le colonie allogene dell'Italia meridionale e della Sicilia. *Genus*. 1940, Vol. IV, No. 3-4, pp. 1-31.
- GINI C. 1955. Méthodes et résultats de l'étude des population primitives. *Genus*. 1955, Vol. XI, No. 1-4.
- TABUTIN D. 1978. La surmortalité féminine en Europe avant 1940. *Population*. 1978, No. 1, pp. 121-148.

SUMMARY

Through a deep analysis of the 1938 survey of CISP, it has been possible to highlight particular demographic, social and economic aspects of the three observed Italian-Albanian populations of Calabria. The analysis has revealed that the examined Calabrian groups show a significant deviation from the Albanians of Kavaja from which they originate. This result confirms the process of assimilation that Corrado Gini intended to explore through the CISP investigations and that Nora Federici has also shown in the conclusions of her first paper about this survey.

Giuseppe DE BARTOLO, Professore Ordinario di Demografia, Università della Calabria,
debart@unical.it

Manuela STRANGES, Ricercatrice di Demografia, Università della Calabria,
m.stranges@unical.it

TRA AUTOIMPIEGO E MICRO-IMPREDITORIALITÀ: TRATTI DELL'IMMIGRAZIONE IN LOMBARDIA

Stefania Della Queva, Daniela Bonardo, Alessio Menonna, Flavio Verrecchia

1. Introduzione

L'uso degli archivi amministrativi finalizzati alla produzione statistica (Martini, 1995), l'innovazione tecnica, l'evoluzione e la valorizzazione dell'infrastruttura informativa sulle imprese (Mancini, 2010) consentono di riscoprire, anche grazie all'apporto informativo di dati inediti, un ambito di ricerca considerato sostanzialmente noto: la partecipazione al mercato del lavoro. L'obiettivo è l'analisi esplorativa della presenza non italiana tra i micro imprenditori (SME¹) mediante l'utilizzo delle informazioni presenti nel Registro delle Imprese (ASIA-ISTAT). In particolare, nel *paper* l'attenzione è rivolta ad un gruppo specifico di lavoratori: i micro imprenditori extra-agricoli di origine straniera (FEAME).

Le informazioni contenute nel Registro delle Imprese sono di interesse per la crescente domanda di statistiche a supporto dell'analisi economica sul mondo dell'imprenditoria. Tali informazioni non possono, tuttavia, essere interpretate senza considerare le trasformazioni esogene – i.e. il cambiamento demografico e socio-economico dovuto alla crescente presenza della popolazione straniera nel Paese. Il trend positivo dei flussi migratori registrato negli ultimi anni a livello nazionale (+173% tra l'1/1/2003 e l'1/1/2011²) avvalorava l'attualità del percorso di analisi intrapreso, volto a considerare la dinamica migratoria e, con particolare riferimento al mondo dell'impresa, la vocazione imprenditoriale dei FEAME. Che sia frutto di una forte propensione all'iniziativa, all'investimento e al rischio, oppure di una strategia di inserimento e di stabilizzazione in Italia, quest'ultima emerge come uno tra i fenomeni di maggior interesse, sempre più centrale nella ricerca socio-economica. Nel *paper* si intendono approfondire i caratteri dell'imprenditoria straniera nella Regione Lombardia, fotografando le principali

¹ Nello specifico: *liberi professionisti, imprenditori individuali* (e lavoratori autonomi) con meno di 10 addetti e con fatturato inferiore a 2 milioni di euro (i.e. rispetto alla definizione EU, si escludono le forme societarie).

² <http://demo.istat.it/>.

dimensioni che caratterizzano l'apporto della compagine straniera alla dinamica produttiva ed economica di uno dei quattro motori regionali europei. È noto, infatti, che l'immigrazione è un fenomeno strutturale che interessa l'Italia nella sua interezza -fattore principale la sua collocazione geografica- e in modo significativo la Regione Lombardia per la sua caratterizzazione settoriale e produttiva (i.e. ancora rilevante la specializzazione industriale della Regione).

2. Metodi statistici e approcci di analisi

Dal trattamento statistico delle informazioni degli archivi del Registro delle Imprese è possibile derivare le caratteristiche anagrafiche degli imprenditori. Considerando il *codice Belfiore*, il codice amministrativo attribuito dal Catasto ad ogni Comune, si può, ad esempio, studiare l'imprenditorialità dal punto di vista del Paese (o località) di origine dell'imprenditore. Il dato fornisce un quadro degli imprenditori di origine non italiana (nel seguito *stranieri*), che possono comprendere sia cittadini stranieri immigrati sia cittadini italiani nati all'estero. Poiché nel Registro non è disponibile il dato sulla cittadinanza, tale informazione viene usata come *proxy* per individuare la popolazione degli imprenditori stranieri. Pur restituendo informazioni sulla provenienza geografica, è possibile una sovrastima degli imprenditori di origine straniera, in quanto sono conteggiati anche i nati all'estero ma da cittadini italiani, o una sottostima perché non tengono conto delle nuove generazioni nate in Italia e non in possesso del titolo della cittadinanza italiana. È evidente che per alcune nazioni di origine diventano necessari approfondimenti come nel caso dei Paesi che, nel secolo scorso, hanno conosciuto importanti flussi migratori di italiani (e.g. Francia, Germania, Svizzera) o delle ex-colonie italiane (i.e. Libia, Etiopia, Eritrea) o ancora degli Stati della ex-Jugoslavia (e in particolare dell'Istria -Fiume e Zara- che italiane dal 1919 al 1945, sono state successivamente annesse alla Jugoslavia di Tito). Nelle applicazioni sono state utilizzate evidenze empiriche su tipologie imprenditoriali (Blangiardo et al, 2011) basate su caratteristiche etniche (i.e. *Shaky, Chinese style*). Gli idealtipi individuati hanno valore empirico se sono strumenti validi di classificazione e se la teoria è in *status nascendi* (Klant, 1984; Boje, 1996) e pur non essendo ipotesi offrono un supporto alla loro costruzione (Weber, 1904).

3. Il contributo dei FEAME al sistema delle imprese in Lombardia

Il monitoraggio sulla presenza straniera realizzato dall'ORIM (Osservatorio Regionale per l'integrazione e la multietnicità, 2002-2010) indica un'evoluzione

del fenomeno migratorio non solo in termini di crescita quantitativa ma anche di mutamento strutturale delle principali caratteristiche socio-demografiche.

È di particolare interesse la presenza dei micro-imprenditori di origine straniera che nel 2007 ammontano a circa 40mila unità e rappresentano il 19,2% del totale dei micro-imprenditori stranieri in Italia e il 7,7% del totale dei micro-imprenditori della Lombardia. In Lombardia tre FEAME su quattro appartengono alla categoria degli imprenditori individuali (76%). Da un punto di vista strutturale, se rispetto alle classi di addetti non si evidenziano differenze distributive tra FEAME ed italiani (in entrambi i casi il 76% non ha ulteriori addetti), considerando, invece, il fatturato nella classe inferiore (<20mila euro) si rileva uno scarto di 10 punti percentuali (35% FEAME vs 25% micro-imprenditori italiani). Grazie alla piramide dell'età (Fig. 1) – da dove si può desumere una prospettiva imprenditoriale di crescita per i FEAME (mediamente più giovani rispetto al totale micro-imprenditori) – e allo studio della dimensione aziendale (Tab. 1 e 2) si hanno conferme sulla crescente rilevanza dei FEAME.

Considerando l'Archivio Asia al 2007 ed osservando dei FEAME la dotazione complessiva (8%) e quella relativa ai nati prima del 2005 (6%) e ancora attivi nel 2007 si può osservare l'evoluzione territoriale che assume rilevanza, ad esempio, nel Bresciano e nel Milanese (Appendice 1). Considerati, infine, i settori, le etnie e le tipologie emergono tratti di economia etnica sia in termini di concentrazione industriale (i.e. *Chinese Style*, Appendice 2) sia in una logica *shaky* dove la specializzazione nelle costruzioni di rumeni e albanesi riporta a sfumature opache sostanzialmente di auto-impiego *-shaky-* di muratori che lavorano con partita iva.

Tabella 1 – SME italiani e stranieri, per classe addetti, Lombardia, 2007 (%)

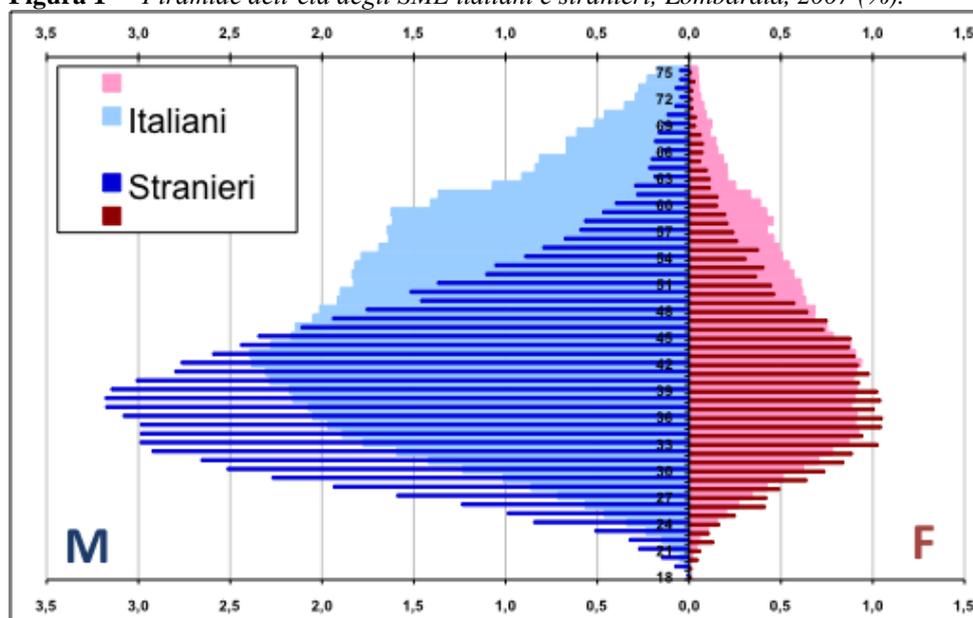
| Classi di addetti | Imprenditore individuale | | Libero professionista | | Totale SME | |
|-------------------|--------------------------|-------|-----------------------|-------|------------|-------|
| | Origine IT | FEAME | Origine IT | FEAME | Origine IT | FEAME |
| 1 addetto | 65% | 69% | 90% | 94% | 76% | 76% |
| 2-4 addetti | 30% | 26% | 9% | 6% | 21% | 21% |
| 5-9 addetti | 5% | 4% | 1% | – | 3% | 3% |
| SME | 100% | 100% | 100% | 100% | 100% | 100% |

Fonte: Elaborazioni su dati ASIA-ISTAT

Tabella 2 – SME italiani e stranieri, per classi di fatturato, Lombardia, 2007 (%)

| Classi di fatturato | Imprenditore individuale | | Libero professionista | | Totale SME | |
|---------------------|--------------------------|-------|-----------------------|-------|------------|-------|
| | Origine IT | FEAME | Origine IT | FEAME | Origine IT | FEAME |
| < 20mila € | 18% | 32% | 32% | 45% | 25% | 35% |
| 20mila-100mila € | 54% | 50% | 53% | 46% | 53% | 49% |
| 100mila-500mila € | 24% | 17% | 14% | 9% | 19% | 15% |
| 500mila-2milioni € | 4% | 1% | 1% | – | 2% | 1% |
| SME | 100% | 100% | 100% | 100% | 100% | 100% |

Fonte: Elaborazioni su dati ASIA-ISTAT

Figura 1– Piramide dell'età degli SME italiani e stranieri, Lombardia, 2007 (%).

Fonte: Elaborazioni su dati ASIA-ISTAT

Tabella 3 – Specializzazione settoriale dei FEAME, Lombardia 2007 (indici e %).

| | B-E (tot=1) | F (tot=1) | G (tot=1) | H (tot=1) | I (tot=1) | J-N (tot=1) | P-S (tot=1) | B-S (%riga) |
|----------------|----------------|--------------|--------------|--------------|--------------|----------------|----------------|----------------|
| Cina | 4,18 | 0,03 | 2,32 | 0,01 | 3,07 | 0,13 | 0,16 | 11% |
| Romania | 0,47 | 2,21 | 0,19 | 0,46 | 0,39 | 0,32 | 0,98 | 10% |
| Albania | 0,49 | 2,48 | 0,14 | 0,79 | 0,34 | 0,23 | 0,34 | 9% |
| Egitto | 0,68 | 1,56 | 0,39 | 0,55 | 2,21 | 1,07 | 0,26 | 9% |
| Marocco | 0,64 | 1,21 | 2,01 | 1,25 | 0,48 | 0,41 | 0,18 | 7% |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| Tunisia | 0,82 | 2,13 | 0,36 | 1,04 | 0,41 | 0,35 | 0,40 | 3% |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| Ex Jugoslavia | 0,74 | 1,49 | 0,60 | 0,64 | 0,40 | 0,99 | 0,91 | 2% |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| Totale | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | |
| Totale (% col) | 9% | 32% | 19% | 5% | 5% | 19% | 12% | 100% |

Fonte: Elaborazioni su dati ASIA-ISTAT

Note: NACE Rev 2.: B: mining and quarrying; C: manufacturing; D: electricity, gas, steam and air conditioning supply; E: water supply; sewerage, waste management and remediation activities; F: construction; G: wholesale and retail trade; repair of motor vehicles and motorcycles; H: transportation and storage; I: accommodation and food service activities; J: information and communication; K: financial and insurance activities; L: real estate activities; M: professional, scientific and technical activities; N: administrative and support service activities; P: education; Q: human health and social work activities; R: arts, entertainment, recreation; S: other service activities.

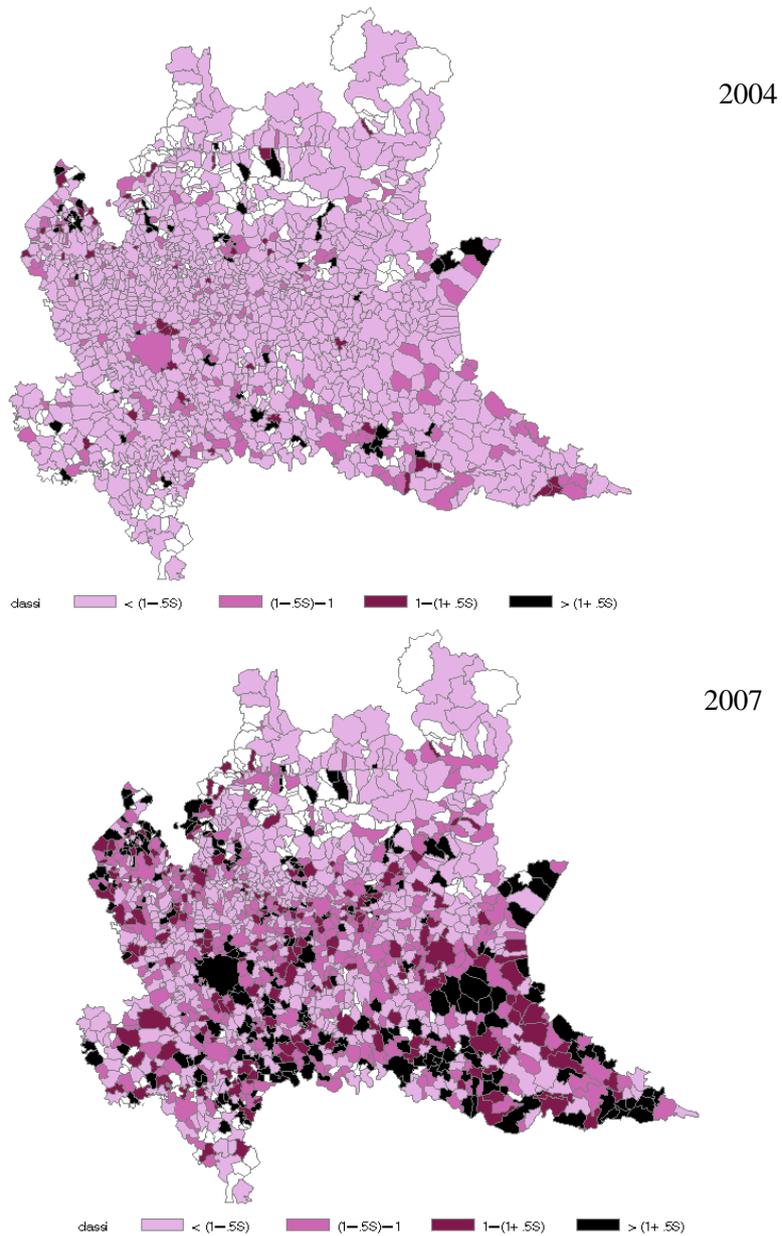
4. Conclusioni

In conclusione adottando un nuovo approccio orientato all'utilizzo di un *sistema di archivi interconnessi* (Mancini, 2010) è possibile ottenere statistiche originali sull'imprenditorialità.

Pur riconoscendo i limiti dell'unità di analisi, i risultati di questo *essay* sui FEAME della Lombardia restituiscono una rappresentazione articolata della presenza di origine straniera tra le fila dei micro-imprenditori, evidenziando aspetti originali nell'ambito occupazionale di grande interesse per lo studio delle potenzialità produttive che i territori esprimono. Emergono, quindi, differenze strutturali, concentrazioni territoriali, specializzazioni settoriali, ma anche tratti di economia etnica: da un lato lo studio rileva caratteri *shaky* (auto-impiego - e.g. muratori slavi, tunisini e albanesi), dall'altro si hanno evidenze di imprenditorialità industriale (i.e. *Chinese style*) con specializzazione settoriale di tipo intermedio nella sub-fornitura del tessile e dell'abbigliamento (Lombardi et al., 2011).

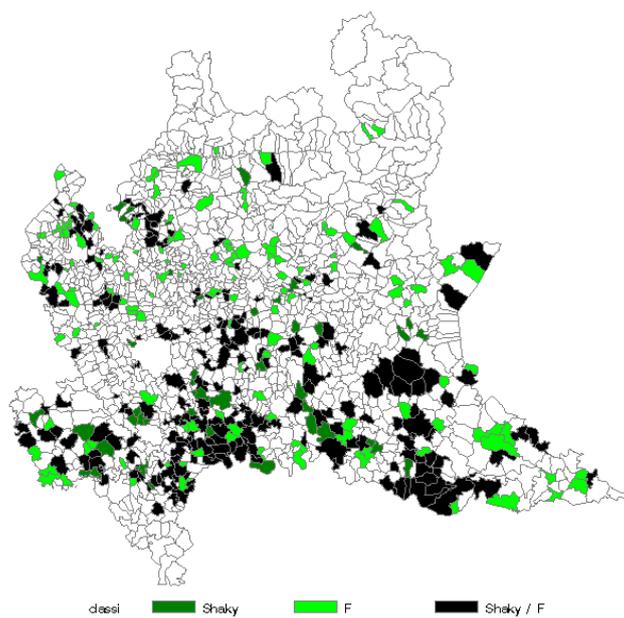
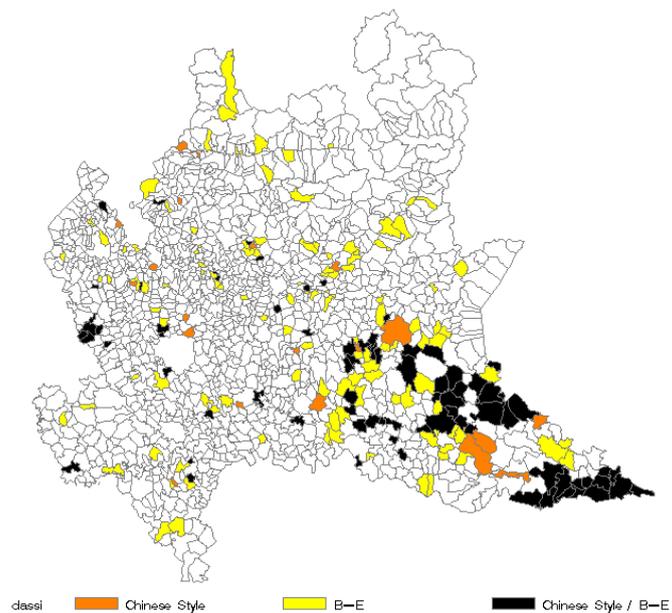
Riferimenti bibliografici

- BLANGIARDO G.C., DELLA QUEVA S., GROSSI P. and VERRECCHIA F. 2011. Foreign Entrepreneurship: The New Pillar of the National Entrepreneurial Vocation vs a Latest Shaky Employment. In *Atti del Convegno Intermedio SIS*, Bologna.
- BOJE T. 1996. Welfare state models in comparative research: do models describe the reality? In GREVE B. (Ed.), *Comparative welfare systems: the Scandinavian model in a period of change*, pp 13-27, London: Macmillan Press.
- KLANT J.J. 1984. *The rules of the game: the logical structure of economic theories*, Cambridge: Cambridge University Press.
- LOMBARDI S., DELLA QUEVA S., FARDELLI D., LORENZINI F., SFORZI F. 2011. Chinese Entrepreneurship in Context: Sector Specialization, Geographical Agglomeration and their Effects on Italian Local Production Systems- In *Atti del Convegno Intermedio SIS*, Bologna.
- MANCINI A. 2010 Il censimento delle imprese: una strategia basata su registri statistici assistiti da rilevazioni campionarie, ISTAT, Roma (9 November 2010) [Slides]
- MARTINI M. 1995 Necessità e possibilità di un registro statistico delle imprese in Italia. In: MARTINI M., BIFFIGNANDI S. (Eds.) *Il registro statistico europeo delle imprese*, Milano: Franco Angeli.
- OSSERVATORIO REGIONALE PER L'INTEGRAZIONE E LA MULTIETNICITÀ 2002-2010. *L'immigrazione straniera in Lombardia. Rapporti 2001-2009*, Regione Lombardia – Fondazione Ismu – Osservatorio Regionale per l'Integrazione e la Multiethnicità.
- WEBER M. 1904. Objectivity in social science and social policy. In SHILS E., FINCH H. (Eds.), *Methodology of the social sciences* (1949), New York: The Free Press., pp 49-112.

Appendice 1 – FEAME: dotazione, Lombardia, al 2004, 2007 (indici, 2007=1).

Note: 1. S: Scarto quadratico medio pari a 0.52.

Appendice 2 – FEAME: specializzazione settoriale / etnica, Lombardia 2007.



Note: 1. B-E (NACE rev. 2): industria in senso stretto; 2. F (NACE rev. 2): Costruzioni.

SUMMARY

Italy is becoming an established immigration country. Lombardy – among the northern regions – historically played a major role in attracting internal migration flows, and nowadays confirms itself as the hub of foreign immigration, in spite of the negative influences of the economic crisis.

In this framework, it's becoming more and more interesting to observe the quantitative structure of foreign business.

Recent works, carried out on a national basis, show that, beside the rise of self-employment, there are some empirical evidences of patterns of ethnic entrepreneurial activity. Other clues of this entrepreneurial spirit are reflected by studies concerning specific regional territories.

In this paper, we present an original deepening work about the features of foreign micro-entrepreneurship in Lombardy.

The methodology used in this essay follows a new approach that emphasizes the use of integrated databases for the recognition of economic aggregates, in the framework of study of the next business census.

The use of data from the ISTAT business statistical archive, indeed, gives the opportunity to deepen our knowledge about the phenomenon of immigrants participation in business-system, in a groundbreaking way. Such topic seems to be one of the emerging cores of the current socio-economic debate.

Stefania DELLA QUEVA, Istat, Rome, dellaqueva@istat.it

Daniela BONARDO, Istat, Rome, bonardo@istat.it

Alessio MENONNA, Fondazione Ismu, Milan, a.menonna@ismu.org

Flavio VERRECCHIA, Istat, Rome, verrecchia@istat.it

UN INDICE DI SVILUPPO UMANO DELLE REGIONI DELL'UNIONE EUROPEA¹

Pasquale De Muro, Matteo Mazziotta, Adriano Pareto, Valentina Talucci

1. Introduzione

Il presente lavoro ha l'obiettivo ambizioso e complesso di misurare un fenomeno multidimensionale quale lo sviluppo umano nell'Unione Europea. Alla luce dei grandi divari territoriali esistenti, non solo tra i paesi europei, ma soprattutto a livello infranazionale, riteniamo sia importante misurare le differenze tra i livelli di sviluppo a livello regionale. Inoltre, non ci limiteremo a calcolare l'Indice di sviluppo umano nella versione tradizionale dell'*United Nations Development Programme* (UNDP), ma affronteremo alcuni problemi metodologici relativi all'aggregazione degli indicatori elementari e confronteremo metodi di aggregazione alternativi.

Sebbene esistano altri lavori che hanno tentato di costruire un indice di sviluppo umano per il contesto europeo, si tratta di misurazioni fatte a livello nazionale (Costantini e Monni, 2005) oppure che non affrontano il problema del metodo di aggregazione (Monni, 1999; Bubbico e Dijkstra, 2011).

Il lavoro è diviso in tre parti. Nella prima (§2) presentiamo brevemente il quadro concettuale di riferimento per la misurazione dello sviluppo. Nella seconda (§3) passiamo in rassegna gli indicatori elementari scelti. Infine, nell'ultima parte (§4) evidenziamo i principali risultati emersi dall'analisi dei dati.

2. Quadro concettuale di riferimento

Esiste ormai un largo consenso internazionale, sia a livello scientifico sia a livello politico, sul fatto che la misurazione e la valutazione del livello di sviluppo o di benessere delle società non possono essere più effettuate basandosi

¹ Il lavoro è frutto della collaborazione congiunta degli autori. In particolare, i paragrafi 1 e 2 sono a cura di Pasquale De Muro, il paragrafo 3 è a cura di Valentina Talucci, il paragrafo 4 è a cura di Matteo Mazziotta e Adriano Pareto.

esclusivamente o principalmente sul prodotto interno lordo (PIL) pro capite.² Infatti, bisogna considerare che:

- lo sviluppo e il benessere (ed altri concetti simili, come il progresso, la qualità della vita) sono fenomeni multidimensionali che non possono essere valutati semplicemente con il PIL;
- il PIL non misura adeguatamente nemmeno il reddito personale, e inoltre costituisce tutt'al più un *mezzo* per raggiungere maggior sviluppo e benessere, e non una misura di questi fenomeni.

Per questi ed altri motivi, negli ultimi decenni, soprattutto a partire dagli anni Novanta, sono stati proposti diversi indicatori allo scopo di correggere il PIL, o di integrarlo, oppure addirittura di sostituirlo (Goossens et al., 2007). Quasi tutti gli indicatori dell'ultimo tipo sono indicatori sintetici, ossia che aggregano un certo numero (a volte anche elevato) di indicatori elementari che rappresentano le diverse dimensioni dello sviluppo: in questo modo si cerca di rappresentare la multidimensionalità del fenomeno. Uno dei problemi principali nella costruzione degli indicatori sintetici è quello dell'aggregazione degli indicatori elementari. Se si usa, come spesso accade, una media aritmetica semplice, si corre il rischio che eventuali valori molto differenti tra i vari indicatori elementari si compensino tra loro: in questo modo si assume implicitamente una perfetta sostituibilità tra le diverse dimensioni di sviluppo. Se si vuole evitare questo inconveniente bisogna dunque utilizzare metodi di aggregazione diversi che rendano imperfetta la sostituibilità. Per questo motivo, nel paragrafo 4 confronteremo i risultati ottenuti utilizzando metodi alternativi.

Fra i vari indicatori sintetici che sono stati proposti, l'Indice di Sviluppo Umano (ISU) elaborato dall'UNDP³ a partire dal 1990 è certamente uno dei più interessanti, poiché, a differenza di tanti altri, è costruito su una solida base teorica – il *capability approach* di Amartya Sen (1985) – ed ha una relativa semplicità di calcolo e interpretazione. Fino al 2009 il metodo di aggregazione usato per l'ISU era la media aritmetica; dal 2010 l'UNDP è passato alla media geometrica, proprio per evitare la perfetta sostituibilità tra le dimensioni.

Tuttavia, bisogna considerare che l'ISU originale è stato costruito allo scopo di fare confronti e classifiche tra nazioni a livello mondiale. Pertanto, volendolo applicare a livello regionale europeo, si rende necessario un duplice adattamento (cosa che del resto è stata prevista e auspicata dagli ideatori stessi dell'ISU):

² Numerosi lavori discutono questo problema. Il riferimento più recente e più importante è certamente il Rapporto Stiglitz-Sen-Fitoussi (2009).

³ Per maggiori informazioni sulla nascita e l'evoluzione di tutta la famiglia di indici di sviluppo elaborati dall'UNDP, nonché per i dati relativi a tutti i paesi si rinvia al sito dei Rapporti sullo Sviluppo Umano dell'UNDP: <http://hdr.undp.org>.

1. un adattamento per tenere conto del contesto europeo, che ha livelli di sviluppo umano piuttosto elevati. Ad esempio, non sarebbe il caso per le regioni della UE prendere in considerazione una variabile come l'alfabetizzazione, ma sarebbe meglio considerare i livelli di istruzione superiore;
2. un adattamento per tener conto della disponibilità dei dati. Non sempre, infatti, i dati che sono disponibili a livello nazionale sono altrettanto disponibili a livello regionale, e questo vale anche per la UE.

3. Gli indicatori di sviluppo umano in Europa

Com'è noto, la costruzione di un indice sintetico si divide nelle seguenti fasi: i) definizione del fenomeno oggetto di studio; ii) scelta e selezione degli indicatori; iii) standardizzazione e aggregazione degli indicatori (OECD, 2008). Concettualmente l'ISU indica i progressi raggiunti da un paese per realizzare beni e capacità essenziali per gli individui, empiricamente i 3 "pillars" entro cui vengono scelti gli indicatori sono: a) condurre una vita lunga e sana; b) conoscenza; c) condizioni di vita dignitose.

L'obiettivo di questo lavoro è proporre una versione dell'ISU applicata alle regioni europee (NUTS2). Come accennato in precedenza, è stato necessario fare un adattamento dell'ISU originario. L'adattamento ha riguardato gli indicatori elementari utilizzati dall'UNDP per rappresentare le 3 dimensioni di sviluppo umano: a) longevità, misurata come speranza di vita alla nascita; b) scolarizzazione, misurata come combinazione di anni medi di scuola e anni attesi di scuola; c) tenore di vita misurato dal prodotto interno lordo pro capite.

Per la prima dimensione abbiamo scelto due indicatori noti e consolidati anche per il calcolo del tradizionale ISU. Il primo fa riferimento alla demografia ed è la speranza di vita alla nascita, il secondo fa riferimento alla salute ed è il tasso di sopravvivenza. Questa prima dimensione si riferisce ad aspetti come l'accesso a servizi sanitari, la disponibilità di cure mediche e farmaci, la qualità dei luoghi di vita, studio e lavoro, che fanno sì che in una regione si viva a lungo e in modo sano. La seconda dimensione è stata rimodulata in virtù del contesto territoriale di analisi. Si tratta della dimensione della conoscenza, ben nota alla comunità scientifica in quanto è una delle *human capabilities* fondamentali, indispensabile affinché ciascun individuo possa esprimere liberamente le proprie attitudini e soddisfare i propri bisogni. Per questa ragione sono stati selezionati due indicatori: il primo, più consolidato e classico, è il tasso d'istruzione secondaria degli adulti; il secondo, di più recente concettualizzazione, è un indicatore di divario digitale che esprime quanto le famiglie europee accedano ad internet. L'impatto dell'ICT nella

vita delle persone non può più essere sottovalutato quando si studiano fenomeni come lo sviluppo umano o la qualità della vita o la disuguaglianza sociale perché, soprattutto nei paesi ad alto sviluppo umano, internet rappresenta un canale di accesso a beni e servizi (e-Commerce, e-Health, e-Government, e-Transport, ecc.) che pone linee di demarcazione molto precise tra coloro che sono “inclusi” o “esclusi” nell’era digitale. Infine, la terza dimensione si riferisce a condizioni di vita dignitose ed è stata rappresentata da due indicatori molto classici. Il primo è il reddito delle famiglie che esprime le capacità o possibilità di spesa e di consumo, in termini economici, di una famiglia. Tale indicatore è stato scelto sebbene si ritenga che in questo tipo di studi sia sempre più urgente sganciarsi da logiche puramente “monetarie” e di privazione materiale, a vantaggio delle altre misure. Il secondo indicatore è il tasso di occupazione.

Tabella 1 – Lista degli indicatori scelti.

| Dimensione | Area | Indicatore | Min | Max | M | S | CV |
|------------------------------|------------------|---------------------------------------------|---------|----------|----------|---------|------|
| Una vita lunga e sana | Demografia | Aspettativa di vita alla nascita | 72,0 | 82,9 | 79,3 | 2,7 | 3,4 |
| | Salute | Tasso di sopravvivenza | 347,9 | 1.846,4 | 992,0 | 185,5 | 18,7 |
| Conoscenza | Istruzione | Tasso di istruzione secondaria degli adulti | 11,2 | 79,4 | 48,3 | 15,6 | 32,2 |
| | Divario digitale | Tasso di accesso delle famiglie ad internet | 26,0 | 96,0 | 67,8 | 14,8 | 21,8 |
| Condizioni di vita dignitose | Economia | Reddito delle famiglie | 1.573,5 | 29.177,3 | 15.666,4 | 7.153,5 | 45,7 |
| | Lavoro | Tasso di occupazione | 40,8 | 80,2 | 64,4 | 7,8 | 12,1 |

Fonte: Eurostat

http://epp.eurostat.ec.europa.eu/portal/page/portal/region_cities/regional_statistics/data/database.

Lavorare rappresenta non solo la fonte primaria di reddito per ciascun individuo e dunque la possibilità di accedere a beni e servizi che possano garantire standard di vita dignitosi, ma anche uno strumento essenziale di inclusione e di coesione sociale. Un alto livello del tasso di occupazione è anche un attributo della regione e del paese di riferimento che ha evidentemente una struttura solida del mercato del lavoro.

Nella Tabella 1, vengono presentati i 6 indicatori scelti per le 3 dimensioni, sono riportati anche il valore minimo (Min), massimo (Max), la media aritmetica

(M), la deviazione standard (S) e il coefficiente di variazione (CV) che ciascun indicatore presenta nella distribuzione.⁴

4. Risultati empirici

In questo paragrafo sono illustrati i risultati relativi alla sintesi degli indicatori di sviluppo umano. Gli indicatori sono stati resi comparabili mediante trasformazione in variabili standardizzate con media 100 e scostamento quadratico medio pari a 10.

Tabella 2 – *Lo sviluppo umano delle regioni europee.*

| Cod. | Regione | Rango | | | Differenza assoluta di rango | | |
|-------------------|--------------------|-------|-----|-----|------------------------------|--------|--------|
| | | Ma | Mg | MPI | Ma-Mg | Ma-MPI | Mg-MPI |
| Prime 10 regioni | | | | | | | |
| NL31 | Utrecht | 1 | 1 | 1 | 0 | 0 | 0 |
| NL23 | Flevoland | 2 | 2 | 2 | 0 | 0 | 0 |
| NL32 | Noord-Holland | 3 | 3 | 3 | 0 | 0 | 0 |
| AT34 | Vorarlberg | 5 | 5 | 4 | 0 | 1 | 1 |
| SE11 | Stockholm | 4 | 4 | 5 | 0 | 1 | 1 |
| NL41 | Noord-Brabant | 7 | 7 | 6 | 0 | 1 | 1 |
| NL21 | Overijssel | 6 | 6 | 7 | 0 | 1 | 1 |
| NL33 | Zuid-Holland | 8 | 8 | 8 | 0 | 0 | 0 |
| NL22 | Gelderland | 9 | 9 | 9 | 0 | 0 | 0 |
| AT33 | Tirol | 10 | 10 | 10 | 0 | 0 | 0 |
| Ultime 10 regioni | | | | | | | |
| HU33 | Dél-Alföld | 226 | 225 | 225 | 1 | 1 | 0 |
| RO31 | Sud - Muntenia | 225 | 226 | 226 | 1 | 1 | 0 |
| RO41 | Sud-Vest Oltenia | 227 | 227 | 227 | 0 | 0 | 0 |
| HU32 | Észak-Alföld | 230 | 229 | 228 | 1 | 2 | 1 |
| BG42 | Yuzhen tsentralen | 228 | 228 | 229 | 0 | 1 | 1 |
| BG34 | Yugoiztochen | 229 | 230 | 230 | 1 | 1 | 0 |
| HU31 | Észak-Magyarország | 232 | 231 | 231 | 1 | 1 | 0 |
| BG33 | Severoiztochen | 231 | 232 | 232 | 1 | 1 | 0 |
| BG32 | Severen tsentralen | 233 | 233 | 233 | 0 | 0 | 0 |
| BG31 | Severozapaden | 234 | 234 | 234 | 0 | 0 | 0 |

La sintesi è stata condotta secondo i seguenti criteri:

a) media aritmetica degli indicatori standardizzati (Ma);

⁴ Si noti che l'uso dei valori estremi (i valori del paese che sta peggio e di quello che sta meglio), per la standardizzazione degli indicatori, rende difficili i confronti temporali.

b) media geometrica degli indicatori standardizzati (Mg);

c) Mazziotta-Pareto Index (De Muro et al., 2010) con penalità negativa (MPI).

Nella Tabella 2 sono riportate le graduatorie delle prime 10 e delle ultime 10 regioni europee, con le differenze assolute di rango tra i diversi approcci adottati.

Nel complesso, i risultati più simili si ottengono con Mg e MPI (approccio non-compensativo), cui corrisponde una differenza media assoluta di rango di 1,4 posizioni. Ciò è dovuto alla capacità intrinseca della media geometrica di “penalizzare” le distribuzioni con valori sbilanciati degli indicatori. Per contro, le divergenze maggiori si osservano con Ma (approccio compensativo) e MPI (differenza media assoluta di rango uguale a 2,6).

La Tabella 3 illustra le posizioni delle regioni italiane nella classifica europea dello sviluppo umano. Si noti che, passando da un approccio compensativo a uno non compensativo, alcune regioni guadagnano posizioni (per es. l’Abruzzo).

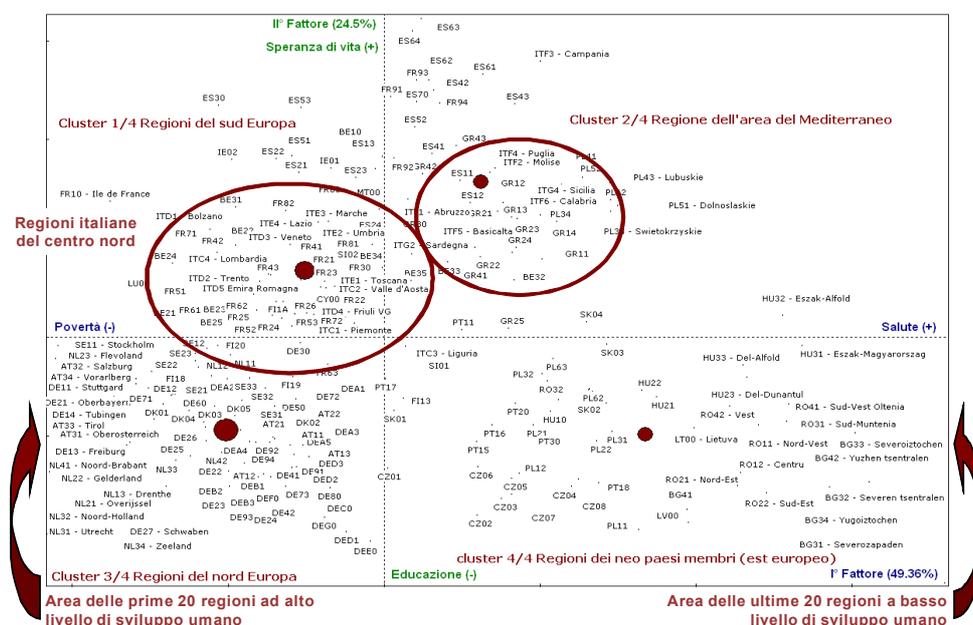
Tabella 3 – *Lo sviluppo umano delle regioni italiane.*

| Cod. | Regione | Rango | | | Differenza assoluta di rango | | |
|------|--------------------|-------|-----|-----|------------------------------|--------|--------|
| | | Ma | Mg | MPI | Ma-Mg | Ma-MPI | Mg-MPI |
| ITD1 | P. A. Bolzano | 49 | 52 | 53 | 3 | 4 | 1 |
| ITD2 | P. A. Trento | 74 | 73 | 74 | 1 | 0 | 1 |
| ITC4 | Lombardia | 78 | 77 | 78 | 1 | 0 | 1 |
| ITD3 | Veneto | 99 | 100 | 98 | 1 | 1 | 2 |
| ITD5 | Emilia-Romagna | 104 | 104 | 103 | 0 | 1 | 1 |
| ITE4 | Lazio | 113 | 113 | 113 | 0 | 0 | 0 |
| ITC2 | Valle d'Aosta | 114 | 114 | 114 | 0 | 0 | 0 |
| ITC1 | Piemonte | 121 | 121 | 121 | 0 | 0 | 0 |
| ITD4 | Friuli - V. Giulia | 125 | 122 | 122 | 3 | 3 | 0 |
| ITE1 | Toscana | 128 | 128 | 126 | 0 | 2 | 2 |
| ITG2 | Sardegna | 123 | 125 | 127 | 2 | 4 | 2 |
| ITE3 | Marche | 131 | 131 | 131 | 0 | 0 | 0 |
| ITE2 | Umbria | 147 | 147 | 145 | 0 | 2 | 2 |
| ITF1 | Abruzzo | 156 | 152 | 150 | 4 | 6 | 2 |
| ITC3 | Liguria | 154 | 155 | 153 | 1 | 1 | 2 |
| ITF5 | Basilicata | 150 | 151 | 154 | 1 | 4 | 3 |
| ITF4 | Puglia | 162 | 164 | 165 | 2 | 3 | 1 |
| ITF6 | Calabria | 166 | 167 | 170 | 1 | 4 | 3 |
| ITG1 | Sicilia | 172 | 175 | 180 | 3 | 8 | 5 |
| ITF2 | Molise | 187 | 191 | 191 | 4 | 4 | 0 |
| ITF3 | Campania | 203 | 204 | 205 | 1 | 2 | 1 |

In conclusione, volendo tracciare una geografia dello sviluppo umano nelle regioni europee si è scelto di effettuare un’analisi in componenti principali degli indicatori elementari e successivamente una *cluster analysis*. In prima battuta, occorre esplicitare il significato dei fattori estratti.

Il primo fattore che spiega il 49,36% dell'inerzia rivalutata, intercetta nel semiasse positivo (+) la dimensione dello sviluppo connessa ad una vita lunga e sana, mentre il semiasse negativo (-) rappresenta condizioni di vita dignitose legate alla povertà monetaria e al reddito. Il secondo fattore che spiega il 24,5% dell'inerzia si caratterizza nel semipiano negativo (-) per la dimensione della conoscenza, mentre in quello positivo (+) per la speranza di vita alla nascita. Proiettando sul piano fattoriale ottenuto i singoli profili riga (regioni) e i cluster ottenuti è possibile apprezzare interessanti configurazioni (Figura 1). Per quanto riguarda la cluster analysis, è stato utilizzato un metodo gerarchico che ha prodotto 4 gruppi che si posizionano in modo chiaro sul piano sopra descritto.

Figura 1 – Un approccio multivariato: la cluster analysis.



In particolare, nel primo cluster ricadono le regioni del Sud Europa che si posizionano nel quarto quadrante, dove si apprezza una speranza di vita elevata e un basso livello povertà. È molto interessante notare che ricadono in questo cluster anche le regioni del Nord Italia, il che significa che a livello nazionale l'Italia risente di un "effetto-Sud", ma a livello regionale sono le regioni del Centro-Nord che vengono "attratte" maggiormente verso le altre regioni dei paesi del Sud Europa, che si caratterizzano per un buon livello di sviluppo. Il secondo cluster ricade nel primo quadrante che intercetta le regioni dell'area del mediterraneo e si caratterizza rispetto alle dimensioni di sviluppo scelte come quello in cui le

condizioni di vita sono soprattutto connesse a un buon livello di salute e longevità. Le restanti regioni del Sud Italia si collocano esattamente in questo cluster. Il terzo gruppo ricade nel terzo quadrante ed intercetta le regioni del Nord Europa che si caratterizzano per un alto livello di sviluppo, soprattutto in relazione a condizioni di vita dignitose, ossia un basso tasso di povertà e un alto livello di conoscenza e, dunque, un tasso di istruzione elevato e un basso divario digitale. Infine, il quarto gruppo ricade nel secondo quadrante e intercetta le regioni dei neo paesi membri e dell'Est europeo. Si tratta di regioni in cui le dimensioni che pesano maggiormente sono la salute e la conoscenza, che vanno probabilmente letti in termini di crescita e di sviluppo rispetto all'informatizzazione e all'istruzione.

Riferimenti bibliografici

- BUBBICO R.L., DIJKSTRA L. 2011. The European regional Human Development and Human Poverty Indices, *Regional Focus*, No. 2, European Union.
- COSTANTINI V., MONNI S. 2005. Sustainable Human Development for European countries, *Journal of Human Development*, Vol. 6, No. 3, pp. 329-351.
- DE MURO P., MAZZIOTTA M., PARETO A. 2010. Composite Indices of Development and Poverty: An Application to MDGs, *Social Indicators Research*, Vol. 104, No. 1, pp. 1-18, doi: 10.1007/s11205-010-9727-z.
- GOOSSENS Y. et al. 2007. Alternative progress indicators to Gross Domestic Product (GDP) as a means towards sustainable development, IP/A/ENVI/ST/2007-10, Study provided for the European Parliament's Committee on the Environment, Public Health and Food Safety.
- MONNI S. 1999. A convergence analysis of human development, *Dipartimento di Economia. Working Paper*, No. 11, Università degli Studi Roma Tre.
- OECD 2008. *Handbook on Constructing Composite Indicators. Methodology and user guide*. Parigi: OECD Publications.
- SEN A. 1985. *Commodities and capabilities*. Amsterdam: North-Holland.
- STIGLITZ J., SEN A., FITOUSSI J.P. 2009. Report of the commission on the measurement of economic performance et social progress, CMEPSP, September 14.

SUMMARY

The aim of this work is to define and measure a multidimensional phenomenon as the development in Europe. The results of the application show that an efficient approach is both selecting few non-compensatory indicators in order to represent the phenomenon and choosing a simple aggregation function that does not permit a compensation among the values of the indicators.

Pasquale DE MURO, Università di "Roma Tre", demuro@uniroma3.it

Matteo MAZZIOTTA, Istat, mazziott@istat.it

Adriano PARETO, Istat, pareto@istat.it

Valentina TALUCCI, Istat, talucci@istat.it

L'UTILIZZO DEGLI ARCHIVI AMMINISTRATIVI PER IL MONITORAGGIO DELLE POLITICHE DEL LAVORO: IL RUOLO DELLE IMPRESE¹

Grazia Di Bella, Daniela Bonardo, Lorena Galiè, Valentina Talucci

1. Strategie di monitoraggio delle politiche del lavoro

Questo lavoro intende esplorare le potenzialità offerte dall'integrazione di microdati di fonte amministrativa con dati statistici disponibili in Istat per lo studio dei fenomeni socio-economici nell'ambito del mercato del lavoro.

Selezionando ed integrando diverse fonti è possibile, infatti, produrre una ricchezza sinergica di analisi mediante l'associazione di attributi specifici delle imprese, dei lavoratori, dei luoghi di lavoro e dei contesti sociali, economici e territoriali.

In particolare viene presentato un approccio metodologico per il monitoraggio delle politiche del lavoro con una applicazione sperimentale alla popolazione dei lavoratori con disabilità assunti sulla scorta della Legge 68 del 1999.

Gli studi e le ricerche applicate nell'ambito della valutazione delle politiche pubbliche costituiscono una frontiera per la statistica in relazione alla disponibilità di dati e alla definizione di opportuni modelli di analisi. La possibilità di misurare gli effetti delle politiche con l'evidenza empirica dei dati è certamente un passo importante verso la valutazione in un'ottica di analisi delle relazioni causali: in tal senso la valutazione comprende l'individuazione della "popolazione obiettivo" della politica, l'esplicitazione dell'azione pubblica intesa come "trattamento" e la misurazione dell'efficacia al netto degli effetti esterni non connessi all'azione politica (Trivellato, 2009).

Nella complessa attività di valutazione delle politiche pubbliche orientate al mercato del lavoro, momento strategico ed essenziale è il monitoraggio del processo di attuazione della politica e dei suoi effetti, finalizzato a determinare i progressi che la politica ha apportato verso gli obiettivi prefissati.

¹ Il lavoro è frutto della collaborazione congiunta degli autori. I paragrafi 1 e 2.1 sono a cura di G. Di Bella, il paragrafo 2.2 è a cura di L. Galiè, il paragrafo 3.1 è a cura di D. Bonardo, il paragrafo 3.2 è a cura di V. Talucci. La sperimentazione si inserisce nell'ambito delle attività della Convenzione stipulata tra Istat e Isfol.

La disponibilità di dati amministrativi di natura fiscale, contributiva e previdenziale relativi alle imprese e ai lavoratori integrati con i dati statistici permette di definire una base informativa anche di tipo longitudinale, di rilevante utilità per l'implementazione di uno strumento analitico di monitoraggio.

Tale strumento deve comprendere, nella fase iniziale, gli elementi descrittivi dell'azione-trattamento e gli eventuali meccanismi attuativi. In particolare occorre evidenziare e formalizzare: la popolazione obiettivo, la durata dell'azione (annuale, pluriennale, ecc.), gli eventuali criteri di obbligatorietà, i possibili livelli geografici dell'azione, l'attivazione o meno di interventi locali di accompagnamento, ed altri possibili aspetti utili alla sua caratterizzazione. Rimane fondamentale tenere in considerazione gli elementi di contesto in cui viene applicata la politica quali la struttura produttiva locale, le dinamiche economiche nazionali e mondiali.

Nel paragrafo 2 vengono presentate le fonti utilizzate e viene brevemente descritto il sistema di integrazione e di trattamento dei dati, nel paragrafo 3 è presentato il caso di studio delle politiche di inserimento dei lavoratori disabili.

2. La base dei dati utilizzata

2.1. Le fonti statistiche e amministrative

Il sistema informativo utilizzato per la sperimentazione è frutto di un processo di integrazione tra fonti amministrative e fonti statistiche. In particolare l'ancoraggio delle informazioni e la fase di validazione si basano principalmente sull'Archivio delle imprese dell'Istat, Asia (Istat, 2011), a cui sono stati agganciati, attraverso tecniche di record linkage, i dati riferiti alle posizioni contributive dei dipendenti delle imprese. Questi dati derivano da fonte Inps ed in particolare dalle dichiarazioni mensili effettuate dai datori di lavoro (dichiarazioni Emens).

Ulteriori informazioni relative agli individui provengono dall'Archivio delle persone fisiche, disponibile in Istat, a sua volta frutto dell'integrazione di più fonti, in cui sono raccolti dati di carattere demografico, economico e sociale delle persone fisiche titolari di codice fiscale.

L'archivio dell'Inps contiene informazioni dettagliate sulle denunce nominative dei contributi versati dai datori di lavoro per ciascun lavoratore dipendente e presenta l'interessante caratteristica di permettere una integrazione di tipo Leed (*Linked Employer-Employee Data*) (Bryson, Forth, Barber, 2006), ovvero sono presenti chiavi di record linkage sia per i lavoratori che per le imprese. Grazie a questa caratteristica le informazioni sono utilizzabili in modo integrato a livello di microdati: dal lato del lavoratore è possibile agganciare le informazioni derivanti dall'Archivio delle persone fisiche; dal lato del datore di lavoro, gli attributi propri dell'Archivio Asia (forma giuridica, classe di addetti, ecc) arricchiscono la base dei

dati. Utilizzando il sistema informativo integrato è possibile: a) osservare la presenza dei lavoratori nel mercato del lavoro; b) analizzare alcune caratteristiche della posizione lavorativa; c) individuare specifiche sotto-popolazioni target; d) tracciare percorsi lavorativi individuali; e) analizzare il ruolo delle imprese.

Occorre tuttavia evidenziare che le finalità dei dati amministrativi sono connesse all'applicazione di una norma legislativa e non hanno un diretto obiettivo statistico; per utilizzare tali dati a fini statistici, e garantire i risultati dell'integrazione, è in corso un processo di "validazione" a partire dall'analisi dei metadati, attraverso la realizzazione di piani di controllo specifici in conformità agli standard di qualità propri della statistica ufficiale. Si sottolinea, inoltre, che queste basi dati contengono un numero di record molto elevato, dell'ordine delle decine di milioni e necessitano, quindi, di particolari trattamenti.

2.2. Le unità e le variabili di analisi

L'unità originaria della fonte amministrativa delle dichiarazioni dell'Inps è stata denominata "profilo contributivo" definito come *l'insieme delle caratteristiche che identificano le tipologie di contribuzione riconducibili all'Inps relative ad un contratto con un datore di lavoro*. Tali caratteristiche, sempre nell'ambito della stessa posizione lavorativa, possono variare nel corso dell'anno, dando luogo ad un nuovo profilo contributivo. L'utilizzabilità degli archivi amministrativi disponibili a livello di microdati presuppone che una popolazione statistica sia contenuta nell'insieme delle unità enumerate nell'archivio amministrativo. Il sistema integrato descritto nel paragrafo precedente permette di definire la popolazione target delle *Posizioni dell'anno t dei lavoratori dipendenti presso le imprese private del registro Asia*². L'individuazione della popolazione deriva dall'analisi delle unità originarie degli archivi amministrativi e dal risultato empirico di integrazione delle fonti. Oltre all'unità statistica della posizione lavorativa sono state prese in considerazione anche le singole unità dell'*Impresa* e del *Lavoratore* da essa derivanti. L'operazione di trattamento dei dati ha l'obiettivo di enumerare le unità delle popolazioni statistiche in esame e di associare correttamente a ciascuna unità gli attributi propri.

² Il registro ASIA è costituito da tutte le imprese localizzate nel territorio italiano attive nell'anno di riferimento. Ad oggi non sono comprese nel campo di osservazione le attività economiche relative a: agricoltura, silvicoltura e pesca (sezione A della classificazione NACE Rev. 2); amministrazione pubblica e difesa; assicurazione sociale obbligatoria (sezione O); attività di organizzazioni associative (divisione 94); attività di famiglie e convivenze come datori di lavoro per personale domestico; produzione di beni e servizi indifferenziati per uso proprio da parte di famiglie e convivenze (sezione T); organizzazioni ed organismi extraterritoriali (sezione U); le unità classificate come istituzioni pubbliche e istituzioni private non profit.

Non viene affrontato in questa sede l'aspetto della copertura delle popolazioni che, per quanto riguarda le imprese è assicurata dal processo di produzione di Asia e per quanto riguarda le posizioni lavorative (e i lavoratori) attiene agli adempimenti legislativi dei datori di lavoro e dipende dalla qualità del processo di raccolta dei dati amministrativi; l'analisi statistica preliminare di trattamento utilizza tutte le informazioni disponibili presso il produttore dei dati amministrativi e sfrutta i possibili confronti con registri di popolazioni o sotto-popolazioni validati.

L'attributo della posizione lavorativa più interessante, ai fini di questo lavoro, è il codice che identifica le particolarità contributive utili alla determinazione dell'ammontare delle contribuzioni dovute dai datori di lavoro all'Inps. Tra queste particolarità è possibile individuare i casi di sgravi contributivi beneficiati dalle imprese in seguito all'assunzione di lavoratori oggetto di specifiche politiche attive del mercato del lavoro. Opportune elaborazioni consentono di costruire indicatori che definiscono le sotto-popolazioni su cui è stata applicata la politica. Altri attributi che descrivono la posizione lavorativa, insieme alle caratteristiche proprie delle imprese e dei lavoratori, completano il quadro informativo.

Per la determinazione della base dati utile al monitoraggio vengono, quindi, costruiti dei *dataset* annuali in cui le sotto-popolazioni sono enumerate e caratterizzate in un'ottica Leed. Attualmente è in esame la serie storica 2005-2009.

3. Il caso di studio: l'inserimento dei disabili

3.1. L'applicazione della Legge 68/1999

Nell'ambito delle numerose iniziative legislative che promuovono percorsi di inserimento o di mantenimento del posto di lavoro, è interessante analizzare l'applicazione della Legge 68/1999 'Norme per il diritto al lavoro dei disabili' (e il successivo regolamento di attuazione D.P.R. 10.10.2000 n. 333). La legge incoraggia l'inserimento lavorativo delle persone con disabilità, prevedendo un sistema integrato di servizi di supporto per il collocamento mirato e azioni volte alla soluzione dei problemi connessi con gli ambienti, gli strumenti e le relazioni interpersonali sui luoghi quotidiani di lavoro. La legge definisce, inoltre, per i datori di lavoro pubblici e privati, le soglie per le assunzioni obbligatorie in base al numero di occupati³, ed include un insieme di facilitazioni per le assunzioni a

³ «I datori di lavoro pubblici e privati sono tenuti ad avere alle loro dipendenze lavoratori disabili in una quota pari al 7 per cento dei lavoratori occupati se occupano più di 50 dipendenti; due lavoratori se occupano da 36 a 50 dipendenti; un lavoratore se occupano da 15 a 35 dipendenti» (Legge 68/1999, art. 3 - Assunzioni obbligatorie. Quote di riserva – comma 1).

tempo indeterminato. La popolazione dei beneficiari della legge, ovvero la “popolazione obiettivo” della politica, è costituita: dalle persone in età lavorativa affette da minorazioni che comportino una riduzione della capacità lavorativa superiore al 45%, dagli invalidi del lavoro con un grado di invalidità superiore al 33% accertato dall’Inail, dai non vedenti o i sordomuti, dagli invalidi di guerra, invalidi civili di guerra, invalidi per servizio. L’inclusione lavorativa è promossa attraverso la concessione di agevolazioni alle imprese che consistono nella riduzione totale o parziale dei contributi previdenziali ed assistenziali per ogni lavoratore disabile assunto. Le agevolazioni concesse hanno una specifica durata: otto anni per i lavoratori con una riduzione della capacità lavorativa superiore al 79% o con handicap intellettuale e psichico; cinque anni per i lavoratori con riduzione della capacità lavorativa compresa tra il 67% e il 79%.

Tali sgravi contributivi, espressamente codificati all’interno dell’archivio Inps, permettono di selezionare la popolazione dei disabili che hanno beneficiato della fiscalizzazione prevista dalla L. 68. Occorre precisare che la L. 247/2007 ha trasformato le “agevolazioni” in “incentivi” per tutte le assunzioni derivate dalle Convenzioni tra il datore di lavoro e il Centro per l’impiego stipulate a partire dal 1° gennaio 2008.

Gli obiettivi della sperimentazione, commisurati alla disponibilità dei dati, sono di verificare la presenza dei disabili nelle imprese nel corso del periodo analizzato e di monitorarne la permanenza nell’ambito della durata dello sgravio contributivo.

Se il primo obiettivo è realizzabile attraverso un’analisi di tipo *cross section* (Bonardo, Di Bella, Galiè, Talucci, 2011), il secondo è possibile solo adottando un’ottica di tipo longitudinale. Una proposta di applicazione longitudinale è presentata nel prossimo paragrafo.

3.2. *Analisi per generazione*

L’applicazione illustrata in questo paper considera la coorte dei lavoratori disabili assunti nel corso dell’anno 2006 in virtù della Legge 68. L’analisi presentata e la sua possibile generalizzazione sono ancora in fase di sperimentazione e di test.

Dal punto di vista concettuale l’identificazione della coorte è stata effettuata nel seguente modo: sono stati selezionati tutti i lavoratori con la presenza del codice di agevolazione contributiva associato alla Legge 68 per il *dataset* dell’anno 2006, ma non aventi questa condizione nell’anno precedente ipotizzando, quindi, che l’assunzione sia avvenuta nel corso del 2006. Per seguire la coorte di lavoratori nei vari anni è stata effettuato un *record linkage* tra i dataset annuali utilizzando come chiave la variabile identificativa del lavoratore. L’arco temporale osservato è 2006-2009. In generale la scelta della lunghezza del periodo da analizzare deve

considerare la durata prevista dell'azione della politica ed eventuali possibili modifiche della normativa di riferimento. Nel caso della Legge 68 la durata degli sgravi contributivi, come descritto precedentemente, varia da 5 a 8 anni; la scelta del periodo osservato è esemplificativa e dipende solamente dalla disponibilità attuale dei dati (2005-2009). La formalizzazione della definizione delle sottopopolazioni è la seguente:

$$Cdis^{2006}(2006) = Pdis(2006) - \{Pdis(2006) \cap Pdis(2005)\} \quad (1)$$

dove:

***Pdis*(t)** è l'insieme dei lavoratori che hanno ricoperto almeno una posizione lavorativa dipendente presso le imprese private del registro Asia nel corso dell'anno t, avente il codice di agevolazione contributiva associato alla Legge 68.

***Cdis*²⁰⁰⁶(2006)** è la coorte dei lavoratori disabili assunti almeno una volta nel corso del 2006 in virtù della Legge 68.

La definizione della coorte osservata negli anni successivi ***Cdis*²⁰⁰⁶(t)** è determinata dalla seguente relazione:

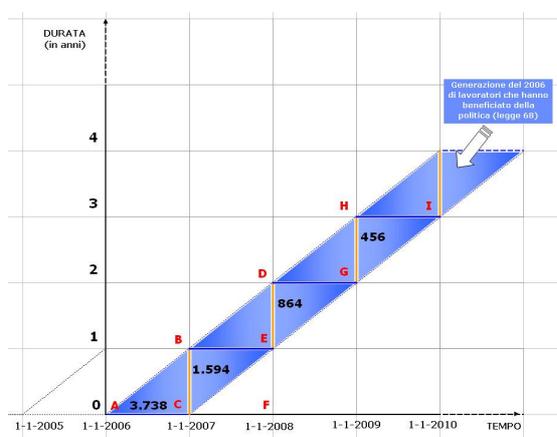
$$Cdis^{2006}(t) = Cdis^{2006} \bigcap_{i=1}^{t-2006} Pdis(2006+i), \quad t = 2007, 2008, 2009 \quad (2)$$

I dati disponibili consentono, quindi, di costruire una struttura longitudinale in grado di seguire lo stesso insieme di lavoratori nell'arco di tempo osservato grazie all'eshaustività del *dataset* nei vari anni. Per presentare i dati della sperimentazione si è preso in prestito il consolidato diagramma di Lexis, ben noto in demografia, che permette di seguire il formarsi e l'estinguersi della generazione di lavoratori beneficiari della politica. Nella Figura 1 sono riportati i dati relativi alle coorte iniziale ***Cdis*²⁰⁰⁶** e alle sottopopolazioni ***Cdis*²⁰⁰⁶(t)**, per t = 2007, 2008, 2009. Come si può osservare il segmento AC rappresenta la generazione di partenza, che è composta dai 3.738 lavoratori disabili assunti dalle imprese grazie alla Legge 68 nel corso del 2006. Il segmento BC rappresenta i lavoratori che passando nel 2007 mantengono la condizione di lavoratore beneficiario della Legge 68, questi sono pari a 1.594 (nel diagramma di Lexis sono "contemporanei"), il triangolo ABC rappresenta, invece, i lavoratori il cui beneficio di legge si interrompe nel corso del 2006 prima del compimento del primo anno. L'area dei triangoli BCE e BDE rappresenta i lavoratori per cui l'azione della politica cessa precocemente nel corso del 2007 prima dello scadere dei due anni. Continuando nell'osservazione del *data set* del 2008, è possibile individuare i lavoratori della coorte che mantengono la condizione di beneficiari nel corso dell'anno (segmento DE) e che risultano pari a 864 unità. Infine il segmento HG rappresenta i lavoratori che alla fine del periodo

di osservazione (2009) sopravvivono nella condizione, ovvero per i quali la policy ha durata continuativa di almeno 2 anni.

A questo punto si ha un quadro del percorso di utilizzazione della Legge 68 nel quadriennio 2006-2009: di anno in anno una quota rilevante di lavoratori si “estingue” dalla condizione; al termine del periodo di osservazione circa 1 lavoratore su 8 sopravvive nell’archivio nella condizione di partenza.

Figura 1 – Adattamento del diagramma di Lexis sulla generazione di lavoratori disabili assunti dalle imprese nel corso del 2006 grazie alla Legge 68/1999.



Per quanto riguarda i 3.282 lavoratori della coorte *Cdis*²⁰⁰⁶ estinti dalla condizione nel corso del periodo di osservazione, è possibile verificare la loro presenza/assenza nella popolazione target delle *Posizioni dei lavoratori dipendenti presso le imprese private del registro Asia*. I possibili “stati” sono: (a) dipendente presso la stessa impresa ma senza agevolazioni fiscali; (b) dipendente presso un’altra impresa; (c) non presente nella popolazione, ovvero non più nella condizione di lavoratore dipendente presso le imprese. Rispetto a quest’ultima modalità, si possono individuare ulteriori segnali di presenza nell’Archivio delle persone fisiche; in particolare è possibile tracciare le seguenti transizioni: (i) permanenza nel mercato del lavoro in altre posizioni (dipendente presso la pubblica amministrazione, lavoratore parasubordinato, imprenditore,...) (ii) assenza dal mercato del lavoro; (iii) decesso.

4. Conclusioni

Questa sperimentazione indica che la linea di ricerca individuata consente interessanti sviluppi. L’ulteriore analisi volta ad individuare le specifiche posizioni

lavorative e, conseguentemente, a ricostruire le storie lavorative a livello longitudinale delinea un potenziale strumento di monitoraggio in grado di seguire le coorti di lavoratori per verificarne lo stato nei vari anni, calcolare la frequenza delle transizioni, calcolare i tempi di permanenza negli stati, valutare il comportamento delle imprese rispetto alla politica in esame.

Riferimenti bibliografici

- BONARDO, D., DI BELLA G, GALIÈ, L., TALUCCI, V. 2011. The labour market participation of people with disability: key challenges for the use of administrative data for statistical purposes. *Proceedings of the Workshop on Enhancement and Social Responsibility of Official Statistics*, Rome: 28-29 April 2011, pp. 21-22.
- BRYSON, A., FORTH, J. and BARBER, C. 2006. (Eds) *Making Linked Employer-Employee Data Relevant to Policy*, DTI Economics Occasional Paper No. 4, Department of Trade and Industry, London.
- ISTAT. 2011. Struttura e dimensione delle imprese - Anno 2009, Archivio Statistico delle Imprese Attive (Asia), *Statistiche Report*. 1 giugno 2011.
- ISTAT. 2009. *La disabilità in Italia. Il quadro della statistica ufficiale*, Roma.
- TRIVELLATO U. 2009. *La valutazione degli effetti di politiche pubbliche: paradigma e pratiche*. Discussion Paper series, IRVAPP.

SUMMARY

The use of administrative data provides new opportunities for monitoring labour market policies. However a treatment process has to be implemented to make data useful for statistical purposes. This paper presents a first experience aimed to investigate the participation of people with disabilities in the labour market, Starting from the linkage between social security data - that has Linked Employer Employee Data characteristic - and the Italian Business Register ASIA - that includes information on structural and demographic characteristics of all active businesses – it is possible to identify persons with disabilities employed in enterprises through the implementation of active policy. A data treatment process, taking into account both the analysis of the populations and statistical variables, was developed. The application presented is based on a longitudinal approach as a result of a record linkage among data referred to different years. The results obtained allow to monitor the effects of policies of job placement of disabled workers.

Grazia DI BELLA, Istat; dibella@istat.it
Daniela BONARDO, Istat; bonardo@istat.it
Lorena GALIÈ, Istat; galie@istat.it
Valentina TALUCCI Istat; talucci@istat.it

UNA NUOVA PROCEDURA DI IMPUTAZIONE DI DATI MANCANTI BASATA SUGLI ALBERI DI DECISIONE

Agostino Di Ciaccio, Giovanni Maria Giorgi

1. Introduzione

In statistica, si parla di *overfitting* quando un modello statistico tende a riprodurre l'errore casuale osservato nei dati campionari. L'*overfitting* si verifica in genere quando un modello è eccessivamente complesso, con un numero di parametri elevato in relazione alla quantità dei dati disponibili. Un modello che è in *overfitting* avrà generalmente una scarsa performance predittiva, in quanto può amplificare lievi fluttuazioni nei dati, soprattutto per variabili esplicative che assumono valori fuori dal range dei valori osservati nel campione.

Per un modello regressivo, possiamo dire che esso è in *overfitting* sui dati analizzati se ricostruisce non solo il valore atteso della variabile dipendente ma anche il suo residuo. Il risultato è che la misura di adattamento del modello ai dati osservati (*apparent error*, si veda Efron 2004) è molto più ottimistica dell'errore di previsione sull'intera popolazione (*prediction error*).

Data una variabile dipendente quantitativa Y e k variabili esplicative X_j , in caso di *overfitting* il modello tenderà ad approssimare $E(Y_i | \mathbf{x}_i) + e_i$ invece del valore atteso $E(Y_i | \mathbf{x}_i)$, ossia tenderà a 'riprodurre' anche i residui osservati, nella previsione della Y , quando siano dati nuovi valori delle variabili esplicative.

Tale comportamento è in genere considerato negativamente nei modelli regressivi, in quanto siamo interessati a prevedere il valore atteso di Y e non a ricostruire una possibile determinazione campionaria. Considerando una funzione di perdita quadratica, il valore ottimale della nostra previsione è appunto il valore atteso $E(Y_i | \mathbf{x}_i)$, ogni altro valore della distribuzione condizionata $(Y_i | \mathbf{x}_i)$, sebbene perfettamente plausibile, fornirebbe, in media, una perdita maggiore.

Dovendo però ricostruire dei dati mancanti, non possiamo utilizzare il valore atteso $E(Y_i | \mathbf{x}_i)$, che sostituito al valore mancante ridurrebbe la variabilità del carattere e altererebbe la sua relazione con gli altri caratteri. Sarebbe più corretto

invece imputare con i valori ottenuti da $E(Y_i | \mathbf{x}_i) + \varepsilon_i$. Tale procedura richiede però la conoscenza delle distribuzioni condizionate di $(Y_i | \mathbf{x}_i)$.

In presenza di diverse variabili qualitative e quantitative, ossia la situazione più comune nella pratica, la definizione della corrispondente distribuzione multipla risulta troppo complessa. D'altra parte, non conosciamo neanche le distribuzioni condizionate di $(Y_i | \mathbf{x}_i)$, che possono solo essere stimate, sotto ipotesi più o meno stringenti.

Avendo a disposizione le distribuzioni condizionate, si potrebbero quindi imputare i dati mancanti estraendo casualmente da queste distribuzioni. Si badi che questo procedimento non implica affatto che il valore imputato sia vicino al valore 'vero' mancante: infatti i due valori potrebbero trovarsi sulle code opposte della distribuzione ed essere molto distanti. D'altra parte, questo procedimento, nel caso le distribuzioni condizionate siano esatte, ci fornirebbe delle variabili imputate con distribuzione esatta, e quindi media e variabilità corretta, e anche la relazione di associazione con le altre variabili corrisponderebbe a quella della popolazione.

Se le distribuzioni condizionate non sono esatte, il risultato è però difficile da prevedere e da valutare. Oltretutto è difficilmente proponibile la stima di tutte le distribuzioni condizionate, per cui si ricorre a degli approcci che permettono di ottenere, sotto ipotesi restrittive, valori estratti dalle distribuzioni di $(Y_i | \mathbf{x}_i)$.

Viceversa, utilizzando un metodo regressivo in overfitting sui dati, come proponiamo in questo articolo, in corrispondenza dei valori mancanti è possibile ottenere direttamente i valori $E(Y_i | \mathbf{x}_i) + e_i$ senza la necessità di stimare le distribuzioni condizionate.

Nel prosieguo assumeremo che i dati sono mancanti in modo casuale (MAR, si veda Rubin 1987), cioè la probabilità che un'osservazione sia mancante può dipendere solo da dati osservati, non su altri valori mancanti o su variabili non osservate. Assumeremo inoltre di essere interessati ad ottenere una matrice di dati "completata" ossia una matrice in cui i valori mancanti sono stati imputati, ma al tempo stesso vogliamo tener conto della variabilità aggiuntiva introdotta dalla procedura di imputazione (Rubin 1987). Infine non faremo ipotesi sul pattern di dati mancanti e considereremo variabili con livello di misura misto.

2. Imputazione singola e multipla

Nei casi in cui ci si trova di fronte a dei dati mancanti è pratica comune ricorrere a metodi di imputazione singola. Invece di non considerare le mancate risposte, che in presenza di più variabili con mancate risposte potrebbe portare a scartare gran

parte del campione, si crea un data-set completato sostituendo a ciascun valore mancante un valore scelto in maniera opportuna.

I metodi più comuni e “classici” di imputazione vengono spesso classificati in 4 gruppi: unconditional mean, unconditional distribution, conditional mean, conditional distribution (Schafer & Graham, 2002).

I primi due metodi producono stime distorte per molti parametri sotto ogni tipo di dati mancanti (MCAR, MAR, MNAR). L'imputazione di una media condizionata può ancora introdurre distorsione. L'imputazione da una distribuzione condizionata è praticamente non distorta sotto MAR e MCAR, ma è distorta sotto MNAR. Comunque il difetto maggiore dell'imputazione singola risiede nel fatto che, pur preservando le distribuzioni marginali e congiunte, tende a sottostimare i livelli di incertezza, in quanto le misure di variabilità tradizionali non considerano che i valori imputati non sono noti ma sono soltanto delle stime. In letteratura, non è stato proposto alcun metodo di imputazione singola che sia in grado di tener conto dell'incertezza aggiuntiva legata ai dati mancanti. Di conseguenza l'utilizzo del data-set completato porterà a delle inferenze non corrette, in quanto la variabilità dei parametri risulterà sottostimata (Rubin 1987). Oltretutto la quasi totalità dei metodi di imputazione singola proposti in letteratura richiedono una struttura monotona oppure un numero basso di missing-data.

L'Imputazione Multipla (MI, Rubin, 1987) non tenta di stimare ogni valore mancante, ma piuttosto cerca di riprodurre un campione casuale dei valori mancanti. Questo processo è esplicitamente rivolto ad individuare valide inferenze statistiche per i test sui parametri e per la stima degli intervalli di confidenza, tenendo conto dell'incertezza aggiuntiva dovuta alla presenza dei valori mancanti.

Il metodo propone la generazione di m data-set completati, che poi sono analizzati utilizzando le procedure statistiche classiche. Successivamente, i risultati delle m -analisi sono combinati, con l'ipotesi che le differenze siano dovute esclusivamente alla diversità delle imputazioni e quindi alla variabilità aggiuntiva.

La procedura può essere schematizzata nei seguenti passi:

- Si definisce la distribuzione multipla delle variabili $P(X | \theta)$
- Si definisce una distribuzione a priori appropriata per θ
- Si calcola la distribuzione a posteriori per θ
- Si estrae un vettore di valori θ^* dalla distribuzione a posteriori
- Si generano le imputazioni tramite estrazioni casuali dalla distribuzione predittiva $p(x^{mis} | x^{obs}, \theta = \theta^*)$

Dove abbiamo indicato con x^{mis} e x^{obs} rispettivamente i valori mancanti e i valori osservati. Gli ultimi due passi vengono ripetuti (dalle 5 alle 10 volte) per ottenere le imputazioni desiderate.

D'altra parte la MI presenta dei problemi che la rendono difficilmente applicabile:

- assenza di una matrice completata, che è conveniente avere in molti casi;
- in genere la distribuzione multivariata è sconosciuta e l'ipotesi di multinormalità irrealistica;
- difficoltà ad analizzare un gran numero di variabili;
- difficoltà ad analizzare dati con livello di misurazione misto;
 - non può considerare i vincoli sulle imputazioni e non può considerare disegni complessi di campionamento.

Per ovviare a questi problemi è stato proposto un diverso approccio alla imputazione multipla, denominato IVEware (Raghunathan et al. 2001, si veda anche MICE, Van Buuren et al., 2000). IVEware costruisce una sequenza di regressioni multiple dove l'imputazione viene effettuata variabile per variabile, estraendo casualmente le imputazioni dalla distribuzione predittiva definita dai modelli di regressione. Viene quindi utilizzato uno schema iterativo di imputazione, che aggiorna via via i valori precedentemente imputati, al fine di preservare al meglio la correlazione tra le variabili.

In sostanza IVEware non richiede di esplicitare una distribuzione multivariata come con MI, ma soltanto le distribuzioni condizionate $P(X_j | X_{-j}, \theta)$. Di conseguenza possiede due importanti vantaggi rispetto alla MI: i modelli regressivi sulle distribuzioni condizionate sono più semplici ed è possibile prendere in considerazione variabili con livello di misura misto.

Non mancano comunque i problemi: IVEware richiede di specificare distribuzioni condizionate per i dati mancanti, sotto l'ipotesi che esista una corrispondente distribuzione multivariata. I risultati sono influenzati dalla forma (essenzialmente lineare) dei modelli regressivi considerati. Fornisce risultati scadenti per i dati qualitativi e la convergenza non è garantita.

3. Il metodo MITREE

Sia MI che IVEware calcolano le imputazioni tramite estrazioni casuali da una distribuzione Bayesiana predittiva stimata. Ci sono altri modi per generare imputazioni multiple? Recentemente, Decision Trees e reti neurali sono stati proposti come metodi efficaci per imputazioni singole (Mesa et al. 2000, Conversano e Siciliano 2004, Bengio e Gingras 1996) ma anche multiple (Di Ciaccio 2011, Di Ciaccio & Vallely 2007).

Per illustrare il metodo MITREE consideriamo innanzitutto un problema di imputazione singola. Dopo un'assegnazione iniziale casuale dei dati mancanti, si

considera ogni variabile contenente dei missing come variabile risposta e si costruisce un modello predittivo attraverso degli alberi di decisione (classificazione per target qualitativa, regressione per target quantitativa). Il modello stimato viene quindi utilizzato per aggiornare, tramite un opportuno coefficiente, il valore mancante (per target quantitativa) o la probabilità di assegnazione (per target qualitativa). La procedura generalmente converge o comunque si ferma dopo un fissato numero di iterazioni. E' noto che gli alberi di decisione con gli opportuni parametri e senza pruning conducono verso una situazione di overfitting, cosa comporta quindi questa procedura?

Consideriamo due unità statistiche che hanno gli stessi valori delle variabili eccetto una (chiamiamola Y), mancante sulla seconda unità. In questo caso l'albero imputerà un valore molto simile (se non uguale) della Y alla seconda unità, mentre MI potrebbe fornire valori abbastanza diversi (dipende dalla varianza stimata dell'errore). In qualche modo, MITREE è quindi vicino ad un metodo K-Nearest Neighbor, o al metodo Hot Deck con la capacità però in più di trattare missing non monotoni, livello di misura misto e un numero elevato di variabili. Occorre infine aggiungere che ripetendo questa procedura con diverse partenze casuali si ottengono imputazioni in generale diverse, da qui la capacità del metodo di generare imputazioni multiple e di tener conto della variabilità aggiuntiva in modo analogo alla MI. Il collegamento tra MITREE e il metodo Hot-Deck suggerisce una migliore performance sulle singole imputazioni, aspetto decisamente rilevante quando si voglia costruire una matrice di dati completata. Da questo punto di vista MITREE presenta dei punti di contatto con il metodo MIDAS (Siddique & Belin, 2008). La variabilità degli stimatori dei parametri viene calcolata allo stesso modo della MI, applicando la regola di Rubin: la varianza complessiva è data dalla somma della *average within imputation variance* e la *between imputation variance* (Rubin 1987).

4. Confronto tramite un'applicazione ai dati EU-SILC

Per valutare le caratteristiche di MITREE abbiamo costruito una simulazione utilizzando i dati cross-sectional EU-SILC 2006 relativi all'Italia e lo abbiamo confrontato con IVEware, che era il metodo di imputazione consigliato da EUROSTAT.

Abbiamo innanzitutto selezionato i 28.418 individui presenti nella fascia di età 17-81, sono state quindi selezionate 36 variabili, 10 quantitative e 26 qualitative.

Su quattro variabili (2 quantitative e 2 qualitative) abbiamo introdotto artificialmente il 25% di dati mancanti in base al valore assunto da una delle variabili senza missing (n. di componenti della famiglia) in modo da creare una

situazione di tipo MAR. Da tali dati abbiamo quindi estratto 300 campioni con 300 unità.

Nella valutazione dei due metodi abbiamo considerato sia la capacità di ricostruire il ‘vero’ dato mancante sia la capacità di costruire delle inferenze corrette su alcuni parametri. Tra questi sono stati considerati gli intervalli di confidenza sulle medie delle variabili e la stima dei coefficienti di correlazione. Per i caratteri qualitativi abbiamo considerato il numero di discrepanze tra la modalità vera e quella imputata. Per il confronto dei singoli valori è stata sempre considerata la prima matrice imputata sia per MITREE che per IVEware. Per il calcolo degli intervalli di confidenza sono state calcolate 10 imputazioni.

Tabella 1 – Alcune misure per il confronto tra MITREE e IVEware: numero di campioni favorevoli ad ogni metodo (su 300 campioni totali).

| | SSE | γ | \bar{x}_1 | \bar{x}_3 | σ_1 | σ_3 | ρ_{12} | ρ_{13} | ρ_{14} |
|---------|-----|----------|-------------|-------------|------------|------------|-------------|-------------|-------------|
| MITREE | 298 | 300 | 278 | 154 | 292 | 166 | 198 | 174 | 244 |
| IVEware | 2 | 0 | 22 | 146 | 8 | 134 | 100 | 122 | 52 |

Nella Tabella 1 abbiamo riportato per ognuno dei due metodi confrontati, il numero di campioni in cui il metodo forniva un risultato più vicino al vero valore. Si può vedere ad esempio che per le variabili quantitative imputate la differenza quadratica tra i valori imputati e i valori veri (SSE) indichi la superiorità di MITREE per 298 campioni su 300. Lo stesso vale per il numero di assegnazioni errate nelle variabili qualitative (γ), le medie aritmetiche delle due variabili quantitative con missing (\bar{x}_1 e \bar{x}_3), le rispettive deviazioni standard, la correlazione lineare tra la prima variabile (con missing) ed altre tre variabili (ρ_{12} , ρ_{13} , ρ_{14}). Quindi a livello di ricostruzione del dato mancante e di stima dei parametri MITREE si è dimostrato nettamente superiore ad IVEware.

Nella Tabella 2 possiamo notare che MITREE è più efficace nella stima dell’intervallo di confidenza di \bar{x}_1 mentre è meno efficace nella stima dell’intervallo di confidenza di \bar{x}_3 .

Tabella 2 – Stima dell'intervallo di confidenza al 95% per la media della variabile X_1 (reddito totale) e X_3 (superficie dell'abitazione). Percentuale di coverage e ampiezza media dell'intervallo su 300 campioni casuali.

| \bar{x}_1 | Coverage | Ampiezza intervallo | \bar{x}_3 | Coverage | Ampiezza intervallo |
|-------------|----------|---------------------|-------------|----------|---------------------|
| MITREE | 94.6% | 2527.0 | MITREE | 89.3% | 8.81 |
| IVEware | 99.0% | 5850.9 | IVEware | 97.0% | 10.87 |

Si può notare che MITREE tende in generale a sottostimare la variabilità, ciò in effetti accade quando le foglie dell'albero mantengono un certo grado di eterogeneità, nonostante l'overfitting. Per correggere tale distorsione si potrebbe comunque applicare un fattore correttivo nella stima della varianza che tenga conto della variabilità 'non spiegata' nelle foglie dell'albero.

Ringraziamenti

Il presente lavoro è stato realizzato nell'ambito del progetto PRIN-2007: "Disuguaglianza, povertà ed esclusione sociale: analisi e coerenza delle fonti informative, nuove misure e modelli interpretativi".

Riferimenti bibliografici

- BENGIO Y., GINGRAS B. 1996. Recurrent neural networks for missing or asynchronous data. In M. E. Hasselmo (eds.) D. S. Touretzky, M. C. Mozer, editor, *Advances in Neural Information Processing Systems* 8, pp. 395-401. MIT Press.
- CONVERSANO C., SICILIANO R. 2003. Incremental Tree-Based Missing Data Imputation with Lexicographic Ordering. In: *Interf. 2003 Proc.*, Minotte M., Swzychak A. (Eds.), Inter-face Found. of North America, Washington.
- DI CIACCIO A. 2011. Bootstrap and Nonparametric Predictors to Impute Missing Data. In *Classification and Multivariate Analysis for Complex Data Structures* (Fishet et al. eds.). Series: Studies in Classification, Data Analysis, and Knowledge Organization. vol. 1, pp. 203-210, Springer-Verlag.
- DI CIACCIO A., VALLELY T. 2007. Use of non-parametric methods for the imputation of missing data. A comparison based on extensive Montecarlo simulations. In: *Proceeding of SCo.o 2007*. Venezia, September 6-8, 2007.
- RUBIN D. 1987. *Multiple Imputation for Nonresponse in Surveys*, John Wiley & Sons.
- EFRON, B. 2004. The estimation of prediction error: Covariance penalties and crossvalidation. *J. Amer. Statist. Assoc.* 99, pp. 619-642.

- EUROSTAT, 2006. Comparative EU statistics on Income and Living Conditions: Issues and Challenges. *Proceedings of the EU-SILC conference* (Helsinki, 6-8 November 2006).
- MESA D., TSAI P., CHAMBERS R.L. 2000. Using tree-based models for missing data imputation: an evaluation using UK census data. *Research note*, Department of Social Statistics, University of Southampton.
- RAGHUNATHAN T.E., LEPKOWSKI J.M., VAN HOEWYK J., SOLENBERGER P. 2001. A multivariate technique for multiply imputing missing values using sequence of regression models. *Survey Methodology*, Vol. 27 No. 1, pp. 85-95.
- SCHAFFER. J.L., GRAHAM J.W. 2002. Missing data: Our view of the state of the art. *Psychological Methods*, Vol. 7, No. 2, pp. 147-177.
- SIDDIQUE J., BELIN T.R. 2008. Multiple Imputation Using an Iterative Hot-Deck with Distance-Based Donor Selection. *Statistics in Medicine*, Vol. 27, pp.83-102.
- SIDDIQUE J., HAREL O. 2009. MIDAS: A SAS Macro for Multiple Imputation Using Distance-Aided Selection of Donors, *Journal of Statistical Software* Vol. 29, No. 9.
- VAN BUUREN S., OUDSHOORN C.G.M. 2000. Multivariate imputation by chained equations: MICE V1.0 User's manual. *Report PG/VGZ/00.038*. Leiden, TNO Preventie en Gezondheid.

SUMMARY

A procedure based on the use of Sequential Regression and Classification Trees for the imputation of missing data is proposed. We dealt with the case of non-monotone patterns of missing data with mixed measurement level of the variables. The aim of the analysis is to obtain a completed data matrix with optimal characteristics (with respect to means, variances and correlations of the variables) which is often the main demand for a statistical office. Moreover we want to obtain a measure of the additional variability due to the presence of missing values. A simulation case with qualitative and quantitative data is analyzed and the results compared with other procedures. In particular, the performance of Multiple Imputation, using IVEWARE, were compared with our proposal using a large simulation with artificial data and with the EU-SILC cross-sectional data. Our non-parametric method showed to be very competitive on these simulations.

Agostino DI CIACCIO, Professore Ordinario, Sapienza Università di Roma,
agostino.diciaccio@uniroma1.it

Giovanni Maria GIORGI, Professore Ordinario, Sapienza Università di Roma,
giovanni.giorgi@uniroma1.it

SCENARI OCCUPAZIONALI DI UN'ITALIA CHE CAMBIA

Maria Grazia Didonna, Gianfranco Servodio

1. Introduzione

Il mercato del lavoro italiano ha subito nell'ultimo decennio profonde modificazioni - anche in seguito alla crisi economico-finanziaria - che hanno evidenziato, tuttavia, la stabilità degli occupati italiani. Obiettivo del presente contributo è quello di analizzare, dal 2004 al 2008, la mobilità del mercato del lavoro italiano considerando le forze lavoro presenti in Italia nella classe di età compresa tra i 15 e 64 anni. Il progetto si propone, quindi, di studiare i tassi di disoccupazione a livello regionale, tenendo conto, altresì, di due fattori socio-economico determinanti: la tipologia e la posizione contrattuale.

2. Dati e metodi

Al fine di effettuare una prima quantificazione del fenomeno in esame sono stati utilizzati i microdati della nuova rilevazione Istat sulle forze lavoro, il cui universo di riferimento è composto dalle famiglie residenti in Italia ed iscritte alle anagrafi comunali, anche se temporaneamente all'estero¹. L'unità di rilevazione è la famiglia di fatto da intendersi come un insieme di persone legate da vincoli di matrimonio, parentela, affinità, adozione, tutela o da vincoli affettivi, coabitanti ed aventi dimora abituale nello stesso comune, anche se non residenti nello stesso domicilio.² L'analisi è stata condotta attraverso la costruzione delle matrici di transizione annuali delle forze lavoro. Ci si riferisce alla popolazione in età (all'origine) 15-64 anni che permane a distanza di dodici mesi, in particolare, dal primo trimestre dell'anno t al primo trimestre dell'anno t+1. Le suddette matrici di

¹ Sono escluse le famiglie residenti in Italia che vivono abitualmente all'estero e i membri permanenti delle convivenze (ospizi, istituti religiosi, etc.).

² Per un approfondimento delle caratteristiche della RFL si veda Gazzelloni (2006).

transizione sono basate sulla classificazione della popolazione in tre stati: occupati, persone in cerca o disoccupati e non forze lavoro o inattivi.³

3. Il mercato del lavoro italiano 2004-2008: tra stock e flusso

Dall'analisi della prima matrice (tabella 1), riferita al biennio 2004-2005, emerge che oltre il 90% degli occupati conserva il proprio posto di lavoro ad un anno di distanza. Il 2%, invece, si ritrova disoccupato, mentre quasi il 6% transita verso la condizione di inattivo. Un dato interessante è quello dei disoccupati che presentano una dinamica evolutiva rilevante considerando che quasi il 60% si sposta nelle altre due categorie, al contrario di quanto avviene per gli inattivi, pari all' 89%, che presentano una forte permanenza nella condizione iniziale.

Tabella 1 – *Matrici di transizione degli stock iniziali e finali della popolazione longitudinale in età lavorativa, flussi in entrata e in uscita tra le condizioni considerate. I trim. 2004-I trim. 2005.*

| 2004 (t-1) | 2005 (t) | | | Totale |
|-------------|----------------|-------------|----------------|--------|
| | Occupati | Disoccupati | Inattivi | |
| Occupati | 26.174 92,1 | 564 2 | 1.668 5,9 | 28.406 |
| Disoccupati | 765 29 | 876 33,1 | 1.001 37,9 | 2.642 |
| Inattivi | 1250 6,1 | 934 4,6 | 18.328 89,3 | 20.512 |
| Totale | 28.189 | 2.374 | 20.997 | 51.560 |

Note: In corsivo le probabilità di transizione.

Fonte: elaborazioni proprie su micro dati longitudinali ISTAT.

Per il biennio 2005-2006 (tabella 2) possono farsi le medesime considerazioni: la quasi totalità di coloro che avevano un'occupazione dodici mesi prima la conservano, vale a dire il 92,7%; i disoccupati tendono a distribuirsi equamente tra occupati e inattivi, e questi ultimi pari all' 87,7% diminuiscono rispetto al biennio precedente poiché aumenta, anche se lievemente, la percentuale di coloro che da inattivi diventano occupati. Gli anni considerati presentano, altresì, una notevole stabilità nel mercato del lavoro degli occupati e degli inattivi documentata dalle probabilità di permanenza negli anni⁴. Nei bienni successivi, anni 2006-2007 e 2007-2008, (tabelle 3 e 4) la situazione rimane sostanzialmente invariata.

³ Per una definizione di occupato, persone in cerca e inattivi si veda dati ISTAT.

⁴ Probabilità che si dispongono lungo la diagonale principale delle matrici.

Tabella 2 – *Matrici di transizione degli stock iniziali e finali della popolazione longitudinale in età lavorativa, flussi in entrata e in uscita tra le condizioni considerate. I trim. 2005-I trim. 2006.*

| 2005 (t-1) | 2006 (t) | | | Totale |
|-------------|----------|-------------|----------|--------|
| | Occupati | Disoccupati | Inattivi | |
| Occupati | 25.866 | 519 | 1501 | 27.886 |
| | 92,7 | 1,9 | 5,4 | |
| Disoccupati | 773 | 804 | 917 | 2.494 |
| | 31,0 | 32,2 | 36,8 | |
| Inattivi | 1.512 | 989 | 17873 | 20.374 |
| | 7,4 | 4,9 | 87,7 | |
| Totale | 28.151 | 2.312 | 20.291 | 50.754 |

Fonte: elaborazioni proprie su microdati longitudinali ISTAT.

Tabella 3 – *Matrici di transizione degli stock iniziali e finali della popolazione longitudinale in età lavorativa, flussi in entrata e in uscita tra le condizioni considerate. I trim. 2006-I trim. 2007.*

| 2006 (t-1) | 2007 (t) | | | Totale |
|-------------|----------|-------------|----------|--------|
| | Occupati | Disoccupati | Inattivi | |
| Occupati | 23.729 | 337 | 1549 | 25.615 |
| | 92,6 | 1,4 | 6,0 | |
| Disoccupati | 615 | 622 | 881 | 2.118 |
| | 29 | 29,4 | 41,6 | |
| Inattivi | 1257 | 631 | 16674 | 18.562 |
| | 6,8 | 3,4 | 89,8 | |
| Totale | 25.601 | 1.590 | 19.104 | 46.295 |

Fonte: elaborazioni proprie su microdati longitudinali ISTAT.

Tabella 4 – *Matrici di transizione degli stock iniziali e finali della popolazione longitudinale in età lavorativa, flussi in entrata e in uscita tra le condizioni considerate. I trim. 2007-I trim. 2008.*

| 2007 (t-1) | 2008 (t) | | | Totale |
|-------------|----------|-------------|----------|--------|
| | Occupati | Disoccupati | Inattivi | |
| Occupati | 24.780 | 454 | 1.547 | 26.781 |
| | 92,5 | 1,7 | 5,8 | |
| Disoccupati | 623 | 583 | 677 | 1.883 |
| | 33,0 | 31,0 | 36,0 | |
| Inattivi | 1514 | 952 | 17.228 | 19.694 |
| | 7,7 | 4,8 | 87,5 | |
| Totale | 26.917 | 1.989 | 19.452 | 48.358 |

Fonte: elaborazioni proprie su microdati longitudinali ISTAT.

Sia nel primo che nel secondo biennio rimane stabile il dato degli occupati con percentuali superiori al 90%, così come gli inattivi che, in egual misura, permangono nella situazione iniziale. Va sottolineato, inoltre, che per il 2006-2007, il 41% dei disoccupati ad un anno di distanza si ritrova inattivo, e lo stesso può dirsi per il 2007-2008 con una percentuale pari al 36%. In generale dunque, l'occupazione sembra crescere nei periodi considerati, nonostante il dispiegarsi della crisi nel 2008. A conferma di quanto detto, si assiste ad un aumento del flusso dei disoccupati in favore dell'occupazione che si attesta al 33% nel 2008, anno in cui dovrebbe diminuire.

4. Mobilità lavorativa per posizione e tipologia contrattuale

Per l'intero periodo 2004-2008 (tabella 5) occorre evidenziare la stabilità degli occupati anche per posizione professionale. La tabella, infatti, conferma la permanenza dei dipendenti e degli indipendenti, a cinque anni di distanza, con rispettivamente il 97,8% ed il 94,6%.

Tabella 5 – *Permanenze e transizioni a 12 mesi tra le posizioni nella professione degli occupati. Anni 2004-2008. Valori percentuali.*

| Condizione al tempo t Condizione al tempo t+1 | Dipendenti | | Indipendenti | |
|--------------------------------------------------|------------|--------------|--------------|--------------|
| | Dipendenti | Indipendenti | Dipendenti | Indipendenti |
| I trim.2004 - I trim.2005 | 98,2 | 1,8 | 5,4 | 94,6 |
| I trim.2005 - I trim.2006 | 98,1 | 1,9 | 6,0 | 94,0 |
| I trim.2006 - I trim.2007 | 97,8 | 2,2 | 6,0 | 94,0 |
| I trim.2007 - I trim.2008 | 97,8 | 2,2 | 5,4 | 94,6 |

Fonte: elaborazioni proprie su microdati longitudinali ISTAT.

La tavola 6 riporta permanenze e transizioni a 12 mesi tra i due caratteri dell'occupazione dipendente in relazione al tempo determinato e indeterminato nel quadriennio preso in esame. È visibile, in questo caso, un lieve aumento della permanenza dei contratti a tempo determinato nel periodo considerato da un 63% ad un 65,9% con un picchi del 70% nei bienni intermedi. Lo stesso può dirsi per il gruppo di coloro che avevano un contratto a tempo indeterminato e che continuano ad averlo con un lieve incremento pari allo 0,5% nell'ultimo trimestre del 2008. Il carattere contrattuale che presenta una maggiore dinamica evolutiva è rappresentato, invece, da coloro che dal tempo determinato passano al tempo indeterminato; se, infatti, la transizione nel 2004-2005 era pari al 37%, nell'ultimo biennio è pari al 34,1%.

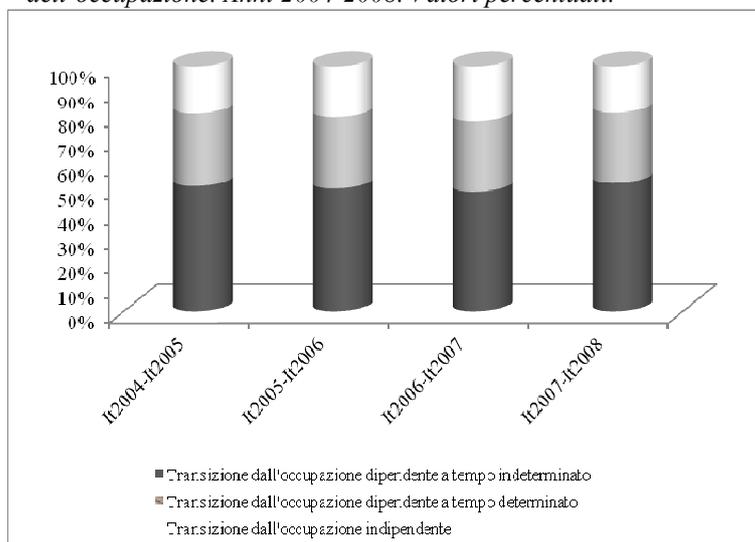
Tabella 6 – Permanenze e transizioni a 12 mesi tra i caratteri dell'occupazione degli occupati dipendenti. Anni 2004-2008. Valori percentuali.

| Condizione al tempo t Condizione al tempo t+1 | Tempo determ. | | Tempo indetermin. | |
|--------------------------------------------------|---------------|-----------|-------------------|-----------|
| | Determ. | Indeterm. | Determ. | Indeterm. |
| I trim.2004 - I trim.2005 | 63,0 | 37,0 | 3,0 | 97,0 |
| I trim.2005 - I trim.2006 | 70,4 | 29,6 | 2,4 | 97,6 |
| I trim.2006 - I trim.2007 | 69,6 | 30,4 | 2,3 | 97,7 |
| I trim.2007 - I trim.2008 | 65,9 | 34,1 | 2,5 | 97,5 |

Fonte: elaborazioni proprie su microdati longitudinali ISTAT.

5. Transizioni verso la disoccupazione per posizione nella professione e carattere dell'occupazione

Nel quadriennio 2004-2008 circa la metà degli occupati, che si ritrovano a dodici mesi di distanza alla ricerca di una occupazione, avevano un contratto a tempo indeterminato. Differente è la tipologia dei contratti a termine, che per loro struttura e per la rinnovabilità annuale a cui sono soggetti, permettono la giacenza nell'occupazione con la conseguente riduzione del transito nella disoccupazione.

Figura 1 – Entrata nella disoccupazione per posizione nella professione e carattere dell'occupazione. Anni 2004-2008. Valori percentuali.

Fonte: elaborazioni proprie su microdati longitudinali ISTAT.

Ne consegue, che il dato rilevante è la perdita del lavoro proprio di chi possedeva una tipologia contrattuale meno flessibile, e perciò in ipotesi più sicura. Al contrario, la flessibilità in quanto tale riesce a “destreggiarsi” più facilmente nelle dinamiche in esame.

6. Trend dei tassi di disoccupazione regionali

I tassi di disoccupazione disaggregati a livello regionale, nel periodo considerato, si mantengono stabili nelle regioni del Nord, ad eccezione della Lombardia che presenta al 2008 un tasso di disoccupazione pari 2,9% rispetto al 4,42% del 2004. Situazione diversa si riscontra nel sud in cui il tasso di disoccupazione diminuisce, ma presenta percentuali nettamente superiori rispetto alle regioni settentrionali, a conferma del dualismo delle situazioni occupazionali esistente tra le regioni del nord, altamente industrializzate, e quelle del sud.

Tabella 7 – *Tassi di disoccupazione per regione di residenza. Anni 2004-2008. Valori percentuali.*

| REGIONI | ANNI | | | | |
|---------------|-------|-------|-------|-------|-------|
| | 2004 | 2005 | 2006 | 2007 | 2008 |
| Piemonte | 3,92 | 4,70 | 4,83 | 4,85 | 4,42 |
| Valle d'Aosta | 2,95 | 2,40 | 2,59 | 1,93 | 1,59 |
| Lombardia | 4,42 | 3,70 | 4,15 | 3,22 | 2,99 |
| Trentino | 2,52 | 3,00 | 2,15 | 3,12 | 3,28 |
| Veneto | 4,56 | 4,10 | 4,43 | 4,17 | 3,74 |
| Friuli | 3,19 | 3,70 | 2,54 | 4,08 | 3,48 |
| Liguria | 6,40 | 7,00 | 6,78 | 5,72 | 5,52 |
| Emilia | 4,16 | 4,70 | 3,39 | 3,23 | 2,64 |
| Toscana | 5,16 | 5,20 | 5,63 | 3,79 | 4,71 |
| Umbria | 7,40 | 7,10 | 5,54 | 4,28 | 2,97 |
| Marche | 5,52 | 3,30 | 4,67 | 5,51 | 3,85 |
| Lazio | 8,70 | 8,60 | 7,33 | 6,91 | 7,94 |
| Abruzzo | 8,71 | 7,60 | 6,80 | 5,62 | 6,18 |
| Molise | 12,34 | 10,50 | 11,03 | 10,41 | 10,33 |
| Campania | 16,02 | 16,30 | 14,66 | 10,44 | 12,80 |
| Puglia | 17,72 | 14,80 | 13,61 | 10,71 | 11,32 |
| Basilicata | 15,50 | 11,20 | 13,27 | 9,50 | 10,94 |
| Calabria | 13,53 | 16,10 | 13,51 | 11,89 | 14,95 |
| Sicilia | 19,34 | 16,60 | 15,49 | 14,20 | 14,74 |
| Sardegna | 15,58 | 11,70 | 11,86 | 11,86 | 10,57 |

Fonte: elaborazioni proprie su microdati longitudinali ISTAT.

Ad avvalorare quanto asserito è il calcolo delle statistiche descrittive dei tassi di disoccupazione (tabella 8). Si nota come gli scarti dalla media diminuiscono nel corso degli anni a dimostrazione che la variabilità tende a diminuire a livello regionale, determinando una riduzione delle disparità tra regioni. Nonostante il netto miglioramento rispetto al 2004, il dualismo tra le regioni del nord e quelle del sud continua a persistere.

Tabella 8 – *Statistiche descrittive. Anni 2004-2008*

| Variabile | Media | S.q.m. (σ) | Minimo | Massimo | Numerosità | Mediana |
|-----------|-------|---------------------|--------|---------|------------|---------|
| 2004 | 8,88 | 5,56 | 2,52 | 19,34 | 20 | 6,9 |
| 2005 | 8,11 | 4,86 | 2,39 | 16,56 | 20 | 7,02 |
| 2006 | 7,71 | 4,54 | 2,15 | 15,49 | 20 | 6,21 |
| 2007 | 6,77 | 3,67 | 1,93 | 14,2 | 20 | 5,57 |
| 2008 | 6,95 | 4,34 | 1,59 | 14,95 | 20 | 5,12 |

Fonte: elaborazioni proprie su microdati longitudinali ISTAT

7. Conclusioni

Nel presente lavoro si è analizzato il recente panorama occupazionale dal quale si evince una riduzione del tasso di disoccupazione nelle varie regioni, a quattro anni di distanza. Emerge, inoltre, una maggior capacità della tipologia contrattuale a termine a rispondere alle esigenze del mercato del lavoro che presenta, altresì, una forte permanenza degli occupati e degli inattivi a dodici mesi di distanza. Ne deriva, concludendo, che l'occupazione rimane stabile ma mutano le condizioni contrattuali, al contrario la disoccupazione sempre più sentita al sud tende ad omogeneizzarsi.

Riferimenti Bibliografici

- BERARDI R., ZACCARIN S. 1991. *La stima dei flussi e di matrici di transizione*. In TRIVELLATO U. (Ed.), *Forze di lavoro: disegno dell'indagine e analisi strutturali*, Annali di statistica, serie IX, Vol.11, Roma, Istat.
- GAZZELLONI S. 2006. *La rilevazione sulle forze lavoro: contenuti, metodologie, organizzazione*. Metodi e Norme, No.32, Roma.
- ISTAT 2009. *Forze di lavoro – media 2008*. Annuari, Roma.

SUMMARY

Our goal in this paper was examine the Italian labor market from 2004 to 2008 in the age group between 15 and 64 years. We explored changes and variation in the employment. By considering two socio-demographic determinants such as type and job title, the project aims is to study the different regional disparities in unemployment rates at regional level. Our results show that a great stability of employment after a year and a stratification of unemployment in the territory.

Maria Grazia DIDONNA, Dottoranda di ricerca, DSSM, Università di Bari,
mariagrazia.dd@alice.it

Gianfranco SERVODIO, Dottorando di ricerca, DSSM, Università di Bari,
gfr82@libero.it

FRAGMENTATION OF PRODUCTION: A COINTEGRATION ANALYSIS

Marco Giansoldati, Claudio Pizzi

1. Introduzione

Le modalità attraverso le quali le imprese si internazionalizzano è un argomento ampiamente dibattuto in letteratura. Gli studi, teorici ed empirici, considerano le implicazioni di tale processo sui flussi di commercio, sugli investimenti diretti esteri (IDE), nonché sulle forme di accordi *non-equity*. Tale interesse appare giustificato anche dall'aumento negli ultimi decenni dei flussi di commercio tra le diverse economie nazionali sia in termini assoluti, che relativi (Zeddies, 2011).

Tale dinamica ha interessato in prevalenza il commercio in parti e componenti, richiedendo il superamento delle teorie classiche e neoclassiche che non considerano l'ipotesi di suddivisione internazionale della produzione e quindi lo scambio di beni intermedi. Infatti, la compresenza di flussi di commercio inter-industriale e intra-industriale ha trovato una sua "sistemazione" teorica solo con lo sviluppo della Nuova Teoria del Commercio Internazionale (Krugman, 1979, 1980; Helpman e Krugman, 1985).

Tale tipologia di flussi è stata originariamente rilevata, sebbene senza un rigoroso impianto teorico, da Grubel e Lloyd (1975), e solo in seguito Greenaway *et al.* (1994) hanno proposto un indice che permette di distinguere il commercio intra-industriale di beni derivanti dalla divisione internazionale del lavoro da quello legato al *love for variety* espresso dai consumatori.

Agli inizi degli anni '90 la "frammentazione della produzione" trova un proprio impianto teorico con il lavoro seminale di Jones e Kierzkowski (1990). In assenza di frammentazione, il bene è realizzato all'interno di un unico blocco produttivo, al contrario, in presenza di frammentazione vi è una suddivisione della manifattura in più blocchi, in cui gli output ottenuti in un blocco diventano input nella fase produttiva successiva eseguita in un altro blocco.

In questa prospettiva i costi per i servizi di connessione non sono solo quelli necessari per fornire il prodotto al mercato finale ma anche prevalentemente quelli necessari richiesti per collegare, tra loro, blocchi produttivi (Jones e Kierzkowski, 1990).

Infatti, lavorare con diversi blocchi produttivi richiede il trasferimento dei beni intermedi tra le diverse unità comportando dei costi di trasporto che dipendono dalla distanza tra gli stabilimenti e dalla complessità del coordinamento tra i blocchi medesimi. Affinché la frammentazione si realizzi è necessario che il processo possa essere suddiviso in blocchi produttivi indipendenti e geograficamente separati e che i costi di connessione siano inferiori ai vantaggi derivanti da una migliore localizzazione dell'impianto.

La frammentazione si può manifestare seguendo due percorsi diversi. Il primo opera attraverso la disintegrazione di un'impresa fordista che passa dalla dimensione nazionale a quella internazionale attraverso strumenti proprietari. Il secondo si manifesta attraverso l'integrazione di piccole e medie imprese e richiama la struttura produttiva dei territori dell'Europa continentale in cui la formazione di distretti industriali rappresenta un esempio di divisione transnazionale del lavoro (Kimura *et al.*, 2007).

Il processo d'internazionalizzazione dei distretti industriali attraverso la frammentazione della produzione è quindi caratterizzato dalla presenza di costi di connessione che variano in base al tipo di bene scambiato, diventando tanto più elevati quanto più i beni scambiati sono input intermedi soggetti a transazioni di tipo *back-and-forth* (Ando, 2006). D'altra parte, la frammentazione, nella sua valenza bidimensionale domestica vs internazionale e proprietaria vs *arm's length* (Kimura e Ando, 2005), genera significativi flussi di beni intermedi tra Paesi.

Da un punto di vista metodologico, la maggior parte dei contributi empirici si sono avvalsi di modelli gravitazionali (Anderson e Van Wincoop, 2004), in cui le importazioni o le esportazioni sono spiegate da un insieme di variabili, macroeconomiche e non, riferite a coppie di paesi. Tale approccio ha permesso di evidenziare le dinamiche d'integrazione produttiva costruite nel corso del tempo. In questo ambito le analisi si sono basate su dati annuali ignorando informazioni infrannuali che potrebbero evidenziare la frammentazione tra coppie di paesi attraverso la presenza di transazioni di tipo *back-and-forth* di semilavorati ed in misura minore di materie prime e prodotti finiti (Ando, 2006; Kimura e Ando, 2005).

L'utilizzo di serie storiche trimestrali o mensili al fine di comprendere le dinamiche dei flussi bilaterali è presente nel lavoro di Fidrmuc *et al.* (2003) che esaminano il nesso tra politica estera e performance esportativa, studiando la relazione di cointegrazione per interpretare le dinamiche di lungo periodo. Alguacil e Orts (2002) analizzano invece le modalità di internazionalizzazione delle imprese spagnole, cercando relazioni di complementarietà o sostituzione tra investimenti diretti esteri (IDE) e flussi di esportazioni. Nel loro lavoro considerano un modello a correzione d'errore (ECM) che permette di evidenziare sia le relazioni di breve sia di lungo periodo. Analogamente Chen (2002) propone l'utilizzo dell'ECM con l'obiettivo di comprendere se il tasso di cambio, le esportazioni e il saldo

commerciale di Taiwan rispetto Stati Uniti e Giappone sono tra loro legati in una dinamica di lungo periodo o se rispondono esclusivamente a shock di breve periodo rispetto alla componente stocastica.

2. Frammentazione internazionale e flussi di commercio: l'esperienza italiana

La formazione di distretti industriali specializzati nella realizzazione di prodotti tradizionali ha originariamente avuto luogo attraverso la condivisione di conoscenza, un elevato grado di socializzazione della produzione e, solo più recentemente, da un processo di delocalizzazione sviluppatosi dalle aree industriali più avanzate, dominate da imprese di grandi dimensioni che soffrivano di diseconomie di scala interne e pressioni sindacali. In generale, il successo dei distretti industriali italiani deriva dalla presenza di economie di scala esterne risultanti da un mix virtuoso di cooperazione, specialmente sul prezzo, e competizione, prevalentemente sulla qualità. Tale struttura, viene meno attorno alla fine degli anni '90, quando la concorrenza esterna, proveniente sia dall'Asia Orientale sia da paesi industrializzati, mette sotto pressione i distretti industriali (Graziani, 2001). Ciascuna impresa distrettuale considerata separatamente non è in grado di assumere una dimensione tale da generare economie di scala interne e raggiungere autonomamente il mercato finale. Di conseguenza, per contrastare la concorrenza domestica e internazionale, le imprese distrettuali perseguono sostanzialmente due strade differenti. Da un lato, si attivano per la costruzione di associazioni e/o organizzazioni il cui scopo è di sviluppare economie di rete e collocare le imprese all'interno delle catene globali di produzione e del valore. Dall'altro le imprese di dimensioni piccole e medie seguono quelle di dimensioni maggiori che agiscono come leader nei processi d'internazionalizzazione, potendo talvolta diventare subfornitori globali.

Coerentemente con quanto esposto in questo e nel precedente paragrafo, la prima ipotesi che si desidera verificare in questo lavoro è:

H1: la dinamica dei percorsi d'internazionalizzazione dipende dal contesto macroeconomico congiunturale e concorrenziale.

Il ventennio analizzato (1991-2010) è suddiviso in due sottoperiodi che riflettono sia la "sofferenza" competitiva domestica evidenziata da Graziani (2001) sia l'inizio di una fase di stagnazione dell'economia dell'area dell'Euro confermata dal Business Cycle Dating Committee del CEPR (CEPR, 2003, p. 4) e di una recessione negli Stati Uniti individuata dall'NBER Business Cycle Dating Committee (NBER, 2001, p. 1). La partizione viene realizzata tra il I e il II trimestre del 2001.

La seconda ipotesi che si desidera testare è:

H2: tra i sistemi locali del Veneto e le aree di destinazione considerate si verificano nel periodo studiato flussi di tipo *back-and-forth*. Transazioni di tale natura implicano che le esportazioni di materie prime e semilavorati, trasformati presso il paese destinatario per sfruttare differenziali nei costi dei fattori, precedono temporalmente le importazioni di prodotti finiti dalle stesse aree (Ando, 2006; Kimura e Ando, 2005).

3. Descrizione e analisi dei dati

Al fine di verificare le ipotesi precedentemente formulate si sono studiate in un settore tecnicamente frammentabile quale quello Tessile-Abbigliamento, le relazioni esistenti tra i flussi commerciali di prodotti finiti, materie prime, e semilavorati tra alcune province del Veneto, (Treviso, Padova, Vicenza e Verona¹) e alcuni paesi caratterizzati da vantaggi comparati nel settore studiato, quali la Romania e la Tunisia.

Le variabili osservate sono state: esportazioni di *filati di fibre tessili* (*EXPRAW*) quale *proxy* dei flussi di materie prime, le esportazioni di *tessuti* (*EXPSEM*) quale *proxy* delle esportazioni di semilavorati, infine, le importazioni e le esportazioni di prodotti finiti (rispettivamente *IMPFIN* e *EXPFIN*) ottenute aggregando le voci *Articoli di abbigliamento, escluso l'abbigliamento di pelliccia* e *Articoli di maglieria*. Si noti che *EXPFIN* viene inserita tra le variabili esplicative vista l'eterogeneità dei contenuti delle voci che, sebbene si riferiscano prevalentemente a prodotti finiti, comprendono anche una parte non del tutto trascurabile di beni intermedi.

I dati trimestrali sono stati reperiti dalle statistiche del commercio estero dell'Istituto Nazionale di Statistica (www.coeweb.istat.it).

La relazione economica discussa precedentemente è descritta dal seguente modello lineare:

$$IMPFIN_t = \alpha + \beta_1 EXPSEM_t + \beta_2 EXPRAW_t + \beta_3 EXPFIN_t + u_t \quad (1)$$

Preliminarmente alla stima del modello si è proceduto a una trasformazione Box Cox (con parametro $\lambda=0.5$) e alla verifica della stazionarietà delle serie storiche mediante test di Kiatkowski-Phillips-Schmidt-Shin. I risultati di questi test hanno

¹ Sono state escluse dall'analisi le province di Rovigo, Belluno, Venezia perché presentano valori estremamente bassi dei flussi di importazioni ed esportazioni nei confronti di Romania e Tunisia all'interno del settore tessile/abbigliamento.

suggerito la necessità di effettuare un'analisi della cointegrazione per ciascuna provincia e per ciascun periodo.

Per i casi in cui il test di Phillips-Ouliaris abbia evidenziato la presenza di cointegrazione il modello (1) è stato riformulato mediante un Unrestricted Error Correction Model (UECM) che può essere utilizzato anche nel caso in cui il grado di integrazione delle serie sia diverso, senza la necessità di seguire la procedura a due stadi suggerita da Engle e Granger (1987) e nel caso in cui la numerosità campionaria sia ridotta (Pesaran *et al.*, 2001). Il modello che viene stimato in caso di cointegrazione è il seguente:

$$\Delta IMPFIN_t = \alpha + \beta_1 \Delta EXPRAW_t + \beta_2 \Delta EXPSEM_t + \beta_3 \Delta EXPFIN_t + \pi_1 IMPFIN_{t-1} + \pi_2 EXPRAW_{t-1} + \pi_3 EXPSEM_{t-1} + \pi_4 EXPFIN_{t-1} + v_t \quad (2)$$

ove π_1 è la stima della velocità di aggiustamento, $-\frac{\pi_2}{\pi_1}$ è l'effetto di lungo periodo per la variabile *EXPRAW*, $-\frac{\pi_3}{\pi_1}$ è l'effetto di lungo periodo per la variabile *EXPSEM*, e $-\frac{\pi_4}{\pi_1}$ è l'effetto di lungo periodo per la variabile *EXPFIN*.

4. Risultati

Nelle tabelle 1 e 2 sono riportate le stime dei parametri dei modelli. Da una prima osservazione è evidente che la struttura degli scambi che caratterizza tutte le provincie nel primo decennio (1991-2001) è descritta da modelli a correzione d'errore in cui il parametro della velocità di convergenza all'equilibrio è sempre negativo e significativo. Esiste pertanto un meccanismo che riconduce il sistema all'equilibrio di lungo periodo. Tale relazione permane nel secondo decennio solo per la provincia di Padova che, delle quattro considerate, è quella che presenta i volumi di import/export più contenuti, mentre sembra venire meno per le altre provincie. Ciò potrebbe avvalorare l'ipotesi H1 secondo la quale la congiuntura e l'assetto competitivo dei sistemi produttivi locali influenzano la dinamica dei percorsi d'internazionalizzazione. In altre parole le provincie con valori d'interscambio più elevati reagiscono alle mutate condizioni economiche e concorrenziali variando la propria modalità di presenza commerciale/produttiva all'estero attraverso forme di frammentazione che si evolvono nel tempo.

Concentrando l'attenzione sulle variabili significative nei diversi modelli stimati si evidenzia la presenza di un fenomeno di frammentazione della produzione che in alcuni casi coinvolge fasi a valle del processo produttivo (in questo caso sono significative le variabili *EXPSEM* e/o *EXPFIN*), in altri le fasi a monte (*EXPRAW*).

In entrambe le circostanze è verificata l'ipotesi (H2) di transazioni *back-and-forth*, seppure con caratteristiche proprie di ogni sistema locale, adattando le proprie strategie in base alle peculiarità del Paese con cui si attivano relazioni di interscambio commerciale.

Tabella 1 – Stime dei parametri del modello standard (1) o UECM (2) per gli interscambi commerciali tra alcune provincie del Veneto e la Romania.

| | VERONA | | TREVISO | | PADOVA | | VICENZA | |
|------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| | 1991-01 (2) | 2001-10 (1) | 1991-01 (2) | 2001-10 (1) | 1991-01 (2) | 2001-10 (2) | 1991-01 (2) | 2001-10 (1) |
| Intercetta | -405.7** | 6277.6** | -795.7 | 9958.9** | 48.5 | -110.2 | -284.1 | 6575.6** |
| ΔEXPRAW | -0.242 | | -0.332 | | -0.56 | 1.353* | 0.181 | |
| ΔEXPSEM | 0.562** | | 0.72* | | -0.12 | -0.162 | 0.198 | |
| ΔEXPFIN | 0.0267 | | 0.31 | | 0.739** | 0.074 | 0.221 | |
| IMPFIN | -0.862 | | -0.408 | | -1.094** | -0.88** | -0.915** | |
| EXPRAW | 0.472 | -1.018* | -0.469 | -0.349 | 0.205 | 1.438** | 0.785* | -0.61* |
| EXPSEM | 1.106** | 0.138 | 0.827 | 0.631** | 0.279 | 0.713** | 0.903** | 0.57* |
| EXPFIN | 0.098* | 0.129 | 0.312 | 0.036 | 1.507** | 0.433 | 0.365 | 0.094 |

Note: significatività *=0.05, ** =0.01

Tabella 2 – Stime dei parametri del modello standard (1) o UECM (2) per gli interscambi commerciali tra alcune provincie del Veneto e la Tunisia.

| | VERONA | | TREVISO | | PADOVA | | VICENZA | |
|------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| | 1991-01 (2) | 2001-10 (1) | 1991-01 (2) | 2001-10 (1) | 1991-01 (2) | 2001-10 (2) | 1991-01 (2) | 2001-10 (1) |
| Intercetta | 169,3 | 3078,4** | -176,3 | 6448,3** | 257,35 | 6700,7** | 73,7 | 5413,5** |
| ΔEXPRAW | 0,721 | | 0,534 | | 0,329 | -0,036 | 0,056 | |
| ΔEXPSEM | 0,045 | | 0,755** | | 0,323* | 0,03 | 0,14 | |
| ΔEXPFIN | -0,06 | | 0,322 | | 0,385** | 0,469* | 0,276 | |
| IMPFIN | -0,452** | | -0,787** | | -0,584** | -0,989** | -0,623** | |
| EXPRAW | 0,06 | 0,196 | 0,267 | 0,06 | 0,437 | -0,389 | -0,09 | 0,195 |
| EXPSEM | 0,455* | 0,543** | 0,8** | 1,018** | 0,621** | -0,065 | 0,459* | 0,302* |
| EXPFIN | 0,261* | 0,14 | 0,713** | -0,361 | 0,021* | 0,271* | 0,679* | 0,096 |

Note: significatività *=0.05, ** =0.01

In altre parole, si tratta di percorsi a livello provinciale, sintesi dei comportamenti dei sistemi produttivi locali presenti nell'area selezionata, che sono eterogenei nelle strategie d'internazionalizzazione, anche all'interno dello stesso

settore di attività economica. Ciò accade perché le aree distrettuali, specialmente quelle più mature, sono spesso guidate nell'accesso ai mercati esteri da imprese leader, soggetti di grandi dimensioni attorno alle quali si è storicamente sviluppata la subfornitura locale.

In sintesi, la struttura degli scambi attivati dalle singole province presenta elementi di diversità le une rispetto alle altre, con percorsi che solo in alcuni casi mutano tra un periodo e l'altro, garantendo però una generale conferma dell'ipotesi di frammentazione della produzione, a diversi livelli del ciclo produttivo.

Riferimenti bibliografici

- ALGUACIL M., ORTS V. 2002. A Multivariate Cointegrated Model Testing for Temporal Causality between Exports and Outward Foreign Investment: The Spanish Case. *Applied Economics*. Gennaio, 2002, Vol. 34, No. 1, pp. 119-132.
- ANDERSON A., van WINCOOP E. 1994. Trade Costs. *Journal of Economic Literature*. Settembre, 1994, Vol. 42, No. 3 pp. 691-751.
- ANDO M. 2006. Fragmentation and Vertical Intra-Industry Trade in East Asia. *The North American Journal of Economics and Finance*. Dicembre, 2006, Vol. 17, No. 3, pp. 257-281.
- CEPR (Centre for Economic Policy Research) 2003. Dating Committee Findings 22 September 2003, *Business Cycle Dating Committee of the Centre for Economic Policy Research*, Londra: CEPR.
- CHEN H.-C. 2002. Taiwan's Exports and Trade Imbalance against US and Japan: An Empirical Investigation Based on Error Correction Model. *Applied Economics*. Dicembre, 2002, Vol. 34, No. 18, pp. 2303-2309.
- ENGLE R., GRANGER C. 1987. Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica*. Marzo, 1987, Vol. 55, No. 2, pp. 251-276.
- FIDRMUC J., WÖRGÖTTER A., WÖRZ J. 2003. Does Foreign Policy Determine Foreign Trade? Cointegration Analysis of Exports from Selected Countries to the Middle East, *Applied Economics*, Marzo, 2003, Vol. 35, No. 5, pp. 565-571.
- GRAZIANI G. 2001. International subcontracting in the textile and clothing industry. In S. ARNDT S.W., KIERZKOWSKI H., (Eds.) "*Fragmentation: New production and trade patterns the world economy*", Oxford: Oxford University Press, pp. 209-230.
- GREENAWAY D., HINE R., MILNER C. 1994. Country-specific factors and the pattern of horizontal and vertical intra-industry trade in the UK. *Review of World Economics/Weltwirtschaftliches Archiv*. Marzo, 1994, Vol. 130, No.1, pp. 77-100.
- GRUBEL H., LLOYD P.J. 1975. "*Intra-industry trade*", London: MacMillan.
- HELPMAN E., KRUGMAN P.R. 1985. "*Market Structure and Foreign Trade*", Cambridge, MA: The MIT Press.
- JONES R.W., KIERZKOWSKI H. 1990. The Role of Services in Production and International Trade: A Theoretical Framework. In JONES R.W., KRUEGER A. (Eds) "*The Political Economy of International Trade. Festschrift in Honour of Robert Baldwin*", Oxford: Basil Blackwell, pp. 31-48.

- KIMURA F., TAKAHASHI Y., HAYAKAWA K. 2007. Fragmentation and parts and components trade: Comparison between East Asia and Europe. *The North American Journal of Economics and Finance*. February, 2007, Vol. 18, No. 1, pp. 23-40.
- KRUGMAN P.R. 1979. Increasing Returns, Monopolistic Competition and International Trade. *Journal of International Economics*. November, 1979, Vol. 9, No. 4, pp. 469-479.
- KRUGMAN P.R. 1980. Scale Economies, Product Differentiation, and Patterns of Trade. *American Economic Review*. December, 1980, Vol. 70, No. 5, pp. 950-959.
- NBER (National Bureau of Economic Research) 2001. The Business-Cycle Peak of March 2001. *Business Cycle Dating Committee*, Cambridge, MA: NBER.
- PESARAN M.H., SHIN Y., SMITH R. 2001. Bounds Testing Approaches to the Analysis of Level Relationships. *Journal of Applied Econometrics*. Maggio-Giugno, 2001, Vol. 16, No. 3, pp. 289-326.
- ZEDDIES G. 2011. Determinants of International Fragmentation of Production in European Union. *Empirica*. November, 2011, Vol. 38, No. 4, pp. 511-537.

SUMMARY

In the last two decades the majority of papers dealing with international fragmentation of production developed around the gravity models. As such, trade flows in parts, components and final products are employed together with a set of macroeconomic variables of partner countries to get the most accurate picture of bilateral relations.

This paper focuses on the textile and clothing industry and examines time series of quarterly trade flows between some provinces of the Veneto Region on one hand, and Romania and Tunisia on the other, first from 1991q1 to 2001q1 and then from 2001q2 to 2010q4.

According to these premises, the present contribution aims at testing two hypotheses. First, different trajectories and modes of internationalization in the province-country trade relationships should vary on the basis of the economic and competitive context. Second, trade flows between local systems of the Veneto Region and the selected partner countries ought to be described through back-and-forth transactions in raw materials, semi-processed and final goods.

In particular, the paper analyses the relationship between imports of final goods and exports of raw materials, semi-processed and final goods. The nonstationarity of some time series suggested to test for the cointegration hypothesis, and the use of an UECM as an interpretative tool of fragmentation itself. Conversely, once stationarity is detected, a standard estimation is employed.

The econometric exercises provide broad support of both abovementioned hypotheses, although sensible heterogeneity is detected across provinces.

Claudio PIZZI, Professore Associato, Università Ca' Foscari Venezia, pizzic@unive.it

Marco GIAN SOLDATI, Assegnista di ricerca, Università Ca' Foscari Venezia,
marco.giansoldati@unive.it

PARTICOLARI FORME DI ECONOMIA DEL CRIMINE: IL BUSINESS DELLA PEDOFILIA ON LINE

Marika La Rosa, Daniela Corso, Paola M. Chiodini,
Francesca Cuppone, Giancarlo Manzi, Bianca M. Martelli, Flavio Verrecchia

1. Introduzione e obiettivi

La pedofilia è considerata una manifestazione psicopatologica (American Psychiatric Association, 2000), appartenente alle parafilie - alterazioni a carico della sfera sessuale. Contestualmente, il comportamento pedofilo è un crimine perseguito dalla maggior parte degli ordinamenti giuridici. L'avvento di Internet, abbattendo ogni frontiera geografica e fisica, ha determinato un'espansione senza precedenti delle comunicazioni e dello scambio di contenuti digitali. Per queste sue caratteristiche la rete, pur nella sua essenza ampiamente positiva, si è mostrata capace di offrire nuovi strumenti per violare i diritti dei cittadini; spesso si tratta di reati estremamente gravi, come quelli nel campo dello sfruttamento sessuale dei bambini. La web-economy ha così aperto un nuovo canale espressivo per la pedofilia; non solo ha consentito una più facile fruizione di materiale pedopornografico e di informazioni per i pedofili (e.g. funzionali al turismo sessuale) ma ha anche consentito che tale materiale potesse essere sfruttato per stabilire relazioni di fiducia con altri pedofili, per ricattare minori vittime degli abusi, per trarre profitti economici. In questo ambito, avendo la maggior parte della offerta pedofila radice commerciale, la deviazione dei pedofili s'intreccia con gli interessi economici di gruppi criminali organizzati.

L'obiettivo del lavoro riguarda lo studio di quella particolare forma di economia del crimine che ruota attorno alla pedofilia on line. Le analisi, sviluppate a partire da basi di dati di fonte NGO¹ e dati ufficiali di indagine giudiziaria, in questa sede avranno come focus l'Offerta e la Domanda di pedopornografia on line. L'offerta di pedopornografia on line viene misurata attraverso l'utilizzo di specifiche

¹ Dati di provenienza istituzionale non governativa. In particolare si tratta dei dati dell'Osservatorio di Telefono Arcobaleno Onlus.

tecniche (i.e. *Hunting*) di reperimento, tracciamento e segnalazione dei siti web pedofili (Telefono Arcobaleno, 2011a, 2011b, 2011c).

L'attività di hunting è finalizzata a contrastare con ogni mezzo la produzione, il commercio e la diffusione sul web delle immagini delle violenze sessuali sui bambini in una logica di assoluta priorità della tutela delle vittime; si tratta di un'opera quotidiana, sistematica e qualificata di monitoraggio del web, realizzata da un'equipe di specialisti IT, tesa all'individuazione, al tracciamento e alla immediata segnalazione dei siti illeciti. L'attività di monitoraggio della rete alimenta quotidianamente una banca dati, costantemente aggiornata, delle attività pedofile online; le informazioni così ottenute, consentono di effettuare analisi sia qualitative, sia quantitative della pedofilia sul web fornendo una preziosa mappatura del fenomeno e contribuendo, quindi, alla conoscenza delle sue reali caratteristiche e dinamiche. Per quanto riguarda l'osservazione della domanda di pedopornografia on line, le tecniche di analisi devono invece essere strutturate in maniera differente in quanto sono evidenti i limiti operativi di chi si accosta al fenomeno per motivi di studio. I possibili strumenti sono costituiti principalmente da: i. strategie di Telefono Arcobaleno (TA); ii. log file dei contatori in chiaro di siti web pedofili; iii. Basi di dati da indagine giudiziaria (Italia).

2. I risultati

2.1. L'offerta di pedopornografia sulla rete internet

Le fotografie e i filmati prodotti attraverso lo sfruttamento dei bambini trovano diffusione sulla rete Internet, utilizzando i diversi canali che la rete offre. Si tratta, principalmente, di siti tradizionali ad accesso libero o a pagamento, attraverso cui il materiale illegale viene venduto, nella consueta cornice corredata da spazi "promozionali" che occupano i livelli più visibili e superficiali della rete. Il pedobusiness, nel tempo, ha affinato metodi e tecniche di promozione e dissimulazione aggiungendo alla già vasta gamma di metodologie di scambio on line dei materiali pedofili, l'utilizzo massivo dei numerosi servizi di *on line storage* presenti sulla rete che consentono di archiviare e scambiare foto e filmati in forma anonima e protetta (i.e. *masked files*). Nel primo trimestre del 2011, TA ha inviato agli *Internet Service Provider* (ISP) coinvolti più di 18mila segnalazioni relative alla presenza di materiale pedopornografico sul web; l'incremento tendenziale registrato rispetto al primo trimestre dell'anno passato (+15%) è di continuità con l'espansione che ha riguardato i periodi precedenti. Attraverso una procedura di tracciamento è possibile risalire alla nazionalità dell'ISP che ospita il materiale illegale e analizzare, quindi, la localizzazione geografica dell'offerta pedofila in

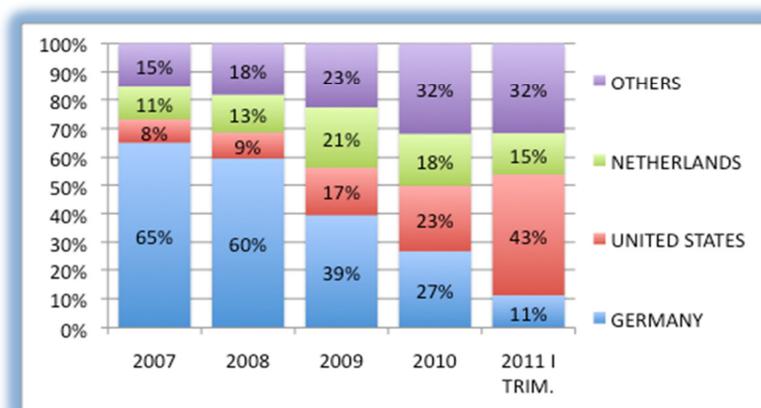
rete. Emerge che, nel primo trimestre del 2011, il 98% dei casi sono riferibili all'Europa (51%) e al Nord America (47%); in particolare, l'offerta della pedopornografia on line si concentra (79%) in 4 paesi: Stati Uniti, Paesi Bassi, Germania, Federazione Russa. Si osserva, inoltre, per i Paesi di origine anglosassone, una espansione congiunturale e tendenziale dei casi di segnalazione di pedofilia online (cfr tabella1). Osservando il fenomeno in una prospettiva storica (2007-2011) si evidenzia un progressivo mutamento della composizione percentuale dei siti di offerta di materiale pedopornografico (cfr. figura 1).

Tabella 1 – Localizzazione dell'offerta di pedopornografia in rete, 1° trim. 2011

| Rank | Paese | Valori assoluti | Var. % | |
|------|--------------------|--------------------|------------|------------|
| | | | t-1 | t-4 |
| 1 | United States | 7.747 | 96% | 111% |
| 2 | Netherlands | 2.661 | -34% | 19% |
| 3 | Germany | 2.027 | -51% | -55% |
| 4 | Russian Federation | 1.934 | 19% | -37% |
| 5 | Cyprus | 1.400 | 1422% | 47% |
| 6 | Canada | 847 | 211% | 256% |
| 7 | Czech Republic | 653 | | 4564% |
| 8 | France | 235 | -38% | 158% |
| 9 | Thailand | 224 | 90% | |
| 10 | United Kingdom | 156 | 290% | 4% |
| | ... | 30 | -91% | -96% |
| | Totale | 18.185 | 20% | 15% |

Fonte: Elaborazioni su dati Osservatorio Internazionale di TA. Note: a. t: 1° trim. 2011.

Figura 1 – Localizzazione dell'offerta di pedopornografia in rete, 2007 - 1° trim. 2011.



Fonte: Elaborazioni su dati Osservatorio Internazionale di Telefono Arcobaleno.

2.2. La domanda di pedopornografia sulla rete internet

Nell'ambito dell'attività di contrasto della pedofilia on line, l'equipe di specialisti di TA, ha effettuato delle rilevazioni periodiche del numero dei visitatori di siti pedofili, in collaborazione con il Nucleo Investigativo Telematico² (NIT).

La maggiore concentrazione di pedofili in rete si registra in Europa e Nord America: dalla rilevazione, emerge che i fruitori di pedopornografia in rete sono prevalentemente di nazionalità americana, tedesca, russa, italiana e inglese.

I primi otto paesi assorbono, insieme, più del 65% della domanda di pedopornografia, che si evolve parallelamente all'offerta, dislocata anch'essa prevalentemente nei paesi europei e anglosassoni. Nel 1% dei casi i visitatori di siti pedofili individuati utilizza delle particolari tecnologie finalizzate a rendere anonima la navigazione in rete -Anonymous proxy- per impedire l'identificazione (Cfr. tabella 2).

I dati rilevati attraverso l'analisi dei contatori in chiaro di alcuni siti pedopornografici confermano, pur con delle differenze, la validità della metodologia di rilevazione utilizzata da TA in termini di geolocalizzazione della domanda di pedofilia che, anche in questo contesto, si concentra negli Stati Uniti, in Germania, nel Regno Unito, in Francia e in Italia.

Rispetto alla situazione italiana, è possibile fornire informazioni più articolate sulle caratteristiche del pedofilo in rete analizzando i dati relativi a tutti i soggetti (circa un migliaio), residenti in Italia, indagati per i reati di produzione, detenzione e divulgazione di materiale pedopornografico dal NIT della Procura della Repubblica di Siracusa, su segnalazione di TA, nel periodo compreso tra il 2001 e il 2009.

Emerge che gli indagati sono tutti di genere maschile e appartengono a tutte le classi di età, con una concentrazione maggiore di soggetti in età compresa tra i 20 e i 39 anni (6 casi per 100mila abitanti di genere maschile).

Dai dati reperiti dal NIT emerge che si tratta prevalentemente di soggetti incensurati (95%), appartenenti a tutti i livelli socio economici e di istruzione; emerge, inoltre, che la maggior parte dei consumatori di pedofilia in rete vive all'interno del nucleo familiare (con i genitori o con un partner). Tali informazioni evidenziano come il pedofilo on line sia, generalmente, una persona ben inserita e integrata nel contesto familiare e sociale, confermando, quindi, la scarsa visibilità di tale patologia.

Rispetto al luogo di residenza la composizione degli indagati è variegata: dai dati disponibili si evince, infatti, che i pedofili on line sono presenti in tutte le

² Nucleo Investigativo telematico (N.I.T.): nucleo interforze di polizia giudiziaria, specializzato in reati informatici, istituito nel 2001 presso la Procura della Repubblica di Siracusa.

regioni con percentuali variabili che riflettono, nella maggior parte dei casi, la diversa densità di popolazione. Nelle prime cinque regioni -Lombardia, Veneto, Sicilia, Emilia Romagna e Lazio- si concentrano più della metà degli indagati (57%).

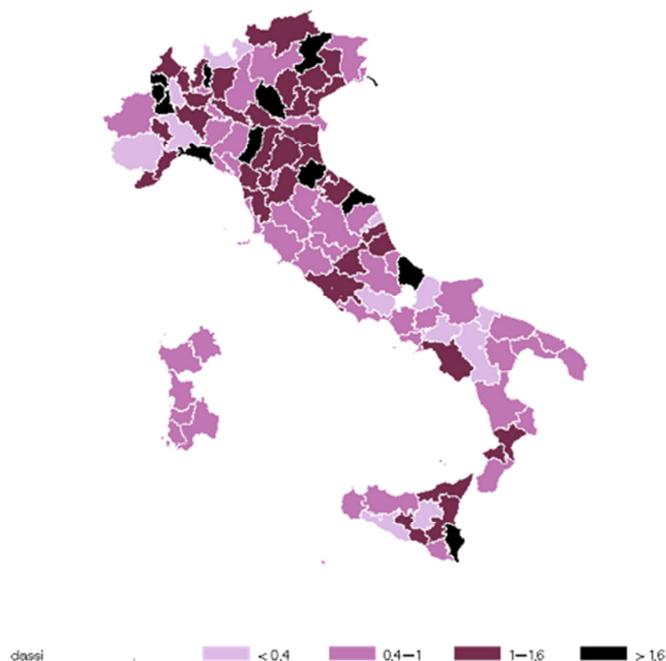
La rappresentazione degli indagati per provincia evidenzia delle concentrazioni maggiori della media italiana (4 casi di pedofilia on line emersi per 100mila abitanti maschi) nell'area della pianura padana e nella Sicilia orientale (cfr Figura 2).

Tabella 2 – Nazionalità dei fruitori di pedopornografia in rete – 1° trimestre 2011.

| Rank | Paese | Media mensile | % |
|------|--------------------|---------------|------|
| 1 | United States | 2.867 | 23% |
| 2 | Germany | 2.099 | 17% |
| 3 | Russian Federation | 745 | 6% |
| 4 | Italy | 692 | 6% |
| 5 | United Kingdom | 499 | 4% |
| 6 | France | 426 | 3% |
| 7 | Ukraine | 398 | 3% |
| 8 | Canada | 376 | 3% |
| 9 | Spain | 270 | 2% |
| 10 | Sweden | 263 | 2% |
| 11 | Netherlands | 213 | 2% |
| 12 | Switzerland | 194 | 2% |
| 13 | Czech Republic | 194 | 2% |
| 14 | Poland | 192 | 2% |
| 15 | Brazil | 189 | 2% |
| 16 | Australia | 175 | 1% |
| 17 | Japan | 147 | 1% |
| 18 | Anonymous Proxy | 146 | 1% |
| ... | ... | — | — |
| | Totale | 12.369 | 100% |

Fonte: Elaborazioni su dati Osservatorio Internazionale di Telefono Arcobaleno. Note: a. si tratta di una stima di minimo del numero di utilizzatori (a partire dagli indirizzi IP). Questo tipo di dati fa riferimento a rilevazioni parziali per costruzione. Quindi, pur essendo necessario un approfondimento metodologico per una migliore quantificazione del fenomeno, è di interesse la possibilità di osservare, pur con i limiti indicati, la presenza e la composizione per paese della domanda.

Figura 2 – Popolazione individuata e incriminata per i reati di produzione, detenzione e scambio di materiale pedopornografico sul web, per provincia, periodo 2001-2009, Italia (casi per 100mila abitanti di genere maschile in età 15-79; indici, 1=Italia).



Fonte: Elaborazioni su dati NIT. Note: a. Rappresentazione basata su un migliaio di casi di indagati dal NIT. b. A livello nazionale emersi 4 casi ogni 100mila abitanti di genere maschile.

3. Conclusioni

Un primo studio dei dati disponibili sembra confermare l'ipotesi che la pedofilia on line sia una vera e propria *economia del crimine* che risponde alle più comuni leggi di mercato, con l'offerta che aumenta al crescere della domanda. L'offerta pedopornografica sul web è, infatti, strutturata e organizzata secondo le più comuni regole del marketing: i. siti *pedo promo* collocati ai livelli più visibili e superficiali del web; ii. siti a pagamento che occupano i livelli meno conosciuti e raggiungibili della rete. I pagamenti delle transazioni avvengono on line e sono gestiti da siti internet molto volatili, per lo più allocati presso server statunitensi, olandesi, russi, ucraini e tedeschi e si avvalgono delle più comuni carte di credito. La matrice commerciale della pedofilia on line è ulteriormente confermata dalla preoccupante circostanza che, sempre più spesso, sullo sfondo delle immagini degli abusi

compaiono le inserzioni pubblicitarie di rinomate aziende che operano sui mercati internazionali. Attraverso questo meccanismo, pertanto, i proventi della pubblicità presente in rete in contesti pedopornografici finanziano, di fatto, buona parte dei siti illegali che diffondono la pedofilia sul web, a testimonianza di come il pedobusiness, rischi progressivamente di essere legittimato come uno dei tanti mercati della *new economy*.

Rispetto alla dimensione geografica, i dati rilevati evidenziano una forte variabilità sia dell'offerta sia della domanda di pedofilia in rete che sembrano evolversi in maniera parallela coinvolgendo prevalentemente i paesi europei e il Nord America.

In particolare, la distribuzione geografica dell'offerta di pedofilia on line sembra riflettere la logica della "convenienza" del pedobusiness di localizzare le proprie attività nei paesi con legislazioni più permissive e presso gli ISP ritenuti, di volta in volta, più tolleranti.

L'analisi della situazione italiana, infine, evidenzia la presenza di una dinamica diversa per la domanda e l'offerta: sebbene, infatti, il nostro paese ospiti una quantità minima di siti pedofili, parte dall'Italia una quota importante delle richieste di materiale pedopornografico in rete.

Restano da approfondire le questioni legate alla stima della domanda di pedopornografia in rete, qui utile solo in termini di quote paese, e le aree di ricerca relative ai flussi finanziari e al circuito della pubblicità sugli spazi del web che diffondono materiale pedopornografico.

Ringraziamenti

Si ringrazia chi è intervenuto al convegno SIEDS contribuendo ad un miglioramento del lavoro complessivo.

Riferimenti bibliografici

- AMERICAN PSYCHIATRIC ASSOCIATION 2000. DSM-IV-TR, *Diagnostic and Statistical Manual of Mental Disorders*, A.P.A., Washington DC.
- OSSERVATORIO INTERNAZIONALE DI TELEFONO ARCOBALENO 2011°. *Short Report Febbraio 2011*, Telefono Arcobaleno onlus, Roma.
- OSSERVATORIO INTERNAZIONALE DI TELEFONO ARCOBALENO 2011b. *Short Report Marzo 2011*, Telefono Arcobaleno onlus, Roma.
- OSSERVATORIO INTERNAZIONALE DI TELEFONO ARCOBALENO 2011c. *Short Report Maggio 2011*, Telefono Arcobaleno onlus, Roma.

SUMMARY

Paedophilia is considered a manifestation of psychopathology. At the same time, paedophile behaviour is a crime that is prosecuted under most legal systems.

The advent of the Internet, while broadly positive in its essence, has made new tools available to violate the rights of citizens and, in particular, has opened up a new channel of expression for paedophiles, allowing for the greater ease of the use and dissemination of child pornography.

In the scenario of the web economy, therefore, the deviation of paedophiles has become intertwined with the economic interests of the criminal organizations that manage the market in online paedophilia. The aim of this work involves the study of this unusual form of economic crime, with particular attention being paid to an analysis of the characteristics and dynamics of the supply of and demand for child pornography on the web. This analysis has been conducted through the use of data from non-governmental institutions (source: *Osservatorio Internazionale di Telefono Arcobaleno onlus - the International Observatory of the Italian non profit organization Telefono Arcobaleno*) as well as administrative data obtained during judicial inquiries.

M. LA ROSA, Centro Studi, Telefono Arcobaleno, marika.larosa@telefonoarcobaleno.org

D. CORSO, Centro Studi, Telefono Arcobaleno, daniela.corso@telefonoarcobaleno.org

P. M. CHIODINI, Università degli Studi di Milano Bicocca, paola.chiodini@unimib.it

F. CUPPONE, ISTAT, cuppone@istat.it

G. MANZI, Università degli Studi di Milano, giancarlo.manzi@unimi.it

B.M. MARTELLI, ISTAT, bmartelli@istat.it

F. VERRECCHIA, Centro Studi, Telefono Arcobaleno,
centrostudi@telefonoarcobaleno.org

LA DISTRIBUZIONE GEOGRAFICA DELLA POPOLAZIONE ITALIANA NEI CENSIMENTI DAL DOPOGUERRA AI GIORNI NOSTRI

Fabio Lipizzi

1. Introduzione

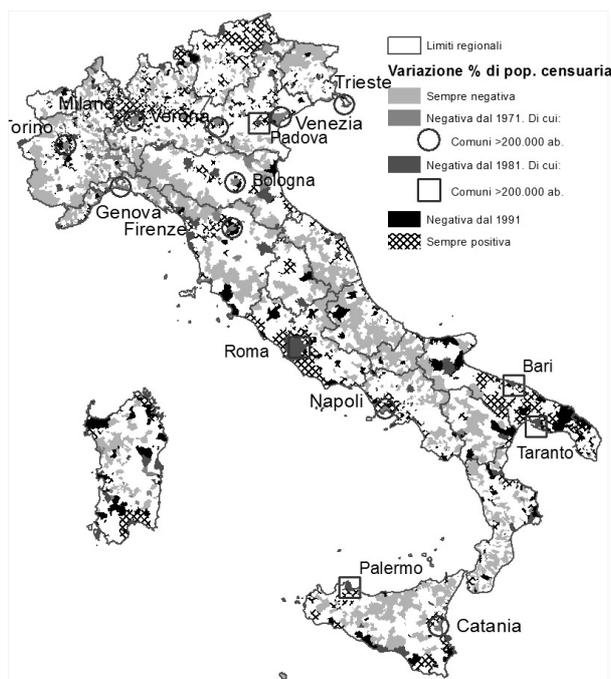
L'obiettivo della comunicazione è la descrizione della geografia della popolazione censuaria dal II dopoguerra ai nostri giorni (1951-2001). La ricomposizione di un quadro complessivo, infatti, può senz'altro essere ricondotto alle diverse fasi dello sviluppo economico che, con diversa intensità, ha interessato nei decenni il nostro Paese. Per sintetizzare i fenomeni demografici stratificati nello spazio e nel tempo, sono stati applicati gli indici di autocorrelazione globale (Indice di Moran) e locale (Local Indicator for Spatial Association - LISA) alle variazioni percentuali di popolazione intercensuarie. Infine, per valutare il complesso fenomeno dell'urbanesimo sono riportati alcuni dati relativi alle basi territoriali. Le potenzialità interpretative insite nelle località abitate sono, infatti, molto promettenti per descrivere i fenomeni in ambito sub comunale.

2. Il fenomeno migratorio nell'Italia repubblicana

L'Italia repubblicana, dal II dopoguerra fino ai giorni nostri, ha attraversato stagioni d'intensi cambiamenti demografici che hanno modificato la geografia del territorio urbanizzato e la struttura della popolazione residente. Fino alla metà degli anni '60 i consistenti flussi migratori, provenienti dalle aree più depresse del paese, si sono indirizzati prevalentemente verso i comuni con una maggiore dimensione demografica. Lo sviluppo della terziarizzazione ha spinto i flussi migratori di piccolo e medio raggio prevalentemente verso i comuni capoluogo di provincia, dando luogo al fenomeno dell'urbanizzazione e al lento abbandono delle campagne. Tuttavia, nei primi anni '60 hanno origine anche imponenti flussi migratori di lungo raggio che si dispongono lungo il triangolo industriale (Genova, Milano e Torino) del nord-ovest della penisola italiana. Nel corso dei decenni esaminati l'Italia muta profondamente la propria vocazione produttiva. Il settore

primario che ha caratterizzato a lungo la storia del nostro Paese, lascia progressivamente spazio all'industrializzazione del II dopoguerra. Il fenomeno si protrae con intensità fino agli inizi degli anni '70, nel periodo del così detto "boom economico" e s'intensifica anche verso la pianura ed i comuni litoranei. Negli anni '80, invece, si registra un radicale mutamento del fenomeno migratorio fin qui delineato. La spinta migratoria della popolazione dell'Italia meridionale si ridimensiona drasticamente, quella verso l'estero si era conclusa già alla fine degli anni '70, e contemporaneamente si segnalano i primi, ma ancora limitati, flussi di migranti provenienti dai paesi in via di sviluppo. Nelle grandi città si assiste ad una perdita di popolazione e ad un aumento di flussi di pendolarismo dai comuni limitrofi, che viceversa, ravvisano un aumento di popolazione tipico della conurbazione urbana delle grandi città con il proprio *hinterland*.

Figura 1 – Comuni per variazione percentuale di popolazione. Anni 1951-2001.



E' stata effettuata una prima analisi quantitativa degli avvenimenti sommariamente descritti, utilizzando le variazioni di popolazione intercensuaria dal II dopoguerra (1951-2001). Per i comuni più grandi (Roma, Venezia e Cagliari) si è proceduto alla ricostruzione della serie storica ai confini attuali, stimando i dati oggetto di variazione territoriale. Nella Figura 1 e nella Tabella 1,

sono rappresentati i comuni che nel corso degli anni presentano variazioni di popolazione intercensuarie:

sempre positive – questo insieme è costituito da 1.169 comuni ed è geograficamente distribuito lungo le coste, nella pianura padana e nelle aree limitrofe ai grandi comuni. In termini di popolazione rappresenta il 25 per cento della popolazione complessiva italiana al 31.12.2009. Oltre il 30 per cento di questi comuni sono rappresentati nella prima classe di popolazione e nella classe tra 10.000 e 50.000 abitanti (Tabella 1).

Tabella 1 – Numero e percentuale di comuni con variazione di popolazione in diminuzione per classi di popolazione, e popolazione al 2009.

| Classi di pop. | Comuni per variazione di popolazione in diminuzione | | | | | | | | | |
|------------------|-----------------------------------------------------|-------|------------|-------|-----------|-------|-----------|-------|-----------|-------|
| | mai | | sempre | | dal 1971 | | dal 1981 | | dal 1991 | |
| | Num. | % | Num. | % | Num. | % | Num. | % | Num. | % |
| < 5.000 | 424 | 36,3 | 1.764 | 95,0 | 45 | 34,1 | 71 | 35,1 | 78 | 26,9 |
| 5.000-10.000 | 333 | 28,5 | 74 | 4,0 | 21 | 15,9 | 26 | 12,9 | 65 | 22,4 |
| 10.001-50.000 | 379 | 32,4 | 19 | 1,0 | 39 | 29,5 | 67 | 33,2 | 123 | 42,4 |
| 50.001-100.000 | 26 | 2,2 | 0 | 0,0 | 13 | 9,8 | 23 | 11,4 | 22 | 7,6 |
| 100.001-200.000 | 7 | 0,6 | 0 | 0,0 | 5 | 3,8 | 10 | 5,0 | 2 | 0,7 |
| > 200.000 | 0 | 0,0 | 0 | 0,0 | 9 | 6,8 | 5 | 2,5 | 0 | 0,0 |
| Totale | 1.169 | 100,0 | 1.857 | 100,0 | 132 | 100,0 | 202 | 100,0 | 290 | 100,0 |
| Popolazione 2009 | 2.902.212 | | 15.088.599 | | 8.183.522 | | 9.283.945 | | 5.258.291 | |

Per avere una misura grossolana dell'ordine di grandezza delle quantità esaminate si pensi che nel 1951 la popolazione media di questo sotto insieme di comuni è di circa 6.000 abitanti, mentre, lo stesso insieme nel 2009 aveva una popolazione media pari a circa 13.000 residenti.

Sempre negative – questo insieme è costituito da 1.857 comuni prevalentemente montani, che si distribuiscono lungo la dorsale appenninica, in Liguria, nella zona delle Langhe, in alcune aree della Sicilia e della Sardegna. Il 95 per cento di questi comuni ha una popolazione inferiore a 5.000 abitanti (Tabella 1) e nel corso dei decenni esaminati la popolazione media per comune, si dimezza, da circa 3.000 a 1.500 abitanti. La caratteristica di questo insieme di comuni è senza dubbio il declino demografico, e il ruolo di marginalità dei piccoli comuni montani che non hanno saputo trasformare il vincolo orografico in opportunità economica.

Positive tra i censimenti 1951, 1961 e 1971 e negative gli altri intervalli – questo insieme è caratterizzato soprattutto dai comuni dell'Italia settentrionale con una popolazione superiore a 200.000 abitanti, cui si aggiungono Napoli e Catania. Questi comuni, che in precedenza trainavano lo sviluppo demografico della Penisola, anche a seguito della crisi del settore industriale degli anni '80, iniziano

un nuovo ciclo urbano caratterizzato da un declino demografico costante in tutti i successivi anni di censimento considerati.

Positive tra i censimenti del 1951, 1961, 1971 e 1981 e negative gli altri intervalli – In questa fase si completa il così detto processo di “deurbanizzazione” verso i comuni della cintura urbana anche nei grandi comuni del centro sud (Palermo, Bari, Taranto e Roma).

Positive tra i censimenti del 1951, 1961, 1971, 1981 e 1991 e negative nell’ultimo intervallo – questo insieme è costituito da 290 comuni con una popolazione al 2009 pari all’8,7 per cento. La classe di popolazione modale è quella tra 10.000 e 50.000 abitanti, dove sono concentrati il 42,4 per cento dei comuni.

3. Una sintesi del fenomeno migratorio nell’Italia repubblicana

3.1. Indici di autocorrelazione spaziali

Per sintetizzare il fenomeno fin qui analizzato, è stato calcolato l’indice di autocorrelazione globale di Moran e l’indice di autocorrelazione locale LISA (Anselin, 1995). I due indicatori sono formalizzati di seguito.

Per il calcolo dell’indice di autocorrelazione spaziale è necessario definire la contiguità tra i poligoni che identificano le n unità amministrative prese in esame, e nel quale è ripartito lo spazio S . Nel caso specifico il criterio che definisce la matrice \mathbf{W} di contiguità, con generico elemento w_{ij} , è il reciproco della distanza euclidea d_{ij} tra due centroidi interni ai poligoni i e j . La matrice \mathbf{W} è standardizzata per riga e la somma di ciascuna riga è pari a 1. Sono stati quindi calcolati gli indici I di autocorrelazione globale di Moran secondo la:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} z_i z_j}{\sum_{i=1}^n z_i^2} \quad (1)$$

dove $z_i = (x_i - \bar{X})$ è lo scarto tra la variabile oggetto di analisi x_i della generica unità i e la sua media, e $S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$ è la somma di tutti gli elementi della matrice di distanza. L’indice I_i di autocorrelazione locale è invece definito dalla seguente:

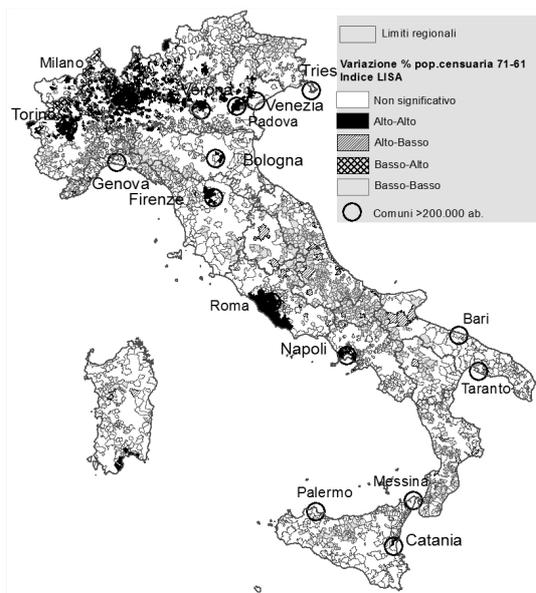
$$I_i = \frac{x_i - \bar{X}}{S_i^2} \sum_{j=1, j \neq i}^n w_{i,j} (x_j - \bar{X}) \quad (2)$$

Dove $S_i^2 = \frac{\sum_{j=1, j \neq i}^n w_{i,j}}{n-1} - \bar{X}^2$

3.2. Un'applicazione degli indici di autocorrelazione spaziali

Gli indici di autocorrelazione globale e locale sono stati applicati ai dati dei 3.650 comuni descritti nei cinque gruppi del precedente paragrafo. La variabile osservata è la variazione percentuale della popolazione intercensuaria. L'indice di Moran assume valori sempre superiori a 0,5 (Tabella 2) ed è sempre statisticamente significativo (con livelli superiori a 0,001), ciò esprime la capacità dell'indicatore di assumere valori simili con una certa regolarità, tra località prossime geograficamente. In altre parole, questo denota un generale effetto contagio tra comuni limitrofi. Tuttavia, questo indicatore non segnala quali sono i comuni che tendono a raggrupparsi sul territorio secondo un preciso *trend* spaziale.

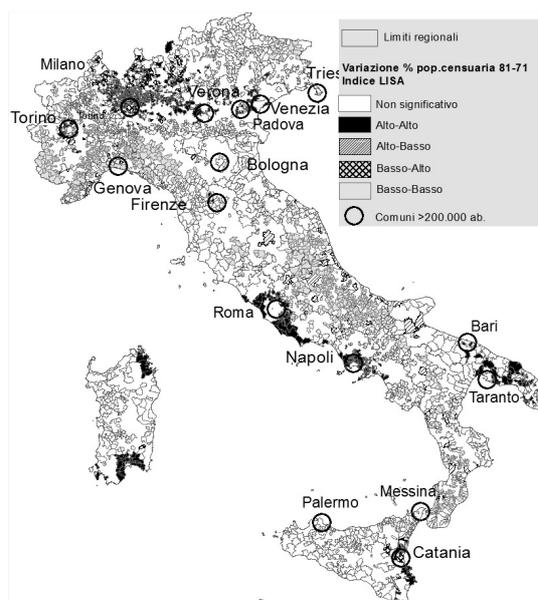
Figura 2 – Indice LISA della variazione intercensuaria di popolazione 1971-1961.



Per avere un'idea più precisa della distribuzione geografica delle variazioni di popolazione sono stati calcolati gli indici di autocorrelazione locali. Due tra le elaborazioni più rappresentative sono cartografate nelle figure 2 e 3. In questo caso i due cartogrammi sono tematizzati per valori statisticamente significativi (con un livello almeno di 0,05) in quattro modalità distinte. Ognuna di queste consente d'identificare *cluster* o, viceversa, *outlier* spaziali (Atzeni *et al.* 2004). In particolare le modalità Alto-Alto o Basso-Basso identificano unità che presentano valori della variabile elevati/bassi associata ad elevati/bassi valori delle unità

vicine. Con le modalità Alto-Basso o Basso-Alto, invece, sono localizzati valori alti/bassi della variabile associata a bassi/alti valori dell'intorno dei comuni.

Figura 3 – *Indice LISA della variazione intercensuaria di popolazione 1981-1971.*



L'aspetto interessante di questa rappresentazione è che si riconoscono i sottoinsiemi di comuni descritti nel paragrafo 2. In particolare: Alto-Alto: descrive la crescita dei grandi comuni e dei loro contermini; Basso-Basso: sono i comuni con una popolazione sempre in diminuzione; Alto-Basso: sono i comuni che attraggono popolazione dai comuni limitrofi; Basso-Alto: sono i comuni che perdono popolazione nei comuni limitrofi.

Il confronto tra i cartogrammi ci consente di apprezzare i cambiamenti che nel corso del decennio sono intervenuti in ambito locale. Per tutti i grandi comuni sono visibili le variazioni di popolazione intercensuaria in diminuzione, se confrontata con la crescita dei comuni delle cinture; più marcata in Figura 3, per Roma e per i comuni del sud dell'Italia, ma anche per Cagliari, Nuoro e alcuni comuni pugliesi che nel periodo 1981-1971 ripercorrono, con un decennio in ritardo rispetto ai comuni del nord, le fasi del ciclo urbano. Particolarmente interessante è anche la configurazione Alto-Basso, dove alcuni comuni del centro-sud di medie dimensioni (Perugia, Ascoli Piceno, L'Aquila, Potenza, Foggia etc.), sono attornati da comuni con popolazione sempre in diminuzione o statisticamente non significativi, presentando le tipiche caratteristiche di *outlier* spaziali. Un'ultima osservazione

riguarda i comuni del nord-ovest situati prevalentemente in Piemonte e in parte in Lombardia. Questi e soprattutto nel periodo 1971-1961 sono classificati con le modalità Basso-Alto. Sono, infatti, i comuni confinanti con le cinture dei comuni di Torino e Milano che in quello stesso periodo presentano variazioni percentuali di popolazione in aumento, anche per effetto del ciclo di vita delle grandi città.

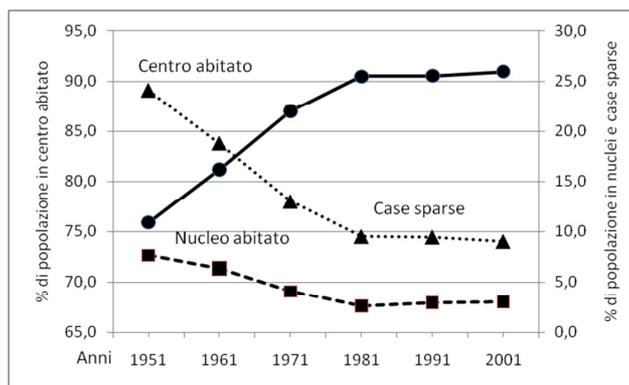
Tabella 2 – *Indice di Moran (con $P\text{-value} > 0.001$) nei 3.650 comuni applicato alle variazioni percentuali di popolazione. Anni 1951-2001.*

| Moran's Index | Variazioni negli anni | | | | |
|---------------|-----------------------|-----------|-----------|-----------|-----------|
| | 1961-1951 | 1971-1961 | 1981-1971 | 1991-1981 | 2001-1991 |
| | 0,573 | 0,602 | 0,549 | 0,554 | 0,535 |

4. Il fenomeno dell'urbanizzazione attraverso i dati delle basi territoriali

Uno studio mirato al fenomeno dell'urbanizzazione, non può prescindere dai dati rilevati ad un livello territoriale sub comunale. L'utilizzo dei dati delle basi territoriali per la rilevazione censuaria, suddivisi in centri abitati, nuclei abitati e case sparse (Lipizzi, 2007) ne sono una delle naturali sedi di analisi.

Figura 4 – *Percentuale di popolazione nelle località abitate. Anni 1951-2001.*



In effetti, per avere un'idea di massima dell'importanza del fenomeno dell'urbanesimo in Italia, ci basta ricordare che nel 1951 la popolazione italiana per centro abitato è pari a 1.496 residenti e nel 2001 si è quasi raddoppiata (2.392 residenti per centro abitato). Come si nota nella Figura 4, che riporta la distribuzione percentuale della popolazione nel corso dei censimenti per località abitata, si segnala un progressivo abbandono delle aree di case sparse di maggiore

intensità nel II dopoguerra fino al censimento del 1981 e minore intensità in tempi più recenti, dove per altro matura il recente fenomeno dell'*urban sprawl*. Di segno opposto è invece la distribuzione della popolazione nei centri abitati che vede un aumento di quote di popolazione sebbene, anche in questo caso, con intensità differenziate negli anni. Segnaliamo infine una fase di rallentamento del fenomeno dell'urbanizzazione a partire dal censimento del 1981.

Riferimenti bibliografici

- ANSELIN, L. 1995. Local Indicators of Spatial Association—LISA, *Geographical Analysis*, Vol. 27, No. 2, pp. 93–115.
- ATZENI S., DETTORI B., USAI S. 2004. L'Econometria spaziale nelle indagini territoriali. Appunti metodologici e un'applicazione alla Sardegna. in *Temi Economici della Sardegna, Quaderni di lavoro, 2004/01*, CEUC.
- LANZANI A. 2003. *I paesaggi italiani*. Roma: Meltemi.
- LIPIZZI F. 2007. Le basi territoriali per i censimenti e le variazioni territoriali tra passato e presente in *Storia e misura. Indicatori sociali ed economici nel mezzogiorno d'Italia (secoli XVIII-XX)*, Frakenco Angeli editore, 139-142, 198-212.

SUMMARY

From a historical point of view, to analyze how the geographic distribution of the Italian population has changed after the war to date is an interesting challenge. The reconstruction of the time series of population census data of the Italian municipalities allows to study the demographic flows during the examined period. In this paper the municipal time series is analyzed with the aim to understand the multiple combination of elements determining the spatial distribution of the population in recent decades. The natural change, the net migration, the repulsive and attractive factors - such as mountain areas exclusion or industrialized localities gravity - contribute in various ways to define the pattern of built-up areas. As well as municipal data, the population census provide spatial details at sub-municipal level for inhabited areas: the time series of the number of localities with the corresponding population from 1951 and 2001, at regional level, is another interesting dataset for investigating the rural-urban migration dynamic. Some geostatistical tools, applied to the Geographic Information System of the census maps, are presented in order to map clusters and to measure the geographic distributions.

IL RUOLO DELLE MICROIMPRESE NEI PROCESSI DI ENTRATA E USCITA DAL MERCATO DEL LAVORO

Silvia Lombardi, Paola Naddeo

1. Introduzione

Il Regolamento CE (2003) 1422 definisce la micro impresa come “un’impresa che occupa meno di 10 persone e realizza un fatturato annuo oppure un totale di bilancio annuo non superiori a 2 milioni di Euro”. È definita invece impresa una qualsiasi unità che esercita un’attività economica, a prescindere dalla forma giuridica adottata. Tra le microimprese rientrano, pertanto, anche i lavoratori autonomi. Le microimprese svolgono in Europa, ed in particolare in Italia, un ruolo economico assai rilevante: secondo i dati Asia-ISTAT, in Italia, nel 2009 le microimprese rappresentano il 95% dell’universo imprese e occupano il 47% degli addetti; il 65% delle imprese non ha dipendenti ed impiega il 15% degli addetti (Istat, 2011).

La letteratura economica ha evidenziato numerose ragioni che possono favorire la diffusione delle microimprese. Dal lato dell’offerta, Blanchflower (2000) e Kim (2007) mostrano, rispettivamente per i paesi Ocse e per gli USA, che la probabilità di autoimpiego è maggiore per gli uomini rispetto alle donne, cresce con l’età ed è più elevata per gli individui con un minor grado di istruzione o per i soggetti con elevata qualifica. Secondo Blanchflower, inoltre, esisterebbe una correlazione negativa tra autoimpiego e tasso di disoccupazione (non rilevata però per l’Italia). Al riguardo Thurik, Carree, van Stel, Audretsch (2008) sostengono che la relazione tra disoccupazione e autoimpiego può essere spiegata attraverso due effetti distinti: un effetto “rifugio” secondo cui alti tassi di disoccupazione spingono verso l’autoimpiego e l’effetto opposto, definito “imprenditoriale”, in base al quale l’autoimpiego determina un calo della disoccupazione.

Bönte, Falck, Heblich (2009) individuano, anche attraverso un’analisi della letteratura esistente, una relazione a forma di U invertita per la scelta di diventare per la prima volta imprenditore, con un picco che si assesterebbe attorno a 40 anni di età. Inoltre, considerano che tanto è maggiore la diffusione della micro imprenditorialità e dell’autoimpiego in una determinata area geografica, tanto maggiore è la probabilità che nascano nuovi imprenditori.

A partire dagli anni ottanta dello scorso secolo, in Europa e negli Stati Uniti si è iniziato a considerare l'autoimpiego come una forma per combattere la disoccupazione, specie giovanile o per permettere agli anziani di permanere nel mercato del lavoro (cfr., ad esempio, Zhang, 2008). La Strategia Europea per l'occupazione, come è noto, si basa su 4 pilastri: occupazione, imprenditorialità, adattabilità e pari opportunità; la capacità di creare impresa viene dunque considerata come una delle strade da perseguire per dare una chance ai giovani e alle donne (Marini, 2008). In Italia, il principale intervento normativo è rappresentato dal D.Lgs. n. 185/2000. Il tema è stato recentemente ripreso nel Libro Bianco sul futuro del modello sociale di maggio 2009. Non si può trascurare come spinte verso forme crescenti di autoimpiego possano determinare una "precarizzazione" dell'occupazione, circostanza che si verifica specie quando il lavoratore autonomo ha un unico committente (Muehlberger, Pasqua, 2009).

2. Obiettivo del lavoro e metodologia

Il presente lavoro ha come obiettivo l'analisi della diffusione territoriale e delle caratteristiche della micro-imprenditorialità nelle province italiane per le classi di età estreme, al fine di individuare anche possibili suggerimenti di policy. Il lavoro si concentra sulle microimprese condotte da giovani di età 25-34 anni e da soggetti di età 55-64 anni, privilegiando un'ottica di genere. La prima classe di età viene considerata rappresentativa dei soggetti che entrano per la prima volta nel mercato del lavoro (gli individui appartenenti alla classe di età 15-24 spesso non hanno ancora compiuto un ingresso definitivo nel mercato del lavoro), mentre la seconda di quei soggetti prossimi all'uscita dalle forze di lavoro.

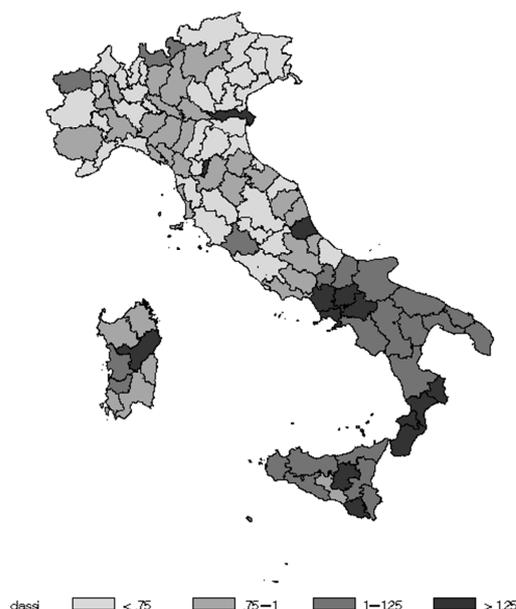
La fonte dei dati utilizzata è rappresentata dall'archivio statistico delle imprese non agricole attive (Asia-ISTAT) per il 2007, che contiene numerose informazioni sulle imprese e i relativi titolari. Dopo un'analisi descrittiva dei microimprenditori a livello provinciale e settoriale (Ateco 2007), si procede ad un'analisi in componenti principali, aggiungendo altre variabili di fonte Istat, che caratterizzano i diversi contesti provinciali. Obiettivo è verificare se esistono pattern di comportamento distinti tra le province italiane.

3. Le microimprese secondo i dati ASIA

In Italia ci sono circa 3 milioni di imprenditori nelle microimprese, di cui il 18,8% sono giovani e il 15,5% anziani. Le province in cui l'incidenza dei giovani è maggiore sono Caserta (23,1%), Prato e Napoli (22,2%), mentre quelle in cui

incidenza è minore sono Trieste e Bolzano (15,4%) e Belluno (15,5%). Alcune delle province in cui è minore l'incidenza dei giovani sono invece quelle in cui è maggiore l'incidenza degli anziani (Udine e Padova - 18,8% - e Bolzano - 18,1%), e viceversa (Caserta - 11,2%, Napoli - 12,1% - e Crotone - 12,9%). In generale si osserva che nelle Regioni meridionali vi è una maggiore presenza di imprese condotte da giovani e minore di anziani, mentre il contrario avviene nelle regioni settentrionali, specie del Nord-Est. Il rapporto tra imprenditori giovani e anziani può essere considerato, con tutti i *caveat* del caso (ad esempio nel tempo si modificano le competenze, le professioni, le condizioni di mercato ecc.), come un indicatore di ricambio generazionale. Il ricambio generazionale risulta quindi superiore nel Mezzogiorno.

Figura 1 – *Indice di ricambio generazionale nella manifattura.*

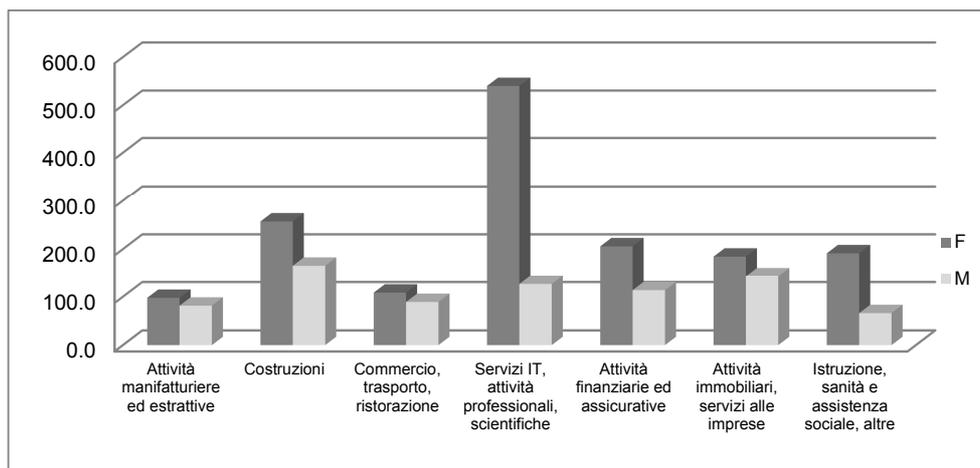


Fonte: Asia-Istat (2007).

Le donne rappresentano il 26,7% del totale degli imprenditori, con una frequenza che decresce all'aumentare dell'età: si passa dal 31,6% dei giovani al 22,2% degli anziani. Anche per gli imprenditori si rileva il dato ben noto secondo cui le penalizzazioni per le donne aumentano al crescere dell'età, come effetto delle decisioni di maternità e delle attività di cura. Il settore di attività economica in cui maggiore è il ricambio generazionale, per le donne, è quello dedicato ai servizi

di informazione e comunicazione e alle attività professionali, scientifiche e tecniche, mentre minore è nel comparto della manifattura.

Figura 2 – Ricambio generazionale per settore di attività economica e genere.



Fonte: Asia-Istat (2007).

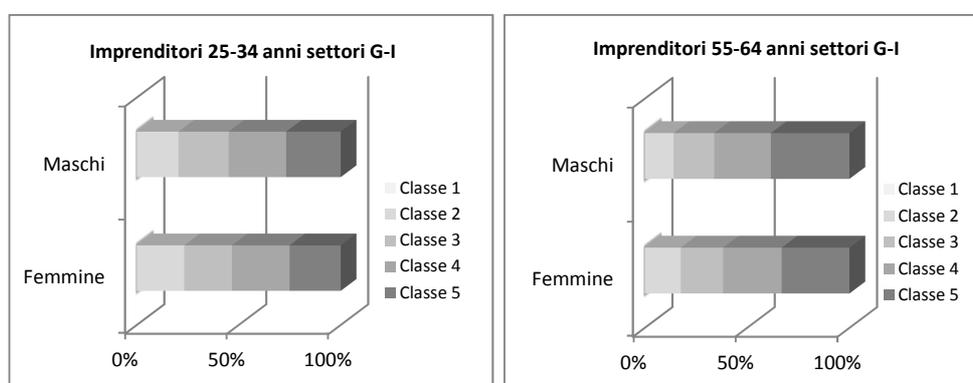
Per quanto concerne la forma giuridica, il 62% dei microimprenditori sono imprenditori, il 25% liberi professionisti e il 13% lavoratori autonomi. Le percentuali non variano significativamente quando si considerano soltanto i giovani microimprenditori sia maschi che femmine, invece che il complesso dei lavoratori.

I microimprenditori si concentrano prevalentemente nel settore del commercio, trasporti e ristorazione (sezioni G-I Ateco 2007), rispettivamente il 41% dei maschi e il 36% delle femmine; elevata è anche la presenza nel settore dei servizi IT e attività professionali, il 22% e il 21%. Una quota consistente di microimprenditori di genere femminile è attiva nei servizi legati all'istruzione, sanità e assistenza sociale (il 24% contro il 10% degli uomini), mentre la loro presenza nel settore delle costruzioni è trascurabile (l'1% contro il 19% dei maschi).

I dati ASIA mostrano che in media i microimprenditori realizzano un volume di affari pari a poco più di 80 mila euro. Tale volume cresce con l'età, passando da poco più di 56 mila euro per i giovani a circa 97 mila euro per gli anziani. Il fatturato dei giovani è inferiore di circa 30 punti percentuali rispetto a quello medio nazionale; tale divario è maggiore per le donne rispetto agli uomini. I microimprenditori anziani realizzano, invece, un fatturato di circa il 21% superiore a quello medio nazionale. In sostanza, indipendentemente dai fenomeni di crescita dimensionale, i microimprenditori tendono a veder crescere sensibilmente il fatturato al crescere dell'età.

Analizzando le sezioni G-I (Ateco 2007) si possono vedere rappresentate nella figura 3 le diverse posizioni dei giovani e degli anziani per classi di fatturato: nella prima classe (il complemento a 100) sono inseriti coloro che dichiarano un fatturato nullo, mentre le altre classi rappresentano i quartili rispetto alla media nazionale. Come si evince dalla figura 3 nel quinto quartile vi è una sotto-rappresentazione delle donne e dei giovani rispetto agli anziani.

Figura 3 – *Giovani e anziani per classi di reddito*



Fonte: Asia-Istat (2007)

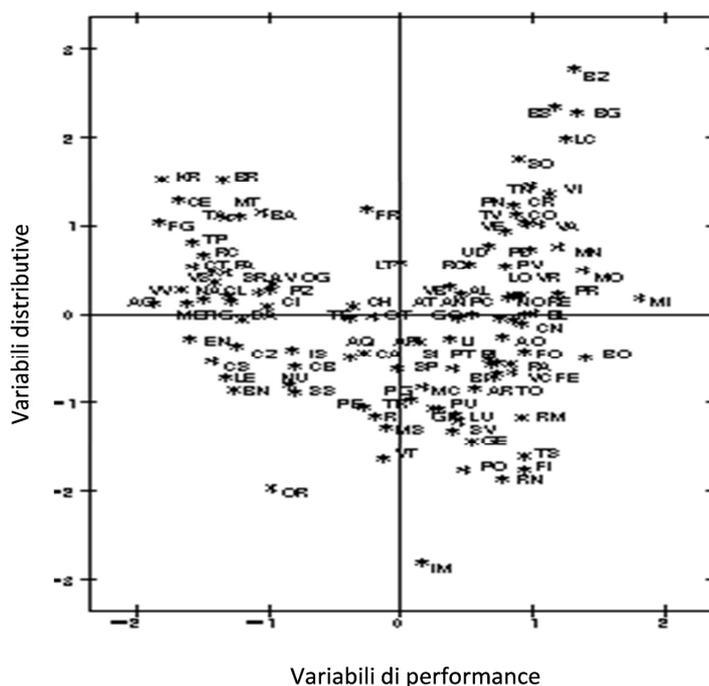
Notevole è lo scostamento del fatturato medio a livello provinciale: si passa dai circa 68 mila di Benevento agli oltre 107 mila di Bolzano. In generale le province in cui il fatturato è inferiore sono nel Mezzogiorno o contengono grandi città (come Roma, Napoli, Firenze e Genova). Il confronto tra realtà estremamente eterogenee risente, oltre della diversa propensione a sotto-dichiarare il fatturato effettivo, anche di elementi quali le diverse strutture produttive e il fatto che in determinati contesti, ad esempio le grandi città, è più facile riscontrare la presenza di imprese marginali che possono sopravvivere ai limiti della legalità.

Il gender gap (calcolato come differenza percentuale tra fatturato maschile e femminile su quello maschile) è pari al 25,9%, riducendosi all'aumentare dell'età (il 28,6% dei giovani contro il 18,8% degli anziani). Ciò farebbe ritenere che la minore persistenza delle donne riguarda soprattutto le microimprenditrici attive in imprese con minore fatturato. A livello provinciale le differenze sono sensibili, oscillando da un gap praticamente assente per la provincia di Oristano (solo lo 0,6%) al 36,3% della provincia di Bolzano. Si osserva una certa correlazione tra gender gap e fatturato pro-capite e con un valore dell'indice di correlazione che assume un valore non trascurabile, pari a 0,4. Si noti che in alcune province del Sud e delle Isole il gender gap tende a divenire negativo per le classi anziane.

4. L'analisi in componenti principali

L'analisi condotta sui dati ASIA sembra indicare che, specie per le donne e nel Mezzogiorno, la micro-imprenditorialità possa rappresentare una modalità di ingresso nel mercato del lavoro, anche se con fatturati piuttosto modesti. Non emerge, invece, una chiara evidenza circa il valore della micro-imprenditorialità come alternativa all'espulsione dal mercato del lavoro per gli anziani. Anche per i microimprenditori sembrano riprodursi forti differenziali di reddito, così come avviene nel mercato del lavoro dipendente. Le donne realizzano in media un fatturato di circa il 30% inferiore rispetto a quello degli uomini; mentre il fatturato degli anziani supera quello dei giovani di oltre il 70%. Si registra, inoltre, un minore attaccamento delle donne alla vita attiva, specie per quelle con fatturati inferiori.

Figura 4 – Il posizionamento delle province italiane



Fonte: Asia-Istat (2007)

Al fine di comprendere meglio le differenze provinciali sono stati incrociati i dati di fonte ASIA precedentemente analizzati con altri dati di fonte Istat, relativi sempre al 2007, idonei a fornire un quadro dei diversi contesti territoriali. In particolare, sono stati considerati anche alcuni dati di contabilità territoriale (il valore aggiunto per abitante e per ULA, la quota del valore aggiunto dell'industria sul totale e la quota di indipendenti su totale), e alcuni relativi al mercato del lavoro (il tasso di disoccupazione, totale e giovanile, il tasso di attività femminile), nonché alcuni dati demografici (la presenza di immigrati e il numero medio di figli per donna). L'analisi in componenti principali sembra spiegare una parte rilevante della variabilità totale: il primo autovalore spiega circa il 76% della variabilità totale, mentre il secondo quasi il 7%.

Il primo autovalore è correlato con segno positivo con variabili quali il valore aggiunto per abitante e i tassi di attività femminili (totali e giovanili), mentre con segno negativo con quelle relative ai tassi di disoccupazione e all'incidenza degli imprenditori sul totale, nonché al gap generazionale di fatturato. Tutto ciò induce a ritenere che alla spiegazione del primo autovalore concorrano soprattutto le variabili di performance. Nei quadranti a destra tendono a posizionarsi le province del Centro-Nord, mentre in quelli di sinistra le province del Sud e delle Isole.

Il secondo autovalore è spiegato positivamente da variabili relative al fatturato, quali la quota del fatturato provinciale sul totale nazionale e gli indicatori di gender gap (totale e giovanile), mentre con segno negativo da variabili relative alla partecipazione femminile. Pertanto, il secondo autovalore è associato principalmente a variabili di tipo distributivo. Il posizionamento delle province italiane rispetto a questa dimensione è meno netto rispetto a quello del primo autovalore. In sostanza le province del Nord appaiono essere meno egualitarie, per quanto riguarda i fatturati dei microimprenditori, rispetto a quelle del Centro Italia, mentre si mostrano sensibili differenze per le province del Sud Italia.

5. Conclusioni

Le analisi condotte hanno mostrato una situazione molto variegata a livello territoriale per il ricorso alla micro-imprenditorialità. Tra i risultati che sembrano emergere vi è quello per cui nei contesti caratterizzati da un minor valore aggiunto pro-capite e da condizioni del mercato del lavoro insoddisfacenti, il ricorso all'autoimpiego può risultare, soprattutto per i giovani, un modo per entrare o rimanere nel mercato del lavoro.

I risultati ottenuti suggeriscono l'opportunità di ulteriori approfondimenti volti a verificare se questa misura di inserimento possa essere considerata uno stepping stone o si traduca, invece, in una trappola per i giovani che la sperimentano (non si

può trascurare, infatti, che tra i microimprenditori vi sono anche molte “false” partite IVA).

Riferimenti bibliografici

- BLANCHFLOWER D.G. 2000. Self-employment in OECD countries, *Labour Economics*, Vol. 7, pp. 471-505.
- BÖNTE W., FALCK O., HEBLICH S. 2009. The Impact of Regional Age Structure on Entrepreneurship, *Economic Geography*, Vol. 85, No.3, pp. 269-287.
- ISTAT 2011. *Struttura e dimensione delle imprese*. Roma.
- KIM G.S. 2007. The analysis of self-employment levels over the life-cycle, *The Quarterly Review of Economics and Finance*, Vol. 47, No.3, pp. 397-410.
- MARINI G. 2008. Le politiche attive del lavoro nei Paesi europei fra flexicurity e mito del “far da sé”: gli incentivi agli start-up, *La cittadinanza europea*, Anno V, No. 1-2, pp. 215-231.
- MUEHLBERGER U., PASQUA S. 2009. Workers on the Border between Employment and Self-employment, *Review of Social Economy*, Vol. 67, No.2, pp. 201-229.
- THURIK A.R., CARREE M.A., VAN STEL A., AUDRETSCH D.B. 2008. Does self-employment reduce unemployment?. *Journal of Business Venturing*, Vol. 23, No. 6, pp. 673-386.
- ZHANG T. 2008. *Elderly Entrepreneurship in an Aging US Economy: It's Never Too Late*, *Series on Economic Development and Growth*. Vol. 2, Hackensack, N.J. and Singapore: World Scientific.

SUMMARY

Micro-enterprises play a very important economic role in Europe, particularly in Italy. We analyse the contribution of micro enterprises in favouring the entry and exit processes in the labour market in Italy. The analysis is carried out in a gender perspective and takes into account the territorial contest. One of the main results is that low per capita value added and poor working conditions could encourage the recourse to self-employment as a way to enter or remain in the labour market, especially for young people. We need more investigations to understand if the transition into micro-enterprises can be considered a stepping stone or a trap for young people.

Silvia LOMBARDI, Istat, lombardi@istat.it
Paola NADDEO, Istat, naddeo@istat.it

IL COMPORTAMENTO RIPRODUTTIVO E LE STRATEGIE MIGRATORIE DEGLI STRANIERI PRESENTI IN LOMBARDIA

Eleonora Mussino, Giuseppe Gabrielli, Anna Paterno,
Salvatore Strozza, Laura Terzera

1. Introduzione

Negli ultimi decenni si è assistito in Italia ad una significativa trasformazione della presenza straniera, sempre più orientata verso un tipo di insediamento definitivo, inserito in un progetto di vita vincolato al contesto di adozione. Le trasformazioni in atto emergono da diversi indicatori, spesso legati alle condizioni familiari (ricongiungimenti familiari, incremento di minori e di nascite ecc.). Tale trend induce, quindi, a porre l'attenzione anche sulle dinamiche riproduttive della popolazione immigrata, che appaiono interessanti pure relativamente al modello italiano di fecondità a cui sempre più contribuiscono le donne immigrate.

Dal punto di vista metodologico, tuttavia, le fonti ufficiali disponibili a livello nazionale non sono in grado di fornire i dati necessari per l'analisi della fecondità e ciò deriva dalle peculiarità della popolazione migrante, caratterizzata da elevata mobilità, invisibilità della componente irregolare, brusche emersioni di ampi contingenti di persone a seguito di regolarizzazioni e temporaneità dello status stesso di straniero per acquisizione della cittadinanza. Inoltre la fecondità delle straniere si realizza in parte anche fuori dai confini nazionali.

Per tali ragioni, ci si concentrerà sulle informazioni fornite dall'Osservatorio Regionale sull'Integrazione e la Multietnicità derivanti dall'indagine campionaria, svolta con tecnica di campionamento per centri, sulla presenza straniera in Lombardia nel 2010 (Blangiardo, 2011). L'indagine, svolta annualmente dall'Osservatorio a partire dal 2001, si è arricchita nell'edizione del 2010 di informazioni retrospettive, qui esaminate, inerenti i percorsi riproduttivi. Pur riferendosi ad un ambito territoriale limitato, i dati utilizzati riguardano un contesto che assume il ruolo di "laboratorio sperimentale" delle tendenze migratorie in atto in Italia, in quanto circa $\frac{1}{4}$ della popolazione immigrata in Italia vive in questa regione e tale consistente presenza è una realtà ormai di lunga data.

Nell'analizzare il comportamento riproduttivo delle donne immigrate in Lombardia si osserverà l'interrelazione con la mobilità internazionale, che distorce

pesantemente, influenzando soprattutto l'età al parto, le misure comunemente utilizzate nell'ambito della fecondità di periodo. In particolare, dopo aver osservato i comportamenti riproduttivi in funzione dell'età (non solo al parto), si porrà attenzione sulla fecondità realizzata nel 2009 cercando, tra l'altro, di individuare le determinanti della nascita di un figlio.

2. Differenti angolazioni nella lettura dei comportamenti riproduttivi

2.1. Gli effetti delle età all'intervista e all'immigrazione

Le analisi proposte in questo articolo riguardano il sottoinsieme delle donne (in totale 3.848 casi) facenti parte del campione complessivo degli stranieri maggiorenni intervistati in Lombardia. La distribuzione per classi di età è sensibilmente differente tra le singole cittadinanze (Tabella 1). Le ucraine sono le meno giovani, quasi il 55% ha 40 anni e più e meno del 14% è al di sotto dei 30 anni. Al contrario le indiane, che in quasi la metà dei casi non hanno ancora compiuto il 30° compleanno, costituiscono il gruppo più giovane. La forte eterogeneità nella struttura per età fa sì che le donne, appartenendo a generazioni differenti, si trovino mediamente in fasi diverse del periodo riproduttivo. Il numero medio dei figli per donna, che sintetizza la fecondità già espressa, va pertanto calcolato distintamente per grandi gruppi di età, poiché i valori che si ottengono senza tale distinzione non consentono di apprezzare le diverse intensità riproduttive a causa di un periodo di esposizione che varia da una cittadinanza all'altra. Ad esempio, albanesi e ucraine hanno praticamente lo stesso numero medio di figli (poco meno di 1,3 per donna), ma una differente struttura per età con le prime nettamente più giovani delle seconde. Tra le donne con 40 anni e più, che prevalentemente hanno completato la fase riproduttiva, le albanesi hanno in media più di 2 figli mentre le ucraine poco più di 1,6. Le indiane hanno in media 1,4 figli per donna, valore innalzato dalla precocità dei comportamenti riproduttivi; considerando solo le donne aventi almeno 40 anni la collettività indiana ha circa 2,2 figli per donna, uno dei valori medi più alti insieme a quelli delle comunità africane. Nel caso delle egiziane, nonostante l'ampia proporzione di donne con meno di 30 anni (42%), il numero medio di figli calcolato senza distinzione per età risulta, insieme a quello delle senegalesi, il più elevato (1,7) poiché la fecondità è abbastanza precoce, come dimostrato dal numero medio di 2,1 figli avuti dalle generazioni che nel 2010 hanno 30-39 anni.

Da un punto di vista metodologico, è noto che il numero medio di figli avuti da donne uscite o prossime ad uscire dal periodo riproduttivo (di 45-54 anni) consente di considerare tutte le nascite indipendentemente dal luogo di procreazione (paese

di origine o di accoglimento), ma sconta oltre ai noti problemi di selezione anche il limite di far riferimento a generazioni con comportamenti riproduttivi sensibilmente differenti rispetto a quelli delle generazioni più giovani e, in genere, di più recente immigrazione. D'altronde, anche il tasso di fecondità totale (TFT) del momento, se utilizzato per le popolazioni immigrate, può portare a valutazioni imprecise dei comportamenti riproduttivi delle comunità straniere poiché non tiene conto della discontinuità che la migrazione può provocare (Toulemon, 2004). Per questa ragione è stato suggerito di considerare nell'analisi l'età alla migrazione.

Tabella 1 – *Distribuzione percentuale per classi di età delle immigrate e numero medio di figli avuti per classi di età, distintamente per cittadinanza. Donne straniere dei Pvs e dell'Est Europa presenti in Lombardia a metà del 2010.*

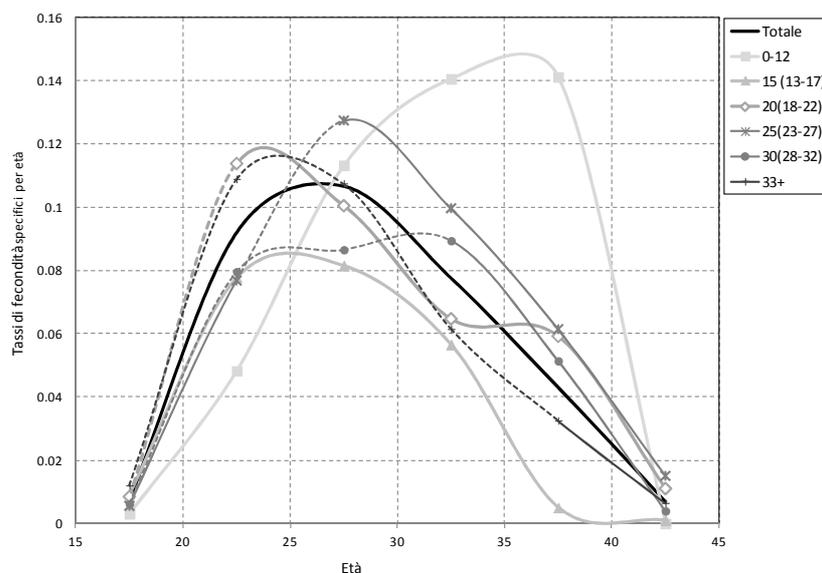
| Paese o aree di cittadinanza | Numero casi | % per classi di età | | | Numero medio di figli | | | |
|---------------------------------|----------------|---------------------|-------|------|-----------------------|-----|-------|------|
| | | <30 | 30-39 | 40 + | Totale | <30 | 30-39 | 40 + |
| Albania | 369 | 41,6 | 31,9 | 26,5 | 1,3 | 0,4 | 1,4 | 2,0 |
| Romania | 493 | 36,7 | 31,9 | 31,4 | 1,3 | 0,4 | 1,5 | 1,8 |
| Ucraina | 224 | 13,9 | 31,6 | 54,5 | 1,3 | 0,3 | 1,1 | 1,7 |
| Altro Est Europa | 303 | 30,7 | 25,5 | 43,8 | 1,1 | 0,3 | 1,0 | 1,6 |
| Cina | 163 | 45,3 | 34,1 | 20,6 | 1,1 | 0,4 | 1,7 | 1,5 |
| Filippine | 207 | 29,6 | 31,1 | 39,3 | 1,1 | 0,2 | 0,7 | 2,0 |
| India | 146 | 48,8 | 36,0 | 15,2 | 1,4 | 0,5 | 1,9 | 2,2 |
| Altro Asia | 275 | 42,4 | 37,3 | 20,3 | 1,2 | 0,3 | 2,0 | 1,7 |
| Egitto | 145 | 42,0 | 37,0 | 21,0 | 1,7 | 0,5 | 2,1 | 2,2 |
| Marocco | 413 | 41,6 | 36,3 | 22,1 | 1,4 | 0,3 | 1,7 | 2,2 |
| Senegal | 69 | 31,2 | 41,8 | 27,0 | 1,7 | 0,3 | 1,7 | 2,7 |
| Altro Africa | 375 | 33,8 | 38,9 | 27,3 | 1,4 | 0,3 | 1,5 | 2,4 |
| Ecuador | 194 | 30,9 | 35,0 | 34,1 | 1,4 | 0,3 | 1,4 | 2,1 |
| Perù | 193 | 35,5 | 24,1 | 40,4 | 1,4 | 0,2 | 1,9 | 2,1 |
| Altro America | 278 | 33,3 | 34,1 | 32,6 | 1,1 | 0,3 | 1,2 | 1,7 |
| TOTALE | 3.848 | 36,0 | 33,4 | 30,6 | 1,3 | 0,4 | 1,5 | 1,9 |

Fonte: ns. elaborazioni su dati della Fondazione ISMU.

I dati in nostro possesso consentono, attraverso le informazioni retrospettive sulla storia migratoria e riproduttiva, di tener conto sia dell'età alla nascita dei figli sia di quella alla migrazione, rendendo possibile l'esplorazione delle relazioni tra mobilità internazionale e procreazione. La migrazione potrebbe avere un impatto negativo sulla fecondità: in base all'effetto *disruption* potrebbe ridurre il comportamento riproduttivo in un periodo più o meno ampio intorno all'evento migratorio. Non va però dimenticato che i comportamenti riproduttivi potrebbero essere essi stessi la causa o la conseguenza dello spostamento migratorio (effetto *interrelation*).

I dati disponibili consentono di analizzare tutto il periodo fecondo (prima e dopo la migrazione) delle donne intervistate distinte in base all'età all'arrivo (Figura 1).

Figura 1 – Tassi di fecondità specifici per età delle donne immigrate per età all'immigrazione in Italia^(a). Donne straniere dei Pvs e dell'Est Europa presenti in Lombardia a metà del 2010.



Nota: (a) La parte tratteggiata è relativa alla fecondità precedente l'età all'immigrazione.
Fonte: ns. elaborazioni su dati della Fondazione ISMU.

I risultati vanno considerati con cautela poiché alcuni gruppi sono numericamente esigui e, in ogni caso, è trascurata la distinzione per cittadinanza. In generale, i tassi di fecondità per età risultano quasi sempre più alti subito dopo l'evento migratorio, tranne che per le donne arrivate dopo i 33 anni. È interessante notare come la distribuzione sia bimodale per le immigrate intorno ai 30 anni, poiché il comportamento riproduttivo risulta più intenso intorno ai 23 anni e nel periodo immediatamente seguente la migrazione. Appare interessante in termini prospettivi il comportamento della seconda generazione e di quelle comprese tra la prima e la seconda. Le donne arrivate intorno ai 15 anni fanno registrare i livelli più bassi di fecondità rispetto a tutti gli altri gruppi considerati, mentre quelle nate in Italia o arrivate entro i 12 anni sembrano adattarsi al modello delle italiane non tanto nell'intensità quanto nel calendario riproduttivo. Si tratta però di dati che meritano approfondimenti e conferme, anche perché il collettivo delle donne di seconda generazione attualmente con 30 anni e più è certamente esiguo e potrebbe risultare anche selezionato. I risultati ottenuti, appaiono però molto simili a quelli di una ricerca condotta sulla fecondità delle donne immigrate nella comunità autonoma della Catalogna (Bueno *et al.*, 2011).

2.2. Una finestra sul 2009: un segmento trasversale della storia riproduttiva

Un ulteriore indicatore della storia riproduttiva è costituito dalla proporzione di donne disaggregate per cittadinanza che hanno avuto un figlio nell'ultimo anno (il 2009). Escludendo, quindi, le migranti arrivate in Italia nel 2010 (381 casi), l'intento è quello di osservare un particolare anno della vita riproduttiva delle intervistate, che è ovviamente condizionato dall'età al momento dell'intervista e dalla storia migratoria e riproduttiva di ciascuna intervistata. Per tener conto dell'età, l'analisi viene effettuata anche dividendo in due grandi classi il collettivo osservato (meno di 30 anni e 30 anni e più); le percentuali dei figli nati nel 2009 sul totale rispettivamente di quelli avuti e di quelli nati in Italia segnano, invece, "le quote osservate" della fecondità già espressa nel complesso e sul nostro territorio.

Tabella 2 – Percentuale di donne immigrate che hanno avuto figli nel 2009 per grandi classi di età e percentuale dei figli nel 2009 sul totale dei figli avuti. Donne straniere dei Pvs e dell'Est Europa presenti in Lombardia a metà del 2010.

| Paese o area di cittadinanza | N. donne in età feconda | % che ha avuto figli nel 2009 | | | % sul totale dei figli avuti | % sul totale di quelli nati in Italia |
|------------------------------|-------------------------|-------------------------------|-----------|----------|------------------------------|---------------------------------------|
| | | Totale | < 30 anni | 30+ anni | | |
| Albania | 326 | 11,3 | 19,6 | 4,9 | 9,5 | 9,9 |
| Romania | 460 | 10,7 | 15,0 | 14,6 | 8,7 | 10,9 |
| Ucraina | 165 | 3,9 | 9,9 | 5,0 | 3,5 | 11,9 |
| Altro Est Europa | 251 | 5,1 | 8,9 | 6,2 | 5,3 | 8,9 |
| Cina | 158 | 20,3 | 19,4 | 26,4 | 18,1 | 28,4 |
| Filippine | 190 | 5,8 | 4,3 | 0,0 | 6,2 | 11,4 |
| India | 143 | 14,2 | 19,6 | 11,7 | 10,3 | 10,7 |
| Altra Asia | 252 | 8,7 | 10,3 | 9,9 | 6,8 | 7,1 |
| Egitto | 134 | 14,6 | 21,8 | 13,0 | 8,5 | 8,7 |
| Marocco | 388 | 15,6 | 17,5 | 18,8 | 12,0 | 12,8 |
| Senegal | 66 | 25,7 | 14,5 | 34,6 | 16,8 | 21,0 |
| Altra Africa | 344 | 13,3 | 14,0 | 15,1 | 10,5 | 11,8 |
| Ecuador | 172 | 8,7 | 7,1 | 11,5 | 7,2 | 8,9 |
| Perù | 165 | 8,5 | 8,0 | 19,1 | 7,2 | 12,5 |
| Altra America | 254 | 7,4 | 10,6 | 9,5 | 7,4 | 9,8 |
| TOTALE | 3.467 | 11,0 | 14,1 | 12,5 | 9,1 | 11,3 |

Fonte: ns. elaborazioni su dati della Fondazione ISMU.

I risultati in tabella 2 mostrano una certa variabilità tra le diverse cittadinanze della fecondità osservata nel 2009. Quest'ultima passa da una quota minima del 3,5% sul totale dei figli avuti complessivamente per le ucraine a una quota massima del 18,1% per le cinesi. La quota osservata della fecondità espressa in Italia varia da un minimo espresso dalle donne provenienti dai paesi asiatici (7,1%, escludendo Cina, Filippine e India), al massimo rilevato tra le cinesi (28,4%). Infine, le ucraine, le donne degli altri paesi dell'Est Europa e le filippine fanno registrare la più bassa percentuale di figli avuti nel 2009; contrariamente, le

senegalesi, le cinesi e le marocchine mostrano i valori più alti. Tali differenze si accentuano considerando le sole donne con trent'anni e più.

2.3. Una finestra sul 2009: alcuni fattori rilevanti della nascita di un figlio

Per includere nell'analisi anche le altre determinanti della nascita di un figlio si è deciso di fare ricorso alla regressione logistica multivariata per stimare gli *odd-ratio* di avere avuto un figlio nel 2009. Appare doveroso precisare che i risultati ottenuti sono di carattere puramente congiunturale e che potrebbero non essere confermati osservando altri anni di calendario. Nella tabella 3a e 3b sono mostrati i risultati di tre diverse applicazioni: il modello 0 considera tra le esplicative solo la cittadinanza della donna; il modello 1 contiene anche i fattori socio-demografici e migratori (presenza o meno di precedenti figli, stato civile, cittadinanza del partner, età, residenza e durata della presenza); il modello 2 contiene le stesse variabili indipendenti del modello precedente ma sono considerate soltanto le donne giunte in Italia prima dei 25 anni. Il modello 0 rappresenta esclusivamente il punto di partenza dell'analisi, poiché mostra la probabilità di aver avuto un figlio nel 2009 per le donne di una data cittadinanza rispetto a quella delle donne albanesi che sono considerate come riferimento (Tabella 3a).

Tabella 3a – *Odd-ratio della regressione logistica binomiale sulla probabilità di avere avuto un figlio nel 2009. Donne dei Pvs e dell'Est Europa in Lombardia, metà 2010.*

| Paese o area di cittadinanza | Modello 0 ^(a) | | Modello 1 ^(b) | | Modello 2 ^(c) | |
|---------------------------------|--------------------------|------|--------------------------|------|--------------------------|------|
| | Odds-Ratio | P> z | Odds-Ratio | P> z | Odds-Ratio | P> z |
| Albania (cat. rif.) | 1,00 | | 1,00 | | 1,00 | |
| Romania | 0,94 | | 1,25 | | 1,28 | |
| Ucraina | 0,32 ** | | 0,62 | | 1,55 | |
| Altro Est Europa | 0,42 ** | | 0,68 | | 0,61 | |
| Cina | 2,00 ** | | 3,18 *** | | 4,07 *** | |
| Filippine | 0,48 * | | 0,77 | | 1,24 | |
| India | 1,30 | | 1,10 | | 0,90 | |
| altra Asia | 0,75 | | 0,91 | | 0,69 | |
| Egitto | 1,34 | | 1,45 | | 1,05 | |
| Marocco | 1,45 * | | 1,57 * | | 1,10 | |
| Senegal | 2,72 ** | | 2,73 ** | | 1,30 | |
| Altra Africa | 1,21 | | 1,92 * | | 1,39 | |
| Ecuador | 0,75 | | 1,62 | | 1,37 | |
| Perù | 0,74 | | 1,53 | | 1,99 | |
| Altra America | 0,63 | | 1,20 | | 1,74 | |

Note: (a) La sola variabile esplicativa è il paese o area di cittadinanza. (b) Le altre variabili esplicative sono quelle riportate nella tabella 4, oltre alla provincia di residenza. (c) Sottoinsieme delle donne arrivate in Italia a meno di 25 anni, stesse variabili esplicative del modello 1. Le differenze statisticamente significative rispetto alla modalità di riferimento (posta uguale a 1) sono indicate con: *** $p \leq 0,001$; ** $p \leq 0,01$; * $p \leq 0,1$.

Fonte: ns. elaborazioni su dati della Fondazione ISMU.

I risultati sono ovviamente in linea con quelli precedenti, mostrando come le ucraine, le donne degli altri paesi dell'Est Europa e le filippine abbiano *odd-ratio* significativamente più bassi rispetto alle albanesi, mentre le senegalesi, le cinesi e le marocchine abbiano espresso probabilità sensibilmente più elevate di avere un figlio. Le cose cambiano nel modello 1, cioè a parità delle caratteristiche socio-demografiche e migratorie: le filippine, le ucraine e le altre donne dell'Est Europa non hanno più rischi significativamente più bassi delle albanesi, mentre si conferma il "vantaggio" delle cinesi e delle africane. La forte eterogeneità tra le principali cittadinanze, già notata in letteratura (Andersson, 2004; Mussino *et al.*, 2009), sembra quindi persistere anche se in maniera meno ampia di quanto atteso. La situazione cambia ulteriormente quando si considerano le sole donne arrivate prima dei 25 anni (modello 2): con la sola eccezione delle cinesi, non si osservano più differenze significative per cittadinanza, sono altri i fattori che entrano in gioco.

È quindi opportuno osservare il ruolo svolto dalle altre covariate (Tabella 3b). Le immigrate straniere senza figli hanno una probabilità più che doppia di avere avuto un bambino nel 2009 rispetto a quelle che ne avevano già almeno uno. Le donne non coniugate hanno un rischio che è meno della metà di quello delle coniugate, quelle con partner italiano poco più della metà di quelle con partner straniero (in genere connazionale). La probabilità decresce al crescere dell'età e risulta più bassa in modo sensibile tra le non residenti. Ancora una volta sembra netta la relazione tra migrazione e fecondità: la probabilità di avere un figlio nel 2009 è significativamente più alta per le donne arrivate da 1-3 anni, confermando l'ipotesi dell'interrelazione tra i due eventi (Andersson, 2004; Milewski, 2007), risultato confermato anche nel sottogruppo delle donne giunte a meno di 25 anni.

Tabella 3b – *Odd-ratio della regressione logistica binomiale sulla probabilità di avere avuto un figlio nel 2009. Donne dei Pvs e dell'Est Europa in Lombardia, metà 2010.*

| Variabili indipendenti | Modello 1 ^(a) | | Modello 2 ^(b) | |
|-------------------------------------------------------|--------------------------|------|--------------------------|------|
| | Odds-Ratio | P> z | Odds-Ratio | P> z |
| Non aveva figli (cat. rif. = aveva figli) | 2,33 | *** | 1,38 | * |
| Non coniugata (cat. rif. = coniugata) | 0,46 | ** | 0,53 | * |
| Partner italiano (cat. rif. = partner straniero) | 0,53 | ** | 0,45 | ** |
| senza partner (cat. rif. = partner straniero) | 0,07 | *** | 0,05 | *** |
| Trenta anni e più (cat. rif. = <30 anni) | 0,45 | *** | 0,61 | ** |
| Non residente (cat. rif. = residente) | 0,26 | *** | 0,56 | |
| 1-3 anni dopo la migrazione (cat. rif. = stesso anno) | 1,57 | * | 1,95 | * |
| 4 anni e oltre (cat. rif. = stesso anno) | 1,04 | | 0,80 | |
| Log Likelihood ^(c) | -983,61 | | -551,46 | |

Note: (a) Le altre variabili esplicative sono il paese o area di cittadinanza e la provincia di residenza. (b) Sottoinsieme delle donne arrivate in Italia a meno di 25 anni, le altre variabili esplicative sono le stesse del modello 1. (c) Log Likelihood Modello 0 = 1.164,04. Le differenze statisticamente significative rispetto alla modalità di riferimento (posta uguale a 1) sono indicate con: *** $p \leq 0,001$; ** $p \leq 0,01$; * $p \leq 0,1$.

Fonte: ns. elaborazioni su dati della Fondazione ISMU.

3. Conclusioni

La presenza straniera in Italia si qualifica sempre più come un fenomeno legato alla formazione della famiglia e, più in particolare, ai comportamenti riproduttivi. La disponibilità di dati dettagliati per il “laboratorio lombardo” ha permesso di eseguire analisi innovative e approfondite che hanno mostrato come non si possa prescindere dall’effettuare considerazioni differenziate per luogo di provenienza e per storia migratoria nello studio di un fenomeno così complesso. Di qui l’urgenza di adeguare il vigente sistema di raccolta dei dati al fine di rispondere a necessità conoscitive aventi importanti implicazioni anche in ambito sociale e politico.

Riferimenti bibliografici

- ANDERSSON G. 2004. Childbearing after migration: Fertility patterns of foreign-born women in Sweden. *International Migration Review*, Vol. 38, pp. 747-774.
- BLANGIARDO G.C. (Eds) 2011. *L’immigrazione straniera in Lombardia. La decima indagine regionale. Rapporto 2010*. Milano: Fondazione Ismu, Osservatorio Regionale per l’integrazione e la multietnicità.
- BUENO X., DEVOLDER D. 2011. Interactions Between Fertility and Migration for Foreign Immigrants in Spain. Presented at *Population Association of America 2011 Annual Meeting*, Washington DC, Draft Version 16/09/2010.
- MILEWSKI N. 2007. First child of immigrant workers and their descendants in West Germany: Interrelation of events, disruption, or adaptation?. *Demographic Research*, Vol. 17, pp. 859-896.
- MUSSINO E., IACCARINO C., PRATI S., STROZZA S. 2009. Short-term reproductive behaviour of foreign women who became mothers between 2002–2006 in Italy, *Vienna Yearbook of Population Research* 2009, pp. 63-82.
- TOULEMON L. 2004. Fertility among immigrant women: new data, a new approach. *Population et sociétés*, No. 400, Paris: INED.

SUMMARY

The aim of the paper is to study fertility of foreign women in Lombardy and its interrelation with their preceding migration to Italy. The results show a strong impact of age at the migration and heterogeneous behaviour by citizenship in the reproductive choice.

Eleonora MUSSINO, Istituto Nazionale di Statistica, mussino@istat.it
Giuseppe GABRIELLI, Istituto Nazionale di Statistica, ggabrielli@istat.it
Anna PATERNO, Università degli studi di Bari, a.paterno@scienze politiche.uniba.it
Salvatore STROZZA, Università degli studi di Napoli Federico II, strozza@unina.it
Laura TERZERA, Università di Milano Bicocca, laura.terzera@unimib.it

IN TEMA DI PAESI A SVILUPPO MINIMO

Marilena Stigliano

Premessa

In una recente occasione (Stigliano, 2011) ell'interessarci della "demografia" dei paesi a sviluppo minimo (PSM) abbiamo avuto occasione di porre in evidenza la macroscopica eterogeneità, esistente sotto molteplici punti di vista (dimensione demografica, ampiezza e distribuzione territoriale, reddito medio procapite, etc.) nel loro ambito.

In questa occasione, invece, andando al di là delle problematiche soprattutto demografiche, concernenti il complesso dei PSM che racchiudono circa 1/8 dell'intera popolazione mondiale, intendiamo proseguire con quanto fatto in precedenza soffermandoci soprattutto ad analizzarne l'origine di tale "categoria" alla quale appartengono principalmente Stati divenuti indipendenti a seguito del processo di decolonizzazione sopravvenuto dopo la seconda guerra mondiale¹ e successivamente talune caratteristiche salienti di ordine prevalentemente socio-economico.

1. L'origine dei PSM

Verso la fine degli anni Sessanta, in occasione della seconda Conferenza sul Commercio e lo sviluppo delle Nazioni Unite, si iniziarono a gettare la basi per la promozione di una azione sovranazionale a favore dei paesi meno sviluppati del pianeta (Monaldi, 1996). A seguire, con l'inizio degli anni Settanta, si pervenne – sempre in ambito Nazioni Unite – all'estensione di una lista di 25 paesi², individuati facendo riferimento a tre criteri base³.

¹ Occorre, però, tener presente che tra di essi figurano anche paesi – quali Etiopia, Haiti e Liberia – la cui autonomia risale a prima degli inizi del XX secolo.

² Sulla base delle denominazioni del tempo (cioè del 1971), alcune delle quali successivamente variate nel corso degli anni, tali paesi erano: a) Alto Volta, Botswana, Burundi, Ciad, Dahomey, Etiopia, Guinea, Lesotho, Malawi, Mali, Niger, Ruanda, Somalia, Sudan, Tanzania e Uganda in *Africa*; b) Afghanistan, Bhutan, Yemen, Laos,

A seguito di una serie di ingressi e qualche uscita – dovuta, come nel caso dello Yemen, all'unificazione dei due precedenti Stati (1990) – nel corso di quaranta anni si è passati da 25 a 48 Stati, con l'Africa che sale da 16 a 33 e l'Asia (con l'Oceania) che passa da 8 a 14 Stati e si sono anche, a più riprese, modificati i parametri di riferimento (Stigliano, 2011).

Un siffatto incremento – concentrato soprattutto nel corso degli anni Settanta ed Ottanta, allorché da più parti si riteneva che soprattutto le migrazioni Sud-Nord potessero condizionare pesantemente gli assetti socio-economici dei paesi coinvolti (Chesnais, 1987) – può essere attribuito ad una molteplicità di valutazioni tanto di ordine politico, quanto di natura socio-economica, oltre che alla constatazione di un sostanziale fallimento, anche in pieno processo di globalizzazione, degli interventi internazionali tesi a promuovere ed agevolare lo sviluppo socio-economico dei paesi meno sviluppati del globo terracqueo.

2. L'eterogeneità dei PSM

Con l'estromissione delle Maldive fanno attualmente parte del coacervo dei PSM quarantotto Stati, sparsi su tre continenti – Africa, Asia ed America (Latina e Caribe) –, che, prescindendo dalla loro ubicazione geografica, presentano sotto molteplici aspetti un forte grado di eterogeneità, tanto da indurre a ritenere che, oltre ai problemi tipici del sottosviluppo, possano avere in comune tra di loro ben poco.

Maldive, Nepal, Samoa occidentale e Sikkim in *Asia e Oceania*; c) Haiti in *America latina e Caraibi*.

³ I parametri di riferimento, originariamente adottati in quanto ritenuti idonei a fornire una nel complesso più che soddisfacente visione globale del grado di sviluppo socio-economico e inoltre in quanto agevolmente quantificabili – e quindi disponibili – per la quasi totalità dei paesi interessati, erano: a) PIL pro-capite inferiore ai 100 \$; b) presenza del settore manifatturiero nel PIL non superiore al 10,0%; c) grado di alfabetizzazione degli adulti non superiore al 20,0%.

Tabella 1 – Indicatori socio-economici per i Paesi a sviluppo minimo (PSM).

| Stati | Superficie (1000 kmq) | Popolazione (in migliaia al 2010) | Reddito medio pro- capite (US\$-2009) | Tasso di occupazione (15 anni ed oltre-2009) | Tasso di istruzione (15 anni ed oltre-2009) | ISU (2010) |
|---------------------|--------------------------|-----------------------------------------|------------------------------------------------|-------------------------------------------------------|------------------------------------------------------|---------------|
| AFRICA | | | | | | |
| Angola | 1.246,7 | 12.531,4 | 1.892 | 75,3 | 69,95 | 0,403 |
| Benin | 112,6 | 8.294,9 | 739 | 72,2 | 41,65 | 0,435 |
| Burkina Faso | 267,9 | 15.264,7 | 517 | 82,3 | 28,73* | 0,305 |
| Burundi | 27,8 | 8.691,0 | 151 | 83,0 | 66,57 | 0,282 |
| Ciad | 1284,0 | 10.111,3 | 610 | 65,4 | 33,61 | 0,295 |
| Comore | 1,9 | 731,8 | 785 | 74,3 | 74,15 | 0,428 |
| Congo (ex Zaire) | 2.344,9 | 66.514,5 | 170 | 67,9 | n.d. | 0,239 |
| Eritrea | 121,1 | 5.028,5 | 369 | 67,6 | 66,58 | n.d. |
| Etiopia | 1.127,1 | 78.254,1 | 345 | 80,7 | 29,82* | 0,328 |
| Gambia | 10,7 | 1.735,5 | 543 | 72,6 | 46,50 | 0,390 |
| Gibuti | 23,2 | 506,2 | 1.076 | n.d. | n.d. | 0,402 |
| Guinea | 245,8 | 10.211,4 | 470 | n.d. | 39,45 | 0,340 |
| Guinea-Bissau | 36,1 | 1.503,2 | 517 | 66,6 | 52,20 | 0,289 |
| Guinea Equatoriale | 28,1 | 616,4 | 17.544 | 60,7 | 93,32 | 0,538 |
| Lesotho | 30,4 | 2.128,2 | 780 | 54,4 | 89,62 | 0,427 |
| Liberia | 99,1 | 3.334,6 | 216 | 67,1 | 59,05 | 0,300 |
| Madagascar | 587,1 | 20.042,6 | 448 | 84,1 | 64,48* | 0,435 |
| Malawi | 118,5 | 13.931,9 | 318 | 70,9 | 73,69 | 0,385 |
| Mali | 1.248,6 | 12.324,0 | 679 | 47,4 | 26,18* | 0,309 |
| Mauritania | 1.030,7 | 3.364,9 | 866 | 46,9 | 57,45 | 0,433 |
| Mozambico | 799,4 | 21.284,7 | 418 | 80,0 | 55,06 | 0,284 |
| Niger | 1.189,6 | 13.272,7 | 343 | 59,7 | 28,67* | 0,261 |
| Rep. Centrafricana | 6622,4 | 4.434,9 | 448 | 73,5 | 55,23 | 0,315 |
| Ruanda | 26,4 | 10.186,0 | 527 | 85,4 | 70,67 | 0,385 |
| Sao Tomé e Principe | 1,0 | 206,2 | 1.302 | n.d. | 88,78 | 0,488 |
| Senegal | 197,0 | 12.853,3 | 1.018 | 68,0 | 49,69 | 0,411 |
| Sierra Leone | 71,8 | 6.294,8 | 393 | 64,1 | 40,92 | 0,317 |
| Somalia | 637,0 | 9.558,7 | 220 | 65,5 | n.d. | n.d. |
| Sudan | 2.460,4 | 40.218,5 | 1.305 | 47,7 | 70,21 | 0,379 |
| Tanzania | 942,8 | 40.213,2 | 526 | 78,1 | 72,90 | 0,398 |
| Togo | 56,8 | 5.858,7 | 480 | 69,4 | 56,89* | 0,428 |
| Uganda | 241,6 | 31.368,0 | 523 | 81,8 | 99,69 | 0,422 |
| Zambia | 752,6 | 11.670,0 | 985 | 60,3 | 70,88 | 0,395 |

(*) *Ultimo anno disponibile: 2008 (Etiopia, Madagascar, Cambogia), 2007 (Timor Est), 2006 (Burkina Faso, Mali, Togo, Haiti) e 2005 (Niger, Bhutan, Laos).*

Fonte: dati ILO, ONU e World Bank.

Ora, tralasciando gli aspetti demografici di cui ci siamo occupati nella precedente richiamata occasione, intendiamo – sempre privilegiando l’aspetto dell’eterogeneità che esiste nell’ambito dei PSM – soffermarci su taluni aspetti

economici e sociali, partendo da un vecchio, tradizionale e tendenzialmente obsoleto indicatore quale il reddito medio procapite⁴.

Tabella 1 (segue) – Indicatori socio-economici per i Paesi a sviluppo minimo (PSM).

| Stati | Superficie (1000 kmq) | Popolazione (in migliaia al 2010) | Reddito medio pro- capite (US\$-2009) | Tasso di occupazione (15 anni ed oltre-2009) | Tasso di istruzione (15 anni ed oltre-2009) | ISU (2010) |
|----------------------------|--------------------------|-----------------------------------------|------------------------------------------------|-------------------------------------------------------|------------------------------------------------------|---------------|
| ASIA & PACIFICO | | | | | | |
| Afghanistan | 645,8 | 32.738,4 | 457 | 54,2 | n.d. | 0,349 |
| Bangladesh | 147,6 | 153.546,9 | 550 | 67,9 | 55,90 | 0,469 |
| Bhutan | 38,4 | 682,3 | 1.783 | 60,8 | 52,81* | n.d. |
| Cambogia | 181,0 | 14.241,7 | 729 | 77,7 | 77,59* | 0,494 |
| Isole di Salomone | 28,4 | 581,3 | 1.366 | 35,8 | n.d. | 0,494 |
| Kiribati | 0,8 | 110,4 | 1.335 | n.d. | n.d. | n.d. |
| Laos | 236,8 | 6.677,5 | 884 | 77,2 | 72,70 * | 0,497 |
| Myanmar | 676,6 | 47.758,2 | 380 | 71,3 | 92,03 | 0,451 |
| Nepal | 147,2 | 29.519,1 | 436 | 70,1 | 59,14 | 0,428 |
| Samoa | 2,8 | 217,1 | 2.926 | 57,5 | 98,78 | n.d. |
| Timor-Est | 14,9 | 1.108,8 | 593 | 68,0 | 50,60* | 0,502 |
| Tuvalu | 0,0 | 12,2 | 2.749 | n.d. | n.d. | n.d. |
| Vanuatu | 12,2 | 215,4 | 2.687 | 83,9 | 82,03 | n.d. |
| Yemen | 528,1 | 23.013,4 | 1.141 | 39,9 | 62,39 | 0,439 |
| AMERICA LATINA | | | | | | |
| Haiti | 27,7 | 8.924,6 | 626 | 64,8 | 48,68* | 0,404 |

(*) *Ultimo anno disponibile: 2008 (Etiopia, Madagascar, Cambogia), 2007 (Timor Est), 2006 (Burkina Faso, Mali, Togo, Haiti) e 2005 (Niger, Bhutan, Laos).*

Fonte: dati ILO, ONU e World Bank.

Soffermandoci ad analizzare i dati riportati nella Tav. 1 emerge con immediatezza l'enorme divario che esiste tra la Guinea Equatoriale, che con un reddito medio pro-capite (2009) di poco superiore ai 17.500 \$ è il più ricco tra i PSM⁵, e alcuni altri Stati (Burundi e Congo), ove non viene raggiunto un livello pari all'1,0% di questa cifra (cioè 175,0\$, grosso modo pari a 0,50\$ al giorno): anche nel ristretto ambito dei PSM si possono osservare, quindi, divari

⁴ Per quel che concerne l'attendibilità dei dati che figurano nella Tav. 1 appare doveroso ribadire che la loro "qualità" non è assolutamente omogenea e conseguentemente la loro affidabilità e la loro comparabilità nel tempo e nello spazio sono ridotte, in quanto in più di un caso – specie per i dati di carattere socio-economico – essi possono presentare gravi deficienze.

⁵ Nonché dell'intero continente africano. Tale, abbastanza recente, situazione è riconducibile all'abbondante produzione petrolifera concentrata su un territorio ristretto e, in termini assoluti, poco popolato: cioè, una situazione simile a quella di alcuni Stati della penisola arabica come Bahrain, Kuwait e Qatar.

paragonabili – e forse anche superiori – a quelli che mediamente esistono tra paesi a sviluppo avanzato (PSA) e paesi in via di sviluppo (PVS). In linea di massima, poi, sempre dagli stessi dati appare evidente che, in media, i paesi africani sono più poveri di quelli asiatici e che, di sovente, i paesi di maggiori dimensioni demografiche presentano valori del reddito medio pro-capite particolarmente bassi.

Globalmente solo tredici paesi avrebbero – sempre sulla base dei dati che figurano nella Tav. 1 – un reddito medio pro-capite superiore ai 1.000\$: sette (tutti di ridotte dimensioni demografiche, cioè con una popolazione che non raggiunge il milione di abitanti) in Asia e sei in Africa (ove, invece, in tre casi su sei si tratta di paesi di dimensioni demografiche piuttosto rilevanti, cioè con più di 10milioni di abitanti).

Alquanto diverso dal quadro di riferimento desumibile da una accorta analisi dai dati sul reddito appare, invece, quello desumibile da un altro indicatore di natura economica, cioè dal tasso di occupazione della popolazione adulta (cioè di età non inferiore ai 15 anni).

Premesso che in alcuni casi abbastanza rari i dati in questione non sono disponibili, tali tassi presentano un campo di variazione estremamente ampio⁶, apparendo in genere notevolmente bassi per i paesi appartenenti al Mondo arabo, tanto in Africa quanto in Asia, a causa essenzialmente di una partecipazione della compagine femminile tradizionalmente abbastanza contenuta e risultando fortemente condizionati dai difformi criteri sui quali si basa, da un paese ad un altro, la “catalogazione” del lavoro femminile nel settore primario.

Una situazione analoga la si può poi osservare per quel che concerne i tassi di istruzione, compresi in un intervallo che va da un massimo pari a 99,69 – cioè praticamente la totalità degli interessati o, in altri termini, analfabetismo zero – dell’Uganda a un minimo di 26,18, cioè poco più di un quarto della popolazione adulta, del Mali.

La macroscopica variabilità che emerge dall’esame delle due serie anzidette è dovuta ad una molteplicità di fattori tra cui primeggia il non sempre soddisfacente grado di attendibilità dei dati statistici relativi a parte dei PSM, ove per una molteplicità di ragioni⁷ i locali sistemi di rilevazione statistica presentano notevoli e croniche carenze strutturali. A determinarla, molto verosimilmente, concorrono inoltre le differenti valutazioni che possono essere fatte in tema di lavoro – soprattutto femminile – nel settore primario e la marginalizzazione della donna che

⁶ Si passa, infatti, da un valore massimo pari a 85,4 per il Ruanda ad uno minimo pari a 35,8 per le Isole di Salomone.

⁷ Tali ragioni vanno, ad esempio, dalla limitata ampiezza demografica di alcuni stati dell’Oceania (vedi Tuvalu), all’estrema povertà di altri ove il reddito medio pro-capite è addirittura inferiore ai 200 \$ l’anno (vedi, da esempio, il Congo) ed allo stato endemico di guerriglia di altri (vedi Somalia).

si può osservare per la maggior parte dei paesi meno “moderni” e “globalizzati” del mondo arabo.

3. Prospettive di sviluppo dei PSM

Le Nazioni Unite iniziarono a prestare un'attenzione speciale ai paesi del Terzo Mondo meno avanzati dal punto di vista socio-economico, cioè in sostanza a quelli che sarebbero stati raggruppati sotto la sigla di PSM, alla fine degli anni Sessanta del secolo scorso, andando così al di là della consolidata distinzione tra PSA e PVS e riconoscendoli, nell'ambito di questi ultimi, come i più vulnerabili della comunità internazionale.

Nel corso degli anni Settanta venne promossa una articolata serie di iniziative sia per richiamare l'attenzione della comunità internazionale su questi paesi, che per realizzare e dare impulso ad una molteplicità di azioni in grado di far fronte al continuo e progressivo deterioramento delle condizioni socio-economiche della maggior parte di loro; furono organizzate tre Conferenze delle Nazioni Unite (Parigi 1981 e 1990, Bruxelles 2001) sui PSM che produssero, di volta in volta, un Programma d'azione di durata decennale. Durante la Conferenza di Bruxelles del 2001 fu istituito un apposito ufficio, l'UN-OHRLLS, creato come strumento per assicurare l'effettiva realizzazione, il monitoraggio e la revisione del Programma di azione per i PSM per il decennio 2001 – 2010⁸. Quest'ultimo ha come obiettivo il miglioramento delle condizioni di vita delle popolazioni dei PSM e svolge, altresì, un ruolo di sistematico quadro di riferimento per la collaborazione tra i PSM e i loro eventuali partner al fine di “determinare una crescita economica sostenuta ed uno sviluppo sostenibile nei PSM, porre fine alla marginalizzazione sradicando la povertà e l'ineguaglianza in questi paesi e rendendoli capaci di integrarsi vantaggiosamente nell'economia globale”, cioè in sintesi migliorare la “qualità della vita” delle loro popolazioni. Inoltre, il Programma di azione anzidetto prevede una trentina di obiettivi di sviluppo⁹ che sono stati formalmente inseriti negli impegni dei PSM e dei loro partner.

⁸ <http://www.unohrlls.org/en/ldc/25/>

⁹ Alcuni degli obiettivi del Programma d'Azione per i PSM per il decennio 2001-2010: incrementare il prodotto interno lordo e gli investimenti, rendere accessibili salute riproduttiva, pianificazione familiare e metodi contraccettivi, assicurare a tutti i bambini (soprattutto alle bambine o ai membri di minoranze etniche) l'accesso ad una educazione primaria di buona qualità, ridurre l'analfabetismo adulto (soprattutto femminile), ridurre la mortalità infantile e materna, dimezzare entro il 2015 l'incidenza della malnutrizione e la percentuale di persone che non hanno accesso a fonti di acqua potabile, ridurre l'incidenza del virus HIV e di altre malattie infettive come malaria e tubercolosi, incrementare la rete

5. Conclusioni

Nel corso degli ultimi quaranta anni, o giù di lì, l'obiettivo di sviluppo socio-economico che da più parti ci si proponeva di conseguire per i PSM non è stato praticamente mai – o quasi mai – raggiunto ed il divario che esiste tra questi ed i PSA ha mostrato sempre più la tendenza ad accentuarsi nel tempo.

In più il prepotente emergere, tanto in Asia quanto in America latina, di alcuni dei PVS quali ad esempio il Brasile da un lato e la Cina e l'India dall'altro ha di fatto allargato notevolmente anche il solco che precedentemente esisteva tra PVS e PSM, in alcuni dei quali – ad esempio, il Bangladesh – vengono attualmente delocalizzate talune lavorazioni ad alto contenuto di lavoro che in passato venivano effettuate nei PVS (ad esempio, Cina).

Riassumendo le prospettive di sviluppo per la maggior parte del PSM appaiono fortemente condizionate, almeno per quel che concerne le prossime generazioni, dall'attuale loro limitatissima dotazione di capitale umano, per cui – salvo casi particolari, come ad esempio quello della Guinea Equatoriale, ove lo sfruttamento delle risorse petrolifere ha determinato nel corso dell'ultimo decennio, anche grazie alla ridotta dimensione demografica del paese, una più che cospicua e fortemente asimmetrica crescita del reddito medio pro-capite – in detti paesi bisognerà, indiscutibilmente, intervenire prioritariamente in questo settore promuovendo un progressivo innalzamento del livello di istruzione delle loro popolazioni ed attenuando i divari che in tale ambito tuttora esistono tra genere maschile e genere femminile.

Riferimenti bibliografici

- CHESSAIS J. C. 1987. *La revanche du tiers-monde*, R. Laffont, Paris; A. SAUVY (1987), *L'Europe submergée. Sud-Nord dans 30 ans*, Paris: Dunod.
- MONALDI V. 1996. I paesi meno avanzati e i programmi di aggiustamento strutturale: un'analisi preliminare. In G. ANCONA – F. BOTTA (Eds.), *Problemi di cooperazione e sviluppo nel Bacino mediterraneo*, Quaderni del Dipartimento per lo Studio delle Società Mediterranee, n° 12, Bari: Cacucci Editore.
- STIGLIANO M. 2011. *Caratteristiche demografiche salienti dei paesi a sviluppo minimo*, "Nike", No 1

stradale e le vie di comunicazione, modernizzare ed ampliare porti, aeroporti e sistema ferroviario, accrescere i servizi nelle aree urbane come per esempio il sistema fognario, aumentare la rete dei sistemi di comunicazione incluso telecomunicazioni e servizi postali, incentivare la diffusione dei telefoni e l'alfabetismo informatico"

(<http://www.unohrlls.org/UserFiles/File/LDC%20Documents/goals%20and%20targets.pdf>)

SUMMARY

Following what we have recently analyzed ("Nike", 2011, n° 1), with this contribution we have intended to continue the study of the so called Least Development Countries especially dwelling upon to consider the origin - that goes up again to the beginning of the seventies -, the marked heterogeneity and the salient socioeconomic characteristics of this grouping. With regard to the demographic characteristics we refer to what we have highlighted in the previous article.

Marilena STIGLIANO, Università di Bari, Dipartimento per lo Studio delle Società Mediterranee dell'Università di Bari e Departamento de Geografía Humana de l'Universidad de Granada, marilena.stigliano@hotmail.it

LA SODDISFAZIONI DEGLI STUDENTI UNIVERSITARI IN MERITO AGLI ESAMI SOSTENUTI: UN'ANALISI STATISTICA SU DUE ATENEI SICILIANI

Agata Zirilli, Angela Alibrandi, Massimiliano Giacalone

1. Introduzione

La qualità del sistema universitario va tenuta costantemente sotto controllo; in tale contesto uno dei servizi primari offerti è la formazione, la cui qualità risulta essere significativamente legata ai principi fondamentali del *Total Quality Management* (Lawrence e Mc Collough, 2004). L'esigenza di introdurre un sistema di valutazione dell'attività didattica in ambito universitario, analogo a quello attivato da tempo in molti paesi europei, è stata soddisfatta anche in Italia tramite le leggi 537/1993 e 370/1999. Per analizzare la qualità di un insegnamento è necessario considerare lo stesso sia dalla prospettiva del docente che da quella dello studente; ciò implica l'individuazione delle principali attività che ciascuna delle due figure è chiamata a svolgere. Focalizzando l'attenzione sulla prospettiva dello studente, particolare importanza riveste la valutazione degli esami. Il voto conseguito a seguito dell'esame finale dell'insegnamento costituisce uno degli output del processo di formazione. Numerosi sono i fattori che possono incidere sull'esito di un esame universitario e, quindi, sulla votazione conseguita: non solo il livello di preparazione raggiunto ma anche altri fattori, sia di natura personale ed emotiva che di tipo didattico (essere esaminati dal docente titolare dell'insegnamento o da eventuali assistenti). La soddisfazione degli studenti in merito agli esami, che rappresenta un importante aspetto della valutazione, permette di ottenere, quindi, utili informazioni sulle caratteristiche dell'insegnamento universitario, sulla qualità del docente e sulla funzionalità della didattica in generale (Tudor, 2006; Wallace, 1999).

2. Il campione

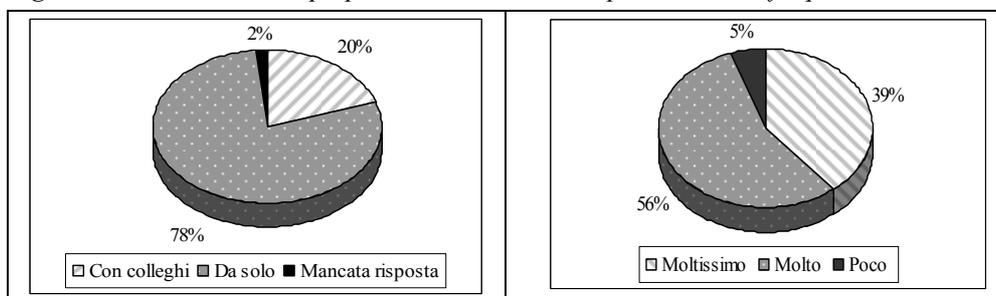
Nell'ambito del Progetto PRIN 2008 "Innovazione nella gestione della qualità dei servizi: approccio statistico ed applicazione in settori di interesse nazionale" (Coordinatore Nazionale: Prof. Pasquale Erto), l'Unità di Ricerca Palermo-

Messina ha perseguito la linea di ricerca relativa alla valutazione della qualità della formazione universitaria. Le Università di Messina e di Palermo hanno somministrato, seguendo un protocollo comune, un questionario volto ad individuare la soddisfazione degli studenti universitari in merito agli esami sostenuti (Marozzi, 2009). Il questionario, organizzato in due sezioni (pre-esame e post-esame) e strutturato con risposte prevalentemente chiuse, è stato distribuito agli studenti il giorno degli esami, per rilevare il grado di preparazione, la modalità di studio, i fattori emozionali, la disponibilità del docente, le aspettative di voto.

La raccolta dei dati è avvenuta nelle sessioni estive di esami (Giugno e Luglio) dell'anno accademico 2009/2010 ed ha interessato tutti gli studenti (sia in corso che fuori corso) iscritti al Corso di Laurea in Scienze Statistiche. Si è consapevoli del fatto che esiste un *selection bias* dovuto al fatto che solo gli studenti che decidono di sostenere l'esame rispondono al questionario e fanno parte del campione osservato, ma si è ritenuto interessante, da un punto di vista investigativo, procedere ad un'indagine effettuata proprio al momento degli esami, con l'ausilio delle tecniche inferenziali che permettono di generalizzare i risultati.

L'indagine ha permesso la rilevazione di informazioni statistiche riguardanti una totalità di 175 studenti (48% per l'Ateneo di Messina e 52% per l'Ateneo di Palermo); la distribuzione degli studenti secondo l'anno di corso risulta la seguente: 19.6% al I anno; 51% al II anno; 17.6% al III anno; 11.8% al I anno fuori corso; la percentuale di studenti che ha superato l'esame è del 90.2%.

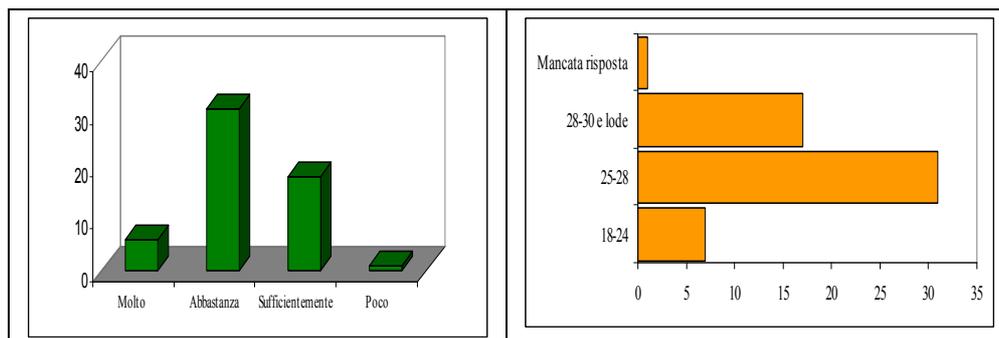
Figure 1 e 2 – Modalità di preparazione all'esame- Importanza della frequenza.



Lo scopo del presente lavoro è stato la valutazione dei fattori a cui è collegata la votazione ottenuta dagli studenti intervistati e la soddisfazione degli stessi al momento degli esami. In particolare, si vogliono individuare i fattori cui il superamento dell'esame risulta essere legato e valutare, altresì, la corrispondenza tra valutazione attesa ed effettiva, tenendo conto degli aspetti individuali quali frequenza ai corsi, stato d'animo con cui ci si appropria all'esame, modalità di preparazione all'esame, disponibilità da parte dei docenti. Le Figure 1-4 riportano la distribuzione degli studenti intervistati, secondo la modalità di preparazione

all'esame, la consapevolezza dell'importanza della frequenza alle lezioni, la percezione personale del livello di preparazione raggiunto e la classe di voto atteso.

Figure 3 e 4 – *Percezione personale del livello di preparazione raggiunto e classe di voto atteso.*



3. Il modello *Cumulative Proportional Odds* per variabili risposta ordinali

3.1. Metodologia

Il *Cumulative Proportional Odds Model* è un modello per dati ordinali, che si ottiene utilizzando un link *logit* per le probabilità cumulate (Agresti, 2010; O'Connel, 2006). Il modello è un'estensione del modello di regressione logistica per dati dicotomici a dati categoriali ordinali. L'evento di interesse è osservare o meno un punteggio; per una variabile ordinale con tre modalità è possibile definire i seguenti *odds*:

$$\Theta_1 = \text{prob. (punteggio di 1)} / \text{prob. (punteggio >1)} \quad (1)$$

$$\Theta_2 = \text{prob. (punteggio di 1 o 2)} / \text{prob. (punteggio >2)} \quad (2)$$

$$\Theta_3 = \text{prob. (punteggio di 1,2 o 3)} / \text{prob. (punteggio >3)} \quad (3)$$

L'ultima categoria non ha un *odds* associato poiché la probabilità che un punteggio sia minore o uguale all'ultimo score è pari a 1.

Tutti gli *odds* sono nella forma:

$$\Theta_j = \text{prob}(\text{punteggio} \leq j) / \text{prob}(\text{punteggio} > j) \quad (4)$$

e possono anche essere espressi nel seguente modo:

$$\square_j = \text{prob}(\text{punteggio} \leq j) / (1 - \text{prob}(\text{punteggio} \leq j)) \quad (5)$$

Il modello logistico ordinale con una sola variabile indipendente è:

$$\ln(\theta_j) = \alpha_j - \beta X \quad (6)$$

dove $j=1, \dots, j-1$, con j numero di categorie. Coefficienti maggiori indicano un'associazione con punteggi più elevati. Per una variabile continua, un coefficiente positivo indica che, non appena i valori di una variabile aumentano, le probabilità di punteggi elevati aumentano. Un'associazione con punteggi più alti significa minore probabilità cumulativa per punteggi inferiori, in quanto essi si verificano meno frequentemente (Soliani, 2004). Ciascun *logit* ha il suo termine α_j , ma lo stesso coefficiente β . Ciò significa che l'effetto della variabile indipendente è lo stesso per funzioni *logit* differenti.

Il modello logistico ordinale si basa sull'assunzione che è presente una variabile latente continua e che il risultato osservato ordinale deriva dalla discretizzazione di una variabile continua sottostante.

La stima del modello è stata effettuata utilizzando la procedura PLUM (Polytomous Universal Model) di SPSS che consente di stimare modelli per dati ordinali, in particolare modelli GLM (Marija Norušis, 2009).

3.2. Applicazione

Da un punto di vista analitico, utilizzando la procedura PLUM, si è proceduto alla stima di due modelli statistici:

- il primo, volto a formalizzare la dipendenza della votazione riportata all'esame (variabile ordinale espressa da 3 categorie ordinate: 1=voto <24; 2=voto compreso tra 25 e 27; 3=voto ≥28) da variabili riferite allo studente, dalla modalità di studio, da fattori emozionali, da fattori legati al docente e dalle aspettative di giudizio;
- il secondo, volto a formalizzare la dipendenza della soddisfazione per l'esame sostenuto (variabile ordinale espressa da 3 categorie ordinate: 1=insoddisfatto; 2=né insoddisfatto né soddisfatto; 3=soddisfatto) da fattori ritenuti potenzialmente predittivi.

La Tabella 1 riporta i risultati dei modelli *logit* ordinali con una sola covariata per la votazione finale (in grassetto sono evidenziati i p-value dei coefficienti statisticamente significativi, al livello $\alpha=0.05$); la stima di tali modelli è stata finalizzata all'individuazione dei fattori che singolarmente possono incidere sulla votazione riportata all'esame; la Tabella 2, invece, riporta i risultati del modello nel quale sono state inserite le covariate scelte in base alle significatività ottenute nei modelli con una sola covariata.

Dai risultati ottenuti si evidenzia come la votazione finale risulti significativamente dipendente in modo diretto da fattori strettamente riconducibili allo studente, quali la frequenza alle lezioni, il grado di preparazione ed il voto

atteso, in maniera inversa dallo studio con i colleghi; lo studio individuale sembra, infatti, essere più premiale ai fini della valutazione finale all'esame.

Tabella 1 – Modelli logit ordinali cumulativi per la votazione finale con una sola covariata.

| Variabile | Coeff. | Err.St. | Wald | p-value | Interv. Confid. Inf | Sup | log veros. |
|---------------|--------|---------|--------|--------------|------------------------|--------|---------------|
| Costante1 | -1.760 | 0.598 | 8.659 | 0.003 | -2.933 | -0.588 | |
| Costante2 | -0.162 | 0.464 | 0.122 | 0.727 | -1.071 | 0.747 | 21.086 |
| Freq.lezioni | 0.672 | 0.323 | 4.317 | 0.038 | 0.038 | 1.306 | |
| Costante1 | 0.777 | 1.406 | 0.306 | 0.580 | -1.978 | 3.533 | |
| Costante2 | 2.424 | 1.421 | 2.911 | 0.088 | -0.361 | 5.209 | 16.416 |
| Grado prepar | 1.283 | 0.546 | 5.526 | 0.019 | 0.213 | 2.352 | |
| Costante1 | -2.964 | 0.609 | 23.713 | 0.000 | -4.158 | -1.771 | |
| Costante2 | -1.301 | 0.376 | 12.006 | 0.001 | -2.037 | -0.565 | 10.847 |
| Stud.colleggi | -1.833 | 0.773 | 5.627 | 0.018 | -3.348 | -0.319 | |
| Costante1 | -2.516 | 0.637 | 15.585 | 0.000 | -3.765 | -1.267 | |
| Costante2 | -1.023 | 0.481 | 4.532 | 0.033 | -1.966 | -0.081 | 12.056 |
| Ansia | 0.672 | 0.629 | 0.021 | 0.884 | -1.325 | 1.141 | |
| Costante1 | -1.576 | 1.405 | 1.258 | 0.262 | -4.329 | 1.178 | |
| Costante2 | 0.030 | 1.296 | 0.001 | 0.982 | -2.510 | 2.569 | 16.356 |
| Empatia doc. | 0.640 | 0.584 | 1.204 | 0.272 | -0.503 | 1.784 | |
| Costante1 | -1.391 | 0.994 | 1.959 | 0.162 | -3.339 | 0.557 | |
| Costante2 | 0.136 | 0.948 | 0.020 | 0.886 | -1.723 | 1.994 | 10.279 |
| Dispon.doc. | 1.213 | 0.999 | 1.473 | 0.225 | -0.746 | 3.173 | |
| Costante1 | 0.052 | 1.096 | 0.002 | 0.962 | -2.095 | 2.200 | |
| Costante2 | 1.698 | 1.095 | 2.404 | 0.121 | -0.448 | 3.844 | 10.048 |
| Voto atteso | 1.252 | 0.524 | 5.694 | 0.017 | 0.224 | 2.280 | |

Tabella 2 – Modello logit ordinale cumulativo per la votazione finale.

| Variabile | Coeff. | Err.St. | Wald | p-value | Interv. Confid. Inf | Sup | log veros. |
|---------------|--------|---------|-------|--------------|------------------------|--------|---------------|
| Costante1 | 0.683 | 1.532 | 0.198 | 0.656 | -2.321 | 3.686 | |
| Costante2 | 2.846 | 1.591 | 3.200 | 0.074 | -0.272 | 5.963 | |
| Freq.lezioni | 0.513 | 0.332 | 2.380 | 0.023 | -0.139 | 1.164 | 46.533 |
| Grado prepar | 0.440 | 0.706 | 0.387 | 0.534 | -0.945 | 1.824 | |
| Stud.colleggi | -1.872 | 0.880 | 4.527 | 0.033 | -3.596 | -0.148 | |
| Voto atteso | 1.394 | 0.625 | 4.972 | 0.026 | 0.169 | 2.620 | |

Successivamente, sono stati stimati i modelli logistici con una sola covariata per la soddisfazione relativa all'esame sostenuto, i cui risultati sono riportati in Tabella 3; la Tabella 4 riporta i risultati del modello nel quale sono state inserite le covariate scelte in base alle significatività ottenute nei modelli con una sola covariata.

Tabella 3 – Modelli logit ordinali cumulativi per la soddisfazione.

| Variabile | Coeff. | Err.St. | Wald | p-value | Interv. Confid. Inf | Sup | log veros. |
|--------------|--------|---------|--------|--------------|------------------------|--------|---------------|
| Costante1 | -3.015 | 0.774 | 15.160 | 0.000 | -4.533 | -1.497 | |
| Costante2 | -0.094 | 0.423 | 0.049 | 0.825 | -0.922 | 0.735 | 12.908 |
| Ansia | 0.367 | 0.560 | 0.429 | 0.512 | -0.731 | 1.465 | |
| Costante1 | -1.721 | 1.236 | 1.938 | 0.164 | -4.143 | 0.702 | |
| Costante2 | 0.996 | 1.123 | 0.786 | 0.375 | -1.206 | 3.197 | 14.133 |
| Empatia doc. | 0.545 | 0.477 | 1.310 | 0.252 | -0.389 | 1.479 | |
| Costante1 | -2.753 | 0.830 | 10.988 | 0.001 | -4.380 | -1.125 | |
| Costante2 | 0.194 | 0.547 | 0.126 | 0.722 | -0.877 | 1.265 | 11.038 |
| Info docente | 0.674 | 0.633 | 1.133 | 0.287 | -0.567 | 1.915 | |
| Costante1 | -0.191 | 1.275 | 0.023 | 0.881 | -2.690 | 2.308 | |
| Costante2 | 2.952 | 1.349 | 4.793 | 0.029 | 0.309 | 5.596 | 17.855 |
| Voto finale | 1.227 | 0.493 | 6.185 | 0.013 | 0.260 | 2.194 | |
| Costante1 | -4.029 | 1.023 | 15.524 | 0.000 | -6.034 | -2.025 | |
| Costante2 | -0.452 | 0.293 | 2.381 | 0.123 | -1.026 | 0.122 | 8.213 |
| Assistente | -3.473 | 1.474 | 5.548 | 0.019 | -6.362 | -0.583 | |
| Costante1 | -3.450 | 0.760 | 20.620 | 0.000 | -4.939 | -1.961 | |
| Costante2 | -0.503 | 0.338 | 2.211 | 0.137 | -1.165 | 0.160 | 11.049 |
| Umore doc. | -0.657 | 0.607 | 1.171 | 0.279 | -1.846 | 0.533 | |

Tabella 4 – Modello logit ordinale cumulativo per la soddisfazione.

| Variabile | Coeff. | Err.St. | Wald | p-value | Interv. Confid. Inf | Sup | log veros. |
|-------------|--------|---------|-------|--------------|------------------------|--------|---------------|
| Costante1 | -1.106 | 1.498 | 0.545 | 0.460 | -4.042 | 1.830 | |
| Costante2 | 2.696 | 1.382 | 3.804 | 0.051 | 0.013 | 5.406 | 18.806 |
| Voto finale | 1.193 | 0.507 | 5.538 | 0.019 | 0.199 | 2.187 | |
| Assistente | -3.349 | 1.529 | 4.796 | 0.029 | -6.346 | -0.352 | |

Dai risultati si evidenzia come la soddisfazione risulti significativamente dipendente da due soli fattori: direttamente dal voto finale conseguito all'esame ed indirettamente dall'interrogazione all'esame da parte dell'assistente; probabilmente, l'essere interrogati dal docente titolare del corso piuttosto che dall'assistente esercita una notevole influenza sulla soddisfazione dell'esame.

4. Conclusioni

Sulla base dell'analisi statistica condotta, è stato possibile giungere ad alcuni interessanti risultati.

La votazione riportata agli esami universitari risulta dipendere significativamente da fattori riconducibili allo studente stesso, ovvero la frequenza alle lezioni ed il grado di preparazione; la tipologia di studio risulta significativa ed il segno negativo del coefficiente denota che lo studio individuale, piuttosto che lo

studio collettivo, favorisce un maggiore rendimento al momento dell'esame stesso. Sembra esserci una buona rispondenza tra voto atteso e votazione conseguita, denotando un buon grado di consapevolezza, da parte degli studenti, circa la propria preparazione.

La soddisfazione degli studenti circa l'esame appena sostenuto risulta dipendere significativamente dal voto finale, com'è lecito attendersi. Sostenere l'esame con l'assistente, piuttosto che con il docente titolare di insegnamento, risulta essere un fattore significativo, che incide negativamente sul livello di soddisfazione. Fattori quali l'umore del docente, le informazioni raccolte dagli studenti circa le domande maggiormente formulate agli esami, l'empatia col docente esaminatore e i fattori emozionali provati al momento dell'esame non risultano, invece, statisticamente significativi. La valutazione della didattica universitaria risulta di particolare importanza in quanto, pur essendo strettamente legata ai risultati ottenuti come *customer satisfaction* da parte dei singoli corsi di laurea, rappresenta anche un *output* del servizio offerto dall'Università.

Si precisa che la stima dei modelli non tiene conto del fatto che le risposte di studenti che sostengono lo stesso esame potrebbero essere tra loro correlate e ciò potrebbe introdurre alcune distorsioni nell'analisi. Inoltre, si evidenzia che l'indagine è stata condotta su un campione formato dalla totalità degli studenti intervistati nei due Atenei siciliani coinvolti nel Progetto PRIN; l'analisi di eventuali differenziazioni tra i profili degli studenti appartenenti ai diversi Atenei costituisce uno degli sviluppi successivi del progetto in questione.

Ringraziamenti

Il presente lavoro è stato realizzato nell'ambito del progetto PRIN 2008 "Innovazione nella gestione della qualità dei servizi: approccio statistico ed applicazione in settori di interesse nazionale" (Coordinatore nazionale: Prof. Pasquale Erto).

Riferimenti bibliografici

- AGRESTI A., 2010. Analysis of Ordinal Categorical Data, 2 edition, Hoboken , NJ: Wiley.
- LAWRENCE J. J., MC COLLOUGH M. A. 2004. Implementing total quality management in the classroom by means of student satisfaction guarantees. *Total Quality Management*, Vol.15, No. 2, pp. 235 - 254.
- MARZZI M. 2009. A composite indicator dimension reduction procedure with application to university student satisfaction, *Statistica Neerlandica*, Vol.63, No. 3, pp. 258-268.
- MARIJA J. NORUŠIS. 2009. *PASW Statistic 18 . Statistical Procedures Companion*. Paerson Education

- O'CONNEL, A.A. 2006. *Logistic regression models for ordinal response variables*. Thousand Oaks: Sage.
- SOLIANI L. 2004. *Manuale di statistica per la ricerca e la professione*, Uninova, Parma.
- TUDOR, G. E. 2006. Teaching Introductory Statistics Online. Satisfying the Students. *Journal of Statistics Education*, Vol. 14, No. 3.
- WALLACE J.B. 1999. The case for student as customer. *Quality Progress*, Vol. 32, No. 2, pp. 47-51.

SUMMARY

Students opinions on educational activities are a particularly important form of assessment in order to obtain useful information on the characteristics of higher education, the quality of teachers and, generalizing, the educational function.

The purpose of this research is to assess the factors that are linked to the vote obtained by the university students and their satisfaction at the time of examinations. In particular, we want to identify factors related to the exam and, also, to assess the correspondence between expected and effective evaluation, taking into account the individual aspects of the student (attendance, anxiety level during the examination, type of study) and aspects relating to other factors (professor availability, empathy with professor, etc.). The used tool for data collection was the questionnaire; it was structured in two sections (pre-examination and post-exam). We used PLUM methodology in order to take into account the ordinal nature of response variable. By means of PLUM methodology, we estimated two models: the first model formalizes the dependence of the examination final mark (ordinal variable expressed by three ordered categories) on variables related to student, study mode, emotional factors, teacher factors and judgment expectations; the second model formalizes the dependence of satisfaction for exam (ordinal variable expressed by three ordered categories) on potentially predictive factors.

Agata ZIRILLI, Ricercatore Universitario Confermato, Università degli Studi di Messina, azirilli@unime.it

Angela ALIBRANDI, Ricercatore Universitario Confermato, Università degli Studi di Messina, aalibrandi@unime.it

Massimiliano GIACALONE, Professore a Contratto, Università della Calabria, maxgiacit@yahoo.it

SOCIETÀ E RIVISTA ADERENTI AL SISTEMA ISDS
ISSN ASSEGNATO: 0035-6832

Direttore Responsabile: Dott. CLAUDIO CECCARELLI

Iscrizione della Rivista al Tribunale di Roma del 5 dicembre 1950 N. 1864



Associazione all'Unione Stampa Periodica Italiana

TRIMESTRALE

La copertina è stata ideata e realizzata da Pardini, Apostoli, Maggi p.a.m. @tin.it – Roma

Stampato da CLEUP sc
“Coop. Libreria Editrice Università di Padova”
Via G. Belzoni, 118/3 – Padova (Tel. 049.650261)
www.cleup.it

ATTIVITÀ DELLA SOCIETÀ

A) RIUNIONI SCIENTIFICHE

- XXXV Politiche e tecniche di valutazione dell'attività della Pubblica Amministrazione e degli interventi sociali (Alghero, 29-31 maggio 1997).
- XXXVI Invecchiamento della popolazione e politiche sociali (Roma, 20-22 maggio 1999).
- XXXVII La mobilità dei fattori produttivi nell'area del Mediterraneo (Palermo, 15-17 giugno 2000).
- XXXVIII Qualità dell'informazione statistica e strategie di programmazione a livello locale (Arcavacata di Rende, 10-12 maggio 2001).
- XXXIX L'Europa in trasformazione (Siena, 20-22 maggio 2002).
- XL Implicazioni demografiche, economiche e sociali dello sviluppo sostenibile (Bari, 15-17 maggio 2003).
- XLI Sviluppo economico e sociale e ulteriori ampliamenti dell'Unione Europea (Torino, 20-22 maggio 2004).
- XLII Sistemi urbani e riorganizzazione del territorio (Lucca, 19-21 maggio 2005).
- XLIII Mobilità delle risorse nel bacino del Mediterraneo e globalizzazione (Palermo, 25-27 maggio 2006).
- XLIV Impresa, lavoro e territorio nel quadro dei processi di localizzazione e trasformazione economica (Teramo 24-26 maggio 2007).
- XLV Geopolitica del Mediterraneo (Bari, 29-31 maggio 2008).
- XLVI Povertà ed esclusione sociale (Firenze 28-30 maggio 2009)
- XLVII Un mondo in movimento: approccio multidisciplinare ai fenomeni migratori (Milano 27-29 maggio 2010).
- XLVIII 150 anni di Statistica per lo sviluppo del territorio: 1861-2011. (Roma 26-28 maggio 2011).

B) GIORNATE DI STUDIO

- Servizi e territorio: il caso della banca, Capri, 3-4-5 maggio 1994
- Mercato del lavoro e migrazione straniera in Italia, Roma, 29-30-31 maggio 1996
- Occupazione e mercato del lavoro, Napoli, 13-14 dicembre 1996 (esaurito)
- L'educazione come processo permanente. Istruzione-formazione-riqualificazione, Milano, 28-29-30 maggio 1998
- Conti pubblici, conti nazionali e welfare state, Teramo, 16-17 ottobre 1998
- Teorie a confronto nella misurazione della povertà, Bologna, 16 aprile 1999
- La qualità dell'informazione statistica, Roma, 6-7 aprile 2000
- Valutazione delle politiche economiche con strumenti statistici. Problemi relativi al disavanzo dello Stato, Roma, 1 dicembre 2000
- Eterogeneità delle dinamiche demografiche dello sviluppo economico nel bacino del Mediterraneo, Foggia, 12-13 ottobre 2001
- Il nuovo Welfare tra riforme e trasformazioni socioeconomiche, Ferrara, 1-2 marzo 2002
- Statistica per l'analisi economica, Campobasso, 2-3 ottobre 2003
- Il ruolo della donna nella mobilità territoriale delle popolazioni, Catania, 1-2 aprile 2005

