

Volume LXVII N. 3/4
Luglio-Dicembre 2013

*Rivista Italiana
di Economia
Demografia
e Statistica*

3/4



SOCIETÀ ITALIANA DI
ECONOMIA
DEMOGRAFIA
E STATISTICA

Organo della
Società italiana
di Economia Demografia
e Statistica

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

SIEDS
SOCIETÀ ITALIANA
DI ECONOMIA DEMOGRAFIA E STATISTICA

CONSIGLIO DIRETTIVO

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

SEDE LEGALE:

C/O Studio Associato Cadoni, Via Ravenna n.34 – 00161 ROMA

sieds.new@gmail.com

rivista.sieds@gmail.com

*Il presente volume è stato realizzato grazie al contributo
dell'Istituto Banco di Napoli*



VOLUME FUORI COMMERCIO – DISTRIBUITO GRATUITAMENTE AI SOCI

VOLUME LXVII – N. 3/4

LUGLIO-DICEMBRE 2013

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



Sede Legale

C/O Studio Associato Cadoni, Via Ravenna n.34 – 00161 ROMA

sieds.new@gmail.com

rivista.sieds@gmail.com

IN QUESTO NUMERO

Questo volume accoglie una selezione delle comunicazioni dei Soci presentate in occasione della 50esima Riunione Scientifica della Società Italiana di Economia, Demografia e Statistica, tenutasi presso l'Università Europea di Roma dal 29 al 31 maggio 2013, sul tema "Trasformazioni economiche e sociali all'inizio del terzo millennio: analisi e prospettive".

Un sentito ringraziamento va ai referee per l'accuratezza e l'importanza del lavoro svolto.

Claudio Ceccarelli

INDICE

Anna Maria Altavilla, Angelo Mazza, Luisa Monaco <i>Incidenza della dinamica demografica sul mercato del lavoro</i>	7
Alberto Arcagni, Gian Carlo Blangiardo, Marco Fattore, Simona Maria Mirabelli <i>Misurare il livello di integrazione della popolazione straniera in ambito economico-lavorativo. Verso l'applicazione di una nuova metodologia</i>	15
Barbara Baldazzi, Alessandro Bianchi, Anna Emilia Martino, Paola Paladini <i>Innovazioni di processo e uso delle variabili testuali: il caso dell'Adult Education Survey</i>	23
Elisa Barbiano di Belgiojoso, Livia Elisa Ortensi <i>Should I stay or should I go? The case of Italy</i>	31
Elisa Barbiano di Belgiojoso, Stefania Rimoldi <i>Debiti e crediti demografici nei paesi europei</i>	39
Antonella Bernardini, Loredana De Gaetano e Matteo Mazziotta <i>Dal censimento al farm register: l'uso delle indagini post censuarie</i>	47
Gian Carlo Blangiardo, Stefania Rimoldi <i>L'invecchiamento "importato" nelle regioni italiane</i>	55

Barbara Boschetto, Antonella Iorio, Carlo Lucarelli, Michele Antonio Salvatore <i>Limitazioni nella partecipazione al mercato del lavoro delle persone con problemi di salute in Italia</i>	63
Claudio Ceccarelli, Anna Pezone, Simona Rosati <i>L'utilizzo delle Liste Anagrafiche Comunali nella statistica ufficiale</i>	71
Eralba Cela, Eros Moretti <i>Gendered remittances of Romanians in Italy</i>	79
Marco Centra, Andrea Cutillo, Valentina Gualtieri <i>I rendimenti occupazionali e retributivi dei differenti percorsi formativi in ottica di genere</i>	87
Andrea Cutillo, Romina Fraboni, Claudio Ceccarelli <i>Classe sociale, coorte e iscrizione universitaria</i>	95
Agostino Di Ciaccio, Giovanni M. Giorgi <i>Statistical analysis of social networks</i>	103
Luigi Di Comite, Simona Giordano <i>Geografia della fame: sessanta anni dopo!</i>	111
Valentina Ferri e Leonardo Palmisano <i>Tipologie di collegamento su ferro tra capoluoghi di provincia: il caso di Trenitalia</i>	119
Marco Fortini, Luca Mancini, Luigi Marcone, Eleonora Mussino, Evelina Paluzzi <i>Chi si stabilisce in Italia? Transizione verso la residenza degli immigrati extracomunitari</i>	127
Maria Carella, Thaís García Pereiro <i>La vulnerabilità post-divorzio: Spagna e Italia a confronto</i>	135
Margherita Gerolimetto, Christine Mauracher <i>Analysis of food consumption in Europe via time series clustering</i>	143
Giovanni Maria Giorgi, Alessio Guandalini <i>A sampling estimator of the Bonferroni inequality index</i>	151
Cinzia Graziani, Silvia Loriga, Michele Antonio Salvatore, Andrea Spizzichino <i>Il mercato del lavoro in Italia dal 1977 ad oggi</i>	159

Sara Grubanov-Bošković, Karra Greenberg <i>Major, initial family-life transitions amongst young Italians.....</i>	167
Mario Mastrangelo <i>Povert� e deprivazione in Italia: un'analisi multilivello.....</i>	175
Silvestro Montrone, Antonella Massari, Paola Perchinunno, Stefania Girone <i>An integrated archive of the lifestyles of families</i>	183
Luciano Nieddu, Cecilia Vitiello <i>A proposal for a semiparametric classification method with prior information.....</i>	191
Paola Naddeo <i>Recenti trasformazioni del lavoro nel settore agricolo.....</i>	199
Michela C. Pellicani, Valeria Moro <i>Age structure evolution in some Sub-Saharan countries: the advantage of ageing.....</i>	207
Roberta Saladino <i>Soddisfazione, aspettative e prospettive degli studenti stranieri dell'Universit� della Calabria</i>	215
Domenico Summo, Tommaso Pepe <i>Dal benessere all'analisi dell'efficienza di un territorio</i>	223
Valentina Talucci <i>Misurare l'Europa che cambia: dalle strategie di Lisbona al piano 2020.....</i>	231

INCIDENZA DELLA DINAMICA DEMOGRAFICA SUL MERCATO DEL LAVORO

Anna Maria Altavilla, Angelo Mazza, Luisa Monaco

1. Introduzione

E' ampiamente riconosciuto che la crescita economica e il benessere sociale di una nazione sono strettamente legati alle trasformazioni demografiche che in essa si realizzano. E' opportuno, pertanto, che i sistemi politici tengano conto dei cambiamenti demografici in atto, se hanno interesse a fronteggiare adeguatamente gli effetti che la dinamica della popolazione produce sull'organizzazione economica e sociale di un paese.

Il fenomeno demografico che in Italia oggi assume particolare importanza, soprattutto per le sue ricadute sul mercato del lavoro, è l'intenso processo d'invecchiamento che, modificando profondamente la struttura per età della popolazione italiana, ha modificato anche la struttura per età della classe lavoratrice (Stranges, 2008), determinando fra l'altro uno squilibrio tra la classe di chi produce all'interno del sistema economico e la classe di chi, non partecipando all'attività produttiva, grava sui primi.

Nel presente lavoro, l'attenzione è concentrata su alcune classi economicamente produttive, in particolare sulle due classi [15-24] e [55-64], all'interno delle quali la prevista differente dinamica demografica determinerà nei prossimi anni una differente evoluzione della potenziale offerta di lavoro (Cesarotto, 2009), con importanti ricadute sulla corrispondente domanda di lavoro, qualora si volesse realizzare il raggiungimento o il mantenimento di determinati obiettivi occupazionali.

Partendo dalla situazione attuale, ipotizzando che il *trend* del mercato del lavoro, osservato nel recente passato nelle diverse ripartizioni in cui si articola il territorio italiano, continui a manifestarsi nel prossimo futuro, alla luce delle previsioni effettuate dall'Istat sulla dinamica della popolazione, è stimata la variazione della domanda di lavoro utile a mantenere i livelli occupazionali ritenuti soddisfacenti o per correggere quelli lontani dagli obiettivi occupazionali fissati, per esempio, dalla strategia UE2020.

2. Analisi della situazione attuale

Prendendo in esame i dati riferiti al periodo 2004-2011 riguardanti le classi 15-24 anni (giovani) e 55-64 anni (adulti), si nota che la popolazione giovanile ha registrato complessivamente una riduzione, mentre la popolazione anziana è aumentata. Tuttavia, come mostra la tabella 1 (ove non diversamente specificato, le tabelle contengono nostre elaborazioni su dati Istat (2011), il dato presenta una forte eterogeneità rispetto alle diverse ripartizioni geografiche; probabilmente, tale fenomeno, oltre che essere influenzato dal declino della natalità nelle regioni meridionali, è effetto, da un lato, dei flussi migratori dall'estero, che tendono a concentrarsi nelle regioni dove vi sono maggiori opportunità occupazionali¹, dall'altro è il risultato del flusso migratorio interno dal Mezzogiorno verso il Nord della penisola².

Tabella 1 – *Variazione % 2004-2011 popolazione residente per ripartizione geografica, (classi di età 15-24 e 55-64).*

Ripartizioni	15-24 anni	55-64 anni
Nord-ovest	3,4	3,4
Nord-est	6,3	5,4
Centro	2,5	7,9
Mezzogiorno	-7,3	13,8
TOTALE	-1,0	8,0

Tabella 2 – *Variazione % 2004-2011 della forza lavoro e dei tassi di attività distinti per ripartizione geografica, (classi di età 15-24 e 55-64).*

Ripartizioni	15-24 anni		55-64 anni	
	Forza lavoro	Tassi attività	Forza lavoro	Tassi attività
Nord-Ovest	-22,09	-24,10	39,12	34,26
Nord-Est	-19,29	-23,54	39,76	32,57
Centro	-15,80	-17,40	32,50	22,77
Mezzogiorno	-31,83	-26,47	26,94	11,90
TOTALE	-24,16	-23,38	33,70	23,80

Con riferimento alle tendenze del mercato del lavoro si rileva che per le coorti di giovani il tasso di attività è stato via via decrescente, mentre per le coorti di anziani è stato crescente. In particolare, nel periodo compreso tra il 2004 e il 2011, a fronte di una riduzione della forza lavoro di età 15-24 anni del 24,16 per cento, la

¹ Peraltro, il saldo migratorio con l'estero, che nel triennio 2008-2010 ha registrato un valore medio del 6,6 per mille, nel 2011 è stato del 5,3 per mille, dunque in significativo calo (dati CNEL 2012).

² Al 2011, il saldo migratorio interno per il Centro-Nord è stato dell'1,3 per mille, a fronte di un saldo negativo registrato nel Mezzogiorno (dati CNEL 2012).

forza lavoro in età 55-64 anni è aumentata del 33,7 per cento (tab. 2). Se si considera, poi, il dato concernente gli occupati nelle diverse fasce di età e riferito alle ripartizioni geografiche (tab. 3), è evidente come l'occupazione giovanile abbia registrato una sostanziale riduzione (meno 29,70 per cento), in modo particolare al Mezzogiorno, mentre l'occupazione anziana è aumentata del 34,06 per cento, e in maggiore misura nelle regioni del Nord Italia.

Tabella 3 – *Variazione % , 2004-2011, degli occupati e dei tassi d'occupazione per ripartizione geografica (classi di età 15-24 e 55-64).*

Ripartizioni	15-24 anni		55-64 anni	
	Occupati	Tassi	Occupati	Tassi
Nord-ovest	-29,55	-31,37	38,11	33,33
Nord-est	-27,47	-31,33	39,16	32,32
Centro	-23,61	-25,19	32,54	22,99
Mezzogiorno	-34,89	-29,84	28,90	13,65
TOTALE	-29,70	-28,98	34,06	24,13

Sempre nello stesso intervallo temporale si rileva che il numero totale di occupati in Italia (classe 15-64 anni) è aumentato complessivamente dell'1,9 per cento, rispetto al 3,7 per cento riferito all'EU27 (Eurostat, 2013). Mettendo a confronto con i rispettivi tassi i dati riferiti alla forza lavoro e agli occupati, emerge l'incidenza della differente dinamica della popolazione all'interno delle due classi nel periodo osservato.

Tabella 4 – *Variazione % , 2004-2011, dei disoccupati e dei tassi di disoccupazione per ripartizione geografica (classi di età 15-24 e 55-64).*

Ripartizioni	15-24 anni		55-64 anni	
	Disoccupati	Tassi	Disoccupati	Tassi
Nord-ovest	23,46	57,45	68,42	24,24
Nord-est	48,89	85,85	66,67	13,64
Centro	14,10	35,05	31,25	0,00
Mezzogiorno	-27,10	7,45	-2,13	-22,22
ITALIA	-6,23	23,66	25,27	-6,30

Per la classe 15-24 si rileva, infatti, che alla diminuzione dei tassi corrisponde una variazione negativa minore della forza lavoro o degli occupati nelle ripartizioni in cui la popolazione ha registrato un aumento, viceversa nel Mezzogiorno dove la classe giovanile ha subito un declino numerico. Con riferimento ai soggetti di età 55-64 che, invece, sono aumentati numericamente in tutte le ripartizioni, all'aumento registrato nei tassi di attività e di occupazione corrisponde sempre una maggiore variazione nella forza lavoro e nel numero di occupati. Con riferimento ai disoccupati, il cui tasso è calcolato rispetto alla forza lavoro, la differente situa-

zione delle due classi sintetizzata nella tab. 4 (nostra elaborazione su dati Istat) va anch'essa interpretata tenendo conto della diversa dinamica della forza lavoro nelle due classi; forza lavoro che si presenta in diminuzione per tutte le ripartizioni nella classe giovanile, in aumento nella classe degli anziani. Le più basse variazioni registrate nel numero dei disoccupati della classe 15-24 rispetto alle variazioni dei relativi tassi va attribuita, pertanto, alla diminuzione della forza lavoro, come pure la diminuzione del numero dei disoccupati nel Mezzogiorno non va interpretata come un segnale positivo, ma come una conseguenza della notevole diminuzione della forza lavoro. Nella classe d'età 55-64 l'aumento della forza lavoro determina l'inversione del rapporto fra tassi di disoccupazione e disoccupati le cui variazioni si presentano sempre maggiori rispetto a quelle rilevate per i rispettivi tassi. Anche in questo caso va segnalata la situazione del Mezzogiorno dove il segno negativo registrato per la variazione del tasso di disoccupazione e dei disoccupati va interpretata alla luce della differente variazione della forza lavoro.

3. Le proiezioni al 2020 sul mercato del lavoro

Dall'analisi effettuata per il periodo 2004-2011 è emerso che la dinamica della popolazione può avere una notevole incidenza sulle grandezze che determinano le caratteristiche del mercato del lavoro e che essa, quindi, non può essere trascurata. Si è osservato, soprattutto, il suo impatto sulla potenziale offerta di lavoro, grandezza alla quale è necessario fare riferimento quando ci si prefigge il raggiungimento di determinati livelli di occupazione e di disoccupazione

Tabella 5 – *Variazione % 2011-2020 della popolazione e dei tassi di occupazione e disoccupazione per ripartizione e per classi d'età 15-24 e 55-64 (1° ipotesi).*

Ripartizioni	15-24 anni			55-64 anni		
	Popol.ne	Occup.ne	Disocc.ne	Popol.ne	Occup.ne	Disocc.ne
Nord-Ovest	10,18	-9,24	-1,36	12,05	-10,75	-3,15
Nord-Est	12,77	-11,32	-1,51	17,08	-14,59	-3,41
Centro	3,92	-3,77	-1,14	13,96	-12,25	-3,01
Mezzog.no	-13,59	15,73	0,94	13,11	-11,59	-2,16

Allo scopo, nell'analisi che segue, con riferimento alle classi di età 15-24 e 55-64 anni e alle diverse ripartizioni geografiche, sono simulate, in assenza d'interventi di *policy*, alcune possibili situazioni. Secondo una prima ipotesi di lavoro si suppone che il numero di soggetti occupati e di disoccupati fino al 2020 rimanga immutato rispetto a quello osservato al 2011. Le previsioni sul tasso di disoccupazione sono state effettuate supponendo che la forza lavoro subisca le stesse variazioni previste per la popolazione

Tenendo conto delle previsioni sulla popolazione (scenario centrale, Istat, 2013), dai risultati raccolti nella tab. 5 (nostra elaborazione su dati Istat) si può valutare l'entità della diminuzione del valore dei tassi di occupazione e disoccupazione nelle classi d'età e nelle ripartizioni nelle quali la popolazione si prevede in aumento e l'entità dell'aumento laddove si prevede un declino della popolazione.

Tabella 6 – *Variazione % media annua 2004-2011 dei tassi di attività, occupazione e disoccupazione per ripartizione (classi d'età 15-24 e 55-64)*

Ripartizioni	15-24 anni			55-64 anni		
	Attività	Occup.	Disoccup.	Attività	Occup.	Disoccup.
Nord-Ovest	-3,86	-5,24	6,70	4,30	4,20	3,15
Nord-Est	-3,76	-5,23	9,26	4,11	4,08	1,84
Centro	-2,69	-4,06	4,39	2,97	3,00	0,00
Mezzogiorno	-4,30	-4,94	1,03	1,62	1,84	-3,53

Una seconda ipotesi prevede, invece, che i tassi di attività, occupazione e disoccupazione seguano fino al 2020 lo stesso *trend* osservato nel periodo 2004-11. Utilizzando i tassi di variazione media annua riportati nella tab.6 (nostra elaborazione su dati Istat), sono estrapolati i tassi per il periodo 2013-2020 e, con l'ausilio delle previsioni Istat, è stimato il numero di soggetti componenti la forza lavoro, gli occupati e i disoccupati necessario per il raggiungimento al 2020 dei livelli dei corrispondenti tassi stimati (tabb.7;8;9, nostra elaborazione su dati Istat).

Tabella 7 – *Forza lavoro (in migliaia) e variazione % 2011-2020 della forza lavoro e dei tassi di attività per ripartizione geografica, nelle classi di età 15-24 e 55-64 (2° ipotesi)*

Ripartizioni	15-24 anni			55-64 anni		
	Forza lavoro	Var. f. lavoro	Var. tassi attività	Forza lavoro	Var.f. lavoro	Var. tassi attività
Nord-ovest	345	-22,70	-29,85	1308	63,64	46,04
Nord-est	271	-20,14	-29,19	985	68,23	43,69
Centro	252	-18,73	-21,80	954	48,34	30,18
Mezzogiorno	328	-41,81	-32,65	1243	30,71	15,56

Seguendo sempre l'ipotesi che rimanga immutato il *trend* osservato nel periodo 2004-11, l'analisi è estesa a tutte le classi lavorative per le quali sono stimati i valori che i tassi di occupazione raggiungerebbero progressivamente dal 2013 al 2020 e, con l'ausilio delle previsioni Istat, il corrispondente numero di soggetti componenti la forza lavoro, gli occupati e i disoccupati.

Tabella 8 – Occupati (in migliaia) e variazione % 2011-2020 degli occupati e dei tassi di occupazione per ripartizione geografica, nelle classi di età 15-24 e 55-64 (2° ipotesi)

Ripartizioni	15-24 anni			55-64 anni		
	Occupati	Var. occupati	Var. tassi occup.	Occupati	Var. occupati	Var. tassi occup.
Nord-ovest	236	-32,10	-38,37	1243	62,20	44,76
Nord-est	189	-30,45	-38,32	959	67,84	43,35
Centro	158	-28,44	-31,14	924	48,68	30,47
Mezzogiorno	184	-45,22	-36,60	1207	33,34	17,88

Tabella 9 – Disoccupati (in migliaia) e variazione % 2011- 2020 dei disoccupati e dei tassi di disoccupazione per ripartizione geografica, nelle classi di età 15-24 e 55-64 (2° ipotesi)

Ripartizioni	15-24 anni			55-64 anni		
	Disocc.	Var. disoocc.	Var. tassi disocc.	Disocc.	Var. disoocc.	Var. tassi disocc.
Nord-ovest	137	38,55	79,25	71	116,32	32,19
Nord-est	118	77,16	121,85	29	98,28	17,86
Centro	107	19,59	47,15	31	48,34	0,00
Mezzog.no	145	-36,18	9,67	44	-5,38	-27,61

Tabella 10 – Tassi di occupazione per classi di età e ripartizioni geografiche. Proiezioni al 2020, tra parentesi tassi 2011.

Ripartizioni	15-24	25-34	35-44	45-54	55-64
Nord-Ovest	(24,5) 15,1	(78,5) 74,4	(83,1) 81,1	(79,6) 85,7	(36,9) 53,8
Nord-Est	(26,3) 16,2	(78,1) 73,2	(85,0) 85,0	(82,0) 89,4	(38,9) 56,3
Centro	(19,9) 13,7	(69,8) 65,2	(79,2) 78,0	(76,8) 79,9	(41,5) 53,7
Mezzog.no	(13,4) 8,5	(47,6) 43,0	(58,4) 54,3	(57,7) 55,1	(35,6) 42,2

Come si evince dai dati raccolti nella tab.10, le diverse classi si troverebbero, in mancanza di interventi di *policy*, in situazioni differenti rispetto all'obiettivo di EU2020. Sono anche estrapolati i tassi riferiti alla popolazione di età 20-64 anni³, per verificare in che misura i tassi di occupazione al 2020 convergerebbero (o divergerebbero) verso gli obiettivi "Europa 2020", che fissano per l'Italia il *target* di un tasso di occupazione del 67-69 per cento da raggiungere entro il 2020. I risultati mostrano come, seguendo il *trend* dei tassi di occupazione passati, il tasso di occu-

³ Si è scelto di considerare questa fascia di età in quanto, la strategia Europa 2020, pur raccomandando che nel raggiungimento dell'obiettivo sia promossa una maggiore occupazione femminile e dei lavoratori anziani, non fissa un *target* per singole classi di età, bensì riferito alla popolazione classe 20-64.

pazione al 2020 si posizionerebbe per l'Italia al 60,9 per cento, dunque ancora distante dal tasso obiettivo.

Tabella 11 – Tassi di occupazione 20-64 anni per ripartizioni geografiche. Proiezioni al 2020

Ripartizioni	Var. % 2004-2011	2013	2014	2015	2016	2017	2018	2019	2020
Nord-Ovest	0,17	68,9	69,0	69,2	69,3	69,4	69,5	69,6	69,7
Nord-Est	0,26	70,9	71,1	71,3	71,5	71,7	71,9	72,1	72,2
Centro	0,11	65,5	65,6	65,6	65,7	65,8	65,9	65,9	66,0
Mezzogiorno	-0,68	47,2	46,9	46,6	46,2	45,9	45,6	45,3	45,0
ITALIA	-0,04	61,1	61,1	61,1	61,0	61,0	61,0	60,9	60,9

Guardando, poi, alle ripartizioni si nota che, mentre nelle regioni del Nord l'obiettivo sarebbe pienamente raggiunto, il Mezzogiorno sarebbe ancora molto lontano dal *target* di crescita occupazionale, presentando un tasso di occupazione del 45 per cento, con un *trend* decrescente. E ancora, guardando alla tabella 10 si rileva che, fatta eccezione per il Mezzogiorno, nelle classi centrali l'obiettivo sarebbe largamente superato, mentre le classi 15-24 e 55-64 presenterebbero un differenziale rispetto al tasso obiettivo ancora significativo.

4. Conclusioni

E' evidenziato dall'analisi dei dati che, nelle varie ripartizioni geografiche, l'Italia presenta attualmente una forte eterogeneità nella struttura del mercato del lavoro, determinata in parte anche dalla diversa evoluzione demografica verificata all'interno del suo territorio. Simulando alcuni possibili scenari sulla futura evoluzione delle grandezze che caratterizzano il mercato del lavoro, e tenendo nella dovuta considerazione la futura dinamica della popolazione, emerge dai risultati che, in mancanza di interventi mirati, lo squilibrio sarebbe crescente tra le ripartizioni e tra le classi d'età, soprattutto con riferimento alla situazione occupazionale. Occorrerebbe, pertanto, adottare per le diverse macro classi d'età e per le diverse aree del paese delle misure di *policy* piuttosto diversificate che tengano conto sia delle differenti posizioni di partenza, sia della futura evoluzione demografica ed economica del Paese

Riferimenti bibliografici

CANTALINI B. RIGHI A. LAURICELLA A. 1993. Invecchiamento e offerta di lavoro, in *Popolazione, tendenze demografiche e mercato del lavoro* IRP-CNR, Roma.

- CESARATTO S. 2009. Invecchiamento della popolazione, mercato del lavoro e welfare: un'introduzione critica. *Note di Economia*, XIV, n. 3, 395-429.
- CNEL 2012. Rapporto Il mercato del lavoro in Italia 2011-2012, Roma
- DEMENY P. 1993. Population, Demographic and Labor Market Trends: The Demographer's Point of View in *Popolazione, tendenze demografiche e mercato del lavoro*, IRP, Roma.
- GESANO G. 2005. Dinamiche di popolazione e mercato del lavoro. *Economia & Lavoro*, 39:1.
- ISTAT 2011. Il futuro demografico del Paese. Previsioni regionali della popolazione residente al 2065.
- MALPAS N. 1993. Démographie et marché de l'emploi: le point de vue de la Communauté Européenne, in *Popolazione, tendenze demografiche e mercato del lavoro*, IRP, Roma.
- STRANGES M. 2008. Dinamiche demografiche e mercato del lavoro in Italia. *Quaderni Europei sul Nuovo Welfare*, n. 10, 192-209.
- TERMOTE M. 1993. Population and Labor Market: The Regional Dimension, in *Popolazione, tendenze demografiche e mercato del lavoro*, IRP, Roma

SUMMARY

The effects of demographic dynamic in the labour market

This paper examines the effects that demographic trends of the next few years may produce to the structure of the labor market in Italy. Two different scenarios, both based on Istat population projections are considered. According to both of them, in the absence of appropriate policy actions, differences among geographical areas and classes of ages, in the composition of the labor market, would be exacerbated. Differentiated policy actions, closely related to the heterogeneity present in the macro groups of ages and in the different areas of the country, are therefore advisable.

Anna Maria ALTAVILLA, Professor, University of Catania,
Department "Economia e Impresa", altavil@unict.it
Angelo MAZZA, Assistant Professor, University of Catania,
Department "Economia e Impresa", a.mazza@unict.it
Luisa MONACO, University of Catania, Department "Economia e Impresa",
lmonaco@unict.it

**MISURARE IL LIVELLO DI INTEGRAZIONE
DELLA POPOLAZIONE STRANIERA
IN AMBITO ECONOMICO-LAVORATIVO.
VERSO L'APPLICAZIONE DI UNA NUOVA METODOLOGIA**

Alberto Arcagni, Gian Carlo Blangiardo, Marco Fattore, Simona Maria Mirabelli

1. L'approccio teorico

La tendenza ad applicare allo studio di fenomeni multidimensionali, strettamente connessi a variabili di tipo ordinale, le medesime strutture logico-formali con cui si analizzano le variabili continue può risultare problematico sotto diversi punti di vista. In primo luogo, tali analisi si basano sull'ipotesi che tutti i fenomeni siano 'misurabili' in termini quantitativi e che come tali possano essere trattati. In secondo luogo, le procedure di aggregazione, sovente utilizzate nella costruzione di indicatori sintetici, divengono inapplicabili, per la natura non numerica delle variabili. Dalla necessità di rispettare la natura dei fenomeni e di superare l'incompatibilità tra la tipologia ordinale dei dati e le procedure di calcolo normalmente utilizzate, nasce l'interesse per la *teoria degli insiemi parzialmente ordinati (poset)* e per la sua applicazione a problemi di valutazione socio-economica. Strutture parzialmente ordinate emergono naturalmente in questi contesti, non appena si considerino sistemi multidimensionali di variabili ordinali che rendono le unità statistiche non completamente confrontabili fra loro. Si pensi, ad esempio, agli studi sul benessere non monetario e sulla deprivazione materiale (Fattore et al. 2011). Dal punto di vista formale, un *poset* è un insieme sui cui elementi sia definito un ordinamento parziale, ovvero una relazione binaria che soddisfi le proprietà di riflessività, antisimmetria e transitività. In pratica, si tratta di un insieme in cui alcune coppie di elementi possono essere ordinate e altre no. Un insieme parzialmente ordinato con un numero finito di elementi può essere efficacemente rappresentato attraverso un diagramma di *Hasse*, cioè attraverso un grafo i cui nodi, che rappresentano le unità parzialmente ordinate, sono connessi da archi discendenti se e solo se è possibile stabilire tra i nodi stessi un ordinamento. Nel presente lavoro, le unità parzialmente ordinate sono costituite dai *profili* di integrazione economico-lavorativa degli immigrati stranieri, cioè dalle possibili configurazioni dei punteggi ordinali sulle dimensioni di valutazione considerate. L'obiettivo è quello di valutare il livello di integrazione di un soggetto immigrato, attraverso un confronto multidimensionale tra il suo profilo e un insieme di profili di riferimento, che rappresentino situazioni di certa integrazione. Trattandosi di

profili multidimensionali, non completamente ordinabili, il confronto non è immediato e richiede l'applicazione di una adeguata metodologia, illustrata nel paragrafo successivo. Come si potrà verificare nel seguito, la metodologia produce una valutazione sintetica del livello di integrazione senza ricorrere ad aggregazioni fra punteggi.

2. Il procedimento di analisi

La base dati cui si fa riferimento è stata resa disponibile dall'Osservatorio Regionale per l'integrazione e la multietnicità della regione Lombardia (ORIM). Per ciascuno degli anni presi in esame, l'ORIM ha fornito i risultati delle indagini campionarie condotte (nei mesi di maggio e giugno) sulla popolazione straniera ultra14enne proveniente da Paesi a forte pressione migratoria e presente in regione a qualsivoglia titolo: stranieri regolarmente soggiornanti (in possesso o meno di iscrizione anagrafica) e stranieri privi di titolo di soggiorno. Per gli anni 2010 e 2011 sono state selezionate, con la tecnica del campionamento dei centri¹, rispettivamente 8mila unità campionarie; per il 2012 circa 7mila.

Nel presente lavoro, l'attenzione è stata focalizzata sulle seguenti variabili:

1. "REG" – *Ha un regolare contratto?*
2. "STA" – *E' un lavoratore stabile?*
3. "RED" – *Reddito netto da lavoro in quartili.*
4. "OVER" – *Corrispondenza tra professione svolta e formazione scolastica acquisita.*
5. "MOB" – *Mobilità lavorativa rispetto agli ultimi 12 mesi.*

Le variabili "REG" e "STA" sono codificate in forma binaria ("0 – NO"; "1 – SÌ"); mentre quella denominata "RED" assume quattro modalità ("1 – 1° quartile. Fino a 800 euro mensili"; "2 – 2° quartile. Da 801 a 1.000 euro mensili"; "3 – 3° quartile. Da 1.001 a 1.300 euro mensili"; "4 – 4° quartile. Oltre 1.300 euro mensili"). La variabile "OVER" ne prevede tre ("1 – Gravemente inadeguata"; "2 – Moderatamente inadeguata"; "3 – Adeguata"); la variabile "MOB" assume tre modalità ("1 – Mobilità discendente"; "2 – Immobilità"; "3 – Mobilità ascendente"). A ciascuna unità statistica, viene associato il corrispondente profilo, sulle variabili selezionate (vi sono in tutto, teoricamente, 144 profili differenti). L'obiettivo della metodologia è assegnare a ogni profilo, e quindi a ciascuna unità statistica, un punteggio di integrazione sul piano economico-lavorativo. Una volta

¹ Riguardo alla metodologia si veda: Baio G., Blangiardo G.C. e Blangiardo M. (2011), Centre sampling technique in foreign migration surveys: a methodological note", in *Journal of Official Statistics*, vol. 27, 3, pp. 1-16.

che le unità statistiche siano state valutate, è possibile calcolare gli usuali indicatori di sintesi, a livello di popolazione e per diverse sottopopolazioni di interesse. L'attribuzione del punteggio ai singoli profili muove dalla determinazione di un insieme di profili *soglia*, esogenamente identificati come rappresentativi di situazioni di integrazione. In pratica, si tratta di determinare i "peggiori" tra tutti i profili che, senza ambiguità, possono essere classificati come integrati: di fatto si tratta di identificare il loro "confine inferiore". Si noti che, data la multidimensionalità del problema, la soglia non è comunque fornita da un singolo profilo (come si è soliti operare nelle indagini monetarie sulla povertà). Le sfaccettature del problema, infatti, impongono la selezione di taluni profili che, seppur diversi, segnano l' "area sociale" di oggettiva integrazione. In questo studio, la *soglia* è composta da 4 profili, precisamente:

- *Profilo 1.* "REG =1", "STA=1"; "REG=3", "OVER=2", "MOB=2"
- *Profilo 2.* "REG =1", "STA=1"; "REG=2", "OVER=2", "MOB=3"
- *Profilo 3.* "REG =1", "STA=0"; "REG=3", "OVER=2", "MOB=3"
- *Profilo 4.* "REG =1", "STA=0"; "REG=4", "OVER=1", "MOB=3"

Esplicitamente, le 4 condizioni minimali espresse dai profili prescelti come soglia corrispondono ai casi in cui:

- il soggetto in età attiva ha un regolare contratto di lavoro, svolge un'attività in modo stabile, percepisce un reddito non inferiore a 1.000 euro mensili e non superiore a 1.300, esercita una professione solo moderatamente inadeguata rispetto al proprio background formativo e negli ultimi 12 mesi non ha subito cambiamenti nella propria carriera lavorativa (*Profilo 1*);

- lo stesso soggetto ha un regolare contratto, svolge un'attività in modo stabile, percepisce un reddito mensile non inferiore a 800 euro e non superiore a 1.000, esercita una professione solo moderatamente inadeguata rispetto alle proprie credenziali formative, ma nell'ultimo anno ha vissuto un percorso di mobilità lavorativa ascendente (*Profilo 2*);

- il lavoratore ha un regolare contratto, svolge un'attività in modo instabile, percepisce un reddito non inferiore a 1.000 euro mensili e non superiore a 1.300, svolge una professione solo moderatamente inadeguata rispetto al proprio background scolastico, ma nell'ultimo anno ha vissuto un percorso di mobilità lavorativa ascendente (*Profilo 3*);

- lo stesso lavoratore ha un regolare contratto, svolge un'attività in modo instabile, percepisce un reddito superiore a 1.300 euro mensili, esercita una professione gravemente inadeguata rispetto al proprio livello di istruzione, ma nell'ultimo anno ha vissuto un percorso di mobilità lavorativa ascendente (*Profilo 4*).

I profili corrispondenti alla soglia ricevono per definizione punteggio di integrazione pari a 1. Lo stesso punteggio viene attribuito ai profili che dominano

almeno uno dei profili soglia, in base al principio che se un immigrato è integrato, un immigrato in condizioni migliori è a sua volta integrato. Dualmente, ai profili che sono peggiori di *tutti* i profili soglia viene assegnato punteggio di integrazione pari a 0. I profili che non ricadono nei casi precedenti, ricevono un punteggio intermedio, in base alla loro posizione nel poset, rispetto alla soglia. Il calcolo viene effettuato considerando l'insieme delle estensioni lineari del poset e stimando la frequenza con cui un profilo si trova sopra o sotto la soglia, in tale insieme (Fattore et al. 2011). Ne risulta una funzione, chiamata funzione di identificazione, che associa a ogni profilo il corrispondente punteggio di integrazione. Infine, a ciascuna unità statistica presente nel campione viene attribuito il punteggio del corrispondente profilo.

Figura 1 - Diagramma di Hasse

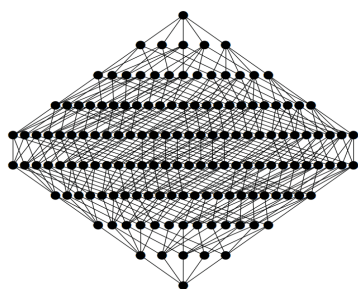
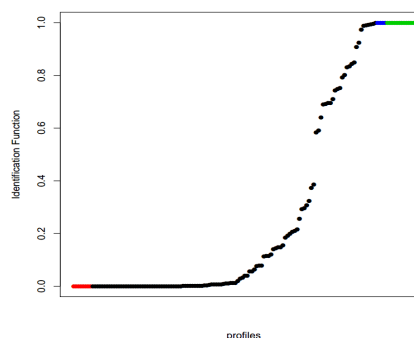


Figura 2 - Funzione di identificazione



3. L'applicazione della metodologia al caso di studio

A partire da tali premesse, è stato innanzitutto analizzato il livello di integrazione della popolazione straniera presente in Lombardia in età 20-64 anni e attivamente inserita (occupata) nel mercato del lavoro, sulla base del campione ORIM dell'anno 2012. Si tratta pertanto di un sub universo che esclude sia i soggetti in condizione non professionale (come studenti, casalinghe o pensionati), sia coloro che, al momento della rilevazione, erano in cerca di una (nuova o prima) occupazione. L'analisi è stata poi estesa al biennio precedente (2010-2011) con l'obiettivo di valutare se, e in quale misura, con l'aggravarsi della crisi economico-occupazionale, il processo di integrazione possa aver subito battute di arresto (Blangiardo, Mirabelli, 2013). I risultati del 2012 mostrano come la componente

maschile, che accentra il 57,3% delle osservazioni, risulti caratterizzata da una presenza di soggetti che sono da ritenersi ‘certamente’ integrati pari al 45,4% dei casi, superando di quasi 25 punti percentuali la corrispondente quota nel sottoinsieme femminile; mentre la condizione di ‘certa’ non integrazione identifica il 4,3% e il 5,4% dei rispettivi collettivi. D’altra parte, il contingente maschile si caratterizza per una maggiore incidenza di soggetti assunti con regolare contratto di lavoro (l’86,9% contro l’84,1% per la componente femminile), quand’anche la quota di lavoratori ‘stabili’ (con contratto a tempo indeterminato) risulti più elevata tra le donne (di circa quattro punti percentuali). Inoltre, un terzo del collettivo maschile percepisce un reddito (netto da lavoro) superiore a 1.300 euro mensili (le lavoratrici con pari entrate solo nell’8% dei casi) e appena un quinto dello stesso contingente si colloca nella fascia di reddito più bassa (non superiore a 800 euro); viceversa, la quota femminile corrispondente accentra il 43,3% del sottogruppo. Riguardo alla corrispondenza tra professione svolta e *background* scolastico acquisito, tra le lavoratrici la quota di coloro che svolgono una mansione per la quale la qualifica richiesta risulta gravemente inadeguata rispetto al livello di istruzione supera di 2 punti percentuali la corrispondente frazione maschile (l’11,4% contro il 9,6%), sebbene la percentuale di donne coinvolte in una attività coerente con il titolo di studio sia leggermente superiore a quella dei maschi (il 50,7% contro il 47,9%). Riguardo alla mobilità lavorativa, gli uomini intraprendono un percorso di transizione ascendente (rispetto alla condizione professionale precedente) nell’11,5% dei casi, mentre le donne solo nell’8,4%; anche se per queste ultime è maggiore la quota di chi non ha mutato la propria condizione lavorativa (l’87,4% contro l’83%). Tra gli aspetti ritenuti più significativi nel determinare il livello di integrazione della popolazione immigrata nella società di accoglienza, la durata di permanenza nel nuovo contesto di insediamento svolge un ruolo di indubbia centralità, come ampiamente documentato dagli studi sull’argomento e come dimostrato dalla crescente attenzione riservata dai decisori pubblici (anche europei) nel campo dell’integrazione. All’aumentare dell’anzianità migratoria migliora il grado di inclusione degli immigrati nella nuova realtà (sociale, politica economica e culturale) e tale relazione sembra valere soprattutto per la dimensione economica, in cui il processo di inserimento appare strettamente dipendente dalla durata della presenza nel paese di immigrazione. Anche alla luce dei risultati ottenuti con l’applicazione della metodologia proposta nel presente lavoro, il legame tra le due variabili trova ampia conferma: la frequenza di inclusioni nell’insieme dei soggetti ‘certamente’ integrati raggiunge il 42,7% dei casi per chi vive in Italia da oltre dieci anni, mentre tra i *new comers* l’analoga quota si colloca al 3,1%. Non a caso, se ci addentriamo nell’analisi delle caratteristiche strutturali emerge un migliore inserimento nel mercato del lavoro da parte dei soggetti che hanno maturato una

presenza ultradecennale nel nostro paese. Tra coloro che vi abitano da oltre dieci anni, si osserva, infatti, una maggiore incidenza di lavoratori con un regolare contratto (il 90,1% del sottoinsieme contro il 50,2% per chi vi dimora da meno di due), una più diffusa stabilità dal punto di vista contrattuale (l'82,8% contro il 63,6%) e il raggiungimento di un maggiore livello di reddito (la probabilità di guadagnare non meno di 1.300 euro mensili riguarda il 32% del collettivo, mentre tra i 'nuovi arrivati' la rispettiva quota si ferma al 2%). Riguardo all'opportunità di svolgere una professione coerente con il proprio *background* formativo, la maggiore anzianità migratoria sembra facilitare un inserimento 'congruo' per oltre la metà dei casi (tra i soggetti di più recente arrivo l'analoga percentuale non raggiunge un terzo). Tuttavia, rispetto alla mobilità lavorativa, i lavoratori da maggior tempo in Italia intraprendono un percorso occupazionale ascendente solo nel 7% dei casi, mentre tra coloro che vi abitano da meno di due anni l'analoga condizione riguarda quasi il 30% dei casi². Passando al livello di scolarità che, insieme all'anzianità migratoria, fornisce un'importante chiave interpretativa della dimensione economica dell'integrazione, la quota di chi è da ritenersi 'certamente' integrato tende inizialmente a crescere all'aumentare del titolo di studio, passando dal 28,8% per chi è ne è totalmente privo (il 3,1% dei rispondenti) al 35,6% per chi ha assolto l'obbligo scolastico (il 26,5% del campione), sino ad arrivare al 41,2% tra i diplomati (il 47%). Al contrario, tra i laureati la quota di chi è 'certamente' integrato scende al 15,9% del sottoinsieme scontando, più pesantemente degli altri, la condizione di sotto-inquadramento presente nel mercato del lavoro. D'altra parte, se è vero che i lavoratori più scolarizzati hanno un regolare contratto di lavoro nell'89,2% dei casi (mentre per chi non ha conseguito alcun titolo la rispettiva percentuale si attesta al 75,4%) e svolgono un'attività in modo stabile nell'80,2% dei casi (a fronte del 70,9% tra chi non alcun titolo formale) percependo un reddito superiore a 1.300 euro mensili nel 25,1% del sottoinsieme (contro il 14,1%), è anche vero che, rispetto sia al percorso di mobilità lavorativa sia alla corrispondenza tra formazione acquisita e professione svolta i vantaggi sul piano dell'integrazione economica, derivanti dal possesso di un titolo di studio elevato, sembrano annullarsi o si rivelano persino controproducenti. Tra i laureati coloro che hanno conseguito un percorso lavorativo di mobilità ascendente concentrano solo il 7,5% del sottoinsieme, mentre tra chi non ha alcuna istruzione l'analoga quota sale al 15,1%; allo stesso modo, tra i lavoratori più scolarizzati, coloro che svolgono una professione adeguata rispetto al titolo conseguito concentrano poco più di un quarto dei casi; mentre tra coloro che ne sono privi le corrispondenti percentuali si attestano su valori nettamente superiori, sino a

² La migliore *performance* è spiegabile con il passaggio dalla condizione di disoccupato a quella di occupato, quand'anche si tratta, in due casi su tre, di 'nuovi' lavoratori assunti senza un regolare contratto di lavoro.

coinvolgere la quasi totalità dei lavoratori che si dichiarano sprovvisti di qualsiasi titolo formale. L'analisi estesa al biennio precedente (2010-2011) conduce, da una prima lettura, a risultati non molto dissimili da quelli cui siamo pervenuti nel 2012. Infatti anche nel 2010, così come nel 2011, i casi di 'certa' integrazione sono più ricorrenti tra gli uomini (in vantaggio di almeno 25 punti percentuali rispetto alla corrispondente quota femminile), tra i soggetti con una anzianità migratoria ultradecennale (che distanziano di 36 punti percentuali chi ha maturato sul territorio una presenza inferiore a due anni) e tra i lavoratori meno istruiti (13 punti in più rispetto a chi conseguito la laurea). Tuttavia, da un esame più approfondito, emergono elementi di valutazione che consentono di cogliere -in maniera più analitica- la dinamica evolutiva del processo di integrazione. A fronte di una minore incidenza di uomini occupati (di oltre dodici punti percentuali rispetto al 2011 e di circa sei rispetto al 2010), la quota corrispondente 'certamente' integrata passa dal 45,7% nel 2010 al 53,5% nel 2011, ritornando ai valori pregressi nell'anno più recente (il 45,4% del sottoinsieme). Per quanto riguarda il collettivo femminile si osserva un analogo andamento: tra il 2010 e il 2011 aumenta l'incidenza di lavoratrici 'integrate' (passate dal 21,1% al 25,6%), mentre negli ultimi dodici mesi la rispettiva quota scende ai livelli osservati nel 2010 (il 20,5%). Sul fronte dell'anzianità migratoria, si osserva tra i lavoratori che vivono in Italia da meno di due anni un sensibile calo di soggetti 'certamente' integrati (-6 punti percentuali rispetto al biennio precedente): nel 2010 essi incidono per il 9,3% del collettivo e nel 2011 per il 12,3%; parallelamente, sale la quota di coloro che 'certamente' non lo sono (dal 7,9% nel 2010 al 9,5% nel 2011 per attestarsi al 12,4% nel 2012). Anche i lavoratori che sono in Italia da oltre 10 anni, in crescita di quasi dodici punti percentuali rispetto al 2010, registrano un analogo *trend*: aumenta nell'anno più recente la quota di soggetti non integrati, mentre diminuisce quella di coloro che 'certamente' lo sono, sebbene la frazione di questi ultimi concentri ancora nel 2012 non meno di due quinti del contingente. Relativamente al titolo di studio, la composizione percentuale di coloro che ne sono totalmente privi non segnala variazioni significative rispetto al biennio scorso: anche nel 2010 quasi il 30% del collettivo si colloca nella posizione più favorevole all'integrazione (nel 2011 poco meno del 40%) e, per quanto riguarda i laureati, raddoppia la quota di chi si trova nella condizione peggiore, passando dal 6% nel 2010 all'11,5% nel 2012, mentre risulta sostanzialmente invariata la frazione di chi si trova nella situazione migliore (dal 16,8% nel 2010 al 15,9% nel 2012).

Alla luce di tali risultati appare evidente come (anche) chi partecipa attivamente al mercato del lavoro (nella maggior parte dei casi con un contratto a tempo indeterminato) subisca gli effetti della crisi economica che perdura, nella realtà italiana (e non solo), ormai da cinque anni. Le difficili condizioni economiche che avversano il mercato del lavoro locale si riflettono, inevitabilmente, sul processo di

integrazione, soprattutto, quando la ‘spinta’ a emigrare è dettata da motivazioni di tipo economico (anche in assenza di realistiche opportunità di inserimento lavorativo nel nuovo contesto). Per questa ragione la partecipazione al mercato del lavoro, come esplicitamente dichiarato dall’Unione Europea e dagli organismi che la rappresentano, svolge un ruolo di crescente importanza, che ritrova nuovo vigore proprio per effetto delle mutate condizioni economiche che espongono a ‘nuovi’ rischi di vulnerabilità e di emarginazione sociale persino chi può contare su una più lunga conoscenza del paese di immigrazione o disporre di adeguati strumenti formativi e culturali.

Riferimenti bibliografici

- BLANGIARDO G.C., MIRABELLI S.M. 2013. Il lavoro. In BLANGIARDO G.C. (Eds) *L’immigrazione straniera in Lombardia. La dodicesima indagine regionale. Rapporto 2012*. Fondazione ISMU, Regione Lombardia, Osservatorio Regionale per l’integrazione e la multietnicità, Éupolis Lombardia, Milano.
- FATTORE M., BRUEGGEMANN R., OWSIŃSKI J. 2011. Usig poset theory to compare fuzzy multidimensional material deprivation across region. In INGRASSIA S., ROCCI R., VICHI M. (Eds.) *New Perspectives in Statistical Modeling and Data Analysis*, Springer-Verlag.

SUMMARY

A new methodology to measure the level of integration of the foreign population in the italian labour market

The aim of this paper is to present a new approach to measure and monitor the integration of immigrants into the labour market, based on partial order theory and previously tested for analysis of multidimensional material deprivation data. The main feature of the methodology is that the information needed for the integration assessment is extracted directly from the relational structure of the dataset, avoiding any kind of scaling and aggregation procedure the study of material deprivation, so as to respect the ordinal nature of the data.

Alberto ARCAGNI, Università di Milano-Bicocca, alberto.arcagni@unimib.it

Gian Carlo BLANGIARDO, Università di Milano-Bicocca,
giancarlo.blangiardo@unimib.it

Marco FATTORE, Università di Milano-Bicocca, marco.fattore@unimib.it

Simona Maria MIRABELLI, Università di Milano-Bicocca,
simona.mirabelli@unimib.it

INNOVAZIONI DI PROCESSO E USO DELLE VARIABILI TESTUALI: IL CASO DELL'ADULT EDUCATION SURVEY¹

Barbara Baldazzi, Alessandro Bianchi, Anna Emilia Martino, Paola Paladini

1. L'Adult Education Survey

L'indagine europea sulla partecipazione degli adulti alle attività formative (Adult Educationa Survey – AES) ha l'obiettivo di fornire dati confrontabili a livello europeo sulla partecipazione degli adulti ad attività di istruzione e formazione e dati di qualità a livello nazionale per permettere politiche efficaci di aggiornamento e riqualificazione del capitale umano. L'indagine viene condotta nello stesso anno in tutti i paesi dell'Unione ed è disciplinata da un apposito regolamento Eurostat.

L'indagine definitiva si è svolta tra settembre e dicembre del 2012 su un campione di circa 6.000 famiglie². Le interviste sono state effettuate con tecnica CAPI (Computer Assisted Personal Interviewing) e tecnica CATI (Computer Assisted Telephone Interviewing). L'Istat ha cercato di migliorare il contatto con le famiglie dei rispondenti, soprattutto nelle indagini sociali (Sabbadini et al., 2000): in questa indagine l'approccio alla famiglia è stato innovativo. La necessità di svolgere un notevole numero di interviste con tecnica CATI e la difficoltà sempre più crescente di avere a disposizione una lista di recapiti telefonici esaustiva, hanno portato all'individuazione di un sistema di “raccolta” dei recapiti telefonici delle famiglie nuovo ed efficace. Parallelamente è stato necessario sviluppare le classificazioni necessarie per il corretto rilevamento delle informazioni: in particolare è stata implementata la classificazione internazionale sugli argomenti della formazione.

Riassumendo le fasi del processo di indagine sono state:

1. Scelta del campione rappresentativo per territorio
2. Invio delle lettere alle famiglie (prima lettera e successivi solleciti)

¹ Questo lavoro nasce dalla riflessione condivisa degli autori. In particolare a B. Baldazzi si deve la redazione dei paragrafi 1 e 4, a A. Bianchi il paragrafo 3, a A. E. Martino il paragrafo 5 ed a P. Paladini il paragrafo 2, mentre le conclusioni sono condivise.

² L'indagine pilota si è svolta nel periodo ottobre-dicembre 2011 intervistando circa 1000 famiglie.

3. Allestimento del sistema web, del sistema IVR (Interactive Voice Response) e del numero verde per la raccolta dei recapiti telefonici
4. Analisi dei risultati del processo di reperimento recapiti telefonici
5. Preparazione delle liste di campionamento per il CAPI e per il CATI
6. Interviste sul campo – Implementazione di un navigatore degli argomenti delle attività di formazione

2. L'utilizzo delle LAC e l'invio delle lettere alle famiglie

La base informativa di riferimento per l'estrazione delle unità di rilevazione (le famiglie) e per il monitoraggio e la gestione della rilevazione sul campo è rappresentata dalle Liste Anagrafiche Comunali (LAC), che l'Istat acquisisce da tutti i Comuni italiani. Ogni lista è costituita dall'insieme degli individui che hanno residenza in un determinato Comune e tengono conto sia delle famiglie che delle convivenze, e sia dei cittadini italiani che stranieri.

Le famiglie che compongono il campione sono state estratte in circa 460 Comuni rappresentativi delle cinque ripartizioni geografiche (Italia nord-occidentale, Italia nord-orientale, Italia centrale, Italia meridionale e Italia insulare) e delle quattro tipologie comunali suddivise secondo un criterio che considera sostanzialmente l'ampiezza demografica del comune stesso (Comune centro dell'area metropolitana, Comuni della periferia dell'area metropolitana, Comuni fino a 10.000 abitanti, Comuni con più di 10.000 abitanti).

Il sistema di raccolta dei recapiti telefonici delle famiglie estratte si è composto di due metodi: 1) l'uso di un software di abbinamento agli elenchi telefonici dei nominativi delle famiglie, 2) l'invio di una lettera informativa a firma del Presidente. Nella lettera veniva richiesta espressamente la partecipazione telefonica all'intervista e veniva data la possibilità di indicare il recapito telefonico dove essere contattati per l'intervista. E' stato, inoltre, previsto un piano di solleciti per quelle famiglie che non avevano contattato l'Istat dopo la ricezione della prima lettera. La sostanziale novità inserita nella lettera è data dalla frase seguente: *“Qualora Lei desiderasse evitare il disturbo di ricevere persone in casa può chiedere che l'intervista avvenga per telefono. In tal caso chiami subito l'apposito Numero Verde (gratuito) per fornire a un operatore o a un sistema di registrazione automatica il Codice Familiare Istat posto in alto a destra e il suo numero di telefono fisso o il suo numero di telefono cellulare che saranno tenuti assolutamente riservati, presso i quali contattare Lei e i suoi familiari. In alternativa potrà accedere al sito...”*. I successivi due solleciti, di forma più breve e più essenziali, ribadivano la possibilità di effettuare l'intervista telefonica e segnalavano i tre sistemi per fornire il recapito telefonico.

Per la mancata consegna della lettera inviata all'indirizzo fornito dalle LAC abbiamo previsto e realizzato una banca dati per registrare le lettere ritornate al mittente. Nella banca dati delle lettere ritornate al mittente, la cui base informativa era costituita da tutte le famiglie facenti parte del campione, venivano memorizzate le seguenti informazioni: riguardo al destinatario se questi era 1) irreperibile, 2) sconosciuto, 3) trasferito, 4) deceduto, e riguardo all'indirizzo se questo era 1) inesatto, 2) insufficiente, 3) inesistente. Infine abbiamo registrato anche il caso di rifiuto, da parte del destinatario, della lettera.

Il monitoraggio conclusivo della banca dati mostra come soltanto lo 0.4% di famiglie non hanno ricevuto tutte e 3 le lettere, l'1.4% di famiglie non hanno ricevuto 2 lettere, e il 2.6% di famiglie non hanno ricevuto una delle 3 lettere, a ribadire come la tempestività negli aggiornamenti anagrafici delle LAC porti ad una presenza di indirizzi errati molto bassa.

3. Il sistema tri-ripartito di raccolta dei recapiti telefonici: Numero Verde, IVR e Portale Web

Tre sono stati i mezzi attraverso i quali la famiglia ha potuto mettersi in contatto con l'Istat per fornire il proprio recapito telefonico: un numero verde, un servizio interattivo di risposta vocale IVR ed un portale Web. Il sistema fa convergere in un Data Base integrato tutte le informazioni raccolte dalle tre fonti diverse e le rende disponibili all'Istat.

L'accesso al data base permette di monitorare costantemente ed in tempo reale l'incremento dei numeri telefonici rilasciati dai rispondenti, consente la distinzione tra le tre tipologie di acquisizione e permette di costruire i relativi indicatori.

Il Codice Familiare Istat univoco, presente nella lettera informativa inviata alle famiglie, è la chiave che permette ai tre sistemi integrati la corretta associazione con il Cognome dell'intestatario della scheda familiare e quindi l'accertamento dell'autenticità del rispondente.

Il servizio di Numero Verde, naturalmente gratuito ed attivo in orario lavorativo dei giorni feriali per l'intera durata delle operazioni di raccolta dei recapiti telefonici, permette alle famiglie facenti parte del campione di parlare con un operatore, che è a disposizione per fornire informazioni sull'indagine. L'operatore, previa verifica della corrispondenza tra Codice familiare e Cognome fornito dal rispondente, è abilitato ad acquisire i recapiti telefonici (massimo 3) ai quali la famiglia vuole essere contattata, registrandoli direttamente nel sistema informatico.

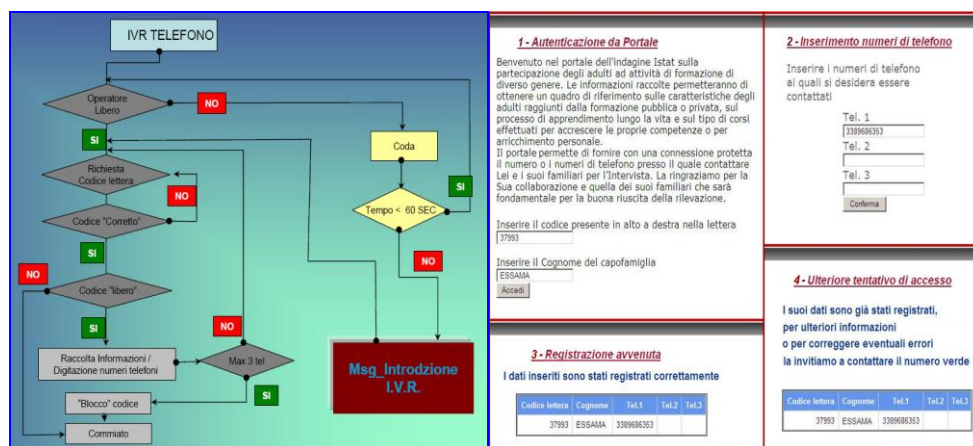
Il sistema telefonico/informatico IVR acquisisce i recapiti telefonici attraverso un impianto fondato su di un sistema interattivo di risposta vocale a toni (Dtmf) operante 24 ore al giorno per tutto il periodo di raccolta delle informazioni. Il

servizio di registrazione automatica è operante solo negli orari e nei giorni in cui il numero verde è inattivo o nei momenti in cui tutti gli operatori sono occupati.

Il rispondente, dopo l'ascolto di un breve messaggio registrato che cita l'Istat e lo scopo dell'indagine, è invitato tramite tastiera telefonica ad autenticarsi con l'inserimento del codice familiare univoco di cinque caratteri numerici. Se l'autenticazione avviene con successo, il rispondente è abilitato a comunicare un massimo di tre recapiti telefonici della propria famiglia, attraverso i quali potrà essere contattato per il rilascio dell'intervista. (Figura 1)

L'Ivr è collegato allo stesso numero verde e la componente applicativa è integrata nel portale Web dedicato.

Figura 1 – Diagramma di flusso IVR e **Figura 2** – Sintesi delle pagine web proposte Numero Verde



Il portale web utilizza un protocollo di sicurezza "https" ed è sempre attivo durante le operazioni di raccolta delle informazioni per la ricezione dei recapiti telefonici delle famiglie campione che hanno a disposizione un accesso ad internet e scelgono di fornire il proprio numero attraverso la rete.

Le pagine web predisposte (Figura 2), propongono dapprima, analogamente agli altri due sistemi di raccolta dei recapiti, una descrizione dell'Istat, dell'indagine e una fase di autenticazione della famiglia attraverso la digitazione in appositi campi del Cognome e del Codice familiare; infine l'inserimento di massimo tre recapiti telefonici negli appositi spazi compilabili. Al termine dell'inserimento dei recapiti viene data una conferma della corretta registrazione.

Il sistema prevede l'inibizione ad un nuovo accesso, sia da portale che da IVR, quando una famiglia ha comunicato uno o più recapiti telefonici, in quanto l'avvenuta associazione del Codice famiglia al Cognome nella fase di

autenticazione e l'esito positivo della procedura completa, prevedono il blocco del codice familiare usato. Eventuali modifiche, aggiunte o sostituzioni dei recapiti forniti possono essere effettuate solo attraverso un operatore del Numero Verde che ha la possibilità di modificare manualmente il Data Base.

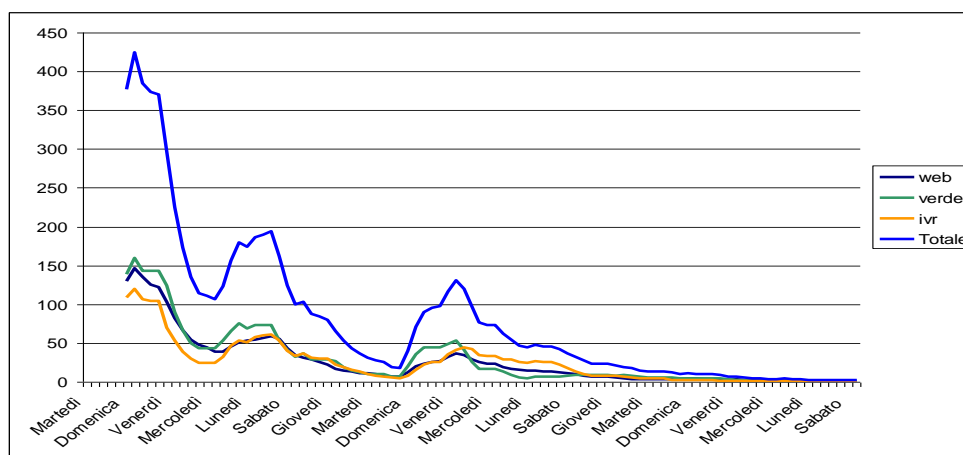
4. I risultati del processo di raccolta recapiti telefonici: preparazione delle liste per le interviste sul campo

Il sistema di raccolta dei numeri di telefono, grazie all'accurata formulazione delle lettere e alla trivariata possibilità per i rispondenti di contattarci, ha funzionato ottimamente.

Tabella 1 – Famiglie che forniscono il loro recapito telefonico rispetto al metodo usato e l'invio della lettera e come incidenza percentuale sul totale delle famiglie.

Lettere	Web	Numero verde	Sistema IVR	% di famiglie
Prima lettera	36,1	37,4	26,5	16,8
Seconda lettera	29,9	37,6	32,5	8,8
Terza lettera	29,4	33,0	37,5	7,7
Totale famiglie	32,9	36,4	30,6	33,3

Figura 3 – Medie mobili settimanali – Famiglie che forniscono i recapiti telefonici in totale e secondo i tre sistemi utilizzati.



Le liste di campionamento sono state, quindi, tre: due liste CATI sovra campionate basandoci sui tassi di risposta avuti nella Indagine Pilota e una lista CAPI composta da una famiglia base e tre sostitute.

Dopo il primo invio il 16,8% di famiglie aveva fornito il proprio recapito (Tabella 1); dopo il secondo invio il totale delle famiglie era del 25,6% per arrivare con il terzo invio al 33,3% di famiglie. La Figura 2 evidenzia, attraverso le medie mobili settimanali calcolate sulle famiglie che hanno fornito il proprio numero di telefono, come il primo invio abbia un effetto sostanzialmente maggiore e come il secondo e il terzo invio inneschino una azione virtuosa di sollecito sui rispondenti³.

La qualità dei recapiti telefonici forniti è stata comprovata in sede di intervista risultando ottima: soltanto il 2,9% di telefoni ricevuti dalle famiglie e contattati è risultato inesistente; ben il 50,9% dei recapiti telefonici erano utenze di cellulari.

La preparazione delle liste di campionamento ha ricalcato la struttura emersa dal sistema di raccolta dei recapiti telefonici: il 33,3% delle famiglie ha comunicato il recapito telefonico; il 34,9% delle famiglie ha avuto abbinato il recapito telefonico attraverso l'uso di un software di abbinamento agli elenchi telefonici; il 31,8% delle famiglie non ha il recapito telefonico abbinato.

5. L'utilizzo della classificazione FOET nella codifica delle attività formative

Le direttive dell'Unione europea hanno indicato la necessità di raccogliere dati statistici dettagliati e di alta qualità sulla partecipazione degli adulti alle attività formative. Si è quindi resa necessaria l'implementazione di una classificazione condivisa per la raccolta di dati comparabili nei paesi in cui è prevista l'indagine AES. La classificazione internazionale FOET (Fields Of Education and Training) ha ovviato a quest'esigenza permettendo l'armonizzazione delle indagini a livello internazionale. E' stata introdotta in Italia a partire dalla corrente edizione, ed utilizzata per la codifica degli argomenti dei corsi di formazione e di istruzione. La classificazione è gerarchica ed è stata costruita a partire dalla classificazione ISCED del 97. E' composta da: 9 Campi (Broad Fields) ad 1 digit, 25 Sottocampi (Narrow Fields) a 2 digit, 77 Voci (Detailed Fields) a 3 digit.

I 9 Campi che raggruppano le attività formative sono: Campo 0 – “Programmi generali”; Campo 1 – “Istruzione”; Campo 2 – “Arte e letteratura”; Campo 3 – “Scienze sociali, affari e diritto”; Campo 4 – “Scienze, matematica e informatica”; Campo 5 – “Ingegneria, produzione e costruzioni”; Campo 6 – “Agricoltura, silvicoltura e pesca”; Campo 7 – “Salute e welfare”; Campo 8 – “Servizi”.

³ Nell'indagine pilota è stato inviata anche una quarta lettera che non ha dato risultati apprezzabili

A livello internazionale sono i 25 Sottocampi, previsti dalla classificazione ISCED97, ad essere comparati; il terzo digit è utilizzato principalmente a livello nazionale, permettendo ulteriori aggregazioni e rendendo il sistema più flessibile ad indagini finalizzate e ricerche ad hoc. Inoltre, le Voci a 3 digit permettono una più facile comprensione dell'attività da codificare, guidando verso una migliore codifica dei Sottocampi a 2 digit.

Per una corretta attribuzione dei codici da parte dei rilevatori è stato implementato un navigatore degli argomenti con chiavi testuali (parole e segmenti chiave). Questo sistema di navigazione è stato costruito raccogliendo le attività di formazione nelle varie edizioni delle indagini e sintetizzandole in un dizionario definitivo. Le fasi per la costruzione del sistema sono state le seguenti:

1. circa 10.000 esempi di argomenti sono stati registrati su supporto informatico desunti dall'Indagine Multiscopo Istat del 2006 "I cittadini e il tempo libero" che aveva previsto una domanda aperta sul questionario cartaceo dove scrivere l'argomento dell'attività di formazione;
2. circa 8.000 esempi di argomenti grazie ad una prima scrematura effettuata accorpando stringhe simili e di uguale significato;
3. nell'Indagine Pilota AES del 2011 è stato testato il navigatore di ricerca tramite stringa o parola chiave sulla lista di 8.000 esempi; in questa fase è stato possibile effettuare la codifica di "attività nuove";
4. revisione completa della lista con seconda scrematura delle ridondanze ed inserimento di argomenti di attività nuove emerse dalla pilota; aggiornamento delle voci con inserimento di acronimi, sigle e argomenti nuovi;
5. stesura del dizionario per la indagine definitiva composto da circa 5.000 esempi di argomenti.

Dal controllo effettuato a fine indagine, confrontando stringhe e codici è emerso l'8% di codici errati; per la gran parte codici di Campo 0 "Programmi generali" ricodificati in sede di correzione dati. Gli errori sono stati più frequenti nelle interviste CAPI; nella modalità CATI la presenza dei supervisori e la possibilità di effettuare la codifica a fine intervista ha limitato gli errori. In definitiva, l'uso della classificazione FOET e del navigatore degli argomenti ha permesso un monitoraggio ed una successiva correzione dell'errore di classificazione, oltre ad una maggiore standardizzazione dei risultati, comparabili a livello internazionale.

6. Conclusioni

Nell'ottica di migliorare la qualità dei dati (Eurostat, 2007, Cuppen et al, 2011) in questa indagine sono state sviluppate alcune innovazioni e miglioramenti nel processo di raccolta dati. Emergono alcuni punti di forza: l'utilizzo delle LAC per

l'estrazione del campione con un miglioramento delle liste di nominativi; la strategia di contatto con le famiglie che ha avuto come risposta un'ampia partecipazione delle famiglie al sistema di raccolta recapiti telefonici e anche all'indagine; il notevole risparmio economico con un maggior utilizzo delle interviste CATI piuttosto che CAPI; la creazione e l'uso, da parte degli intervistatori, del navigatore degli argomenti delle attività di formazione con esempi calzanti alla realtà. Alcune criticità sono: l'eccessiva gravosità del sistema di invio lettere; la dilatazione dei tempi di contatto con la famiglia; la necessità di avere, a priori, i tassi di risposta per effettuare il sovracampionamento.

Riferimenti bibliografici

- CUPPEN M., VAN DER LAAN P., VAN NUNSPEET W. 2011. Re-engineering Dutch social surveys: From single-purpose surveys to an integrated design, presented at World Statistics Congress of the ISI.
- EUROSTAT 2007. Handbook on Data Quality Assessment: Methods and Tools
- SABBADINI L.L., CAMPORESE R., QUATTROCIOCCHI L. 2000. Le Strategie di prevenzione delle mancate risposte e degli errori non campionari nelle indagini multiscopo, *Rivista Italiana di Economia Demografia e Statistica*, Vol. LIV, No. 2, pp. 165-180.

SUMMARY

Process Innovation and use of textual variables: the case of AES Survey

We present the process of AES survey: 1) choice of the representative sample with LAC, 2) dispatch of letters to households, 3) Setting up the web-based system, the IVR system and toll-free for the collection of phone numbers, 4) Analysis of results of collection of phone numbers, 5) preparation of lists CATI/CAPI of sampling, 6) use of FOET classification to codify fields of learning activities.

Barbara BALDAZZI, Istat, baldazzi@istat.it
Alessandro BIANCHI, Istat, abianchi@istat.it
Anna Emilia MARTINO, Istat, annmartin@istat.it
Paola PALADINI, Istat, papaladi@istat.it

SHOULD I STAY OR SHOULD I GO? THE CASE OF ITALY¹

Elisa Barbiano di Belgiojoso, Livia Elisa Ortensi

1. Introduction

Recent official statistics have indicated that the ongoing economic crisis has slowed down inflows into Italy. In addition, nearly one million foreigners included in the Population Register until 2011 were missing at the last census, and have presumably left the country (Istat, 2012; Blangiardo, 2012). Such a scenario suggests that re-emigrations² have become an Italian issue as well.

There is no clear consensus among scholars on the causes for re-emigration, and to the best of our knowledge, there are few studies about return migration as regards Italy. However, some investigations into migrant assimilation in the labour market (Strom, Venturini and Villosio, 2013; Dell'Aringa and Pagani, 2011) have found a selection process in the decision to return: migrants with higher wages or higher experience are more likely to leave Italy, and according to Coniglio et al. (2005), this finding can be extended to include irregular migrants.

Given this background, we make the following research hypotheses about the possible determinants of the new waves of re-emigration from Italy.

1. First, we speculate that there may be differences between onward migration flows and return flows to the countries of origin (e.g. Nekby, 2006). In fact, while the decision to return to the country of origin marks the conclusion of the migration project, onward migration is more likely to represent a quest for better opportunities outside Italy.
2. Following the New Economics of migration theory, we hypothesize that the family has a central role in decisions about migration (as stated by Dustmann, 2003). As a consequence, our models include information about the possible presence of family members in emigration.
3. We also formulate the hypothesis that, except for forced migrants (such as asylum seekers or refugees), migration to Italy is still for the most part

¹ Although they share responsibility for the work, sections 1 and 3 were written by Elisa Barbiano di Belgiojoso and paragraphs 2 and 4 by Livia Elisa Ortensi.

² Re-emigrations include both emigration to the country of origin and to a third country (onward migration).

economically driven. We expect covariates about employment to be highly related to future intentions regarding mobility.

4. Lastly, we speculate that the economic crisis plays an important role in the intention to leave the host country. We therefore expect the likelihood of leaving Italy to increase in the more recent waves of the surveys included in the analysis, in line with the economic cycle.

2. Data and Methods

The data for this analysis come from the Italian ORIM survey on immigrants in Lombardy, carried out routinely by the Lombardy Regional Observatory for Integration and Multiethnicity. The figures are based on an annual cross-sectional survey consisting of face-to-face interviews carried out using the Centre Sampling statistical procedure (Baio et al., 2011). These surveys are of particular interest as they also include information on irregular migrants. In our analysis, we used a pooled dataset of the last three waves of the survey (2010-2012) in which information about future intentions of mobility was available. The final sample consists of 22,399 migrants aged 14 and over from heavy pressure emigration countries, and includes undocumented migrants, naturalized citizens and second-generation migrants.

For our specific purposes, these data have two limitations, one connected to study design and the other to the fact that, although a dedicated question about return intention was included in each wave, these surveys were not conceived for the sole purpose of studying migrants' intentions about returning to their home country or about secondary migration. The first kind of limitation is the selection effect produced by the use of a retrospective survey, since data was obtained only from migrants who had not returned to their countries of origin or moved to other countries up to the time of interview. This bias is likely to be higher for long-term migrants who are a selected sub-population of survivors of secondary migration. The second limitation is the lack of crucial information about the family's migration history and economic condition, since detailed information is only available regarding the person interviewed. We speculate that, especially for family migrants, information about the year of arrival and the breadwinner's work situation would provide a clearer explanation of their return intentions. For this reason, only models for female workers are given here. In fact, housewives decisions about their future in emigration are also a consequence of the male breadwinner's work situation, and since this information is not available, housewives might bias the female model. As the analysis is based on cross-sectional data, the direction of causality has to be explained with caution.

Finally, we consider the respondent's intention to leave Italy in the 12 months following the survey as a proxy of re-emigration. Of course, detailed data on true re-migrants would be preferable. However, in our view, the specification of a short, well-defined, fixed term is of more use in helping us to identify real future emigrants. On the positive side, these data represent a precious, up-to-date source of information in the Southern European context about determinants of re-emigration during the current economic crisis. The inclusion of second-generation migrants, undocumented migrants and overstayers, along with the large size of the sample, makes it an invaluable tool for trying to work out migrants' re-emigration intentions. As our dataset has a hierarchical structure, with migrants (level 1) nested in communities (level 2), we used a multilevel approach. Due to this structure, the odds of experiencing the outcome of interest are not independent, because migrants from the same country of origin share a common exposure to observed and unobserved community characteristics. We think that the multilevel approach is always to be preferred when dealing with datasets like ours, which include information on migrants from different communities, and we therefore chose a generalized linear mixed-model approach, fitting a two-level random intercept logistic regression. This model accounts for the non-independence of observations within groups. We fitted different models for the intention to re-emigrate to another country (model 1) and for the intention to return to the respondent's country of origin (model 2), taking as our baseline the intention to stay in Italy. We also fitted different models for men and women; for the latter, we provided a model only for those women who said they were not housewives.

3. Results

How widespread is the intention to leave Italy among migrants? Nine out of ten say that they intend to stay in the host country for at least one more year. Among those who wish to leave, there is a slight preference for returning home, which increases over time. Both models 1.1 and 2.1 indicate that women have lower odds of expressing the intention to leave Italy compared to men, and the odds of expressing the intention to re-emigrate are also considerably lower (0.38).

In line with hypothesis 1, the two flows do indeed show differences in their driving factors. Onward migration would appear to be related more to a search for better opportunities on the part of graduate migrants, and these are primarily a solution to unemployment: the odds of expressing the intention to move to another foreign country is more than twice as high among those who lost their job in the year before the interview and among the long-term unemployed. In the case of working women, having a job which is considered as a niche for foreigners (such

as working in a family as a housemaid, baby sitter or caregiver) reduces the odds of re-emigration.

Table 1 – *Model 1: two-level random intercept logistic regression move to a third country (baseline intention to stay in Italy) Odds Ratios and significance.*

	Model 1.1 ALL		Model 1.2 MALE		Model 1.3 FEMALE		Model 1.4: FEMALE (no housewives)	
	OR	sign.	OR	sign.	OR	sign.	OR	sign.
Female [ref. male]	0,3805	**						
Age	0,9811	**	0,9862		0,9741	*	0,9886	
Partner citizenship [ref. The same]								
Italian	1,0049		1,3395		0,7361		0,8536	
Other citizenship	2,1604	***	2,3383	***	1,8630		1,6264	
No partner	1,3383	*	1,6246	**	0,8714		0,7283	
Educational level [ref. none]								
Compulsory	0,8657		1,1110		0,4722	*	0,3891	*
High school	0,9530		1,2217		0,5211	*	0,5184	
University degree	1,7841	**	2,4039	**	0,8995		0,6635	
Homeowner [ref. not]	0,6973	**	0,7083	*	0,6815		0,6504	
Employed	0,8926		0,8347		1,2700		0,5003	*
Loss of the work in the last year	2,8059	***	2,9074	***	2,2605	*	0,8334	
Long-term unemployed	2,6032	***	2,5121	***	3,0336	**	0,4227	
Cohabitation with at least a family member [ref. no]	0,7399	*	0,7184	*	0,8344		0,8042	
Possibility of stay [ref. temporary]								
Permanent	1,3671	**	1,3262	*	1,4950	*	1,5916	
No valid permit	1,8148	**	1,8633	**	1,2872		1,4334	
Children [ref. All in Italy]								
All in the country of origin	1,4635	**	1,4652	*	2,1669	*	2,4221	*
No children	1,3248	*	1,2429		1,5116		1,6900	
Some in the country of origin some in Italy	1,8190	**	1,6790	*	2,1915	*	2,0899	
Ethnic niche	1,0065		1,1137		0,6462		0,6087	*
Year of the survey [ref. 2010]								
2011	1,2749		1,3327		1,1657		1,0871	
2012	1,7473	***	1,7317	***	1,7659	*	1,3598	
Savings	0,7629	**	0,7969	*	0,6882	*	0,5220	**
Constant	0,1920	***	0,0474	***	0,0546	**	0,1241	***
σ_u	0,5028		0,5066		0,5972		0,6962	
ρ	0,0714	***	0,0724	***	0,0978	***	0,1284	***

Note: Significance ***<0.001; **<0.01; *<0.05

The type of residence permit held is also significant: undocumented migrants are more likely to express the intention to migrate, but so are those with a legal status allowing permanent residence (EU or naturalized citizens, holders of long-term EC residence permits). The latter in particular are allowed to travel and work legally in any EU country without restriction. Self-reported ability to set aside monthly savings, as well as house ownership, is also related to lower odds of onward migration.

Those who have all or some of their children in the country of origin are more likely to express the intention to re-emigrate: in this sense, having a transnational family indicates the need to continue the experience of emigration, given the fact

that part of the family has not yet migrated. The same relation is observed in those who are not cohabiting with any family member. Although we obviously have no information about non-cohabiting family members, we can make the hypothesis that when the partner and small children are not cohabiting, they are highly likely not to be present in Italy, being a proxy for transnational family ties.

Table 2 – *Model 2: two-level random intercept logistic regression return home (baseline to stay in Italy) Odds Ratios and significance.*

	Model 2.1 ALL		Model 2.2 MALE		Model 2.3 FEMALE		Model 2.4: FEMALE without housewife	
	OR	sign.	OR	sign.	OR	sign.	OR	sign.
Female [ref. male]	0,8159	*
Age	1,0289	***	1,0339	***	1,0247	***	1,0299	***
Partner citizenship [ref. The same]								
Italian	0,7099		1,0712		0,6608		0,6395	
Other citizenship	1,1733		0,7772		1,6998	*	1,8856	*
No partner	1,2796		1,4424		1,0783		1,0190	
Educational level [ref. none]								
Compulsory	0,5797	***	0,5630	**	0,5760	*	0,5011	*
High school	0,5907	***	0,5914	**	0,5740	*	0,5538	*
University degree	0,5761	**	0,5669	*	0,5830	*	0,4553	*
Homeowner [ref. not]	0,8050		0,9860		0,6278	*	0,6855	
Employed	0,5120	***	0,4801	***	0,6516	*	0,4048	**
Loss of the work in the last year	1,2925		1,3859		1,3249		0,7869	
Long-term unemployed	1,5216	**	0,3425		1,7629	*	0,3363	
Cohabitation with at least a family member [ref. no]	0,6240	***	0,6323	**	0,5667	**	0,5412	**
Possibility of stay [ref. temporary]								
Permanent	1,0510		1,2088		0,9055		0,7601	
No valid permit	1,6413	**	1,5511	*	1,7395	*	1,7179	*
Children [ref. All in Italy]								
All in the country of origin	2,1358	***	2,2476	***	2,4000	***	2,2462	***
No children	1,1709		1,4553	*	0,8843		0,8560	
Some in the country of origin some in Italy	1,9605	***	1,6235	*	2,3896	***	1,8282	*
Ethnic niche	1,2882	**	1,4442	**	0,9731		0,9324	
Year of the survey [ref. 2010]								
2011	1,2131		1,3668	*	1,0458		1,1602	
2012	1,4207	**	1,3838	*	1,4692	*	1,5163	*
Savings	0,7731	**	0,7889	*	0,7346	*	0,6487	**
Constant	0,0488	***	0,0272	***	0,0522	***	0,0935	***
σ_u	0,5157		0,44161		0,4441		0,4840	
ρ	0,0747	***	0,05596	***	0,0565	***	0,0664	***

Note: Significance ***<0.001; **<0.01; *<0.05

Following the worsening of the economic situation, a stronger intention of onward migration is observed in migrants interviewed in 2012 compared to 2010.

Model 2 deals with intention to return to the country of origin. Education plays a significant role, indicating that the odds of non-educated migrants returning are greater than that of all migrants with a higher level of education. Older migrants are also more likely to express their intention to end the migration experience.

The role of the family is the same as that observed in the case of re-emigration intentions, but children left behind are a greater draw for women than for men. This is not surprising, since the perceived emotional costs of transnational families are higher for women. On the economic and legal side, intention to return to the country of origin is related to more precarious situations such as long-term unemployment and not having a valid permit.

The role of the ethnic niche is also interesting: men employed in sectors considered to be male ethnic niches (construction industry, agriculture, factory work and transportation) are considerably more likely to express the intention to return to their country of origin, while this relation is not observed in women. This confirms the findings of Paggiaro (2013), who noted that male immigrant workers are at greater risk of unemployment because their ethnic niches are the economic sectors most badly hit by the economic downturn. Female niches, on the other hand, are comparatively less affected by the crisis, as the need for care services provided by foreign women is driven by demographic factors and shortcomings in welfare. As with future onward migration intentions, a higher intention of return to the country of origin is observed in migrants interviewed in 2012 as compared to 2010, following the worsening of the economic situation, but for men the odds are significantly higher for the 2011 wave as well.

4. Discussion

Our work confirmed our initial hypotheses. The first interesting finding is that the two flows are different. The odds of onward migration are higher for male and graduate migrants and are higher both for the long and short term unemployed. Re-emigration is also favored by legal conditions such as EU citizenship and a long-term EC residence permit making it possible to work in another EU member state. Flows toward countries of origin, on the other hand, are less closely related to unemployment: only in cases of long-term female unemployment are the odds of expressing the intention to return home significantly higher.

Unlike the situation for onward migration flows, migrants with no education are considerably more likely to express the intention to return home: in this sense, if onward migration appears to be a quest for better conditions on the part of the migrants with the best potential, concluding the migration experience would seem to be related to a giving up on the part of those individuals with more limited capabilities. In fact, the odds of giving up and going home were higher among those who were working or had previously worked in male niches characterized by low-skilled jobs. In line with hypothesis 2, the role of the family network is significant (as stated by Dustmann, 2003). When children and family members live together in emigration, demonstrating the existence of a settlement project

involving more than one family member, the odds of moving are lower. Transnational families act both as push and pull factors. For those who choose to continue migrating, giving financial support to those left in the country of origin might be the reason for seeking better opportunities outside Italy, while those who choose to go home have the chance to reunite the family after migration. As we predicted in hypothesis 3, economic conditions prove determinant: in fact, the ability to save money is a significant factor in reducing the intention to move, and proved even more important than unemployment, being an indicator at the family level, especially in determining the decision to end the migration experience. Finally, in line with hypothesis 4, the odds of moving are indeed significantly higher in 2012 than 2010, and a clear relation emerges with the worsening of the Italian economic situation. For males, the odds of returning are significantly higher for 2011 as well.

In conclusion, a dual selection process may be observed, based on gender and human capital. As a consequence of the persisting economic downturn, men and women with less human capital might be more likely to give up on, or conclude, their migration project, being less well-equipped to tackle the worsening job market and the limited opportunities in a segmented Italian job market which is particularly hard on men. At the same time, highly educated men might be more likely to leave Italy, following the same pattern as educated young Italian citizens, as Nekby (2006) found in Sweden. Something similar was observed with regard to legal status. Not surprisingly, undocumented migrants have higher odds of having the intention both to emigrate to a third country and to return home. If migrants actually behave in accordance with their short-term intentions, we should expect a reduction in illegal migration. But just as for education, men with a greater possibility of free circulation in EU member states express their intention to use this possibility to re-emigrate elsewhere. Interestingly, such intentions are not expressed by women with the same education level or legal status. Women's lesser propensity to re-emigrate is interesting and worthy of further analysis. One possible explanation could be a higher perceived cost of re-emigration, and giving up the idea of migration because of the relatively better opportunities for women in Italy. A stronger commitment to the family and the higher costs (also psychological) of first migration could also determine women's greater reluctance to change the original migration project.

Interestingly, in our analysis, the number of years elapsing since migration is not significant, indicating that the mere length of time since migration does not automatically determine a higher or lower propensity to settle in Italy. Instead, a crucial role is played by the state of the economy, the job situation, and the presence of family ties in emigration, along with the migrant's personal capabilities.

References

- BAIO G., BLANGIARDO G.C., BLANGIARDO M. 2011. Centre Sampling Technique in Foreign Migration Surveys: a Methodological Note, *Journal of Official Statistics*, vol. 27, No.3, pp.451-465.
- BLANGIARDO G.C. 2012. Se tre indizi (statistici) possono fare una prova, *Il Sole 24 Ore*, 21 May 2012.
- CONIGLIO, N., DE ARCANGELIS G., SERLENGA L. 2005. *Return Intentions of Irregular Migrants and Skill: the Perverse Effect of the Brain Waste*. Paper presented at the Annual Congress of the Italian Association of Labour Economists (AIEL). Rome, September 22-23 2005.
- DELL'ARINGA C., PAGANI L. 2011. Labour Market Assimilation and Over-Education: The Case of Immigrant Workers in Italy, *Economia Politica*, vol. 2, pp. 219-240.
- DUSTMANN C. 2003. Children and return migration, *Journal of Population Economics*, vol. 16, pp. 815-830.
- ISTAT, 2012. Migrazioni internazionali e interne della popolazione residente, Statistiche Report.
- NEKBY L. 2006. The emigration of immigrants, return vs onward migration: evidence from Sweden, *Journal of Population Economics*, vol. 19, pp. 197-226.
- PAGGIARO A. 2013. How do immigrants fare during the downturn? Evidence from matching comparable natives, *Demographic Research*, Vol.28, No. 8, pp. 229-258.
- STROM S., VENTURINI A., VILLOSIO C. 2013. *Wage assimilation: migrants versus natives and foreign migrants versus internal migrants*, European University Institute Working Papers 30.

SUMMARY

In this paper we analyse re-emigration intentions using a pooled dataset on a sample of more than 22,000 migrants. The results show that while onward migration appears to be a quest for better conditions on the part of migrants with the best potential, going home seems to be related to giving up on the part of individuals with lower capabilities. Family network, gender, ethnic niche, legal status and the economic crisis also emerged as having a significant role in the decision process.

Elisa BARBIANO DI BELGIOJOSO, Ricercatore di Demografia, Università degli Studi di Milano - Bicocca, elisa.barbiano@unimib.it

Livia Elisa ORTENSI, Assegnista di Demografia, Università degli Studi di Milano - Bicocca, livia.ortensi@statistica.unimib.it

DEBITI E CREDITI DEMOGRAFICI NEI PAESI EUROPEI¹

Elisa Barbiano di Belgiojoso, Stefania Rimoldi

1 Introduzione

Il potenziale demografico di un paese in un certo istante consiste nell'ammontare di anni di vita residua, stimata sulla base dei profili di sopravvivenza degli individui che ne fanno parte. Le diverse popolazioni si distinguono per la quantità di potenziale demografico di cui dispongono, per come esso si forma nel tempo e, soprattutto, per la sua struttura. Le poste aggiuntive derivano dalle nascite e dal saldo migratorio netto (che può nondimeno essere negativo) mentre le componenti detrattive si riferiscono al semplice consumo di tempo vissuto e ai decessi². Pertanto, l'equilibrio tra le diverse componenti di variazione determina la solidità/fragilità dell'ammontare di futuro che ciascuna popolazione possiede. Elementi di differenziazione qualitativa tra i potenziali demografici derivano, invece, dal rapporto tra le quote di futuro destinate alla formazione e al pensionamento (che in un ipotetico bilancio rappresentano il potenziale "debito"), e la quota di futuro destinata al lavoro (il potenziale "credito"). Il presente contributo si propone di esaminare i processi di formazione del potenziale demografico e le sue caratteristiche strutturali, relativamente alle popolazioni dei (27) paesi dell'Unione Europea e con riferimento al periodo 1.1.2007-1.1.2012, con l'obiettivo di evidenziarne i punti di forza e di debolezza, utili a fornire indicazioni circa l'opportunità e la sostenibilità strutturale delle politiche demografiche indirizzate alla fecondità e alle migrazioni.

¹ Pur condividendo la responsabilità dell'intero lavoro, i paragrafi 1, 2, 3.1 sono dovuti a Stefania Rimoldi e i paragrafi 3.2 e 4 a Elisa Barbiano di Belgiojoso

² Per le definizioni delle diverse poste e per la metodologia di calcolo adottata si consulti Blangiardo, 2012 e Blangiardo e Rimoldi, 2013.

2 Dati e metodi

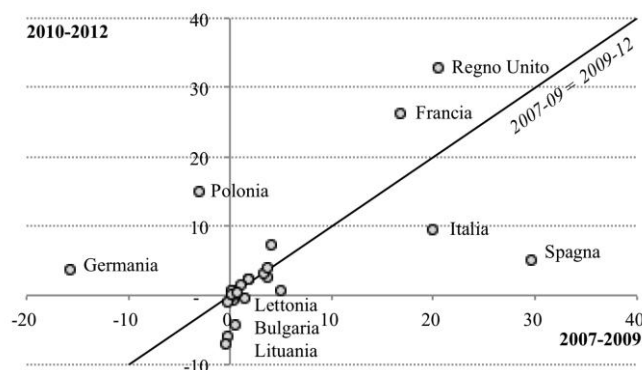
I dati utilizzati (di fonte Eurostat) per il calcolo del potenziale demografico, della sua variazione nel tempo e della sua struttura fanno riferimento alla popolazione per età e genere agli istanti 1.1.2007 – 1.1.2012, alle tavole di mortalità per genere degli anni 2006-2011, ai decessi per genere e singolo anno di età (in anni compiuti) 2006-2011 e, infine, ai movimenti anagrafici 2006-2011 in corrispondenza di ciascuno dei 27 paesi dell'Unione Europea.

3 Risultati

3.1. La formazione del patrimonio demografico

Nell'orizzonte temporale osservato il patrimonio demografico complessivo dell'Europa, e dunque l'ammontare del suo potenziale di futuro, in termini di anni-vita, è cresciuto: da 20,2 miliardi al 1.01.2007 a 20,7 miliardi al 1.01.2012, pari ad un tasso di incremento medio annuo composto del 4,7 %. Tale crescita, tuttavia, si manifesta con un andamento irregolare: il tasso d'incremento annuo appare in diminuzione lungo tutto il periodo ad eccezione dell'ultimo anno (da 4,9 a 4,2‰, rispettivamente nel 2007/2008 e 2010/2011, a 5,6‰ nel 2011/2012).

Figura 1 - *Variazioni medie annue dei patrimoni demografici dei paesi dell'Unione Europea. Milioni di anni-vita, trienni 2007-2009 e 2010-2012*



Fonte: elaborazioni su dati Eurostat

Nel quadro del generale aumento del patrimonio demografico dell'Europa, per alcuni paesi si osserva un andamento tendente alla stagnazione o addirittura al regresso (Figura 1). Si tratta di Germania, Lituania, Lettonia, Romania, Bulgaria e

del Portogallo. In particolare, per la Germania si rilevano perdite piuttosto intense nel triennio 2007-2009, nell'ordine di 16 milioni di anni-vita l'anno, solo parzialmente compensate dai recuperi osservati nel triennio successivo + 4 milioni di anni vita. Sul fronte opposto, Francia, Regno Unito, Italia e Spagna si fanno notare per l'ammontare degli incrementi annui assoluti, anche se solo per i primi due paesi gli incrementi sono crescenti nei due intervalli. Mentre in Italia i 20 milioni annui aggiuntivi registrati nel primo triennio risultano dimezzati nel secondo e per la Spagna, per la quale si era registrato il massimo ammontare nel primo triennio, si osserva nel secondo triennio una drastica riduzione a poco più di 5 milioni di anni-vita aggiuntivi l'anno.

In termini relativi, il paese più dinamico è il Lussemburgo, per cui si osservano tassi di crescita pari a 28 e 24 per 1000 nei due periodi considerati. Seguono Cipro e Irlanda (rispettivamente con tassi di crescita pari a 12, 2 e 16,3 ‰). In complesso questi tre paesi, tuttavia, contribuiscono annualmente alla crescita del potenziale di futuro dell'Europa con soli 3 milioni di anni-vita. Ben l'84% della variazione del patrimonio demografico dell'Unione Europea si deve, infatti, a Francia, Regno Unito, Italia e Spagna. L'entità del loro apporto alla formazione del potenziale futuro europeo suggerisce un percorso di analisi delle componenti che ne determinano la variazione e alcune riflessioni in merito alla "qualità" del loro contributo.

In generale, la variazione del patrimonio demografico deriva dai contributi addizionali delle nascite e dei saldi migratori (eventualmente negativi), al netto della perdita di patrimonio da parte dei soggetti deceduti nel corso dell'anno o del semplice "consumo" da parte dei soggetti ancora in vita alla fine dell'anno.

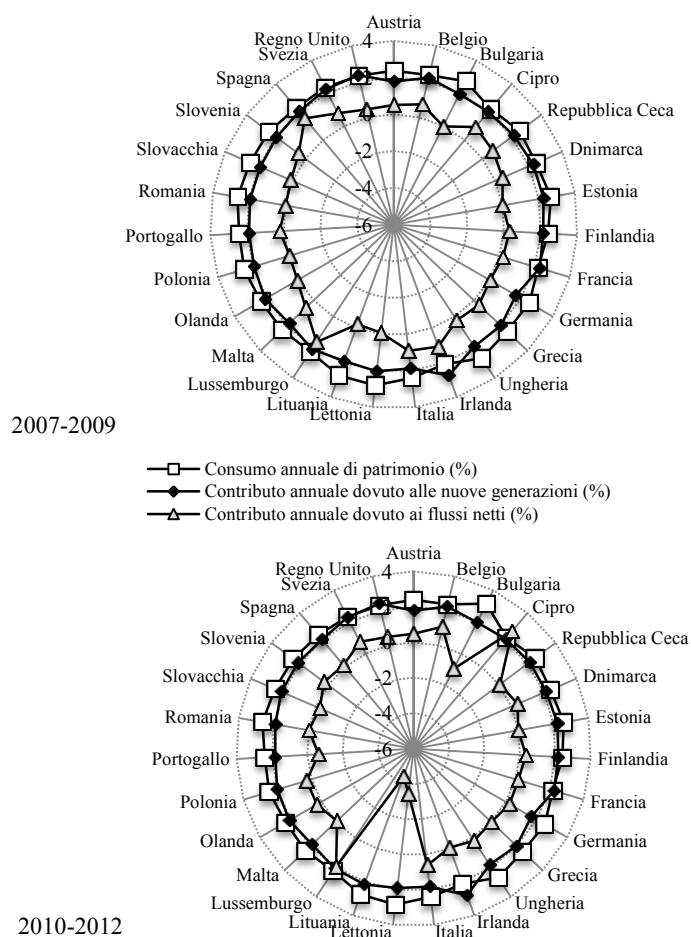
La proporzione di patrimonio consumato annualmente esprime sinteticamente l'incidenza degli effetti congiunti della struttura per età e dei profili di sopravvivenza nella formazione del patrimonio demografico in ognuno dei 27 paesi europei (Figura 2): tale proporzione, stabilmente pari al 2,4% nel periodo considerato per il complesso dell'Unione Europea, raggiunge i valori minimi in corrispondenza di Irlanda e Francia (2,1% e 2,2%, rispettivamente), mentre supera il 2,7% in Bulgaria, Ungheria e nelle Repubbliche Baltiche.

D'altro canto, il contributo delle nuove generazioni, espresso dalla proporzione di anni-vita guadagnati per mezzo delle nascite (anche qui, condizionato dai profili di sopravvivenza), pari al 2% del patrimonio per il complesso dell'Europa, appare particolarmente importante in Irlanda (mediamente stabile attorno al 2,8%) in Svezia (in crescita dal 2,2% al 2,4%), nel Regno Unito (stabile al 2,4%) e in Francia (2,3%). Valori particolarmente bassi, invece, si osservano in corrispondenza di Germania (1,7%) e Austria (1,8%).

Senza dubbio la componente più variabile nella formazione del patrimonio demografico è quella derivante dai flussi migratori netti: per il complesso

dell'Unione Europea essa vale solo lo 0,3% del patrimonio, in media nel periodo considerato, e la sua incidenza diminuisce nel tempo. L'estrema variabilità tra i valori relativi ai 27 paesi e in corrispondenza di uno stesso paese nei diversi anni può essere ricondotta a due cause principali: anzitutto al fatto che una larga quota dei movimenti migratori è "interno" al sistema europeo e che, pertanto, risultano compensarsi a livello aggregato; in secondo luogo, al fatto che le crisi economiche (del 2008 e del 2010) hanno nettamente influenzato i flussi (di emigrazione e di ritorno), ribilanciando l'equilibrio tra costi e opportunità dell'emigrazione e/o

Figura 2 - Contributo delle diverse componenti alla formazione del patrimonio demografico: medie 2007-2009 e 2010-2012



Fonte: elaborazioni su dati Eurostat

riorientando i flussi verso nuove destinazioni. Rispetto all'andamento medio europeo, si distinguono tre tipologie di percorsi. Un primo gruppo è costituito da Belgio, Lussemburgo, Cipro e Svezia: il loro patrimonio demografico appare particolarmente determinato da flussi netti positivi in aumento nel periodo (in media da 0,8% a 1% del patrimonio demografico nel passaggio tra i due trienni). Un secondo insieme di paesi, composto da Italia, Irlanda, Repubblica Ceca e Spagna, appare anch'esso caratterizzato da flussi migratori positivi netti, ancorché decrescenti (in media da 1,2% nel 2007-2009 a 0,4% nel 2010-2012). Infine, per il terzo gruppo di paesi si osserva l'intensificarsi di flussi netti negativi, a partire dal 2010: si tratta di Bulgaria, Malta, Portogallo, Lettonia, Lituania e Slovacchia (da 0,1% nel 2007-2009 a -1,2% nel 2010-2012).

3.2. Aspetti strutturali: il contributo per età

Anche per le popolazioni, così come accade nella formazione del "capitale" del singolo individuo, il numero di anni dedicati alla formazione e al pensionamento costituiscono un "debito" mentre gli anni di lavoro ne formano il "credito". In prospettiva, dunque, in un certo istante, il patrimonio demografico di ogni popolazione, per diretta conseguenza della sua struttura per età, è costituito da un ammontare di anni da contabilizzare come (potenziale) debito e di un ammontare di anni di (potenziale) credito. Va tuttavia osservato come, diversamente rispetto a quanto avviene per il singolo individuo, per cui necessariamente le tre fasi sono subordinate l'una all'altra nel tempo, per una popolazione l'equilibrio tra le parti di vita destinate alla produzione e quelle destinate al consumo dipende dalla dinamica delle sue componenti.

Applicando, quindi, alla struttura per età e genere della popolazione di ciascun paese il medesimo modello di calcolo, che tiene conto dei corrispondenti livelli di sopravvivenza e ipotizza come soglie d'ingresso e uscita dal mondo del lavoro rispettivamente 20 e 67 anni, risulta come un cittadino europeo disponga mediamente³ – in ipotesi di un'equa spartizione tra individui dei 20,4 miliardi di vita residua – 41,2 anni di vita residua: di questi, 24,9 saranno da trascorrere in condizione di lavoro – e dunque da mettere a bilancio come poste a "credito" – 14,1 in pensione e 2,3 in formazione – per un totale di 16,4 anni da contabilizzare come "debito". Pertanto nel bilancio complessivo delle diverse fasi della vita di ciascun individuo, a ogni anno residuo di potenziale debito corrispondono circa 1,5 anni di potenziale credito; tale vantaggio appare in lieve diminuzione nel tempo.

³ Ci si riferisce alla media calcolata sull'intero intervallo 2007-2012.

Tabella 1 - Anni medi di vita residua in formazione, lavoro e pensione nei 27 paesi europei. Triennio 2007-2009 e 2010-2012.

	2007-2009 (valore medio)				2010-2012 (valore medio)			
	formazione	lavoro	pensione	totale	formazione	Lavoro	pensione	Totale
Austria	2,1	24,6	15,3	42,0	2,1	24,1	15,5	41,7
Belgio	2,4	24,7	15,0	42,1	2,4	24,6	15,2	42,2
Bulgaria	1,9	23,3	11,3	36,5	1,9	22,7	11,7	36,3
Cipro	2,4	27,5	15,6	45,5	2,4	27,5	16,2	46,1
Rep. Ceca	2,1	24,8	13,2	40,1	2,1	24,3	13,6	40,1
Germania	1,9	23,1	14,9	39,9	1,9	22,6	15,1	39,6
Danimarca	2,6	25,0	14,1	41,7	2,5	24,8	14,7	42,0
Estonia	2,2	24,1	12,0	38,3	2,2	24,2	12,9	39,3
Spagna	2,1	25,1	15,8	43,0	2,2	24,7	16,2	43,0
Finlandia	2,4	24,3	15,1	41,8	2,4	24,0	15,4	41,8
Francia	2,6	25,3	15,9	43,7	2,6	24,9	16,1	43,7
Grecia	2,0	24,0	15,3	41,3	2,1	23,5	15,6	41,1
Ungheria	2,1	23,8	11,5	37,4	2,1	23,6	11,8	37,5
Irlanda	3,0	28,8	15,1	46,9	3,1	28,3	15,6	47,0
Italia	2,0	23,3	15,8	41,1	2,0	23,0	16,0	41,0
Lituania	2,2	24,2	11,3	37,6	2,1	23,5	12,0	37,6
Lussemburgo	2,5	26,2	15,2	43,9	2,5	26,1	15,6	44,1
Lettonia	2,0	23,6	11,1	36,7	2,1	23,1	11,6	36,8
Malta	2,3	25,8	15,0	43,0	2,1	25,0	15,6	42,8
Olanda	2,5	25,5	15,3	43,3	2,4	25,0	15,6	43,1
Polonia	2,2	25,6	12,7	40,4	2,2	25,1	13,2	40,5
Portogallo	2,2	24,5	14,9	41,5	2,1	23,7	15,3	41,1
Romania	2,2	25,1	11,5	38,8	2,1	24,8	11,9	38,7
Svezia	2,4	24,8	15,6	42,8	2,4	24,8	15,9	43,1
Slovenia	2,0	24,3	14,3	40,6	2,0	24,0	14,8	40,8
Slovacchia	2,2	26,2	11,9	40,4	2,2	25,7	12,5	40,5
Regno Unito	2,5	25,5	15,0	43,0	2,5	25,4	15,4	43,3
EU-27	2,3	24,9	14,1	41,2	2,2	24,6	14,5	41,3

Fonte: elaborazioni su dati Eurostat

Secondo l'approccio proposto, i paesi con il patrimonio demografico pro-capite più elevato – Irlanda, Cipro, Francia, Olanda, Malta, Spagna, Regno Unito e Svezia – sono quelli connotati sia da una struttura per età più giovane, sia da più elevati livelli di sopravvivenza. In quest'ottica le popolazioni di Italia e, soprattutto, Germania appaiono penalizzate dall'invecchiamento e pertanto risultano disporre di un patrimonio demografico (pro-capite) inferiore ai paesi caratterizzati da analoghi livelli di mortalità. Al contrario i paesi con un patrimonio demografico più basso sono caratterizzati da livelli inferiori di sopravvivenza e/o una struttura per età più anziana. Considerando in aggiunta la distinzione rispetto alle tre fasi di vita (tabella 1), tali paesi sono tra l'altro caratterizzati dalla più ampia quota di tempo residuo da vivere nella fase del lavoro: si tratta di popolazioni connotate da una forte componente lavorativa, combinata a livelli di sopravvivenza più bassi. D'altro canto, Austria, Germania, Spagna, Francia, Grecia, Svezia e Italia mostrano una maggiore proporzione di anni di vita residua da trascorrere in pensione, nuovamente per effetto combinato della struttura per età e di una sopravvivenza più elevata. Considerando, poi, il rapporto tra credito e debito, l'Italia risulta fanalino

di coda nella graduatoria, con quasi 1,3 anni di credito per ogni anno di debito, preceduta da Francia, Germania, Svezia, Grecia, Finlandia e Spagna. In cima alla graduatoria si collocano invece i paesi dell'est Europa, per i quali il divario tra le componenti attiva e inattiva è decisamente più elevato (circa 1,8).

L'incremento del potenziale demografico registrato a livello europeo è determinato dall'aumento del numero di anni vita residua in pensione (+243 milioni), compensato solo parzialmente dal calo della componente lavorativa (-64 milioni). Tale andamento globale, tuttavia, cela alcune importanti differenze tra i paesi, sia in termini di contributo alla crescita del potenziale demografico di ciascuna componente, sia in termini di trend delle stesse. Se, ad esempio, in Italia, Francia, Danimarca e Spagna l'80% circa della crescita è imputabile all'aumento della vita residua nella fase della pensione, in Belgio, Svezia, Cipro e Lussemburgo si riscontra una quasi equa distribuzione tra le fasi di lavoro e pensione mentre per Regno Unito ed Irlanda si osserva una situazione intermedia (il 60% è ascrivibile alle pensioni e circa il 30% al lavoro).

Osservando l'andamento delle tre componenti nei restanti paesi si possono rilevare alcune importanti specificità. La prima è relativa alla Germania per la quale il numero di anni di vita residua in formazione e soprattutto in lavoro subiscono una forte contrazione a fronte di una minor crescita per il tempo da trascorrere in pensione. Tali cambiamenti determinano una perdita consistente di anni vita tra il primo e il secondo triennio. Un andamento simile – seppure con qualche differenza – è riscontrabile per molti dei paesi est europei (Bulgaria, Ungheria, Romania), per le Repubbliche Baltiche e per il Portogallo. In questi casi si registra una perdita di potenziale demografico per quanto concerne le prime due fasi del ciclo di vita (formazione e lavoro), in parte controbilanciato dall'aumento del tempo residuo per pensione. Fanno eccezione Lettonia ed Estonia per le quali si registrano diminuzioni anche per gli anni di vita residua in pensione. Inoltre da segnalare la particolare situazione della Polonia per la quale, a fronte di una buona crescita del potenziale negli ultimi anni, si registra una non trascurabile perdita di potenziale demografico in corrispondenza della fase lavorativa.

4 Conclusioni

I risultati ottenuti dimostrano come non sia tanto importante se il potenziale demografico di una popolazione aumenti o diminuisca, ma quanto siano solide le poste che ne determinano la dinamica e, soprattutto, quanto siano equilibrate le diverse componenti che ne qualificano la struttura. In definitiva, se il trend osservato nel periodo considerato dovesse confermarsi in futuro, in assenza di variazioni, si assisterà a una progressiva contrazione del vantaggio (pur lieve) del

numero medio di anni di potenziale credito sul potenziale debito demografico, per l'effetto congiunto tra il continuo aumento della vita residua attribuibile alla pensione e la continua diminuzione della vita residua nel lavoro, con ineludibili conseguenze sulle poste finanziarie degli stati. Politiche demografiche indirizzate al sostegno del futuro demografico della popolazione europea appaiono dunque maggiormente efficaci se indirizzate alla creazione di "nuovo potenziale" e in misura minore al richiamo di potenziale già esistente altrove; tali considerazioni esulano evidentemente da valutazioni qualitative circa il potenziale demografico acquisito, dai costi legati alla formazione e al pensionamento, dai vantaggi economici derivanti dalla forza lavoro. L'incentivazione di politiche demografiche a sostegno della fecondità, inoltre, non potrà che considerare l'ulteriore aggravio derivante dall'ipotecare i primi 20 anni delle nuove generazioni in formazione. Nel bilancio di medio periodo appare dunque auspicabile il mantenimento del contributo di flussi migratori extraeuropei.

Riferimenti bibliografici

BLANGIARDO G.C. 2012. Discovering the demographic *GDP*, *Rivista Internazionale di Scienze Sociali*, vol.1, pp. 45-58.

BLANGIARDO G.C., RIMOLDI S.M.L. 2013. The potential demography: a tool for evaluating differences among countries in the European Union, *Genus*, Vol. LXVIII, No. 3, pp. 63-81.

SUMMARY

Demographic Debts and Credits of the European Countries.

The Demographic Asset (DA) of a country consists of the amount of future for its population: it is calculated as the sum of the future of its members. Considering the DA of the European Union, this paper aims to evaluate the quality of its dynamics (whether it depends from natural and/or migration balance) and of its structure (the balance between the years of "debt" – devoted to education and retirement - and the years of "credit" – devoted to work). The results show that it is not important whether the demographic potential of a population increases or decreases, but how the items forming the dynamic are consistent, and moreover how the components of "debt" and "credit" balance.

Elisa BARBIANO DI BELGIOJOSO, Ricercatore di Demografia, Università Milano Bicocca, elisa.barbiano@unimib.it

Stefania RIMOLDI, Ricercatore di Demografia, Università Milano Bicocca, stefania.rimoldi@unimib.it

DAL CENSIMENTO AL *FARM REGISTER*: L'USO DELLE INDAGINI POST CENSUARIE¹

Antonella Bernardini, Loredana De Gaetano e Matteo Mazziotta

1. Introduzione

L'impostazione metodologica e organizzativa del 6° Censimento dell'Agricoltura 2010 è stata caratterizzata da numerose innovazioni nelle diverse fasi dell'indagine, a partire dalla lista pre-censuaria che è stata prodotta mediante l'integrazione di numerosi archivi amministrativi provenienti da fonti informative agricole differenti.

Tale metodologia innovativa ha tracciato la strada verso la futura (ma non troppo) costituzione del *Farm Register*; ossia di un registro di aziende agricole continuamente alimentato che, oltre a decretare la fine del Censimento tradizionale, costituisce una risorsa di inestimabile importanza per le statistiche del settore e, quindi, per le relative politiche di intervento. In tale contesto di riferimento, è semplice comprendere l'importanza assunta dalle due indagini di qualità condotte dopo la fine del Censimento e quale rilevanza debba assumere l'analisi finale dei risultati delle stesse.

Il presente lavoro ha l'obiettivo di illustrare le modalità di utilizzo dei risultati delle indagini post censuarie verso la costituzione del *Farm Register*.

2. Il 6° Censimento dell'Agricoltura e le sue innovazioni

I dati censuari consentono la definizione di un quadro informativo statistico su una molteplicità di fenomeni agricoli tradizionali e nuovi (sviluppo rurale, sostenibilità ambientale, rapporti con il territorio, ecc.) ai fini della programmazione e valutazione delle politiche agricole europee, nazionali e regionali. Il Censimento 2010, inoltre, è caratterizzato da alcune innovazioni metodologiche e organizzative di particolare interesse e rilevanza.

¹ Il lavoro è frutto della collaborazione congiunta degli autori. In particolare, i paragrafi 3 e 4 sono curati da Antonella Bernardini, il paragrafo 2 da Loredana De Gaetano, e i paragrafi 1 e 5 da Matteo Mazziotta.

Dal punto di vista organizzativo, le Regioni e Province Autonome hanno potuto optare tra due modelli alternativi. Un modello ad “*alta partecipazione*”, in cui la Regione si assumeva il coordinamento e l’organizzazione delle operazioni censuarie compresa la registrazione dei dati sul sistema informatico della rilevazione (SGR). Un secondo modello, denominato a “*partecipazione integrativa*”, in cui tali attività erano coordinate dall’Istat e alla Regione erano assegnate funzioni limitate. In entrambi i casi l’Istat ha definito gli aspetti tecnici e metodologici, sovrintendendo alle operazioni censuarie, assicurandone il monitoraggio e adottando le misure necessarie per garantirne il buon andamento (Mancini, 2010).

L’acquisizione delle informazioni si è basata su una tecnica di rilevazione multicanale ovvero, accanto alla tradizionale intervista “faccia a faccia”, era presente la rilevazione via web come strumento di raccolta opzionale. Il Sistema informatico di Gestione della Rilevazione (SGR) ha consentito di effettuare il monitoraggio di tutte le fasi del processo produttivo ai vari livelli di operatività della rete di rilevazione, fino alla chiusura delle attività censuarie (Mancini, 2010).

Inoltre, per la prima volta, la lista pre-censuaria, ossia la lista delle aziende agricole da censire, non è stata predisposta dai singoli comuni, come in passato, mediante aggiornamento degli elenchi aziendali ottenuti con il censimento precedente, ma è stata realizzata direttamente dall’Istat, mediante l’integrazione di archivi amministrativi e statistici. Tale lista pre-censuaria, redatta utilizzando sia le informazioni contenute nel Sistema Informativo Agricolo Nazionale (SIAN) sia altri archivi delle amministrazioni ed Enti interessati all’agricoltura, ha permesso di identificare, a priori, il potenziale universo delle aziende agricole (Giovannini, 2010). Da evidenziare che l’uso statistico delle fonti amministrative specifiche del settore agricolo consentirà all’Istat di ottenere le informazioni necessarie a costruire ed aggiornare costantemente il futuro registro statistico delle aziende agricole (*Farm Register*).

3. L’indagine di misura verso il *Farm Register*

Le indagini di misura sono utilizzate come strumento per la stima degli errori di risposta nelle indagini esaustive, come i censimenti, e in quelle campionarie; l’errore di risposta si verifica nella fase di raccolta dati per una serie di cause.

Sono state proposte diverse tecniche (Forsman e Schreiner, 1991) al fine di misurare le componenti dell’errore di risposta, tra le quali, particolarmente utilizzata è la misurazione replicata della stessa unità di rilevazione; ossia ad un campione di unità, già intervistate, viene sottoposta nuovamente una selezione di domande provenienti dal questionario dell’indagine originale. Nel caso in cui la risposta ottenuta

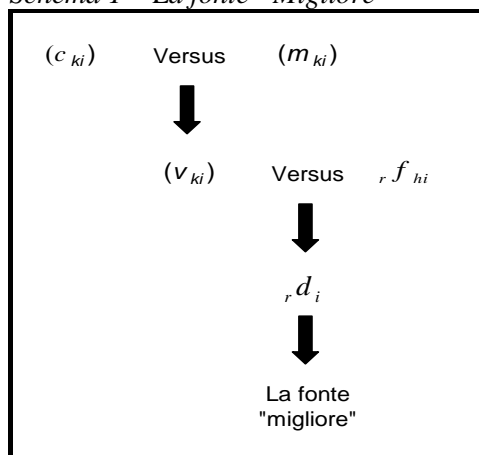
alla reintervista differisca da quella ottenuta all'indagine originale, si attiva la cosiddetta procedura di riconciliazione (Bernardini et al. 2011).

L'indagine di misura del 6° Censimento Generale dell'Agricoltura, realizzata tra maggio 2011 e gennaio 2012, con tecnica di rilevazione telefonica (CATI), è stata condotta su un campione di circa 50.000 aziende già rilevate al Censimento. Le variabili sottoposte a re-intervista sono le principali colture e i principali allevamenti. Obiettivo della rilevazione era stimare la distorsione, la varianza di risposta e le sue principali componenti (Bernardini et al. 2013).

L'utilizzo dei risultati in ottica *Farm Register* deve essere inteso come confronto tra il "valore vero" o riconciliato, individuato attraverso l'indagine di misura, e i valori delle variabili di interesse contenute nelle differenti fonti amministrative. Certamente, per ciascuna variabile, la distorsione più bassa tra il valore riconciliato e il dato amministrativo permette di definire la fonte "migliore".

Sia k la generica azienda agricola e i la variabile rilevata all'indagine di misura, l'indagine mette a confronto il valore censuario (c_{ki}) con il valore rilevato all'indagine stessa (m_{ki}) ricavandone il valore vero (v_{ki}) attraverso la riconciliazione. Il valore v_{ki} può essere quindi utilizzato per il calcolo delle distorsioni rispetto ai valori assunti dalla variabile i nelle r fonti amministrative (${}_r f_{hi}$). Tali distorsioni (${}_r d_i$), ordinate in ordine crescente, consentono facilmente di individuare la distorsione più piccola, ossia la fonte amministrativa le cui variabili di interesse sono più vicine al "valore vero". Tale procedimento è ovviamente valido se le fonti amministrative posseggono le variabili (SAU, superficie totale, coltivazioni legnose agrarie, ecc.) che sono state sottoposte a controllo di qualità dall'indagine di misura (Schema 1).

Schema 1 – La fonte "Migliore"



4. L'indagine di copertura (PES) verso il *Farm Register*

L'indagine di copertura è una rilevazione post-censuaria utile a valutare l'errore di copertura del Censimento.

In particolare la componente di errore, derivante dal fatto che alcune unità facenti parte del campo di osservazione sfuggono alla rilevazione, non essendo individuabile durante la fase di raccolta dati, e non potendo essere corretta nella fase di controllo e correzione, viene valutata attraverso l'indagine di copertura.

L'indagine di copertura, o PES – *Post Enumeration Survey* del 6° Censimento dell'Agricoltura – ha l'obiettivo di stimare, a livello regionale, il numero di aziende agricole realmente esistenti alla data di riferimento del Censimento (24 ottobre 2010) e il tasso di copertura definito come rapporto tra il numero di aziende rilevate al Censimento e il numero di aziende realmente esistenti sul territorio nazionale.

Per stimare il tasso di copertura si definisce un modello statistico basato essenzialmente sulle ipotesi adottate nel modello di Petersen (Wolter, 1986) che rientra nella classe dei modelli del tipo *dual-system* o cattura ricattura e che rappresenta uno dei più diffusi tra quelli volti a misurare gli errori di copertura del Censimento.

Un'indagine pilota è stata realizzata nei mesi di giugno e luglio 2011 (nella Regione Veneto) e da ottobre 2011 fino a gennaio 2012, la rilevazione, è stata condotta nelle altre regioni e province autonome (Bernardini et al. 2012). L'indagine si basa su uno schema campionario di tipo areale in cui sono stati coinvolti circa 1.500 fogli di mappa del Catasto dei terreni² e per le province autonome di Trento e Bolzano³, sono state estratte rispettivamente 1.641 e 776 particelle catastali⁴ che hanno costituito le unità finali di campionamento. La modalità di rilevazione ha previsto una rigorosa procedura di individuazione delle aziende agricole a partire dalle informazioni presenti nell'archivio del Catasto dei terreni. L'individuazione delle aziende agricole è avvenuta mediante una procedura articolata in due distinti passi (Mazziotta et al. 2011):

passo 1) individuazione delle aziende agricole; in tale fase, il rilevatore ha preso contatto con gli intestatari della particella catastale al fine di conoscere se, al 24 ottobre 2010, il terreno di tale particella fosse incluso in un'azienda agricola. In

² Foglio di mappa catastale: insieme di particelle catastali continue, rurali o urbane, che appartengono a precise porzioni di territorio, predeterminate dal catasto dei terreni.

³ Nelle province di Trento e Bolzano, non sono presenti i fogli di mappa; quindi la rilevazione si svolge direttamente sulle particelle catastali.

⁴ Particella catastale: poligono chiuso di una porzione di territorio situata in uno stesso Comune appartenente ad un'unica proprietà (la quale può appartenere a più soggetti), avente la stessa qualità o classe e la stessa destinazione. Una particella è individuata da un numero identificativo, da un foglio di mappa all'interno del quale è inserita, dal Comune e Provincia in cui il foglio ricade.

caso di risposta affermativa, il rilevatore ha acquisito il nome del conduttore dell'azienda;

passo 2) intervista al conduttore, individuato al passo precedente, al fine di accertare, con riferimento al 24 ottobre 2010, l'appartenenza al campo di osservazione di tale azienda e le sue principali caratteristiche (SAU, superficie totale, ecc.).

Al termine di questi due passi, il rilevatore ha definito l'elenco delle aziende agricole individuate sul territorio con le loro principali caratteristiche. Conclusa la rilevazione sul campo a cura dei rilevatori e degli Uffici Regionali di Censimento (URC), è iniziata, una delle fasi più delicate di tutto il processo: il *record linkage* a cura dell'Istat. Tale fase ha lo scopo di effettuare l'abbinamento delle aziende individuate nei fogli di mappa, all'indagine di copertura, con quelle individuate al Censimento; la complessità risiede nel fatto che è necessario che l'abbinamento (con l'utilizzo anche di *linkage* probabilistico) sia esatto, ossia ad una azienda rilevata al Censimento deve corrispondere la medesima unità all'indagine di copertura.

Il *record linkage* è stato effettuato con il censimento e con tutte le fonti amministrative che hanno concorso alla composizione della lista pre-censuaria. Tale approccio ha consentito, non solo, di stimare il tasso di copertura (abbinati con il censimento) ma di poter dettagliare la sottocopertura (non abbinati con il censimento). Infatti, tutte le aziende agricole rilevate alla PES e non abbinati con le aziende enumerate al censimento sono state linkate con la lista pre-censuaria e con tutte le fonti amministrative di partenza. Ad esempio, una generica azienda agricola, rilevata all'indagine di copertura, non si è abbinata con il Censimento ma, ha trovato abbinamento con una azienda presente nella lista pre-censuaria non enumerata dal Censimento, la PES ha invece dimostrato che sarebbe stato corretto farlo in quanto eleggibile ad azienda agricola. Un altro esempio: una generica azienda rilevata alla PES, e non abbinata con il censimento, si è abbinata con una unità agricola presente in una delle fonti amministrative prese in considerazione; in questo caso, l'Istat ha scartato tale unità in fase di preparazione della lista pre-censuaria perché ritenuta 'non eleggibile' e, invece, la PES ha dimostrato che sarebbe dovuta essere inserita nella lista pre-censuaria ed enumerata dal Censimento. In definitiva è possibile scomporre la sottocopertura del censimento nelle seguenti componenti:

- azienda non presente in alcuna fonte amministrativa;
- azienda presente in lista pre-censuaria;
- azienda presente in fonti amministrative (esclusa la pre-censuaria);
- azienda esitata al censimento come mancata risposta.

Grazie a tale operazione, svolta per la prima volta in Italia per un censimento, è possibile studiare dei criteri oggettivi di eleggibilità delle unità agricole, contenute nelle diverse fonti amministrative, ad aziende e quindi progettare e realizzare, con molte più informazioni, il *Farm Register*.

Anche l'abbinamento tra le aziende agricole rilevate alla PES e il Censimento ha consentito la disaggregazione in più componenti del tasso di copertura :

- azienda rilevata al censimento;
- demografia di azienda agricola;
- azienda che presenta legami multipli con il censimento.

Con il termine “demografia di azienda agricola” si deve intendere un abbinamento tra le due liste in cui è rispettata una continuità tra le unità agricole; ossia è la stessa azienda agricola che, tuttavia, nel corso del tempo intercorso tra le due rilevazioni, ha subito un cambiamento nelle variabili anagrafiche o nella forma giuridica. Per “legami multipli con il censimento” si intende una unità rilevata alla PES che trova abbinamento con più aziende rilevate al censimento o viceversa; in questo caso si è trattato di affrontare il problema dell'abbinamento univoco tra terreno e azienda agricola, ad esempio è il caso di terreni divisi tra più persone della stessa famiglia.

Anche questa specificazione delle unità abbinate con la lista censuaria risulta particolarmente informativa al fine di comprendere i legami tra le unità agricole, i mutamenti di stato nel tempo e, quindi, le eventuali metodologie da applicare nella progettazione ed esecuzione del *Farm Register*.

5. Conclusioni

L'obiettivo dei prossimi anni è lo studio e la progettazione del passaggio dal censimento tradizionale al Registro delle aziende agricole. L'invito degli organismi internazionali ad armonizzare le statistiche sulle imprese anche attraverso l'utilizzo di fonti amministrative sta caratterizzando gli studi e le prospettive future dell'Istituto Nazionale di Statistica.

Il Regolamento (CE) N. 177/2008 del Parlamento Europeo e del Consiglio del 20 febbraio 2008 istituisce un quadro comune per i registri utilizzati a fini statistici e abroga il Regolamento (CEE) N. 2186/93, relativo al coordinamento comunitario dello sviluppo dei registri di imprese a fini statistici e, inoltre, è parte di una serie di norme destinate ad armonizzare le Statistiche di business delle infrastrutture europee. In definitiva, gli Stati Membri realizzano uno o più registri armonizzati a fini statistici quale strumento per la preparazione e il coordinamento di indagini, nonché quale fonte di informazioni per analisi statistiche della popolazione di imprese e della sua demografia, per l'utilizzo dei dati amministrativi e per l'individuazione e la costruzione di unità statistiche.

Come già descritto, la progettazione e la realizzazione del 6° Censimento generale dell'Agricoltura sono state caratterizzate da numerose innovazioni organizzati-

ve e metodologiche, tra le quali, la creazione della lista pre-censuaria mediante l'integrazione di diverse banche dati provenienti da differenti fonti amministrative ne è certamente l'esempio più importante.

Questa innovazione può radicalmente modificare l'impostazione di tutte le indagini agricole in un'ottica di costituzione di un *Farm Register*. In tale contesto di riferimento, normativo e statistico, si inseriscono le indagini di controllo del Censimento che, assumono, in questa edizione censuaria, una elevata importanza al fine di misurare il grado di bontà della metodologia adottata dal Censimento stesso nella costituzione della lista pre-censuaria. I risultati dell'indagine di misura potrebbero, essere usati per valutare la qualità dei dati di natura agricola contenuti nelle differenti banche dati amministrative. L'applicazione di modelli statistici per la stima della distorsione relativamente ad alcune variabili significative contenute nelle fonti amministrative consente di testare la significatività e la bontà delle fonti stesse nell'ottica di valutarne quella maggiormente completa, aggiornata e "affidabile". Una maggiore capacità informativa, in ottica *Farm Register*, proviene dall'utilizzo dei risultati dell'indagine di copertura. La possibilità di poter scomporre l'errore di sottocopertura nelle sue componenti (mancata risposta, errori di lista, ecc.), a seguito di una attenta e indispensabile fase di *record linkage*, risulta fondamentale per comprendere le modalità di costruzione del registro e l'utilizzo *ad hoc* delle fonti che lo alimentano costantemente. Sembra altresì indispensabile la progettazione di indagini denominate C-sample e D-sample al fine di verificare, in corso, la qualità del registro.

La PES del Censimento dell'agricoltura costituisce il primo caso, in Italia per un censimento economico, di scomposizione del tasso di copertura e dell'errore di sottocopertura nelle diverse componenti. Tale informazione deve essere il punto di partenza per la progettazione e la costruzione del Registro statistico delle aziende agricole.

Riferimenti bibliografici

- BERNARDINI A., DE GAETANO L., MAZZIOTTA M., SORIANI L. (2012). The Coverage Survey of the 6th Agricultural Census. In atti del Convegno della XLVI Riunione Scientifica Società Italiana di Statistica, Roma 20 – 22 giugno 2012.
- BERNARDINI A., MAZZIOTTA M., ROMEO R., SORIANI L. (2011). The Developments for the Quality Evaluation of the Italian Agricultural Censuses. In atti del Convegno Intermedio della Società Italiana di Statistica, Bologna, 8-10 giugno 2011.

- BERNARDINI A., DE GAETANO L., MAZZIOTTA M., SORIANI L. (2013). La qualità dell'informazione del 6° Censimento generale dell'agricoltura: le indagini post censuarie. *Rivista Italiana di Economia Demografia e Statistica*, Volume LXVII n. 1.
- FORSMAN G., SCHREINER I. (1991). The Design and Analysis of Reinterview: an Overview, in: *Measurement Errors in Surveys*, Biemer, P. Groves, R.M. Lyberg, L. Mathiowetz, N. and Sudman, S. (Eds.), New York, Wiley, 279-302.
- GIOVANNINI E. (2010). Il censimento generale dell'agricoltura del 2010 , *Agri-regioneuropa*, Anno 6 , Numero 22.
- MANCINI A. (2010). Le principali novità del 6°Censimento dell'agricoltura, *Agriregioneuropa*, Anno 6 , Numero 22.
- MAZZIOTTA M., BERNARDINI A., SORIANI L. (2011). Indagine di copertura del 6° Censimento generale dell'agricoltura. *Manuale delle istruzioni*, Istat.
- WOLTER K. (1986). Some Coverage Error Models for Census Data, *Journal of the American Statistical Association*, Vol. 81, No. 394, pp. 338-346.

SUMMARY

From census to farm register: the use of post census survey

The methodological and organizational approach of the last Census of Agricultural Holdings was characterized by several innovations in the different phases of the survey. The pre-Census list has been implemented by the integration of many administrative data from different agricultural sources. This innovative methodology has paved the way for the future establishment of the Farm Register. It is a register continuously updated that represents a relevant resource for the agricultural statistics. In this context, it is easy to understand the importance of the results of the two post Census quality surveys. The aim of the present paper is to show how the use and the analysis of the quality surveys influence the Farm Register construction.

Antonella BERNARDINI, Ricercatore, Istat, anbernar@istat.it
Loredana DE GAETANO, Tecnologo, Istat, degaetan@istat.it
Matteo MAZZIOTTA, Ricercatore, Istat, mazziott@istat.it

L'INVECCHIAMENTO "IMPORTATO" NELLE REGIONI ITALIANE

Gian Carlo Blangiardo, Stefania Rimoldi

1. Introduzione

Il ruolo delle componenti naturali nel processo di invecchiamento della popolazione italiana è stato ampiamente trattato in letteratura già dagli anni '70 del secolo scorso. Ancorché secondario rispetto a quello della dinamica naturale, il contributo della dinamica migratoria all'invecchiamento demografico va assumendo un'importanza sempre maggiore, in conseguenza dei costanti flussi di ingresso dall'estero e in assenza di un'analogica controparte in uscita dal nostro paese (Gesano e Strozza 2011). In teoria, le migrazioni tendono a produrre un invecchiamento nelle popolazioni di origine e un ringiovanimento in quelle d'arrivo, data la marcata caratterizzazione giovanile dei migranti. Tuttavia, per entrambe le aree, si tratta di un apporto generalmente modesto e di un fattore di ringiovanimento/invecchiamento destinato a manifestarsi solo in via transitoria (Paterno 2011) e a trasformarsi, nel corso degli anni, in ciò che si può definire "invecchiamento importato/esportato" (Blangiardo 2012), ossia un più accentuato/attenuato flusso di accesso all'universo degli anziani che, come si avrà modo di documentare per molte regioni italiane, non è che l'effetto dei cambiamenti strutturali acquisiti nel passato attraverso la mobilità della popolazione.

2. Dati e metodi

Nel presente lavoro il contributo delle migrazioni nell'alimentare i flussi di accesso alla popolazione anziana, indicato come invecchiamento "importato (o esportato)", è quantificato attraverso il confronto tra lo stock di individui che ogni anno si affacciano alle età anziane e il corrispondente ammontare di nascite registrate 65 anni prima, adeguatamente depurato dagli effetti della mortalità. Per ciascuna regione la dimensione di tale invecchiamento è calcolata come differenza tra la popolazione 65enne stimata dall'Istat al 31 dicembre degli anni 2017-2064

(Istat 2011) e l'apporto teoricamente atteso dalle corrispondenti generazioni nate in loco negli anni dal 1952 al 2000.

3. Dinamica e struttura demografica: la persistenza del divario nord-sud, vecchie diversità e nuove somiglianze

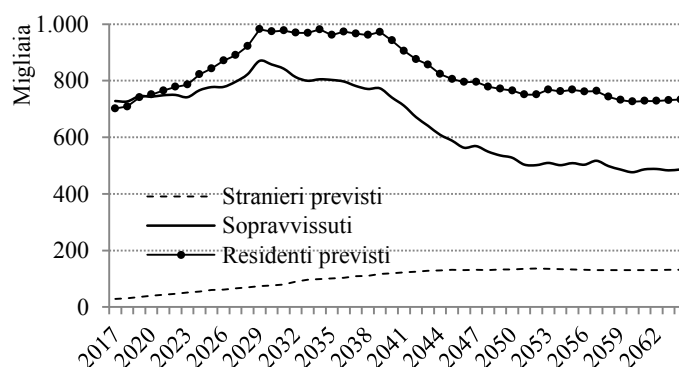
Al Censimento 2011 la popolazione italiana si è attestata poco al di sotto dei 60 milioni di abitanti. Dopo una fase di rallentamento osservato negli ultimi due decenni del secolo scorso, nella prima decade del nuovo millennio essa sembra aver ripreso la crescita, spinta da sostenuti flussi migratori dall'estero. L'andamento non è stato uniforme sul territorio nazionale: nell'arco dell'ultimo mezzo secolo (1961-2011) i tassi di crescita medi annui sono stati globalmente più elevati per le popolazioni delle regioni del Centro-Nord (4,9 contro 3,8 per 1000, registrato per l'Italia) e decisamente più bassi per quelle meridionali e insulari (2,7 e 2,4 per 1000, rispettivamente). Tra le regioni più dinamiche spiccano la Lombardia (6,6 per 1000) e il Lazio (8,5 per 1000) mentre, al contrario, il Molise, la Basilicata e la Calabria fanno segnare perdite anche consistenti (-4,3, -1,3 e -0,7 per 1000). Nel corso dell'ultimo ventennio l'asse di crescita appare maggiormente orientarsi verso Est, con Trentino A.A, Emilia Romagna e Veneto nel ruolo di protagonisti. Com'è noto, la dinamica della popolazione osservata fino agli anni '80, pur con i divari e ritardi nelle diverse aree del territorio, è stata principalmente il frutto della componente naturale, che fino alla fine degli anni '60 ha assunto saldi così elevati da compensare le pur sostanziali emorragie provocate dagli intensi flussi migratori verso l'estero. Già dai primi anni '80 la forza della componente migratoria (questa volta in ingresso nel nostro paese) ha però preso il sopravvento, con la complicità di un livello di fecondità in netto calo ovunque. Alla stazionarietà demografica che ha caratterizzato gli anni '80 fa seguito, dall'inizio del decennio successivo, l'esplosione delle immigrazioni dall'estero con il picco raggiunto nel 2002 (+11 per 1000, un'intensità mai toccata prima di allora neppure dalla componente naturale). Gli anni più recenti sembrano invece caratterizzati da un progressivo ridimensionamento delle immigrazioni, e da una leggera ripresa della crescita naturale che appare, comunque, sostanzialmente "tirata" proprio dalle immigrazioni pregresse e dai loro conseguenti riflessi sulla fecondità.

4. L'invecchiamento "importato" nelle regioni italiane

Al 1 gennaio del 2017 si valuta che saranno circa 728.000 gli individui nati in Italia nel 1952 e "teoricamente sopravvissuti" nel corso del periodo 1952-2017 in base al rischio di mortalità cui sono stati sottoposti nel corso della loro esistenza,

(figura 1). Secondo le stime Istat più aggiornate sono, invece, circa 702.000 i 65enni previsti, con una differenza pari a circa 26.000 unità. Se le previsioni Istat fossero confermate, per il contingente di nuovi ingressi nell'età anziana si tratterebbe di una perdita rispetto ai sopravvissuti teorici pari ad almeno 50.000 unità qualora si tenesse conto anche del contributo additivo della componente straniera nelle previsioni, che già nel 2017 comprende circa 28.000 65enni. A partire dal 2019, tuttavia, il numero di ingressi previsti nell'età anziana supererà quello dei sopravvissuti, con uno scarto che aumenterà progressivamente fino a raggiungere, nel 2054, il picco massimo di oltre 260.000 unità, per poi successivamente attestarsi sulle 250.000. Inoltre, il contingente di 65enni (italiani) previsto, depurato dalla componente straniera, continuerebbe a rimanere inferiore a quello dei sopravvissuti fino al 2023, scontando l'effetto delle perdite legate al nostro passato di "emigranti"¹.

Figura 1 – Numero di 65enni previsti e "teoricamente sopravvissuti" in Italia. 2017-2064



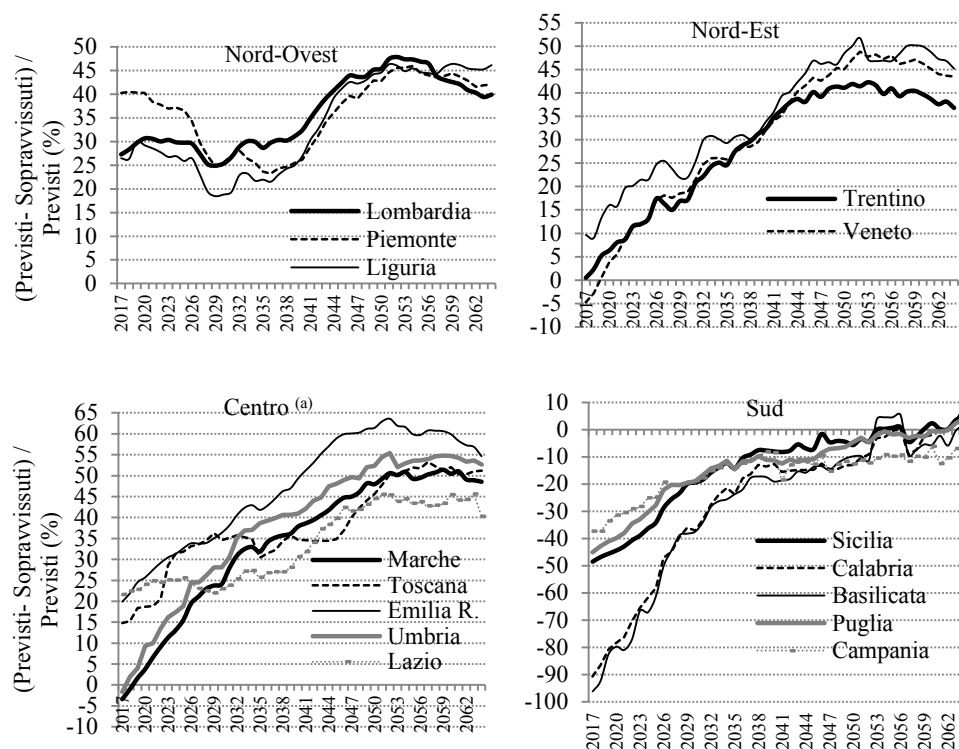
Fonte: n/elaborazioni su dati Istat

Con riferimento al dettaglio regionale, nel cui ambito l'invecchiamento importato riflette anche (per non dire soprattutto) i flussi di mobilità interna (di cui il secondo dopoguerra ha fornito straordinari esempi) si possono distinguere quattro modelli evolutivi differenti: nord-occidentale, nord-orientale, centrale e meridionale (figura 2). Nel modello nord-occidentale, che si realizza in Piemonte, Lombardia e Liguria, la popolazione 65enne (nel complesso delle 3 regioni) fornisce già nel 2017 un surplus rispetto al numero teorico di sopravvissuti, pari a

¹ Va precisato che nel contingente di 65enni italiani si conteggia anche il saldo tra gli stranieri che hanno acquisito la cittadinanza italiana (i più) e i connazionali che l'hanno persa (una modesta minoranza), per cui il divario è, di fatto, un'approssimazione per difetto.

circa 58mila unità (pari al 31% del contingente di 65enni previsto). Nel 2054 tale componente “importata” dell’invecchiamento raggiungerà, per il complesso delle tre regioni, le 93.000 unità, pari al 47% di coloro che presumibilmente entreranno per la prima volta nel contingente degli anziani; negli anni successivi tale incidenza si ridimensionerà rimanendo comunque superiore al 40%.

Figura 2 – Scarto 65enni previsti – sopravvissuti (% previsti)



Note: (a) Inclusa l'Emilia Romagna
Fonte: n/elaborazioni su dati Istat

Negli anni di picco massimo (2053-54) ben un quarto dell’invecchiamento importato nel bilancio nazionale sarà attribuibile a Piemonte, Lombardia e Liguria. D’altro canto, il modello nord-orientale (osservato in Veneto, Friuli e Trentino) mostra, negli anni più prossimi, un saldo tra 65enni previsti e sopravvissuti, quasi nullo (o, nel caso del Veneto, addirittura negativo), destinato tuttavia a incrementarsi considerevolmente nei decenni successivi: nel 2054 raggiungerà le 44.000 unità (47% dei previsti) per poi ridiscendere nel 2064 a 42.000 (43%).

Nell’ultimo decennio di previsione, il contributo di queste regioni all’ammontare dell’invecchiamento importato per il complesso dell’Italia, si attesta

intorno al 17%. Anche il modello che accomuna le regioni del Centro (allargato a Emilia Romagna e comprensivo di Toscana, Umbria, Marche e Lazio) appare caratterizzato da un'iniziale parità fra l'ammontare dei 65enni previsti e sopravvissuti. Tuttavia è qui che si registra il più elevato incremento del divario tra 65enni sopravvissuti e previsti nel corso degli anni successivi, sia perché il numero dei previsti rimane elevato più a lungo, sia a causa di una progressiva diminuzione del numero di sopravvissuti (dovuta a un calo più repentino delle corrispondenti coorti di nascite negli anni '70²). L'incidenza dell'invecchiamento importato raggiungerà, per questo insieme di regioni, il 52% dell'invecchiamento previsto nel 2054 (118.000 unità, ben il 45% dell'invecchiamento importato dell'Italia in complesso). Alla fine del periodo di previsione tale incidenza scenderà al 48%, con 108.000 65enni addizionali previsti rispetto ai sopravvissuti. Nel modello di riferimento delle regioni del Mezzogiorno (che qui riguarda unicamente Basilicata, Puglia, Campania, Calabria e Sicilia) il numero di 65enni previsti è (quasi) sempre inferiore a quello dei sopravvissuti. Il saldo per il complesso di quest'area, pari a circa -98.000 unità (oltre il 50% dei 65enni previsti) nel 2017 ma appare destinato a diminuire drasticamente nel corso dei decenni successivi fino ad annullarsi nel 2063. Dal quadro appena disegnato restano escluse alcune realtà (Valle d'Aosta, Abruzzo, Molise e Sardegna) per le quali si notano singolarità specifiche. La Valle d'Aosta, pur trattandosi di una regione nord-occidentale, si va allineando al modello meridionale, con un numero di 65enni previsto inferiore a quello dei sopravvissuti lungo quasi tutto l'arco di previsione (solo dal 2047 lo scarto si fa positivo ma non supera l'11% dei previsti). Abruzzo, Molise e Sardegna, tradizionalmente connotate da ampi flussi in uscita, sembrano destinate a trasformarsi (almeno per quanto riguarda la popolazione anziana) in aree di attrazione. In Abruzzo, dal 2024 il numero di previsti supererà quello dei sopravvissuti e tale differenza arriverà al 47% dei 65enni previsti nel 2052. Per il Molise il mutamento sarà meno rapido ma più intenso: ancora nel 2017 la prevalenza dei sopravvissuti sui previsti sarà pari all'80% di questi ultimi ma dal 2032, anno in cui le poste si pareggeranno, il numero dei 65enni previsti sopravvanzerà quello dei sopravvissuti, con uno scarto che raggiungerà il 25% alla fine del periodo considerato. Infine, per la Sardegna si osserva un cambiamento più equilibrato: da un differenziale negativo tra previsti e sopravvissuti, pari al 25% dei primi, registrato nel 2017 si arriverà a un analogo (26%) differenziale positivo nel 2064. In definitiva, comunque, l'incidenza complessiva del contributo di queste tre regioni alla formazione del surplus di 65enni previsti nel 2064 non supera le 12.000 unità (4,5% del surplus complessivo italiano).

² Ad esempio, per i nati in Emilia Romagna si registra nel decennio 1970-1980 un tasso di decremento medio annuo pari a -51 per 1000, contro il -34 per 1000 registrato per il complesso dei nati in Italia.

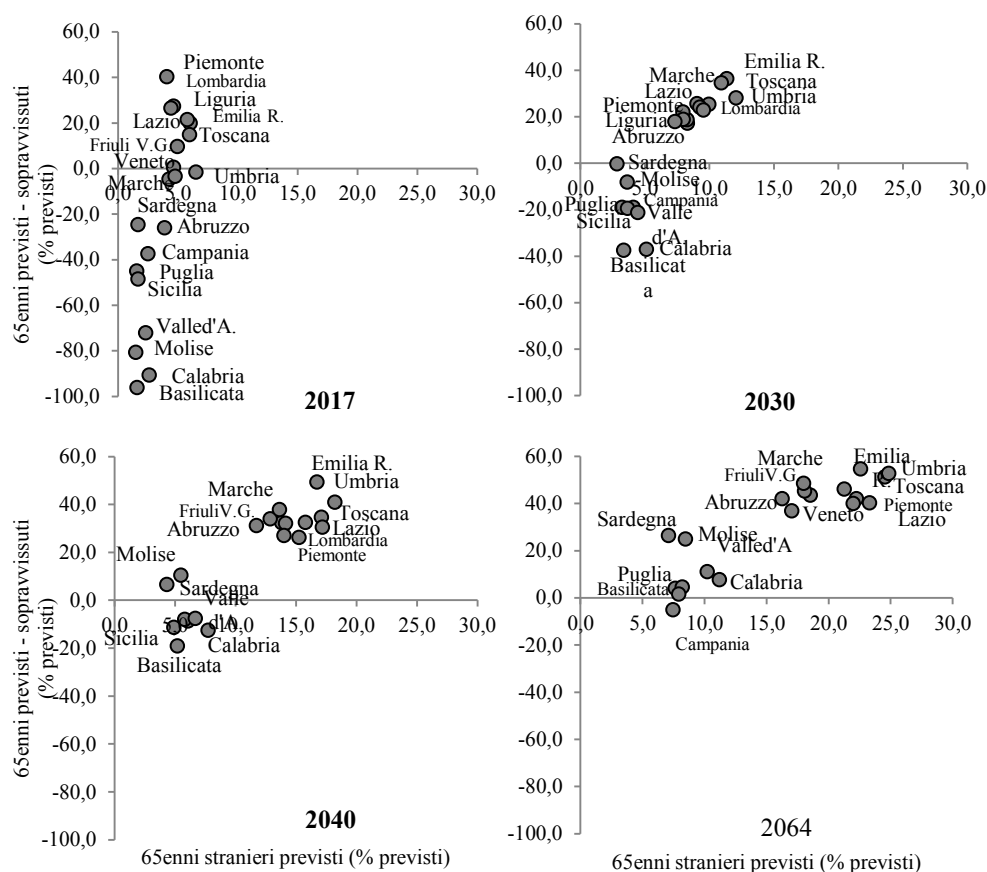
5. Invecchiamento e stranieri

L'evoluzione dell'invecchiamento importato dipende sia dalla dinamica delle migrazioni interne (e dagli effetti strutturali a esse correlati), sia dalla proporzione di stranieri che ogni anno si aggiungono al contingente di anziani.

In corrispondenza del primo anno disponibile per l'analisi, il 2017, la dimensione relativa dell'invecchiamento importato è scarsamente influenzata dalla proporzione di stranieri tra i 65enni previsti (figura 3). I differenziali regionali, positivi per le regioni del nord e negativi per le regioni del sud, sono attribuibili, essenzialmente, ai grandi spostamenti dal sud verso il nord che hanno caratterizzato l'Italia del dopoguerra. Le regioni del sud sono caratterizzate da saldi negativi che, nel caso della Calabria e della Basilicata, ad esempio, sfiorano il 100% dei 65enni previsti. Nel nord-ovest, al contrario, la quota di invecchiamento importato varia tra il 20% (Liguria) e il 40% (Piemonte e Lombardia). Saldi positivi, seppure di minore entità, si registrano nelle regioni centrali: Emilia Romagna, Umbria e Marche. Bisognerà attendere oltre un decennio affinché si possa apprezzare il cambiamento: nel 2030 l'incidenza della componente straniera sull'invecchiamento comincia a farsi sentire nel centro-nord: tuttavia in Piemonte e Lombardia l'effetto "stranieri" non è ancora arrivato a compensare il venir meno dei flussi provenienti dalle altre regioni italiane, pertanto entrambe godono di circostanze contingenti relative alla struttura per età temporaneamente più favorevoli. In Veneto, al contrario, come in gran parte delle regioni centrali, la componente straniera determina un severo aumento dell'invecchiamento "importato". Le regioni del sud continuano a mantenere cospicui saldi negativi (-40% circa), nonostante sostanziali recuperi da parte di Puglia e Sicilia, mentre l'Abruzzo comincia a manifestare saldi positivi. Dieci anni più tardi, nel 2040, gli effetti dell'invecchiamento importato dall'estero si fa sentire quasi ovunque: nelle regioni del nord-ovest e centro (tra il 15% e il 20% dei 65enni) esso riporta l'invecchiamento importato sopra la soglia del 20% (ad esempio in Lombardia e Piemonte) e fino al 50% (nel caso dell'Emilia R.). Sul fronte opposto, nelle regioni del sud, si osserva una generale riduzione del credito di 65enni, imputabile significativamente alla componente straniera, solo per la Calabria (con circa l'8% dei 65enni stranieri).

Il termine estremo dell'intervallo di previsione, il 2064, vede il gruppo delle regioni del centro-nord caratterizzato da una quota di invecchiamento importato tra il 37% (Trentino A.A.) e il 55% (Emilia Romagna) e una proporzione di 65enni stranieri tra il 16% (Abruzzo) e il 25% (Umbria). L'invecchiamento importato sarà positivo anche per le regioni del sud (ancorché non superiore al 26% della Sardegna), per cui la proporzione di stranieri tra i 65enni sarà compresa tra il 7% (Puglia) e l'11% (Calabria), con l'unica eccezione della Campania, che in proposito fa registrare ancora un saldo negativo.

Figura 3 – Scarto 65enni previsti – sopravvissuti (% previsti) e 65enni stranieri previsti (% previsti) nelle regioni italiane. 2017-2064



Note: Alcuni nomi di regioni sono omessi per sovrapposizione
 Fonte: n/elaborazioni su dati Istat

5. Conclusioni

Per l'Italia in complesso, lo scarto fra ingressi previsti nell'età anziana e sopravvissuti, positivo dal 2019, aumenterà fino a raggiungere nel 2054 il picco di 260.000 unità, per poi attestarsi a 250.000. Fino al 2019, dunque, l'Italia potrebbe ancora approfittare della "finestra demografica" aperta (questa volta sull'invecchiamento) da un saldo migratorio con l'estero che è stato a lungo negativo. I modelli evolutivi che contraddistinguono il dettaglio regionale ricalcano

l'usuale suddivisione del territorio in macro-aree, secondo una scala dell'intensità dell'invecchiamento importato decrescente da nord verso sud. Tale scala, in principio determinata dalle migrazioni interne risulta via via sempre più influenzata dai saldi migratori con l'estero. In questi termini, il processo di polarizzazione delle regioni, osservabile chiaramente già a partire dal 2030, rischia di riportare alla ribalta l'immagine di un'Italia spaccata a metà: la linea di demarcazione separerebbe l'Abruzzo, aggregato al gruppo delle regioni del centro-nord (che ha come punta d'avanguardia l'Umbria), dal Molise che invece rimarrebbe ancorato ai modelli evolutivi delle regioni meridionali (che trovano l'estrema espressione nel comportamento della Campania).

Riferimenti bibliografici

- BLANGIARDO G.C. 2012. Gli aspetti statistici. In FONDAZIONE ISMU (Ed.) *Diciottesimo Rapporto sulle migrazioni 2012*, Milano: Franco Angeli, pp.37-54.
- GESANO G., STROZZA S. 2011. Foreign migrations and population aging in Italy. *Genus*. LXVII, n. 3, pp. 83-104.
- ISTAT 2011. Il futuro demografico del Paese: previsioni regionali della popolazione residente al 2065, <http://demo.istat.it/>
- PATERNO A. 2011. Is immigration the solution to population aging?. *Genus*. LXVII, n.3, pp. 65-82.

SUMMARY

The imported aging in the Italian regions

The gap between the projected number of 65 years old and the hypothetical survived population will be positive starting from 2019 and it will reach the amount of 250,000 by 2064. The regional evolutionary paths follow a decreasing intensity from North to South. This scale, in principle determined by internal migrations is increasingly influenced by net migration with foreign countries.

Gian Carlo BLANGIARDO, Università Milano Bicocca,
giancarlo.blangiardo@unimib.it
Stefania RIMOLDI, Università Milano Bicocca, stefania.rimoldi@unimib.it

LIMITAZIONI NELLA PARTECIPAZIONE AL MERCATO DEL LAVORO DELLE PERSONE CON PROBLEMI DI SALUTE IN ITALIA

Barbara Boschetto, Antonella Iorio, Carlo Lucarelli, Michele Antonio Salvatore

1. Introduzione

L'inserimento nel mondo del lavoro rappresenta un fattore determinante per l'integrazione sociale delle persone con problemi di salute o limitazioni dell'autonomia personale.

In conformità alla concezione biopsicosociale di salute, la nuova Classificazione Internazionale del Funzionamento, della Disabilità e della Salute (ICF), elaborata dall'Organizzazione Mondiale della Sanità (OMS) nel 2001, individua, come dimensione fondamentale per la definizione di disabilità, le restrizioni nella partecipazione alla vita collettiva e dunque anche lavorativa, dando particolare rilievo all'influenza dei fattori contestuali quali mancanza di servizi, presenza di barriere architettoniche, mancanza di agevolazioni nell'accesso all'istruzione e al lavoro.

La presenza di limitazioni funzionali risulta avere un forte impatto sulle difficoltà di inserimento lavorativo, determinando quote consistentemente più elevate di esclusi rispetto alla popolazione complessiva; tali quote risultano, inoltre, crescenti in presenza di maggiori livelli di gravità delle limitazioni (Istat 2012).

Dal punto di vista normativo, a partire dagli anni '90 diverse misure sono state adottate nel nostro Paese al fine di garantire l'inserimento professionale e condizioni di lavoro agevolate per chi presenta problemi di salute e limitazioni fisiche o mentali.

Tra queste assumono particolare rilevanza le leggi a favore delle persone con disabilità come la legge 104/92, definita legge-quadro per l'assistenza, l'integrazione sociale e i diritti delle persone handicappate, e la legge 68/99, diretta a promuovere servizi di sostegno e di collocamento mirato.

L'obiettivo del presente studio è quello di confrontare, nell'ambito delle forze di lavoro in Italia, comprensive anche dei cosiddetti inattivi appartenenti alla zona

grigia¹, l'occorrenza della disoccupazione e dell'appartenenza a tale zona tra coloro che dichiarano di avere problemi di salute e/o difficoltà funzionali e coloro che dichiarano di non averne, nonché identificare i principali fattori di rischio tra gli individui con problemi e/o difficoltà bisognosi di forme di assistenza per poter lavorare.

2. Dati e metodi

Nel presente lavoro sono stati utilizzati i dati della Rilevazione sulle Forze di Lavoro in Italia relativa al secondo trimestre 2011.

L'indagine, condotta dall'Istituto Nazionale di Statistica, fornisce le stime ufficiali dei principali aggregati dell'offerta di lavoro rappresentati da occupati, persone in cerca di occupazione e persone inattive, rilevando le principali caratteristiche socio-demografiche e lavorative. Le unità di rilevazione sono rappresentate dalle famiglie residenti in Italia, all'interno delle quali vengono intervistati tutti i componenti di 15 anni e più; sono escluse le persone che vivono permanentemente in convivenze. L'indagine è di tipo continuo, con cadenza trimestrale, e utilizza un disegno di campionamento a due stadi comuni-famiglie con stratificazione dei comuni a livello provinciale sulla base dell'ampiezza demografica.

Nel secondo trimestre 2011 è stato incluso un modulo *ad hoc* diretto ad approfondire, tra gli individui di età 15-64 anni, il tema della partecipazione al mercato del lavoro delle persone con problemi di salute o difficoltà funzionali in grado di condizionare le abituali attività della vita quotidiana (Istat 2013a). Tramite tale modulo è stato innanzi tutto possibile individuare coloro che presentano questo tipo di problemi e rilevare le limitazioni sul lavoro che ne derivano, nonché l'assistenza offerta o di cui avrebbero bisogno.

Il campione dell'indagine è composto da 77.720 famiglie per un totale di 142.094 individui di 15 anni e più. Tra gli individui di età 15-64 anni cui è stato rivolto il modulo hanno accettato di rispondere in 92.898 (per un tasso di risposta del 93,2% sul totale degli individui di 15-64 anni).

Ai fini dell'analisi relativa al presente studio, il campione è stato ristretto alle forze di lavoro (occupati e persone in cerca di occupazione), a cui è stata aggiunta

¹Per zona grigia dell'inattività si intende quel gruppo di persone che, pur non disponendo dei requisiti per rientrare nell'aggregato dei disoccupati, mostrano un certo grado di *attachment* al mercato del lavoro. Tale categoria comprende coloro che cercano lavoro pur non avendo fatto azioni attive di ricerca nelle ultime quattro settimane, coloro che cercano lavoro ma non sono immediatamente disponibili ad accettarne uno entro le due settimane successive e coloro che non cercano lavoro ma sarebbero disponibili ad accettarne uno se venisse loro proposto.

la zona grigia dell'inattività (di cui alla nota 1). Si parlerà dunque in seguito di forze lavoro estese, ossia comprensive degli inattivi appartenenti alla suddetta zona.

A seguito di tale selezione, sono stati considerati nell'analisi 62.902 individui, di cui 52.838 che dichiarano di non avere problemi di salute o difficoltà funzionali e 10.064 che invece presentano una o più malattie croniche o problemi di salute della durata di almeno sei mesi e/o una o più difficoltà funzionali. Questi ultimi sono stati classificati per livello di gravità sulla base del bisogno o meno di assistenza per poter lavorare (usufruiscono o avrebbero bisogno di almeno una delle tre seguenti forme di assistenza: persone che forniscono assistenza; attrezzature speciali o adattamenti dell'ambiente lavorativo; modalità di lavoro flessibile quali lavori sedentari, telelavoro, lavoro meno faticoso, ecc.).

L'associazione tra la condizione occupazionale e le condizioni di salute è stata analizzata mediante modelli di regressione multinomiale, che hanno permesso di calcolare gli odds ratio (OR) con i relativi intervalli di confidenza al 95%, aggiustati per i potenziali confondenti.

La variabile *outcome* è rappresentata dalla condizione occupazionale (comprendente le seguenti modalità: occupati, persone in cerca di occupazione e zona grigia); i fattori confondenti considerati nell'analisi sono l'età, il genere, la ripartizione geografica di residenza e il titolo di studio.

L'analisi è stata condotta prima sulla popolazione complessiva considerata e successivamente sui soli individui con problemi di salute o difficoltà funzionali bisognosi di assistenza al fine di indagare il ruolo dei fattori di rischio considerati all'interno di tale sottopopolazione.

3. Risultati

Nella tabella 1 è riportata la distribuzione delle forze di lavoro considerate per le variabili in studio, secondo le condizioni di salute.

Oltre un quinto degli individui con problemi di salute o difficoltà funzionali (il 23% tra i bisognosi di assistenza per poter lavorare e il 21% tra i non bisognosi) appartiene alla classe di età 55-64 anni contro il 10% di coloro che dichiarano di non avere problemi o difficoltà.

La quota di donne varia tra il 43% nella popolazione senza problemi o difficoltà e il 48% in quella bisognosa di assistenza.

Per quanto riguarda il livello di istruzione, oltre la metà dei bisognosi di assistenza (52%) e il 43% degli individui con problemi o difficoltà non bisognosi, presentano al più la licenza media inferiore, contro il 37% tra i "sani".

Tabella 1 - Distribuzione percentuale delle forze di lavoro estese per le variabili in studio, secondo le condizioni di salute. Rilevazione sulle Forze di Lavoro, Italia, secondo trimestre 2011

Variabili	Condizioni di salute		
	Senza problemi di salute o difficoltà funzionali	Con problemi di salute e/o difficoltà funzionali	
		Senza bisogno di assistenza	Con bisogno di assistenza
	[n=52.838]	[n=8685]	[n=1379]
Età			
15-29 anni	19,9	7,1	8,0
30-54 anni	70,4	72,1	69,4
55-64 anni	9,7	20,8	22,6
Genere			
Maschio	56,9	53,8	51,9
Femmina	43,1	46,2	48,1
Livello di istruzione			
Al più licenza media inferiore	36,8	42,9	52,2
Diploma di scuola superiore	46,6	42,4	37,7
Almeno la laurea	16,7	14,7	10,1
Ripartizione geografica residenza			
Nord	47,2	45,7	47,8
Centro	19,2	24,5	15,9
Mezzogiorno	33,6	29,8	36,3
Condizione professionale			
Occupato	82,2	84,6	60,9
In cerca di occupazione	7,2	5,7	12,1
Zona grigia di inattività	10,6	9,7	27,1
Presenza di problemi e/o difficoltà			
Almeno un problema, nessuna difficoltà	-	59,5	25,7
Nessun problema, almeno una difficoltà	-	6,3	3,5
Almeno un problema, almeno una diffic.	-	34,2	70,8
Totale	100,0	100,0	100,0

Le quote di disoccupati e inattivi appartenenti alla zona grigia sono nettamente più elevate tra coloro che necessitano di assistenza, pari rispettivamente a 12% e 27%. Tra gli individui senza problemi o difficoltà, i disoccupati rappresentano il 7% delle forze di lavoro considerate, mentre gli inattivi della zona grigia sono l'11%; tali quote non si discostano da quelle osservate per la popolazione con

problemi non bisognosa di assistenza, per la quale sono pari rispettivamente a 6% e 10%.

Tabella 2 – Odds ratio (OR) di disoccupazione (a sinistra) e di appartenenza alla zona grigia di inattività (a destra) mutuamente aggiustati per le variabili selezionate. Modello di regressione multinomiale. Rilevazione sulle Forze di Lavoro, Italia, secondo trimestre 2011 - Individui con problemi di salute e/o difficoltà funzionali bisognosi di assistenza

Variabili	Disoccupazione		Zona grigia di inattività	
	OR	IC 95%	OR	IC 95%
Presenza di problemi e/o difficoltà				
(Rif. Almeno un problema, nessuna difficoltà)	1,00	-	1,00	-
Nessun problema, almeno una difficoltà	1,54	(0,61-3,88)	0,42	(0,15-1,15)
Almeno un problema, almeno una difficoltà	1,45	(0,93-2,27)	0,95	(0,69-1,30)
Età				
(Rif. 55-64 anni)	1,00	-	1,00	-
15-29 anni	4,13	(2,18-7,81)	1,49	(0,84-2,66)
30-54 anni	1,16	(0,73-1,85)	0,75	(0,54-1,04)
Genere				
(Rif. Maschio)	1,00	-	1,00	-
Femmina	0,91	(0,63-1,30)	2,11	(1,58-2,81)
Livello d'istruzione				
(Rif. Almeno la laurea)	1,00	-	1,00	-
Al più licenza media inferiore	1,26	(0,72-2,21)	10,65	(5,16-22,02)
Diploma scuola superiore	0,84	(0,47-1,49)	4,38	(2,09-9,15)
Ripartizione geografica				
(Rif. Nord)	1,00	-	1,00	-
Centro	1,17	(0,72-1,91)	2,43	(1,61-3,67)
Mezzogiorno	1,15	(0,76-1,73)	5,43	(3,95-7,47)

Tra gli individui che necessitano di assistenza il 71% presenta sia problemi che difficoltà, contro il 34% di coloro non bisognosi di assistenza, tra i quali il 60% presenta almeno un problema ma nessuna difficoltà.

In entrambe le sottopopolazioni i problemi di salute considerati più gravi sono rappresentati da problemi alla schiena o al collo (inclusi artrite o reumatismi), problemi alle gambe o ai piedi (inclusi artrite o reumatismi) e malattie del cuore, ipertensione o problemi circolatori. Le principali attività in cui vengono dichiarate

le maggiori difficoltà, in entrambe le sottopopolazioni, sono legate al movimento (sollevare e trasportare qualcosa, camminare e salire le scale, piegarsi e chinarsi).

Per quanto riguarda l'associazione tra condizioni di salute e condizione occupazionale, il rischio di disoccupazione, in termini di odds ratio aggiustati per età, genere, titolo di studio e ripartizione geografica di residenza, è risultato, per gli individui con problemi bisognosi di assistenza, pari a 3 volte quello dei sani (IC95%:2,5-3,6), mentre il rischio di rientrare nella zona grigia di inattività è risultato pari a 4,2 volte (IC95%:3,7-4,9) (dati non riportati in tabella).

L'occorrenza della disoccupazione e quella della condizione di appartenenza alla zona grigia non differiscono significativamente per gli individui con problemi o difficoltà non bisognosi di assistenza, assumendo sempre i "sani" come riferimento (dati non riportati in tabella).

La tabella 2 riporta i risultati dell'analisi condotta all'interno della popolazione con problemi o difficoltà bisognosa di assistenza. Le variabili considerate risultano tutte associate alla condizione occupazionale, ad eccezione della variabile "Presenza di problemi e/o difficoltà". L'occorrenza della disoccupazione tra gli individui in età 15-29 anni risulta pari a 4,1 volte quella degli individui in età 55-64 anni. Non si rilevano differenziali significativi in corrispondenza delle altre variabili considerate.

L'occorrenza della condizione di inattivo appartenente alla zona grigia non differisce significativamente tra le diverse classi di età. Le donne presentano un rischio (in termini di odds ratio) maggiore rispetto a quello degli uomini (OR pari a 2,1). I residenti nelle regioni del Centro e del Mezzogiorno presentano odds pari a 2,4 e 5,4 volte quello di chi risiede nel Nord. Per quanto riguarda il livello d'istruzione, coloro che possiedono al più l'obbligo scolastico e coloro che hanno conseguito il diploma di scuola media superiore presentano rispettivamente OR pari a 10,7 e 4,4 con riferimento a chi possiede almeno la laurea.

4. Conclusioni

Lo studio ha evidenziato, nell'ambito delle forze di lavoro estese in Italia, la presenza di forti criticità nell'inserimento lavorativo delle persone bisognose di forme di assistenza per poter lavorare. Tra queste ultime sono state rilevate quote consistentemente più elevate di disoccupati e soprattutto di inattivi appartenenti alla zona grigia. Gli inattivi considerati comprendono principalmente individui disponibili a lavorare ma scoraggiati nella ricerca del lavoro (Istat 2013b). La maggiore occorrenza della disoccupazione e dell'inattività rappresentata dalla zona grigia, tra i bisognosi di assistenza, è stata rilevata anche a parità di fattori di rischio.

Il bisogno di assistenza per poter lavorare determina la maggiore occorrenza della disoccupazione e non si rilevano, tra i bisognosi, i tradizionali differenziali di genere, istruzione e territorio che caratterizzano il nostro Paese; solo tra i giovani di 15-24 anni si rileva un maggiore svantaggio, in linea con gli elevati livelli generali di disoccupazione giovanile.

Per quanto riguarda l'occorrenza dell'inattività rappresentata dalla zona grigia, invece, i differenziali emergono e si osserva, tra i bisognosi di assistenza, il maggior svantaggio delle donne, dei residenti nel Mezzogiorno e soprattutto dei meno istruiti.

Gli individui bisognosi di assistenza residenti nelle regioni del Mezzogiorno risultano con molta probabilità maggiormente scoraggiati, oltre che per le minori opportunità offerte da una più bassa domanda di lavoro che caratterizza questa zona del Paese, per la minore attenzione volta all'attivazione dei servizi previsti dalla legge riguardo all'assistenza e all'inserimento lavorativo delle persone con problemi di salute (Ministero del Lavoro e delle Politiche Sociali e Isfol 2012).

Tra i bisognosi di assistenza i meno istruiti risultano, con molta probabilità, demoralizzati dalle maggiori difficoltà di trovare lavori non manuali e dunque compatibili con le loro limitazioni fisiche; è da sottolineare che allo stesso tempo il più basso livello di istruzione potrebbe essere conseguenza delle cattive condizioni di salute (Istat 2011).

Le persone bisognose di assistenza presentano molto probabilmente disabilità o comunque problemi di salute fortemente invalidanti, di qui dunque la necessità di monitorare e garantire l'attuazione delle norme in materia di assistenza e inserimento lavorativo. Favorire di fatto l'accesso al mercato del lavoro di questo segmento della popolazione rappresenterebbe, oltre che un obiettivo di giustizia sociale, un sostegno all'occupazione in un paese che tradizionalmente mostra criticità sotto questo aspetto.

Riferimenti bibliografici

- ISTAT. 2013a. Limitazioni nello svolgimento dell'attività lavorativa delle persone con problemi di salute. Anno 2011. *Statistiche Report*. Roma.
- ISTAT. 2013b. Disoccupati, inattivi, sottoccupati: indicatori complementari al tasso di disoccupazione. Anno 2012. *Statistiche Report*. Roma.
- ISTAT. 2012. Inclusione sociale delle persone con limitazioni dell'autonomia personale. Anno 2011. *Statistiche Report*. Roma.
- ISTAT. 2011. L'integrazione degli alunni con disabilità nelle scuole primarie e secondarie di I grado, statali e non statali - Anni scolastici 2008/2009 e 2009/2010. *Statistiche in breve*. Roma.

Ministero del Lavoro e delle Politiche Sociali, ISFOL. 2012. *Relazione sullo stato di attuazione della legge recante norme per il diritto al lavoro dei disabili 2010-2011*. Roma, Senato della Repubblica.
<http://sbnlo2.cilea.it/bw5ne2/opac.aspx?WEB=ISFL&IDS=18878>.

SUMMARY

Limitation in work participation related to health conditions in Italy

Among the Italian labour force including the gray zone of inactivity, we compare the occurrence of unemployment and the occurrence of belonging to the gray zone between those who report health problems and/or difficulties in basic activities and those who do not report any of them. Moreover we evaluate the role of selected risk factors among the subgroup of people with health problems and/or difficulties in need of special assistance at the workplace.

Using data from the 2011 Italian Labour Force Survey conducted by the Italian National Institute of Statistics, adjusted odds ratios (ORs) for socio-demographic variables are estimated among a nationally representative sample of 62,902 persons aged 15-64 years.

The occurrence of unemployment and belonging to the gray zone of inactivity are higher among individuals with health problems and/or difficulties in need of assistance at the workplace compared to individuals without problems or difficulties: adjusted ORs are 3.0 and 4.2. There are no significant differences for individuals with problems and/or difficulties who do not require assistance compared to individuals without problems or difficulties.

Among people who need assistance, young persons aged 15-24 years show a higher occurrence of unemployment, while women, those living in southern Italy and those with a low level of education denote a higher occurrence of belonging to the gray zone.

The study highlighted the need for monitoring and ensuring the implementation of laws which support assistance at work and job placement of people with work-limiting health problems.

Barbara BOSCHETTO, Istituto Nazionale di Statistica, boschett@istat.it

Antonella IORIO, Istituto Nazionale di Statistica, iorio@istat.it

Carlo LUCARELLI, Istituto Nazionale di Statistica, calucare@istat.it

Michele Antonio SALVATORE, Istituto Nazionale di Statistica, salvatore@istat.it

L'UTILIZZO DELLE LISTE ANAGRAFICHE COMUNALI NELLA STATISTICA UFFICIALE¹

Claudio Ceccarelli, Anna Pezone, Simona Rosati

1. Introduzione

L'impiego della LAC (Lista Anagrafica Comunale) costituisce una delle principali innovazioni dell'ultimo Censimento generale della popolazione e delle abitazioni (2011). Pur mantenendo inalterate le caratteristiche di esaustività e simultaneità rispetto al censimento tradizionale, il 15° Censimento generale della popolazione appare totalmente rinnovato nei metodi e nelle tecniche. È stato realizzato, infatti, secondo la tecnica assistita da lista.

Gli errori di copertura, che inevitabilmente derivano dalla tecnica di rilevazione assistita da lista, sono stati affrontati ricorrendo all'utilizzo di liste ausiliarie, rispetto alle quali è stato effettuato il confronto, al fine di recuperare i dati di chi, pur avendo dimora abituale in Italia, non risulta iscritto in anagrafe. Analogamente si è tenuto conto dei cambiamenti di stato e di residenza, avvenuti tra la data di riferimento della LAC e quella del censimento, per aggiornare la lista. In tutti questi casi, gli operatori censuari di ciascun comune venivano inviati presso le famiglie o convivenze per la consegna e la compilazione del questionario.

Dopo una rilevazione sperimentale condotta su ben 1.850 comuni, l'Istat ha acquisito le liste anagrafiche comunali sia a gennaio 2011 (con riferimento al 31 dicembre 2010), per poter predisporre il piano di stampa e spedizione dei questionari e caricare i dati nel Sistema di Gestione della Rilevazione (SGR), sia in prossimità della data di riferimento della rilevazione censuaria, per poter aggiornare i dati di SGR con le variazioni anagrafiche intervenute nell'intervallo di tempo considerato. Attualmente l'Istat ha completato la rilevazione delle liste anagrafiche comunali con riferimento alla data del 1° gennaio 2013.

Lo scopo di questo lavoro è duplice: da un lato descrivere lo sforzo compiuto, in termini di correzione, integrazione e standardizzazione dei dati, per trasformare la

¹ L'articolo è stato realizzato dal lavoro congiunto degli autori. Ai soli fini dell'attribuzione: i paragrafi 1 e 2 sono stati redatti da Anna Pezone, il paragrafo 3.1 da Simona Rosati e il paragrafo 3.2 da Claudio Ceccarelli.

LAC in un archivio di dati confacente alle esigenze del censimento; dall'altro illustrare il potenziale impiego della LAC nell'ambito delle rilevazioni su popolazione e famiglie.

2. La Lista Anagrafica Comunale

La LAC è un archivio che contiene informazioni relative a tutti gli individui residenti nel territorio di ciascun comune italiano, distinti per famiglia o convivenza. Queste informazioni, presenti nelle liste anagrafiche dei comuni, sono state trasmesse all'Istat per mezzo di STARLAC, un'applicazione *web* che consente di: i) acquisire i dati in formato elettronico, nel rispetto dei più elevati *standard* di sicurezza; ii) supportare i comuni durante la fase di trasmissione dei dati; iii) monitorare lo stato di avanzamento della rilevazione; iv) produrre la reportistica sulla qualità dei dati trasmessi. Gli *standard* tecnici richiesti per la produzione del *file*, determinati in base alle esperienze acquisite, sono stati scelti per essere: sostenibili, chiari e imprescindibili per tutti i comuni. Gli *standard* prevedono il rispetto delle seguenti specifiche tecniche: a) nome, struttura e formato del *file*, incluse le istruzioni di codifica di alcune variabili (relazione di parentela, stato civile e sesso); b) tipo, formato e lunghezza delle variabili.

STARLAC consente inoltre di eseguire operazioni quali: il controllo dei dati trasmessi; il supporto per la riclassificazione di alcune variabili; l'acquisizione di informazioni aggiuntive (per esempio: il *software* utilizzato dai comuni per la gestione dei dati e l'elenco degli indirizzi fittizi presso i quali sono iscritti i senza fissa dimora); il monitoraggio dello stato di avanzamento delle attività; il resoconto dettagliato sull'analisi di qualità dei dati trasmessi.

Durante la fase di trasmissione STARLAC consente di individuare due tipi di errore: a) errori dovuti al mancato rispetto dei requisiti *standard* richiesti; b) errori per mancata conformità alle disposizioni di legge in materia di anagrafe (Regolamento anagrafico della popolazione residente²). Le liste inviate secondo le specifiche tecniche richieste, o con errori di tipo b) entro soglie di tolleranza prestabilite, si considerano acquisite; la correttezza dell'avvenuta ricezione viene confermata da un messaggio di posta elettronica.

Successivamente le LAC acquisite al 31 dicembre 2010, necessarie alla rilevazione censuaria, sono state sottoposte ad ulteriori elaborazioni mediante l'applicazione TRASLAC. Ciò al fine di verificare: la congruità dei dati aggregati

² Il *Regolamento Anagrafico* (D.P.R. n.223 del 30 maggio 1989) stabilisce per i comuni le regole di compilazione delle schede anagrafiche. Tuttavia, a causa della diversa tempistica con cui i comuni sono passati dal sistema cartaceo a quello elettronico, il processo di armonizzazione delle liste è ancora in corso.

rispetto a fonti di informazione esterne; la correttezza e coerenza delle variabili non controllabili in fase di acquisizione. Il trattamento post-acquisizione ha inoltre permesso di: individuare le unità di censimento (individui, famiglie o convivenze); imputare la “relazione di parentela” in caso di “capofamiglia” mancante o multiplo; standardizzare e geo-codificare gli indirizzi; estrarre il campione di famiglie alle quali sottoporre il questionario in forma abbreviata; individuare il modello di questionario in base all’ampiezza del nucleo familiare (3 o 6 persone); assegnare il codice identificativo unico per questionario; creare le *password* per l’accesso al questionario *online*; stampare e spedire i questionari; caricare i dati in SGR.

L’archivio finale delle LAC riferite al 31 dicembre 2010 contiene oltre 61 milioni di individui, 25 milioni di famiglie e circa 300.000 convivenze. Oltre il 93% dei comuni ha trasmesso i dati con successo al primo tentativo; soltanto pochi comuni hanno dovuto ritrasmettere a causa di errori.

Per aggiornare tempestivamente SGR durante le operazioni di censimento, tenendo conto delle variazioni anagrafiche (nascite, morti e trasferimenti di residenza) avvenute tra la fine del 2010 (data di riferimento della prima acquisizione delle LAC) e la data di riferimento del censimento (9 ottobre 2011), l’Istat ha adottato tre diverse strategie: i) i comuni con meno di 5.000 abitanti dovevano inserire manualmente tutti i cambiamenti demografici avvenuti fino all’8 ottobre 2011; ii) i comuni con popolazione compresa tra 5.000 e 20.000 abitanti potevano scegliere se registrare manualmente le variazioni anagrafiche o inviare una seconda LAC con dati riferiti all’8 ottobre³; iii) i comuni con più di 20.000 abitanti dovevano obbligatoriamente inviare una nuova LAC. Per tutti i 2.438 comuni con oltre 5.000 abitanti (per un totale di circa 49 milioni di individui) l’aggiornamento di SGR è stato effettuato dall’Istat, che ha identificato tutte le variazioni demografiche mediante procedure di *linkage* a livello individuale e familiare. Le chiavi di *linkage* utilizzate sono state nell’ordine: 1) codice identificativo individuale attribuito dal comune; 2) codice fiscale e stringa completa composta da: cognome, nome, sesso e data di nascita. Per determinare, invece, i cambiamenti di residenza è stata utilizzata una procedura di abbinamento più sofisticata che ha messo a confronto gli indirizzi grezzi, quelli normalizzati e i codici identificativi del toponimo.

L’insieme dei dati individuali contenuti nella LAC, i risultati della rilevazione registrati su SGR e il confronto censimento-anagrafe, eseguito contestualmente alla rilevazione, hanno consentito, per la prima volta in Italia, di pubblicare a distanza di un anno dalla data di riferimento del censimento, la popolazione legale dei comuni italiani per sesso, età e cittadinanza (italiana e straniera).

³ Il 90% dei comuni di questa classe demografica ha scelto di inviare una nuova LAC e lasciare che fosse l’Istat a individuare i cambiamenti avvenuti.

Dall'analisi della distribuzione dei comuni rispetto alle date di acquisizione delle LAC, con riferimento alle tre principali rilevazioni (riferite al 31 dicembre 2010, all'8 ottobre 2011 e per la rilevazione in corso al 1° gennaio 2013), si evince che il processo di trasmissione/acquisizione è ormai consolidato ed è un'attività che oltre l'85-90% dei comuni riesce a svolgere entro il termine stabilito.

La rilevazione delle LAC è prevista nel Programma Statistico Nazionale fino al 2015; i dati della rilevazione sono, fra l'altro, utilizzati per modulare le strategie campionarie che precedono il Censimento permanente del 2016 e che saranno applicate nelle sperimentazioni del 2014-2015. All'interno di questo quadro di attività, riferito esclusivamente all'Istat, si dovrà tener presente che, in base al D.L. n. 179 del 18 ottobre 2012, convertito in legge n.221 - art.2 - del 17 dicembre 2012, dal 1° Gennaio 2015 si prevede a regime il funzionamento dell'Anagrafe Nazionale della Popolazione Residente (ANPR), cioè l'archivio anagrafico centralizzato, sul quale avrà il pieno controllo e la gestione il Ministero dell'Interno, e sul quale i comuni opereranno tutti gli aggiornamenti. Tale archivio, che nel corso del 2015 sarà presente simultaneamente ai dati della rilevazione delle LAC del 2015, potrà rappresentare in futuro la base informativa unica dalla quale l'Istat potrà attingere le informazioni necessarie per il campionamento delle unità di rilevazione del censimento permanente e delle indagini campionarie correnti.

3. Impiego della LAC nelle rilevazioni statistiche su popolazione e famiglie

3.1. L'estrazione dei campioni

L'archivio delle LAC, costituito in occasione del 15° Censimento generale della popolazione e delle abitazioni, e da allora in periodico aggiornamento, rappresenta la principale base informativa per la fase di estrazione dei campioni di famiglie o individui per le indagini dell'Istat. Per semplicità di esposizione faremo riferimento a un campione a due stadi in cui le unità di primo stadio sono i comuni, mentre le unità di secondo stadio sono le famiglie anagrafiche, che a loro volta rappresentano le unità di rilevazione. Una volta selezionato il campione dei comuni e determinato il numero di famiglie per ciascun comune, occorre estrarre i nominativi delle famiglie.

In precedenza, anche nelle indagini maggiormente evolute, la selezione dei nominativi veniva effettuata da ogni singolo comune campione che, in base alla metodologia indicata dall'Istat, estraeva dalla propria lista anagrafica i nominativi delle famiglie da intervistare. Gli elenchi dei nominativi venivano poi inviati all'Istat, talvolta anche in formato cartaceo, e sottoposti a controlli di correttezza prima di essere assegnati ai rilevatori.

Da maggio 2011 il processo di estrazione dei campioni è svolto interamente all'interno dell'Istituto tramite l'impiego della LAC. Si tratta senza dubbio di un'innovazione nel processo di produzione delle indagini statistiche che comporta notevoli vantaggi, soprattutto per quanto riguarda la metodologia di estrazione adottata. Tra questi è da sottolineare la rapidità con cui si ottiene la lista completa dei nominativi, essendo il procedimento centralizzato e applicabile a qualsiasi indagine. Va poi evidenziato che un maggior controllo su tutto il processo di estrazione garantisce anche un miglior risultato in termini di qualità. Inoltre, vale la pena osservare che tale procedimento rende attuabile la definizione a priori di regole specifiche per il sovra-campionamento. Ciò al fine di evitare che alcune caratteristiche della popolazione di riferimento non rappresentino adeguatamente il fenomeno oggetto di indagine: è il caso, ad esempio, della cittadinanza. È noto, infatti, come il fenomeno delle mancate risposte totali sia molto più diffuso fra i cittadini stranieri, che a causa dell'incomprensione della lingua o per la più elevata mobilità sul territorio, più frequentemente non rispondono o sfuggono all'intervista.

La presenza del codice fiscale per ciascun individuo rende più semplice l'integrazione e il confronto con dati provenienti da fonti amministrative, al fine di sopperire, da un lato, alla mancanza di informazioni, dall'altro di verificare la validità di un insieme di dati rilevati. Infine, sempre l'uso del codice fiscale consentirebbe di individuare, tra coloro che hanno già partecipato a una data rilevazione, sottoinsiemi specifici di famiglie o singoli individui da sottoporre a nuove interviste, con lo scopo di effettuare studi ad hoc longitudinali di approfondimento.

Per contro si evidenziano principalmente due inconvenienti:

- a. la discrepanza temporale tra la data di acquisizione delle LAC e quella riferita al momento dell'estrazione, che implica maggiori probabilità di trovare nel campione famiglie che, in quell'intervallo di tempo, hanno cambiato abitazione e che, dunque, non sarà possibile contattare;
- b. la complessità intrinseca delle elaborazioni, soprattutto per quanto riguarda l'intero processo di acquisizione e aggiornamento delle LAC, incluse le procedure di trattamento, controllo e correzione dei dati.

3.2. *La definizione delle stime campionarie*

La mancata risposta totale nelle indagini statistiche è divenuta nel tempo un attributo sempre più frequente. Di conseguenza, anche da un punto di vista teorico, è necessario che se ne tenga conto sin dalla fase di progettazione e realizzazione

del disegno di campionamento fino alla definizione del processo di stima (Särndal e Lundström, 2005).

Le indagini statistiche risentono di possibili distorsioni legate ai meccanismi che generano la mancata risposta totale. Sono molteplici i fattori che incidono e condizionano i comportamenti sia dei rilevatori sia dei potenziali rispondenti, che vanno a incrementare da un lato il numero dei mancati contatti e dall'altro il numero di rifiuti alla partecipazione.

La complessità delle indagini, la concomitanza di differenti tecniche di rilevazione, con l'utilizzo di differenti reti di rilevazione, impongono risposte metodologiche più articolate che prevedono l'utilizzo di informazioni più definite anche per i non rispondenti. In altri termini, lo sfruttamento di fonti aggiuntive può, se non risolvere, perlomeno limitare gli effetti distorsivi dei meccanismi di mancata risposta totale.

A livello di costruzione delle stime, già dall'utilizzo degli stimatori post-stratificati e soprattutto dopo l'introduzione degli stimatori calibrati (Deville e Särndal, 1992), l'utilizzo di informazioni ausiliarie provenienti da fonti esterne, anche campionarie, è ormai diventata una consuetudine per il trattamento delle mancate risposte totali (Kott, 2006; Ceccarelli, Cutillo, 2007; Ceccarelli et al. 2008).

È noto che i passaggi che portano dal campione teorico alla determinazione delle stime finali, o meglio, alla determinazione dei coefficienti di riporto all'universo, possono essere riassunti in due macro fasi ben distinte:

1. raccolte le informazioni dal campione oggetto di indagine e determinato il coefficiente di riporto all'universo da disegno, si procede ad adeguare il campione osservato a quello teorico in modo da tener conto della mancata risposta e ripristinare sia la numerosità campionaria sia la struttura del campione inizialmente estratto;
2. grazie all'uso di stimatori, più o meno complessi, che sfruttano informazioni ausiliarie, si arriva alla determinazione dei coefficienti di riporto all'universo finali che eliminano la distorsione da sottocopertura delle liste di campionamento.

Proprio nel contesto della prima fase si inseriscono le sperimentazioni sull'utilizzo della LAC per la determinazione dei coefficienti di correzione della mancata risposta totale. Senza perdere di generalità possiamo considerare che il numero di interviste realizzato si discosti dalla numerosità campionaria teorica a causa di: errori dovuti alla lista, mancati contatti, rifiuti a partecipare all'indagine.

Ognuna di queste grandezze è funzione di vari fattori. In particolare, gli errori di lista sono funzione del grado di aggiornamento delle liste stesse, da cui sono

estratti i campioni, e del disegno di indagine⁴. Il mancato contatto è funzione della capacità della rete di rilevazione di garantire l'adeguata capillarità sul territorio e di riuscire a contattare le famiglie rispetto alle consuetudini e agli stili di vita delle famiglie stesse. La mancata partecipazione, infine, è funzione del lavoro di sensibilizzazione da parte degli Istituti e delle Agenzia di statistica nonché della capacità di ogni singolo intervistatore di coinvolgere la famiglia nell'intervista.

Una metodologia proposta, in corso di sperimentazione, si basa essenzialmente sull'assunto che il meccanismo che genera la mancata risposta totale si possa considerare ragionevolmente *missing at random* (MAR) all'interno di gruppi omogenei di famiglie. Le informazioni contenute nella LAC, infatti, offrono la possibilità di costruire una serie di distribuzioni sulle caratteristiche delle famiglie, o degli individui, tali da poter essere utilizzate per la costruzione di sottogruppi della popolazione per i quali la mancata risposta sia MAR, a meno di fattori non controllabili, in modo da ridurre la distorsione dovuta a meccanismi non casuali di generazione delle mancate risposte (Särndal e Lundström, 2005, pp.94-96).

Nella fattispecie, per ridurre gli effetti distorsivi delle mancate risposte si sta sperimentando l'utilizzo della tipologia familiare, costruita in base al sesso e alla classe di età dei componenti. Come è noto, infatti, uno dei possibili meccanismi che producono la mancata risposta totale è senza dubbio legato al numero di componenti della famiglia: è più difficile, infatti, riuscire a contattare famiglie costituite da una sola persona che famiglie più numerose. Ed è per questo motivo che si prevede di inserire questa caratteristica anche nel procedimento di estrazione dei nominativi, al fine di individuare gruppi di famiglie omogenee, rispetto alla tipologia, per meglio governare il delicato processo di sostituzione delle unità non rispondenti e ridurre, quindi, gli effetti delle mancate risposte totali.

Riferimenti bibliografici

CECCARELLI C., COPPOLA L., CUTILLO A., DI LAUREA D. 2008. Combining survey and administrative data in the Italian Eu-Silc experience: Positive and critical aspects. In *Conference of European Statisticians - Working Session on Statistical Data Editing, 1-10*. UNECE. URL: <http://www.unece.org/fileadmin/DAM/stats/documents/2008/04/sde/wp.14.e.pdf>.

⁴ Ci si riferisce sostanzialmente al fatto che se un'indagine è continua per tutto l'anno e l'estrazione avviene una sola volta ad inizio anno, coloro che vengono intervistati al termine del periodo hanno probabilità maggiore di essersi trasferiti ad altro indirizzo rispetto a coloro che sono intervistati all'inizio del periodo.

- CECCARELLI C., CUTILLO A. 2007. Il Trattamento della Mancata Risposta Totale nell'Indagine Eu-Silc: Una Valutazione Tramite una Misura del Cambiamento. *Congiuntura.*, Vol.1° Trimetre, Udine: CREF, pp. 91-112.
- DEVILLE, J.C., SÄRNDAL, C.E. 1992. Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87 (418): 376–382.
- KOTT P.S. 2006. Using calibration weighting to adjust for nonresponse and coverage errors. *Survey Methodology*, 32 (2): 133–142.
- SÄRNDAL C.E., LUNDSTRÖM S. 2005. *Estimation in Survey with nonresponse*. New York: John Wiley & Sons Ltd.
- PEZONE A., PELLIZZARO F. 2012: *The use of population registers in the 15th italian census: challenges and preliminary evidence*. Joint UNECE/Eurostat Meeting on Population and Housing Censuses Expert Group Meeting on Censuses Using Registers, Geneva, Switzerland, 22-23 May 2012
- PEZONE A., SCAVALLI M. 2010: *Using Population Registers as a Population List: Some Results from the Pilot Survey 2009*. Joint UNECE/Eurostat Expert Group Meeting on Register-Based Censuses, The Hague, Netherlands, 10-11 May 2010.

SUMMARY

The use of municipal registers in official statistics

One of the major innovations introduced by the Italian national institute of statistics in the last Population Census round held in October 2011 was the use of administrative registers to assist census operations. Millions of households were no longer visited by enumerators but received their questionnaires by post at the address printed in the municipal population registers (henceforth LAC) of the municipality where the household head resided. Coverage errors due to births and deaths as well as changes of address occurred between the mailing-out date and the Census Date were tackled by instructing enumerators to deal with the variations during field operations.

In addition to census operations, the LAC are also used to support the process to realize the sample surveys on households and individuals. In practices, the LAC are used to drawing samples of households and individuals and to support the studies to test methods to reduce nonresponse bias.

Claudio CECCARELLI, Istituto nazionale di statistica, clceccar@istat.it

Anna PEZONE, Istituto nazionale di statistica, pezone@istat.it

Simona ROSATI, Istituto nazionale di statistica, sirosati@istat.it

GENDERED REMITTANCES OF ROMANIANS IN ITALY

Eralba Cela, Eros Moretti

1. Introduction

The collapse of the communist regimes in Central and Eastern European countries led to new forms of mobility (transborder, shuttle migration, circular migration) and a growing feminization of migrant flows. Indeed, for many Romanians, short-term migration to Western countries became a life-strategy meant both to supplement their insufficient incomes and to enhance social status in their communities of origin. In this framework Italy represents one of the main destination countries for Romanian migrants, hosting the largest Romanian community out of Romania.

However, these migration movements have developed in gendered ways because of many factors which differently affect the propensity to migrate of men and women. For sure the characteristics of the labor market in the receiving countries, especially in Southern Europe countries, where female immigration is partly the result of gendered division of labor, have shaped the migration paths of men and women (King et al. 2000). In Italy the existence of a gender specific demand in the labor market, in particular for women in the domestic sector, represents so far a strong pull factor that shapes women's migration trajectories. The establishment of female networks of mutual assistance constituted a further support to women migration processes; as far as Romanian migrants are concerned, after the Romania's admission in Europe in 2007, women became more involved in circular migration paths and acquired an important role in the process of fostering the "transnational family life" (Boccagni and Banfi 2011)

Nevertheless, women are almost invisible in the migration studies because of the theorization of their transnational links within the frame of transnational families, where families are considered more often as groups of kin-related people sharing the same values and interests rather than people with competing agendas depending on their gender, ages, and statuses (Hondagneu-Sotelo 1992). The household unit is indeed very important in understanding the migration process (Massey 1990), but the inequalities among family members, and the conflicts over,

material and symbolic resources within the household should be considered as well. The challenges triggered by migration in relation to the redefinition of male and female identities within transnational households may be better examined through a gender lens.

For sure, women become more visible in the transnational networks when they acquired a breadwinning role and thanks to remittance and transnational care networks (Aranda 2003). Many researchers indeed have focused on migrants' gendered behaviors toward remittances¹. Because the term 'transnationalism' refers to the migrants' simultaneous embeddedness in their societies of origin and destination (Levitt and Glick Schiller 2004), one way of studying transnationalism is to examine the remittances that migrants send back to their relatives left behind, in order to maintain active ties with them. We therefore focus on the differences and similarities between transnational practices of men and women, within the social field linking Romania and Italy.

Romanians represent the largest migrant population in Italy, with almost 1 million people (Istat 2011), and women outnumber men (55% and 45% respectively). Italy is also the most important sending country for Romanian migrant remittances. De Sousa and Duval (2010), based upon the data reported by the National Bank of Romania covering the period 2005-2009, estimated an average of remittances from Italy per semester of about 473.79 million U.S. dollars, that is 1.7 higher than the estimates of Romanian migrants' remittances from Spain, which is the second destination country of Romanian migrants. There are very few studies on Romanian cross border ties, which have shed light on many aspects of Romanian transnationalism, but they have failed to consider gender as an analytical category. Our aim is to fill in part this gap by examining the different patterns of Romanian men's and women's remittances. Our paper, therefore, focuses on the specific characteristics of the socio-economic integration of Romanian men and women in Italy, and their role in determining or not remittance behavior.

2. Data, methods and results

In order to address Romanian men's and women's remittances we use quantitative data from the Integrometro survey, coordinated by the ISMU Foundation and carried out in 2009. The aim of the research was to understand the degree of integration of migrants into the Italian society (Cesareo and Blangiardo

¹ For a detailed review of the theoretical models of remittances see Fokkema, Cela and Ambrosetti (2013).

2009). The survey collected information on a total sample of 12,047 individuals aged between 19 and 71 years old. Respondents were selected using the center sampling technique (Baio et al. 2011). The questionnaire was composed of two sections. The first addresses questions related to the four dimensions of integration identified in the survey: cultural, social, economic and political. The second section contains socio-demographic information. Only analyses for Romanians citizens are presented in this article. Therefore, the pooled sample is reduced to N=1,774. Due to missing values on relevant variable our final sample was reduced to N=1,135.

Our dependent variable is the decision to send remittances (1=yes; 0=no). We verified how migrant behavior is influenced by demographic characteristics and by different aspects of integration in Italy. We used logistic regressions to model the likelihood of sending remittances to the country of origin. In order to analyze whether the likelihood to remit depends on gender specific characteristics we first used the whole sample and run a model with gender interaction terms (not shown here) for the demographic variables as well as for some other variables of interest. The gender interaction terms are significant and this illustrates that gender explains the different impact of the determinants on the likelihood to remit for men and women. Therefore, the analyses were separately run for men and women.

We run 4 different models (table 1). In the first model, we included only socio-demographic factors as independent variables. In the second model, we added variables capturing the dimension of socio-cultural integration. Then, in the third model, besides the control variables, we included variables referring to the economic integration. All variables were considered in the last model.

The independent variables were socio-demographic characteristics generally known to have an impact on remitting behavior; the first one was civil status, with married being the reference group. We presumed that marital status differently shapes the probability of sending remittances for Romanian men and women owing to different social expectations in the society of origin. Women from traditional societies might have more incentives to leave the labor market once married or when children were born. Furthermore, marital status, as a proxy for having children in Romania, may induce especially women to remit as a way to assert the motherhood at distance. To examine whether remitting behavior depends on educational attainment, the variable educational qualification was introduced ranging from 1=no school to 4=university degree (reference group). Other variables controlled for were age and length of residence in Italy (as years in Italy).

Eight variables capturing the migrants' socio-cultural integration were included in the analysis: a dummy for Italian language proficiency ("bad" reference group), which we expected to have a positive effect on the decision to remit because it facilitates integration into the Italian labor market; a dummy for having an Italian partner (0=no reference group); friendship ranging from 1=more migrants to

3=more Italians (reference group); and a dummy for participation in associations in Italy (1=yes). Previous findings on the topic of remittances show that stronger ties to the host country do not necessarily mean weaker ties with the sending country, and that assimilation and transnationalism may simultaneously occur (Portes et al. 2007). Bearing this in mind, we sought to explain the gendered variation of the influence of social determinants (i.e. having more migrant friends, membership in associations) on the likelihood to remit. Feelings toward the host society ranged from 1=very good/good enough to 4=bad/bad enough (reference group). Other variables referred to whether or not men and women were interested in current affairs in Italy (0=no, 1=yes), and whether or not they had decided to settle in Italy permanently, ranging from 1=uncertain about future to 3=yes (reference group). Since many more Romanian women than men choose an Italian spouse, this might have different bearings on the likelihood to remit of men and women. Furthermore, we expected that the presence of children (and partner) in Italy (ranging from 1=living alone, to 2=living with children and partner and 3=living with others - the reference group) would have a negative effect on the likelihood to remit, especially for women. The presence of spouse and children in Italy, might decrease the likelihood to remit as a consequence of the shift of family life and interests from Romania to Italy.

Finally three variables referring to economic integration were employed in the analysis: migrant's accommodation, ranging from 1=ownership to 3=rent shared with others (reference group); occupational status ranging from 1=entrepreneur to 4=unemployed (reference group); and perceived difficulties with current income, ranging from 1=do not declare to 3=save (reference group). Romanian men generally perform better economically than their female counterparts in Italy, and based on others' predictions that transnational behavior is positively related to economic integration in the host country (Portes et al. 2007), we expect to find that patterns of remittances are further gendered.

With regard to the variables capturing the influence of demographic characteristics on the likelihood to remit, the probability of sending remittances increases with age (the older the migrant, the more likely s/he is to send remittances). For women, the results are similar and significant in all four models, while for men it appears that age is not a strong determinant. In all the models, the length of stay reduces the probability of sending remittances. This result is consistent with the "remittance decay hypothesis" that reflects the possibly negative trend in remittances over time. In our case, even after controlling for co-residence of close family members, such as children and spouse, remittances tend to decrease with the time spent abroad, especially for women. Married women are more likely to send remittances than are unmarried ones (except for the first model, where the relation is negative).

Table 1 – Effects (odds ratios) of variables on likelihood of sending remittances for women and men

Variables	Women			Men		
	Demographic	Socio-cultural	Economic	Demographic	Socio-cultural	Economic
Age (in years)	1.069***	1.060***	1.053***	1.001*	1.000	1.029**
Length of stay in Italy (years)	.867***	.897***	.906***	.919***	.944*	.907***
Civil Status (not married)	.570***	2.204***	2.012***	.535***	1.664*	1.585**
Education (univ)	.562	.619	.310***	.126***	.250*	.171**
no education	.998	.950	.676	.141***	.275*	.136***
primary education	1.143	1.103	.785	.200***	.355	.188***
secondary education		1.232	1.174		1.135	.864
Knowledge of Italian (no)		1.691**	1.871**		2.047***	4.194***
Italian Partner (no)		1.705**	1.471		.889	1.045
interest in what happens in Italy (no)						
Friends (more Italians)						
more migrants		1.336	1.265		2.976***	2.859***
both		1.081	1.102		3.892***	3.683***
Association Participation (no)		.710	.804		3.580***	3.727***
How do you feel in Italy (bad)						
very good		2.467**	2.093*		.960	1.657
good enough		2.573**	2.404**		2.608*	.806
neither good nor bad		2.386**	2.428**		2.150	2.926
Future projects (never)						
Uncertain about the future		1.681**	1.647*		3.112***	2.646***
To stay in Italy for a period		2.039***	2.072***		5.404***	5.809***
Accommodation (share with others)						
Home ownership			.459***			.580**
rent			.429***			.497***
Employment (unemployed)						
entrepreneur			2.787			16.381***
regularly employed			2.379***			3.262***
irregular			3.479***			3.539***
Perceived difficulties with current income (monthly)						
do not declare			.587			1.766
Difficulties to make ends meet			.851			3.18***
spend all the monthly income			.619**			.647*
Total						
			1.049***			1.008
			.906***			.937*
			2.256***			1.360
			.351**			.489
			.716			.315*
			.345			.346
			1.174			.864
			1.871**			4.194***
			1.471			1.045
			1.265			2.859***
			1.102			3.683***
			.804			3.727***
			2.093*			1.657
			2.404**			.806
			2.428**			2.926
			1.647*			2.646***
			2.072***			5.809***
			.523*			.580**
			.444***			.497***
			2.787			16.381***
			2.379***			3.262***
			3.082***			3.539***
			.739			1.766
			.851			3.18***
			.539**			.647*

Note. In brackets we put the reference group. *p<0.1, **p<0.05, ***p<0.01

The civil status of men appears to be less significant, from a statistical point of view, in explaining this economic behavior.

With regard to the socio-cultural variables, there is a positive relation between education and the probability of remitting especially for men (not significant for women). Having no education, or only a primary-level one, compared with graduates, reduces the likelihood of remittances. Hence more educated men are more likely to send money home, especially when economic variables are considered (third model). Likewise, having an Italian partner increases the probability of sending remittances for both men and women (but the effect is stronger for men), while being interested in what happens in Italy increases the likelihood of remittances only for women when socio-cultural aspects are considered. Men's friendships with both migrants and Italians, as well as men's membership in associations are strong determinants of their remittance behavior, but have no impact on women's remittances. However, in relation to the variable "how do you feel in Italy?" women replying they feel comfortable are more likely to remit than others. By contrast, this variable does not affect men's decisions to remit (except for the second model, where a positive relation is found). Overall, those migrants who have decided to settle permanently in Italy are less likely to remit, which is consistent with the relationship found between the length of stay and remittances. Unlike for women, temporary projects of permanence in Italy increase the probability that men will send money home, which could be partly explained by men's plans to invest at home, as shown by other empirical studies.

As one might expect, home ownership reduces the likelihood of remittances for men and women, compared to those who share accommodation with others, because both settlement projects and higher living expenses in Italy may occur for home owners. Another meaningful gender difference concerns the relationship between occupational status in Italy and remittances. In general, being employed increases the probability of remittances for both men and women; but for men being an entrepreneur has a strong positive impact on the decision to remit.

Finally, living with the partner and/or children (compared with those living with friends and/or other persons) has opposite effects on sending money home: it decreases women's likelihood to remit (as found in the model considering the socio-cultural variables), but it increases men's probability of remitting, which could be consistent with the results of the empirical evidence on the purpose of men remittances, that is, investment.

3. Discussion and conclusion

Gender plays a significant role in shaping Romanians likelihood of remittances. For women, socio-economic integration substantially increases the probability to send remittances. Their transnational engagement is closely related to their increased incorporation into the Italian society and lifestyle and, at the same time, to the children left behind in Romania.

The probability of sending remittances for men seems to be clearly related to their economic integration in the labor market, as well as to their projects of return home. In this framework, remittances probably are needed to set up their own businesses at home and thus to prepare their return (Vlase 2012).

We therefore conclude that women and men do not hold the same views and interests when sending remittances in their society of origin. Probably this is due to the interaction between social expectations in the home country and specific characteristics of their socio-economic integration in Italy.

References

- ARANDA E. 2003. Global care work and gendered constraints: The case of Puerto Rican transmigrants, *Gender and Society*, Vol. 174, pp 609-626.
- BAIO G., BLANGIARDO G.C., BLANGIARDO M. 2011. Centre Sampling Technique in Foreign Migration Surveys: A Methodological Note, *Journal of Official Statistics*, Vol. 27, N. 3, pp. 451-465.
- BOCCAGNI P., BANFI L. 2011. Transnational family life and female migration in Italy: One or multiple patterns? in KOFMAN E, KRALER A, KOHLI M, SCHMOLL C. (eds.), *Gender, generations and the family in international migration*, Amsterdam: Amsterdam University Press, pp. 287-311.
- CESAREO V., BLANGIARDO G. 2009. *Indici di integrazione. Un indagine empirica sulla realtà migratoria italiana* Milano: Francoangeli.
- DE SOUSA J., DUVAL L. 2010. Geographic distance and remittances in Romania: out sight, out of mind?, *Economie internationale*. Vol. 1121, pp. 81-97.
- FOKKEMA T., CELA E., AMBROSETTI E. 2013. Giving from the Heart or from the Ego? Motives behind Remittances of the Second Generation in Europe. *International Migration review*, Vol. 74, Issue 3, pp. 539-572.
- HONDAGNEU-SOTELO P. 1992. Overcoming patriarchal constraints: The reconstruction of gender relations among Mexican immigrant women and men, *Gender and Society*. Vol. 63, pp.393-615.
- KING R, LAZARIDIS G., TSARDANIDIS C. 2000. *Eldorado or Fortress? Migration in Southern Europe*. Basingstoke: Macmillan.

- LEVITT P., GLICK SCHILLER N. 2004. Conceptualizing simultaneity: A transnational social field perspective on society, *International Migration Review*, Vol. 383, pp. 1002–1039.
- MASSEY D. 1990. Social structure, household strategies, and the cumulative causation of migration, *Population Index*. Vol. 561, pp. 3-26.
- PORTES A., ESCOBAR C., RADFORD A.W. 2007. Immigrant Transnational Organizations and Development: A Comparative Study, *International Migration Review*, Vol. 411, pp. 242-281.
- VLASE I. 2013. My Husband Is a Patriot!?: Gender and Romanian Family Return Migration from Italy, *Journal of Ethnic and Migration Studies*, DOI:10.1080/1369183X.2013.756661

SUMMARY

Gendered remittances have become an important research topic capturing the attention of many scholars worldwide. From this point of view Romania represent an interesting study case of feminization of migration and remittances flows. This paper aims to explore the patterns of transnational socio-economic activities of Romanian men and women in Italy, by analyzing the gender specific determinants of remittances. Using recent quantitative data from the Integrometro survey, the findings suggest that gender plays a significant role in shaping the likelihood of remittances.

I RENDIMENTI OCCUPAZIONALI E RETRIBUTIVI DEI DIFFERENTI PERCORSI FORMATIVI IN OTTICA DI GENERE

Marco Centra, Andrea Cutillo, Valentina Gualtieri

1. Introduzione

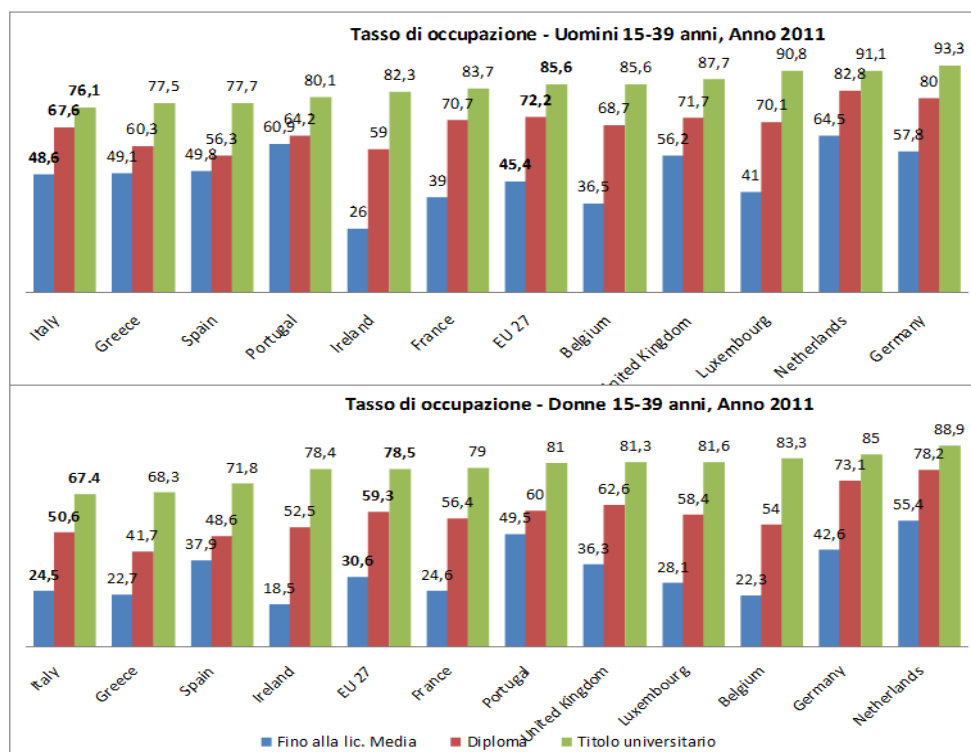
Il ruolo cruciale del capitale umano nel favorire i processi di sviluppo e di crescita è ampiamente condiviso sia dalla letteratura in materia (Becker, 1964; Gori, 2004; Cipollone e Sestito, 2010) che dagli orientamenti dei governi e delle organizzazioni sovranazionali. Nel dibattito europeo sulle strategie per favorire il rilancio dell'economia viene ribadita la funzione decisiva dell'investimento in capitale umano, per garantire i livelli occupazionali e la competitività delle imprese. Le evidenze empiriche nazionali e sovranazionali sui rendimenti del capitale umano indicano che un aumento della scolarizzazione genera minore rischio di disoccupazione, redditi più elevati e carriere più rapide. Inoltre, il capitale umano sostiene i processi di innovazione del sistema produttivo, aumentandone i margini di efficienza e i livelli di competitività. Un meccanismo virtuoso viene attivato incrementando la propensione allo sviluppo di capitale umano da parte dell'offerta di lavoro e garantendo, dal lato della domanda, una dinamica delle innovazioni del sistema produttivo in grado di assorbire le nuove conoscenze e competenze di cui sono portatori i lavoratori. Tale processo, che rende incentivante l'investimento in capitale umano sia per i lavoratori che per le imprese (e quindi per l'intero sistema), sembra nel nostro paese aver subito un rallentamento, con ricadute sulla sostenibilità del mantenimento dei livelli di competitività, rischiando di allontanare l'Italia dai principali *competitors* dell'area continentale. In Italia si rilevano elementi di debolezza di carattere strutturale sia nei livelli di istruzione della popolazione che nell'allocazione delle competenze, che impediscono uno sviluppo e un impiego ottimale delle conoscenze e delle abilità nel processo produttivo.

Nonostante gli ultimi decenni siano stati caratterizzati da un continuo incremento del livello d'istruzione della popolazione italiana, il nostro Paese continua a mostrare un forte ritardo in ambito formativo, anche nelle generazioni più giovani. Nel 2011 appena il 14,5% dei 15-39enni possiedono un titolo

universitario, contro una media EU27 di 24,6%. L'Italia presenta una quota di giovani di 15-39 anni con istruzione terziaria inferiore alla maggior parte dei paesi dell'Unione Europea (20,1% per la Germania, 27,1% per l'Olanda, 22,3% per la Grecia, 31,2% per la Francia, 33,9% per il Regno Unito, 32,0% per la Spagna).

Contrariamente a quanto sarebbe lecito attendersi, la scarsità di offerta di forza lavoro altamente scolarizzata, si associa sia ad un basso rendimento del titolo di studio terziario in termini di occupabilità sia ad una ridotta remunerazione. Oltre a presentare i più bassi livelli di occupazione, i laureati italiani, mostrano un divario rispetto all'occupabilità dei diplomati tra i più bassi rispetto agli altri paesi europei (Figura 1), sia per gli uomini che per le donne.

Figura 1 – Tasso di occupazione dei 15-39enni per sesso e titolo di studio, 2011.

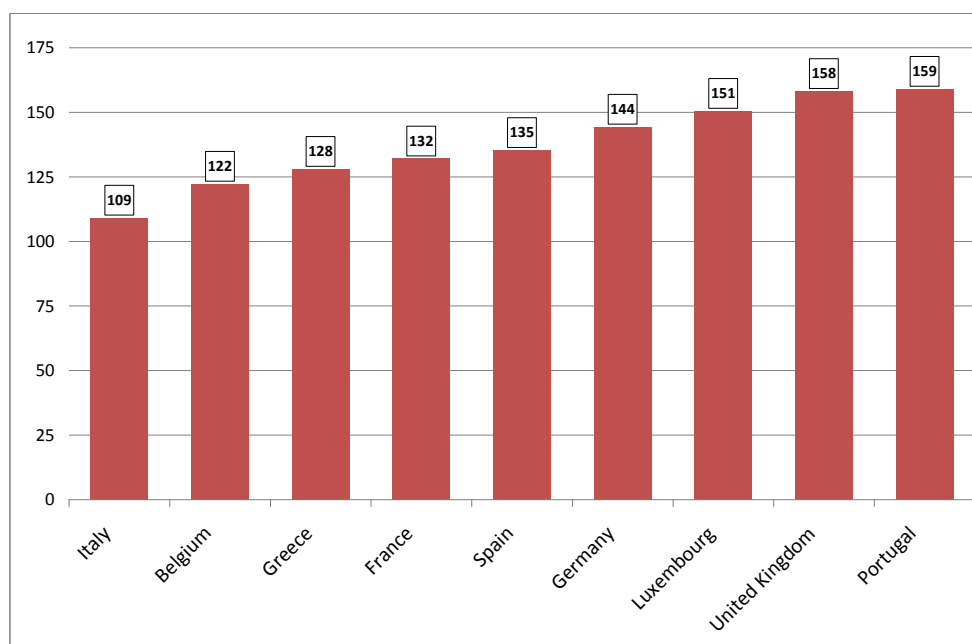


Fonte: Eurostat, Labour Force survey.

La scarso rendimento dell'investimento in istruzione si evidenzia per i laureati anche sul piano retributivo: il divario tra retribuzioni di laureati e diplomati in età compresa tra 25 e 34 anni è particolarmente basso: il salario medio di un laureato

supera del 9% quello di un giovane diplomato nella stessa classe di età (Figura 2). Tali evidenze rivelano che il vantaggio connesso al conseguimento di un titolo di studio universitario non è altrettanto adeguato all'investimento formativo al pari di quanto avviene nei principali paesi europei.

Figura 2 – Retribuzione degli occupati in età 25-34 anni con istruzione terziaria numeri indici: upper secondary and post-secondary non-tertiary education = 100, 2012.



Fonte: OECD, *Education at a glance*, 2012.

Agli scarsi premi, in termini di occupabilità e di livello retributivo, associati all'investimento in istruzione delle giovani generazioni, si associa in Italia un marcato differenziale di genere a svantaggio della componente femminile. Al sensibile aumento della propensione alla scolarizzazione e alla formazione delle donne, registrato nel medio-lungo periodo, non è corrisposto un altrettanto rapido sviluppo del rendimento economico dell'investimento in capitale umano. Nonostante i livelli di istruzione delle donne superino quelli degli uomini, i premi in termini occupazionali e retributivi per la componente femminile, a parità di altre caratteristiche, risultano sensibilmente inferiori. Le differenze di genere e il ridotto rendimento del capitale umano non si attenuano, peraltro, nel tempo, dal momento che le generazioni più giovani ne sono particolarmente colpite.

In questo lavoro si intendono analizzare alcuni aspetti relativi al rendimento dell'istruzione terziaria nel mercato del lavoro italiano, limitatamente alla componente giovanile, prestando particolare attenzione alle differenze di genere. Il principale obiettivo è valutare l'efficacia delle diverse discipline universitarie con riferimento sia alle probabilità di accedere all'occupazione sia al premio economico in termini di reddito da lavoro. Il contributo ha quindi un duplice scopo: il primo, è volto a comprendere in che misura le differenti scelte di genere in termini di percorsi formativi incidano sui livelli di occupabilità. Si intende quindi qualificare il rendimento dell'investimento in capitale umano verificando quali sono le discipline che rendono maggiormente sul mercato del lavoro, sia in termini di probabilità di trovare un'occupazione che in termini di reddito da lavoro. In secondo luogo si intende verificare se e in che misura le differenze di genere continuino a persistere a parità di percorsi formativi, evidenziando quali siano le discipline che assicurano differenze di genere meno marcate. La popolazione di interesse è limitata ai laureati in età compresa tra 25 e 34 anni; le analisi sono condotte sulle stime ricavate Rilevazione sulle forze di lavoro dell'ISTAT, nell'edizione del 2011.

Per rispondere agli obiettivi sopra esposti sono stati sviluppati due modelli multivariati atti ad isolare l'effetto del genere e della disciplina universitaria sulla probabilità di essere occupati e sul reddito da lavoro dipendente.

Nel primo modello la variabile di interesse osservata è una variabile dicotomica che assume valore 1 nel caso l'individuo sia occupato e 0 altrimenti; la probabilità è stimata tramite un modello di regressione logistica. Le variabili esplicative utilizzate sono in primo luogo il genere e la disciplina (a 12 modalità); tali variabili sono comprese tra i regressori sia come effetti semplici che come termini di interazione. Le ulteriori variabili di controllo utilizzate sono: la ripartizione geografica di residenza, l'età, in forma semplice e quadratica, il ruolo dell'individuo all'interno della famiglia e la cittadinanza.

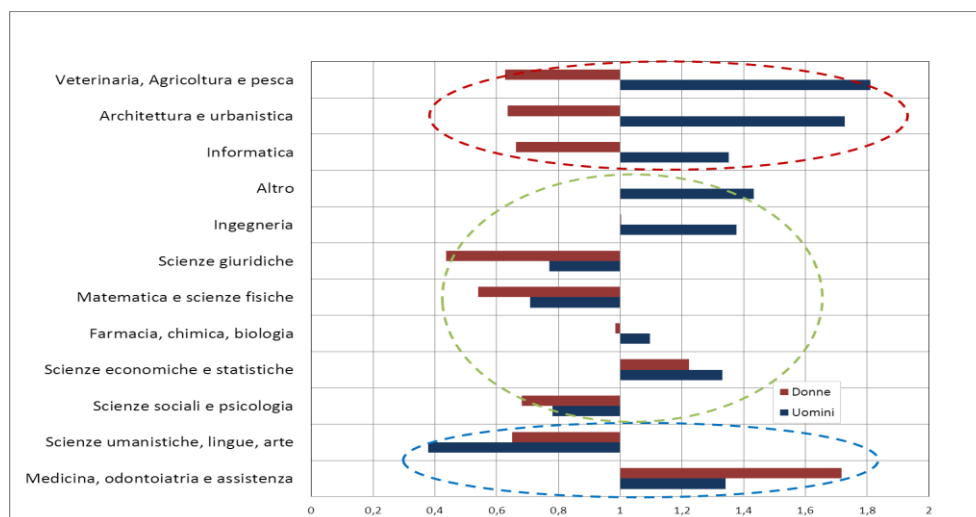
Nel secondo modello, il collettivo d'interesse è ristretto ai soli laureati 25-34enni occupati come dipendenti. L'esclusione dei lavoratori autonomi deriva dall'assenza dell'informazione sulla retribuzione per tale categoria di lavoratori. Il modello utilizzato è un modello di regressione lineare, dove la variabile dipendente è il salario mensile (in termini logaritmici). Le variabili utilizzate come regressori, oltre a quelle utilizzate nel modello precedente, comprendono informazioni relative al lavoro svolto, consentendo di controllare l'eterogeneità dovuta al profilo occupazionale: il numero di anni nell'attuale posizione (tenure), la professione svolta (articolata in quattro modalità che aggregano i 9 gruppi professionali della

classificazione delle professioni¹), la tipologia contrattuale (tempo determinato o indeterminato) e l'orario di lavoro (tempo pieno o parziale).

2. Le evidenze empiriche

Nella Figura 3, sono riportati gli *odds ratio*, cioè la misura relativa degli effetti stimati per le discipline in etichetta rispetto alla categoria di riferimento (le donne laureate in altre discipline). Le discipline sono ordinate in base a differenziali uomo-donna decrescenti.

Figura 3 – *Stime logit per la probabilità di essere occupati, odds ratio (laureati in età 25-34 anni, 2011)*



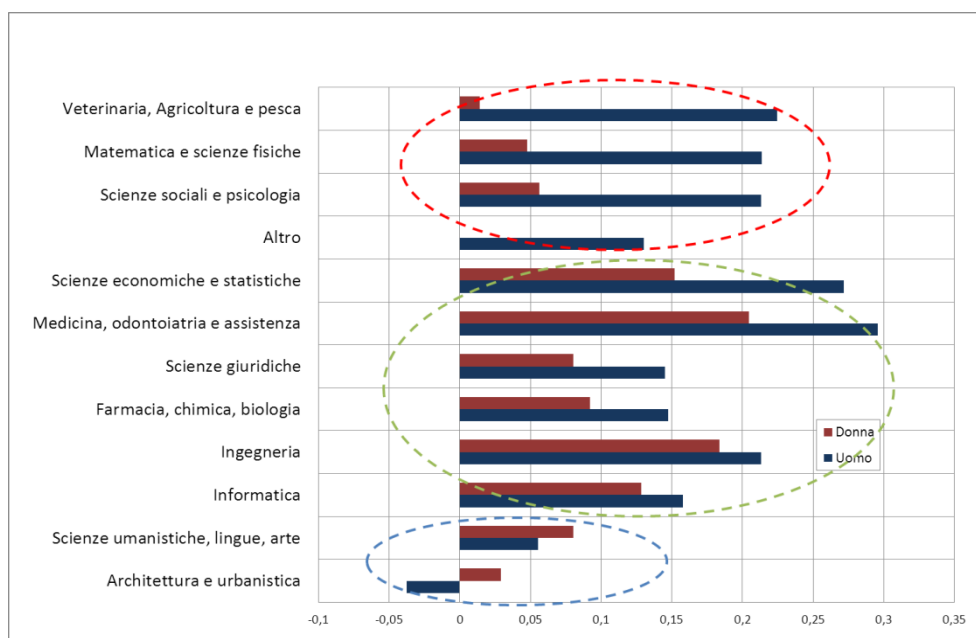
Fonte: elaborazioni su dati Istat, Rilevazione forze lavoro, 2011.

Per gli uomini, le discipline che consentono l'accesso più probabile all'occupazione sono Veterinaria, agricoltura e pesca; Architettura e urbanistica; Informatica; Ingegneria; Medicina, odontoiatria e assistenza. Quelle che invece non presentano probabilità elevate sono le Scienze umanistiche, lingue e arte; Scienze sociali e psicologia; Scienze giuridiche; ma anche Matematica e scienza fisiche. Tra le donne, invece, le migliori discipline sono Medicina, odontoiatria e

¹ La classificazione delle professioni CP2001 dell'Istat al primo digit è stata raggruppata nel seguente modo: nelle professioni altamente qualificate appartengono le professioni 1 e 2; le professioni appartenenti al gruppo 3 sono definite come professioni tecniche, nelle professioni medie sono comprese le categorie dalla 4 alla 7, mentre le professioni non qualificate fanno parte del gruppo 8. Le professioni 9 (forze armate) sono escluse da tutte le analisi.

assistenza; Scienza Economiche e Statistiche; Ingegneria; Farmacia, chimica e Biologia. Osservando i differenziali di genere, quelle che consentono un posizionamento relativo migliore per le donne rispetto ai colleghi uomini sono Medicina, odontoiatria e assistenza e Scienze umanistiche, lingue e arte. Quelle che invece presentano differenziali molto elevati sono tra le migliori per gli uomini e peggiori per le donne in termini di probabilità di occupazione: Veterinaria, agricoltura e pesca; Architettura e urbanistica; Informatica.

Figura 4 – Stime OLS sul log-reddito da lavoro dipendente (laureati 25-34 anni, Anno 2011)



Fonte: elaborazioni su dati Istat, Rilevazione forze lavoro, 2011.

In figura 4 sono invece riportate le stime degli effetti marginali ottenuti nel modello sul salario. Come in precedenza, le discipline sono ordinate in base ai differenziali di genere decrescenti. Rispetto alla variabile reddito, le discipline migliori per gli uomini sono Medicina, odontoiatria e assistenza; Scienze Economiche e Statistiche; Ingegneria; Informatica; Veterinaria, agricoltura e pesca;

Architettura e urbanistica; Informatica; Ingegneria; Medicina, odontoiatria e assistenza; Matematica e scienza fisiche; Scienza sociali e psicologia. Tutte queste discipline consentono rendimenti più elevati rispetto alla disciplina che permette i migliori rendimenti economici tra le donne, cioè Medicina, odontoiatria e

assistenza. Altre discipline con buoni rendimenti economici femminili sono Ingegneria; Scienza economiche e statistiche; Informatica.

Come già per la probabilità di occupazione, i differenziali di genere più elevati si trovano in Veterinaria, agricoltura e pesca, alla quale si aggiungono Matematica e Scienze fisiche e le Scienze sociali e psicologia. All'opposto, le donne riescono a ottenere retribuzioni migliori degli uomini in due sole discipline, le Scienze umanistiche, lingue e arte ed Architettura e urbanistica.

3. Conclusioni

In questo lavoro sono stati analizzati alcuni aspetti relativi al rendimento dell'istruzione universitaria, osservati nella fase di ingresso nel mercato del lavoro dei giovani laureati italiani. Il differenziale di genere in favore della componente maschile viene confermato dai risultati dell'analisi, sia in merito all'occupabilità che ai livelli retributivi. L'analisi secondo la disciplina rivela che le discipline dove le donne hanno maggiori livelli di occupabilità sono quelle che garantiscono un più basso differenziale salariale di genere. Il differenziale retributivo di genere medio è molto elevato, mostrando un aumento del livello della differenza all'aumentare della retribuzione della disciplina di studio.

Alcune discipline (scienze economiche, scienze sociali), mostrano un basso differenziale in termini di occupabilità, mentre altre (scienze umanistiche e medicina) presentano livelli di occupabilità più alti per le donne.

Le discipline che presentano differenziali retributivi a vantaggio della componente femminile (architettura e scienze umanistiche) assicurano parallelamente un reddito sensibilmente più basso rispetto alla media. Le laureate in materie che non garantiscono elevata occupabilità, mostrano, una volta occupate, un differenziale salariale a svantaggio degli uomini. Il superamento di più elevate barriere all'ingresso comporta per le donne un vantaggio maggiore in termini retributivi, sia sui livelli assoluti che in riferimento al gap di genere.

L'analisi secondo la disciplina rivela, in ultima analisi, un contesto in cui alcuni percorsi garantiscono alle donne una migliore occupabilità; tuttavia gli stessi percorsi presentano un rendimento, in termini retributivi, più basso rispetto al totale e con un più basso differenziale (Scienze umanistiche), oppure con elevati livelli retributivi per i soli uomini, e non per le donne (Veterinaria e Matematica).

Riferimenti bibliografici

- BECKER G.S., 1964. *Human capital*, New York, Columbia University Press.
- CIPOLLONE P., SESTITO P. 2010. *Il capitale umano*, Collana "Farsi un'idea", Il Mulino, Bologna.
- GORI E. 2004. L'investimento in Capitale Umano attraverso l'istruzione, in G. Vittadini (a cura di), 2004, *Capitale Umano. La ricchezza dell'Europa*. Guerini e ass. ed., Milano.

SUMMARY**Occupational and wage returns for different university subjects: a gender analysis**

In order to understand different outcomes on the Italian labour market for young graduates, this paper provides analysis using a representative sample of Italian graduates 25-34 years old. A specific aim is to measure the gender gap on occupational probabilities and on wages. The results confirm high gender gaps over the high skilled population in favor of men. These differences, however, vary depending on the university subject. Few subjects also show gaps in favour of women: Humanistic and Medicine concerning the occupation probability as well as Humanistic and Architecture concerning wages.

Marco CENTRA, ISFOL, Servizio Statistico, m.centra@isfol.it
Andrea CUTILLO, ISTAT, Dipartimento delle Statistiche Sociali e Ambientali, cutillo@istat.it
Valentina GUALTIERI, Servizio Statistico, v.gualtieri@isfol.it

CLASSE SOCIALE, COORTE E ISCRIZIONE UNIVERSITARIA

Andrea Cutillo, Romina Fraboni, Claudio Ceccarelli

1. Introduzione

Nel corso del tempo, diverse misure sono state adottate per incoraggiare la partecipazione della popolazione con background familiare più svantaggiato ai percorsi di istruzione superiori, e favorire così l'uguaglianza delle opportunità di istruzione. Si ricordano, tra le altre, le riforme del 1962 (obbligo scolastico a 14 anni), del 1999 (obbligo elevato a 15 anni) e del 2003 (obbligo scolastico a 16 anni ed obbligo formativo a 18 anni). Molto rilevanti, per una maggiore partecipazione universitaria, sono state la riforma del 1969, che ha permesso anche agli studenti di istituti tecnici e professionali, tradizionalmente frequentati dai giovani delle classi meno abbienti, il libero accesso all'università (a condizione di aver concluso un ciclo superiore di cinque anni) e la riforma del 1999, del 3+2, che aveva tra gli obiettivi principali proprio un aumento del numero di iscrizioni e una diminuzione del drop-out universitario, grazie all'avvicinamento dell'orizzonte di conseguimento del ciclo breve, accompagnato da una maggiore offerta formativa che meglio si adattassero alle esigenze più variegata della popolazione studentesca.

Il conseguimento di elevati titoli di istruzione, peraltro, è uno degli strumenti principali in grado di ridurre le disuguaglianze esistenti all'interno della società e in grado di favorire la mobilità sociale. Inoltre, gli effetti dell'accumulo di capitale umano sono pervasivi dell'intero ciclo di vita delle persone (nell'accesso sul mercato del lavoro, nei successivi sviluppi di carriera, nel mantenere stili di vita e di salute a minore rischio, nel grado di partecipazione sociale), oltre ad essere determinanti per la crescita economica e sociale di un paese nel suo complesso. Nonostante i citati interventi normativi, però, la trasmissione intergenerazionale del titolo di studio rimane molto elevata e la possibilità di portare avanti gli studi è fortemente legata alla classe sociale (tra gli altri, Comi, 2004; ISTAT, 2012; Checchi et al., 2013; Pisati, 2002).

Questo lavoro analizza l'iscrizione all'università per quanti hanno conseguito un diploma di scuola secondaria superiore per classe sociale di origine nel corso delle generazioni. Ci si sofferma su questo passaggio perché è ormai questo,

insieme all'abbandono della scuola secondaria superiore, il momento in cui permangono forti differenze tra le classi sociali (ISTAT, 2012).

2. Il metodo

I dati utilizzati per l'applicazione empirica provengono dall'indagine multiscopo ISTAT "Famiglie e soggetti sociali" del 2009, condotta con tecnica PAPI su un campione di circa 24mila famiglie e rappresentativa dell'intera popolazione italiana. In questa terza edizione di indagine è infatti presente un modulo sul percorso di istruzione degli individui di 16 anni e più, dalla scuola elementare in avanti, con la data di ciascun evento di iscrizione, abbandono o conseguimento di un ciclo scolastico e universitario. Inoltre, attraverso le informazioni sulle caratteristiche della famiglia di origine degli intervistati quando questi avevano 14 anni, e cioè quando si compiono le scelte sulla scuola secondaria superiore che influenzano poi fortemente i successivi percorsi universitari, è anche possibile ricostruire la classe sociale di origine secondo la declinazione indicata da Cobalti e Schizzerotto (Cobalti, Schizzerotto 1994)¹.

Questa fonte ha il grande pregio di poter analizzare i percorsi di istruzione delle differenti generazioni a seconda della classe sociale di origine. D'altra parte, non essendo un'indagine prettamente incentrata sull'acquisizione della formazione, presenta anche alcuni indubbi svantaggi, quali la mancanza di informazioni molto rilevanti (ad esempio, il voto scolastico e il tipo di scuola secondaria frequentata e il tipo di facoltà prescelta).

Nel modello, la variabile dipendente è una dicotomica che assume valore 1 se l'individuo si è iscritto all'università e 0 altrimenti. Il modello è stimato separatamente per generazione e classe sociale di origine, che sono quindi variabili di classificazione. Le variabili esplicative sono riferite alle caratteristiche dell'individuo, quali il sesso, l'età al conseguimento del titolo di scuola secondaria superiore, la ripartizione geografica di residenza; inoltre si includono delle variabili relative alle caratteristiche della famiglia di origine quali il titolo di studio più elevato tra i genitori, per tenere sotto controllo la fortissima associazione tra questo e la probabilità di intraprendere e concludere con successo un percorso universitario (Checchi, 2000; Di Pietro, 2004) e il numero di fratelli avuti, di modo da tenere conto delle maggiori o minori aspettative familiari che ricadono sul

¹ La classificazione di Cobalti e Schizzerotto divide la popolazione in borghesia (imprenditori, liberi professionisti e dirigenti); classe media impiegatizia (impiegati di concetto e impiegati esecutivi); piccola borghesia urbana (artigiani e commercianti, con e senza dipendenti); piccola borghesia agricola (proprietari agricoli); classe operaia urbana (operai qualificati e non qualificati, dell'industria e dei servizi); classe operaia agricola (braccianti). In questo lavoro si utilizza una classificazione più aggregata, a quattro modalità: borghesia, classe media impiegatizia, piccola borghesia e classe operaia.

singolo individuo, oltre che di maggiori o minori vincoli finanziari. La diversa propensione al proseguimento degli studi viene inoltre messa in relazione alla maggiore o minore capacità scolastica (l'essere stati bocciati tra elementari e medie e l'essere stati bocciati alle superiori²). Infine, attraverso un link tra i micro-dati individuali e il tasso di disoccupazione per genere, a livello macro³, nell'anno di fine delle scuole superiori e cioè in corrispondenza del momento in cui si decide se proseguire negli studi o meno⁴ si vuole cercare di indagare il costo-opportunità dell'investimento in capitale umano relativamente all'offerta del mondo del lavoro: è infatti vero che peggiori opportunità occupazionali possono incoraggiare l'iscrizione universitaria, sia per mancanza di alternative professionali immediate, sia per acquisire capitale umano da poter spendere in maniera più proficua sul mercato del lavoro (Di Pietro, 2006; Dellas e Koubi, 2003).

In questo lavoro, siamo interessati all'evoluzione delle probabilità effettive di iscriversi all'università (il collettivo è quindi composto da quanti hanno concluso con successo la scuola secondaria superiore).

Senza perdere di generalità, formalizziamo il modello: sia s_{ijc}^{*u} la variabile latente continua e non osservabile associata alla scelta u , l'iscrizione universitaria, per il generico individuo i della generica generazione j della generica classe sociale c . Sebbene s_{ijc}^{*u} non sia osservabile, si può osservare la variabile binaria s_{ijc}^u che assume valore 1 se l'individuo si iscrive all'università e valore 0 se non si iscrive. La probabilità di iscriversi all'università, e cioè di osservare $s_{ijc}^u = 1$ ($P(s_{ijc}^u = 1)$) viene stimata tramite regressione logistica. Tramite l'utilizzo di un individuo rappresentativo medio per ogni classe sociale all'interno di ogni generazione (un individuo che abbia quindi come esplicative il valor medio delle variabili esplicative di ogni gruppo) è possibile stimare le probabilità previste di iscriversi per ciascuno di questi individui rappresentativi:

$$P(\bar{X}_{jc}, \beta_{jc}) = \frac{1}{1 + e^{-(\bar{X}_{jc}\beta_{jc})}} \quad (1)$$

Se consideriamo per la generica generazione j due specifiche classi sociali, A e B, è quindi possibile calcolare le differenze nelle probabilità di iscriversi

² In realtà, questa informazione non è direttamente disponibile, ma è ricavata tramite l'utilizzo degli scarti tra la data di inizio e fine dei differenti cicli scolastici.

³ ISTAT, Serie storiche (seriestoriche.istat.it).

⁴ Poiché l'iscrizione universitaria può avvenire anche in un periodo sfasato rispetto al conseguimento del titolo di secondaria superiore, a rigore il tasso di disoccupazione più corretto sarebbe in corrispondenza del momento di effettuare la scelta. Poiché questa data è ricavabile solamente per quanti si sono iscritti, si è scelto di utilizzare questa approssimazione. La stragrande maggioranza di quanti si iscrivono, comunque, lo fanno al completamento del ciclo scolastico.

all'università tra l'individuo rappresentativo della classe sociale A e l'individuo rappresentativo della classe sociale B.

$$P(\bar{X}_{jA}, \beta_{jA}) - P(\bar{X}_{jB}, \beta_{jB}) = \frac{1}{1+e^{-(\bar{X}_{jA}\beta_{jA})}} - \frac{1}{1+e^{-(\bar{X}_{jB}\beta_{jB})}} \quad (2)$$

Seguendo l'approccio di Oaxaca e Ransom (1994), stimiamo lo stesso modello logistico all'interno di una generazione j per il collettivo formato da tutti gli individui delle due classi sociali in analisi, A e B. In questa maniera si ottiene il vettore dei coefficienti β_{jAB} , che rappresenta la struttura dei coefficienti che avrebbe prevalso in assenza di differenze dovute alla classe sociale. Il vettore dei coefficienti *pooled* rappresenta quindi la struttura neutrale rispetto alla classe sociale. Utilizzando il vettore dei coefficienti *pooled* si possono stimare le due probabilità $P(\bar{X}_{jA}, \beta_{jAB})$ e $P(\bar{X}_{jB}, \beta_{jAB})$ che, aggiunte e sottratte nella (2), permettono la scomposizione del differenziale osservato in tre termini:

$$P(\bar{X}_{jA}, \beta_{jA}) - P(\bar{X}_{jB}, \beta_{jB}) = \{P(\bar{X}_{jA}, \beta_{jAB}) - P(\bar{X}_{jB}, \beta_{jAB})\} + \\ + \{P(\bar{X}_{jA}, \beta_{jA}) - P(\bar{X}_{jA}, \beta_{jAB})\} + \{P(\bar{X}_{jB}, \beta_{jAB}) - P(\bar{X}_{jB}, \beta_{jB})\} \quad (3)$$

Il primo termine in parentesi graffe a destra dell'equazione, che utilizza lo stesso vettore dei coefficienti stimati β_{jAB} su due collettivi differenti, mostra il differenziale nelle probabilità di iscrizione spiegata dalle differenze nelle caratteristiche medie del collettivo A e B (componente spiegata del differenziale). Il secondo e terzo termine in parentesi graffe rappresentano invece la parte di differenziale che può essere invece attribuita a differenti rendimenti delle stesse caratteristiche osservate. Questa parte di differenziale, la componente non spiegata del differenziale, può quindi in qualche misura associarsi a differenze nel comportamento dei figli delle differenti classi sociali.

3. I risultati

In tabella 1 si riportano i tassi di iscrizione universitaria per classe sociale e generazione e le differenze assolute tra classi sociali all'interno delle generazioni⁵. In generale, tutte le classi sociali mostrano un incremento dell'iscrizione per la generazione più giovane, i 25-29enni al momento dell'intervista. Tra i figli dei borghesi, il tasso di iscrizione all'università è molto elevato, su valori intorno al

⁵ L'analisi è svolta sul solo collettivo utilizzato nei modelli. I tassi di iscrizione possono quindi essere leggermente diversi da quanto ricavabili dal file dei microdati a causa delle osservazioni eliminate per mancanza di variabili di interesse.

60%, già a partire dalla generazione dei nati tra il 1940 e il 1949, fino ad un massimo del 67,9% tra i più giovani, nati tra il 1980 e il 1984. I figli della classe media impiegatizia mostrano un incremento di circa 10 punti nel corso delle due generazioni di cui sopra, i figli della piccola borghesia di circa 15 punti, mentre i figli della classe operaia di circa 6 punti. Questo fa sì che, nel corso delle generazioni, si riduca la differenza tra la borghesia e il ceto medio (rispettivamente di 4 punti con la classe media impiegatizia e di 8 con la piccola borghesia), per effetto di un lento ma graduale avvicinamento all'interno di quest'ultimo. Viceversa, pur con l'incremento descritto, i figli della classe operaia si allontanano dagli altri, con differenze nei tassi di iscrizione stabilmente nell'intorno di 40 punti verso i borghesi e un aumento del distacco dalla classe media impiegatizia (da 24,6 a 29,9) e dalla piccola borghesia (da 8,4 a 17,1).

Tabella 1 – Tassi di iscrizione universitaria per classe sociale e generazione (per 100 persone con diploma superiore della stessa generazione) e differenze tra classi sociali per generazione

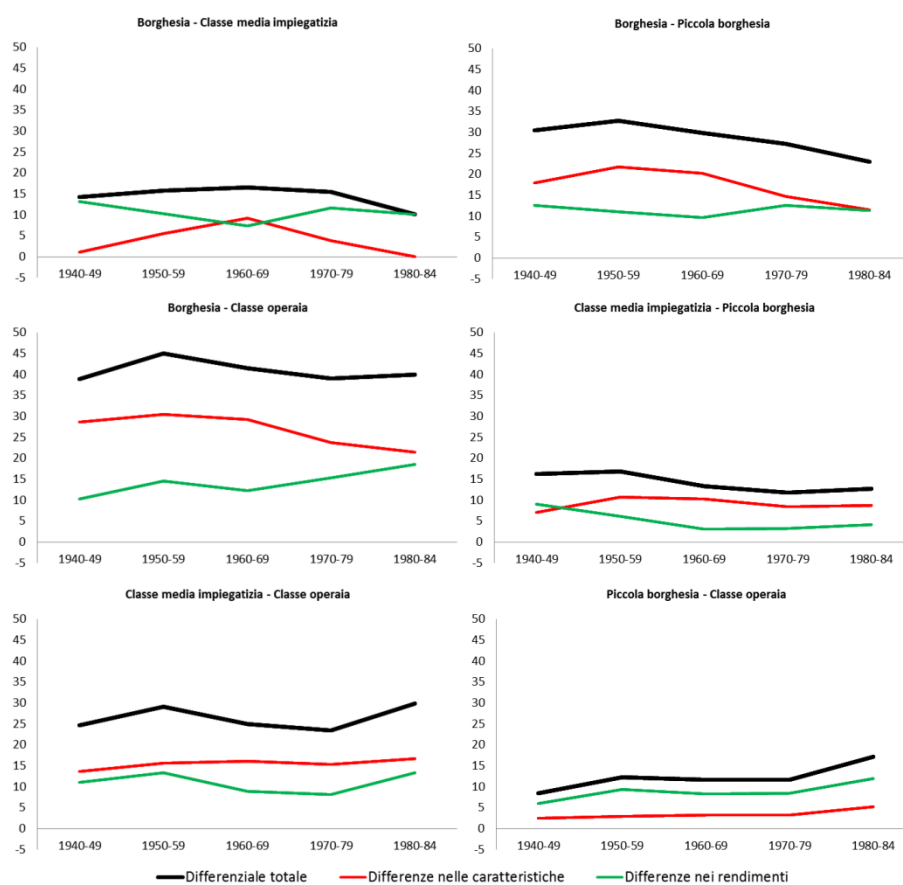
Classe sociale	GENERAZIONE					
	Prima del 1940	1940-49	1950-59	1960-69	1970-79	1980-84
	<i>Tasso effettivo di iscrizione</i>					
Borghesia	49.1	60.9	65.2	58.7	62.6	67.9
Classe media impiegatizia	34.5	46.6	49.3	42.2	47.1	57.8
Piccola borghesia	27.9	30.4	32.4	28.9	35.3	45.0
Classe operaia	17.0	22.0	20.2	17.2	23.6	27.9
Totale	31.0	37.2	35.4	31.7	39.1	47.9
	<i>Differenze nei tassi di iscrizione tra classi sociali</i>					
BOR – CMI	14.6	14.3	15.8	16.5	15.5	10.1
BOR – PB	21.2	30.5	32.7	29.9	27.3	22.9
BOR – CO	32.1	38.9	45.0	41.5	39.0	40.0
CMI – PB	6.6	16.2	16.9	13.3	11.8	12.8
CMI – CO	17.5	24.6	29.1	25.0	23.5	29.9
PB – CO	10.9	8.4	12.2	11.6	11.7	17.1

Applicando la scomposizione dell'equazione (3), si nota come diminuisce la parte del differenziale dovuto alle caratteristiche tra i figli dei borghesi e gli altri (figura 1)⁶. Questo perché la variabile più rilevante nello spiegare l'iscrizione e il successo nell'istruzione dei figli è il titolo di studio dei genitori, e nel corso del tempo il gap su questa variabile si è ridotto considerevolmente tra le differenti

⁶ La generazione dei nati prima del 1940 è esclusa poiché trattandosi di un collettivo aperto a sinistra, gli individui possono vivere situazioni completamente differenti tra di loro. E' quindi più complicato riferirsi a cambiamenti nelle caratteristiche e nei comportamenti, anche a parità di generazione.

classi. Mentre infatti i figli dei borghesi già nelle generazioni più anziane avevano elevatissime quote di genitori altamente istruiti, il generale innalzamento del livello di istruzione all'interno della popolazione ha fatto sì che quote sempre più ampie di popolazione delle altre classi sociali arrivassero a conseguire elevati titoli di studio.

Figura 1 – *Scomposizione dei differenziali osservati nel tasso di iscrizione tra classi sociali nel corso delle generazioni.*



Nel collettivo in analisi (quanti hanno conseguito un diploma di scuola secondaria superiore), i figli di borghesi con almeno un genitore con titolo di scuola secondaria superiore o oltre passano dal 78,4 per cento della generazione dei nati prima del 1939 all'83,9 di quella dei nati nel 1980-1984; i figli della classe media impiegatizia dal 44,4 al 75,9; i figli della piccola borghesia dal 12,2 al 48,4;

e i figli degli operai dal 3,5 al 32,1. Questo implica che, dove il differenziale totale è rimasto grossomodo invariato (tra figli dei borghesi e figli degli operai), si è andata ad allargare la parte di differenziale dovuta al rendimento delle caratteristiche, in qualche modo associata a diversi modelli comportamentali, denotando quindi un ulteriore elemento di divergenza tra la classe più svantaggiata e le altre.

4. Conclusioni

In questo lavoro abbiamo analizzato l'iscrizione universitaria per classe sociale e generazione. Anzitutto, si osserva un generale innalzamento dei livelli di istruzione post-secondaria: il tasso di iscrizione tra quanti sono in possesso di un titolo di scuola secondaria superiore passa dal 31,7 per cento dei nati prima del 1940 al 47,9 per cento tra i nati del 1980-1984. I figli dei borghesi presentano elevati tassi di iscrizione universitaria già a partire dalle generazioni più anziane, fino ad arrivare a due giovani su tre dei diplomati nati tra il 1980 e il 1984. Nel corso delle generazioni si riduce la differenza tra la borghesia e la classe media e si registra un lento ma graduale avvicinamento all'interno delle due classi appartenenti alla classe media. Viceversa, pur con un incremento dell'iscrizione all'università, i figli della classe operaia si allontanano dagli altri, con differenze stabilmente nell'intorno di 40 punti verso i borghesi e un aumento del distacco verso la classe media impiegatizia e la piccola borghesia. Applicando una tecnica di scomposizione sulle differenze osservate, diminuisce la parte del differenziale dovuto alle caratteristiche tra i figli dei borghesi e gli altri. E' infatti il titolo di studio dei genitori la variabile più rilevante per spiegare i percorsi di istruzione: con il generale innalzamento del livello di istruzione all'interno della popolazione, il gap su questa variabile si è ridotto nel tempo tra figli dei borghesi, che già in larga parte discendevano da genitori altamente istruiti nelle generazioni più anziane, e gli altri, che discendono in misura crescente da famiglie più istruite. Si è invece andata ad allargare la parte di differenziale dovuta al rendimento delle caratteristiche, in qualche modo associabile a diversi modelli comportamentali, tra i figli degli operai e gli altri, denotando un ulteriore elemento di divergenza tra la classe più svantaggiata e le altre.

Riferimenti bibliografici

CHECCHI D. 2000. University Education in Italy, *International Journal of Manpower*, No. 21, pp. 177-205

- CHECCHI D., FIORIO C., LEONARDI M. 2013. Intergenerational Persistence in Educational Attainment, *Economics Letters*, No. 118, pp.229-232.
- COBALTI A., SCHIZZEROTTO A. 1994. *La mobilità sociale in Italia*, Bologna, Il Mulino.
- COMI S., 2004. *Intergenerational mobility in Europe: evidence from ECHP*, CHILD Working Paper, No. 18.
- DELLAS H., KOUBI V. 2003. Business cycle and schooling, *European Journal of Political Economy*, No. 19, pp. 843-859.
- DI PIETRO G. 2006. Regional labour market conditions and university dropout rates: evidence from Italy, *Regional Studies*, Vol. 40, pp. 617-630
- DI PIETRO G. 2004. The determinants of university dropout in Italy: a bivariate probability model with sample selection, *Applied Economics Letters*, Vol. 11, pp. 187-191.
- ISTAT 2012. *Rapporto Annuale sulla Situazione del Paese*, Roma, ISTAT.
- OAXACA R., RANSOM M. 1994. On discrimination and decomposition of wage differentials, *Journal of Econometrics*, No. 61, pp. 5-21.
- PISATI M. 2002. La partecipazione al sistema scolastico. In SCHIZZEROTTO (Ed) *Vite Ineguali. Diseguaglianze e corsi di vita nell'Italia contemporanea*, cap.4, pp.141-186, Bologna il Mulino.

SUMMARY

Social class, generation and university enrolment

Using a representative sample of the Italian population, we study the association between social origins, birth cohort and university enrolment in Italy. We also decompose the observed enrolment gap between different social class in explained and unexplained components. We show a slow convergence in enrolment rates between high class and middle class, while the low class is increasing the gap with the middle class. Moreover, even if the gap between high class and low class is stable, the unexplained part of the gap is increasing, showing a certain degree of divergence between people from low class and the others.

Andrea CUTILLO, Istat, Dipartimento delle Statistiche Sociali e Ambientali,
cutillo@istat.it

Romina FRABONI, Istat, Dipartimento delle Statistiche Sociali e Ambientali,
fraboni@istat.it

Claudio CECCARELLI, Istat, Dipartimento delle Statistiche Sociali e Ambientali,
claudio.ceccarelli@istat.it

STATISTICAL ANALYSIS OF SOCIAL NETWORKS

Agostino Di Ciaccio, Giovanni Maria Giorgi

1. Introduction

The marked increase in the use of social networks, especially among younger age groups, offers a new opportunity for statistical surveys. According to a recent study in Italy, of young people between the ages of 18 and 30, who grew up during the boom of social networks and web 2.0 (14% of the population), 91% are enrolled in a social network, 55% in a forum, 34% constantly follow one or more bloggers and 17% have their own blog.

The advent of smartphones and tablets will tend to accentuate this phenomenon: in Asia, even now, 60% of the people that frequent social networks already use a primary tool. According to our estimates, in Italy about 55% of the messages on Twitter are sent and received on a mobile phone, 54.8% of young people between 14-29 years of age have a smartphone (2012) and services that require the use of a mobile phone, for example real-time information about train delays in certain sections, are already being offered on Twitter.

Many companies are already engaged in extracting information from the social networks: to back up the launch of a new product or carry out a political poll, for example.

The potentialities of analysis are considerable: you can succeed in analysing millions of posts with costs and time extremely low compared to a traditional survey. However, analysing this information calls for special techniques that combine textual analysis with advanced statistical techniques and suitable software tools. In fact, through the social networks it is possible to “listen” to the opinions expressed by thousands or even millions of people concerning a wide range of subjects. However, these opinions are expressed textually, with language and procedures typical of the social network being used.

The BuzzMetrics application by Nielsen, which has been on the US market for 10 years now, is meeting with success on the international market and its strongpoint is the large number of sources that information can be extracted from, over 180 million blogs and 100 thousand forums throughout the world. Another example is Sysomos, a Canadian company founded in 2005 as the result of an

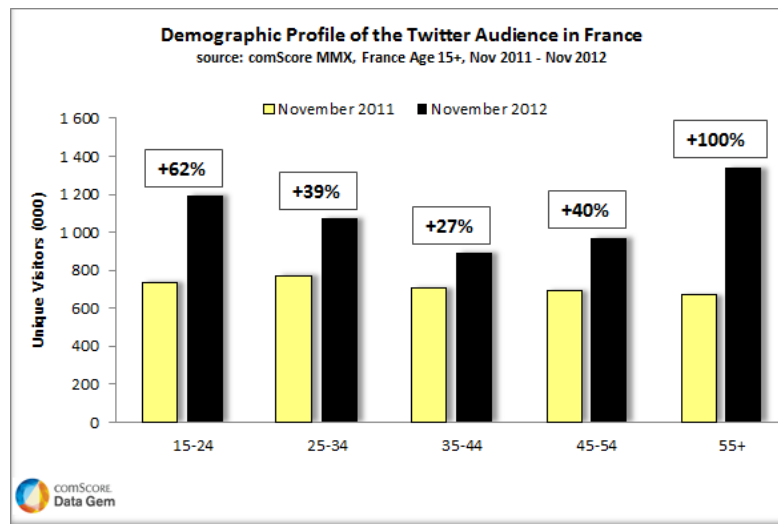
advanced research programme of the University of Toronto. Thanks to an analysis of the language combined with data mining techniques, Sysomos identifies the important subjects, the problems and the sentiments of the discussions and it activates automatic spam filtering.

In Italy an interesting example is *Voices from the Blogs*, an initiative that started off as a research project of the Milan University in October 2010 and now provides services to authorities, firms and news agencies. It uses advanced statistical techniques developed by G. King and D. Hopkins of Harvard University. Also in Italy, we have *Cogito* by Expert Systems, based on a database of millions of concepts and relations but the technology is not so suitable for analysing very short texts. The company *Blogmeter* has a similar approach.

A list of the most common applications of surveys carried out on the social networks is: web brand reputation, brand protection, analysis of the competition, market research, monitoring of social phenomena, opinion surveys, analysis and evaluation of services.

It could be particularly interested to identify the “opinion leaders”, that is to say the people or organisations that can influence the online world. many people listen to them and, above all, act on their advice.

Figure 1 - Users of Twitter in France (2011-2012).



Unlike a traditional survey, which involves a maximum of 1-2 thousand contacts, by means of the social networks we can analyse even millions of posts. For instance, about 140 million tweets are entered on Twitter every day

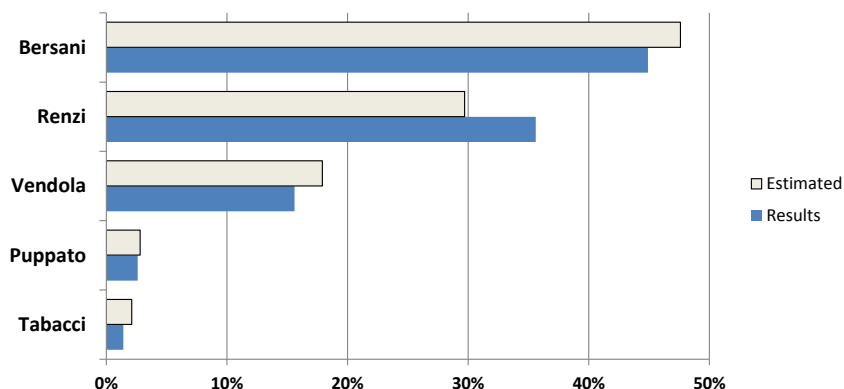
(throughout the world, 2011 survey) and the phenomenon is increasingly markedly, thanks to the fact that it does not require the use of a computer, and is even spreading to the not-so-young groups (especially women). This would make studies carried out on this social network more reliable.

A recent survey on Twitter, made in France by comScore (figure 1), shows that within the twelve months from 2011 to 2012 the users of Twitter increased by 53%, reaching 5.5 million users in November 2012 and making France the seventh largest market for Twitter. The number of visitors over 55 years of age doubled compared to the previous year and represent the most important segment of the French public, with 1.3 million users, followed by the 15-24 years group, which represents 1.2 million users (62%).

2. Advantages and disadvantages of the use of social networks

Political polls, prior to an election, are an ideal testing ground for evaluating the reliability of the use of social networks. By now almost all political elections are preceded by surveys carried out on Twitter and on blogs. This was the case, for example, in the US presidential elections, with very satisfactory results, and is now being done in Italy also. Take, for example, the primary elections of the centre-left coalition in Italy on 25 November 2012. We show in figure 2 the analysis made two days before the elections, using Sentiment Analysis on Twitter, published in the *Corriere della Sera* newspaper, the results of which were, on the whole, satisfactory.

Figure 2.- Comparison of sentiment analysis with the real results of the primary elections of the PD (democratic party)



Obviously, the tools used and the ability to use them properly are, as usual, vital for obtaining good results and the advantages and disadvantage related to this type of survey must always be taken into account.

Among the advantages we can mention:

- “Real-time” survey: an analysis can be made within a few days.
- Possible retrospective analysis: by analysing the tweets related to a given period of time it is possible to analyse the sentiment corresponding to events and actions that can be taken (advertising campaigns, promotion works and so on).
- Georeferencing: in some cases information about the place of origin of the tweets can be obtained.
- No questionnaires, and low costs: since no questionnaires have to be completed there are considerable savings in carrying out the survey.

Some of the disadvantages are:

- The sample observed may be distorted, especially if the target population is one that makes little use of internet.
- We cannot ask, we can only “listen” to what is being said on the social networks, and so we cannot put the questions we think most important.
- The analysis calls for expert and reliable researchers, familiar with the phenomenon being investigated, the social networks, the language used, the software, and the statistical techniques that have to be used.

3. The characteristics of Twitter

In addition to the text message a great deal of other information is present: 128 fields, which can come to as much as 250 if a retweet (a large number of fields is empty). Certain relevant information is present in the fields usually compiled, for example:

- *source*: indicates the method by which the tweet was sent
- *user.created_at*: date the account was created
- *user.description*: a string in which the user describes his or her account.
- *user.favourites_count*: the number of tweets the user entered as favourites since the time of registration
- *user.followers_count*: the number of followers
- *user.friends_count*: the number of users of which the reference user is a follower
- *user.name*: the name of the user
- *user.statuses_count*: the number of tweets written by the user

Taking all the tweets into consideration, we find that about 29% are retweets while 26% are replies. On the other hand, taking only tweets on political subjects into account, we find that as many as 42% are retweets and only 16% replies. Note that if we only consider “volume data”, such as the number of followers or the number of retweets, we get conflicting and not very significant information. It can be seen that the number of times a user is retweeted is not necessarily related to the number of followers.

4. Sentiment Analysis methodology and software

The term Sentiment Analysis means techniques that can automatically extract, analyse and classify opinions expressed, on the basis of a written text, usually present on the WEB.

In this article we only take into account the Twitter social network and only distinguish between positive opinions and negative opinions. It must also be said that quite often it is difficult to classify a tweet as a positive opinion or a negative opinion. This is due to the fact that, because of their brevity, the texts are not very structured and are often ironical or allusive, with links to other documents. Users generally use a language rich in metaphors and references. The texts often do not contain complete sentences and are often closely connected with the latest news or, in any case, with very recent news.

In some cases it is also possible to define neutral opinions, even though for some subjects, like politics for instance, this type of message is not very frequent.

A commercial software available for this purpose is SAS Sentiment Analysis, which offers the user three different methods of analysis:

- a supervised classification statistical model,
- a set of rules for defining sentiment,
- a hybrid system, combining the preceding options.

The statistical model basically consists of the estimation of a feedforward neural network with a binary target (but, unfortunately, the documentation does not explain this aspect). Estimation of the model is done starting with a set of “training” documents (training corpus) already classified as positive or negative. The user cannot interfere with the model used, which makes it very simple but not very flexible for an expert user.

As an alternative or, better still, in addition, a sophisticated language can be used, capable of defining a set of very complex textual rules for finding the positivity or negativity of the texts being reviewed. On the other hand, the characteristics of the tweets and their brevity, makes the use of rules in classifying

them ineffective. Included in the program we have a dictionary of terms, a dictionary of ontology, a list of synonyms, a list of stop-words (terms we do not consider useful in the analysis). It is also possible to add, to the set of rules, words identified as positive or negative in the statistical model. This list can also be manually edited subsequently. The model created “imports” rules from the statistical model but does not create a true hybrid model, which can instead be explicitly created when both the statistical model and the set of rules have been defined. In short, the classification of the sentiment of a text is obtained by combining, by means of predefined weights, the answer obtained from the statistical model and from the set of rules.

In our application on the tweets, the set of rules we defined was able to classify only a small percentage of tweets and therefore the solution is in any case dominated by the classification of the statistical model.

Instead of that software, Enterprise Miner by SAS can be used, where there is a Text Mining module with which it is possible to construct a complete and sophisticated analysis using machine learning methods and models: Neural Networks, Support vector machine, Gradient boosting, Naïve Bayes, Ensemble methods.

After a small percentage of tweets for each day have been classified and the training data-set then created, the model is constructed as follows:

- 1) filtering and cleaning the texts;
- 2) tokenizing the texts and constructing the documents x words frequency matrix;
- 3) analysing and filtering the most relevant terms;
- 4) singular value decomposition of the frequency matrix;
- 5) splitting of the texts into training, validation and testing;
- 6) applying the supervised statistical classification models, tuning the models;
- 7) comparing the models and selecting the model with the best performance;
- 8) making a new estimation of the selected model, using all the available text;
- 9) assessing the sentiment on all the available texts.

With an analysis of this kind, a percentage of between 70% and 80% of correct classification can be achieved on new tweets. What is more, since we are not interested in the classification of the single tweets but only in an overall estimate of the number, this procedure can obtain very reliable estimates, with an error very small in the forecast of the number of positive and negative tweets.

Lastly, the SAS Sentiment Analysis module is easy to apply and fast and very effective in analysing medium sized texts, thanks to ontological and grammatical dictionaries. Enterprise Miner Text Mining is more flexible and more effective for short texts like tweets but requires a good knowledge of the statistical tools used

(Neural Networks, Support vector machine, classification trees). Non-commercial software can also be used with success to obtain these analyses but it requires more work and a certain amount of computer skills.

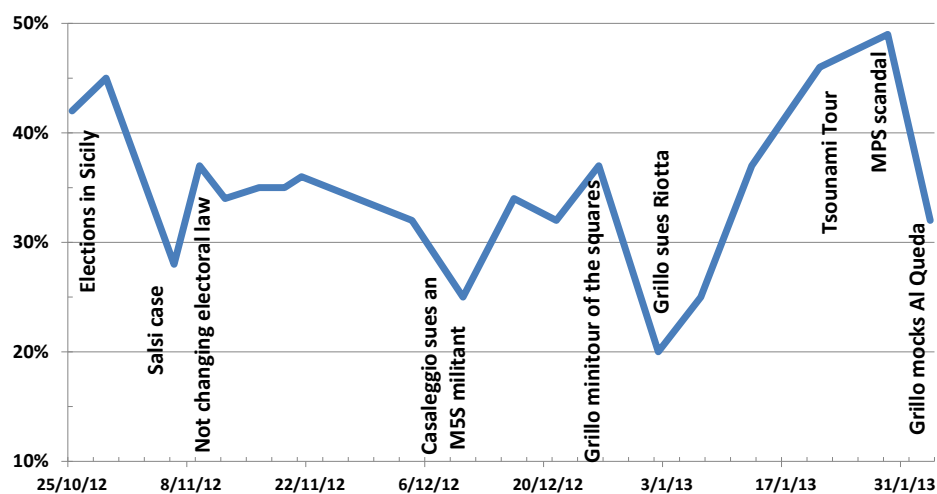
5. Results of analysis of the tweets concerning Beppe Grillo

As an example of application, we analysed the sentiment about Beppe Grillo expressed in the tweets. We fixed ourselves the aim of classifying the tweets as positive or negative.

We extracted 29646 tweets by searching for the hashtag #Grillo on 20 different days between 25 October 2012 and 3 February 2013. The first 100 tweets were classified manually for each of the 20 days.

There were many aspects that could have been taken into account in the analysis but that we neglected in this example. We did not consider the presence of retweets, of links to other pages, we did not consider information about the authors of the tweets or evaluate the presence of opinion leaders.

Figure 3. Pattern of positive sentiment about Beppe Grillo during the period 25/10/2012 – 3/2/2013



In figure 3 we show the results of the analysis made with SAS Enterprise Miner. We also included in the graph the presence of particular events that occurred during the period and that had great influence on discussion on Twitter. Note that the graph actually represents the evolution of the discussion and does not

give the number of supporters of Grillo. In the graph we show the percentage of favourable tweets, for each day analysed. It can, however, be noted that the average of positive tweets (around 33%) is very close to an estimate of the votes obtained by Grillo in the elections of February in the youngest age group (18-24).

6. Conclusions

The analysis we showed, made as part of a degree thesis (Claudia Proia 2013), shows the potential inherent in the statistical analysis of textual data taken from the WEB. It would, in fact, be unthinkable to conduct a similar survey using traditional tools without having enormous resources. In our opinion, the results of a sentiment analysis can be considered very interesting if one takes into account the necessary statistical methodologies and the characteristics of the reference group. The analysis can be conducted in more depth than shown in the previous paragraphs, including the analysis of: followers, opinion leaders, geolocation data, other information from blogs and from other information channels. Taking these data into account certainly will make the analysis more complex but also more complete.

References

- SAS Sentiment Analysis Studio: Building models. Course notes (2011). SAS Institute Inc., Cary, USA.
- PROIA C. 2013. Sentiment Analysis for the statistical classification of tweets. *Graduate thesis*. Department of Statistics, Sapienza, University of Rome.

SUMMARY

Analysis of information, expressed in a textual and unstructured manner on the web, is becoming increasingly common in web-marketing or political polls. These analyses require an advanced statistical methodology that combines text-mining with machine learning. This article analyses the potential of this type of analysis and gives an example of sentiment analysis application to a political poll.

Agostino DI CIACCIO, Full Professor, "Sapienza" University of Rome,
agostino.diciaccio@uniroma1.it

Giovanni Maria GIORGI, Full Professor, "Sapienza" University of Rome,
giovanni.giorgi@uniroma1.it

GEOGRAFIA DELLA FAME: SESSANTA ANNI DOPO!¹

Luigi Di Comite, Simona Giordano

1. Introduzione

La contrapposizione tra potenziale di crescita delle popolazioni e potenziale di crescita delle sussistenze (alias, delle risorse alimentari) è questione vecchia di secoli. Agli inizi degli anni Cinquanta del XX secolo apparve un più che interessante volume dal significativo titolo “*Geografia da fame*”; in esso l’autore, Josuè De Castro, faceva il punto sui molteplici quesiti che possono emergere a livello planetario dai problemi posti dalla malnutrizione, con specifico riferimento agli anni immediatamente successivi al secondo conflitto mondiale. Da allora, cioè nel corso di circa un sessantennio, pur avendo la realtà subito notevoli modificazioni, il problema della fame nel mondo non è mai venuto meno. Eloquenti risultano i dati del Report pubblicato dalla FAO, *The State of food and agriculture 2012*, e le cifre contenute nel Report “Indice Globale della Fame (Global Hunger Index – GHI)” presentato dal Cesvi, in collaborazione con Link 2007, ISPI, Comune di Milano, e con il Patrocinio di Expo 2015. Nel Report 2012 vengono analizzati oltre 120 Paesi, 20 dei quali presentano un “Indice Globale della Fame” allarmante o estremamente allarmante, con il 50% ed oltre della popolazione in stato di denutrizione. Fra essi si ritrovano Burundi, Eritrea e Haiti. Per il calcolo del GHI vengono combinati tre indicatori: la percentuale di popolazione denutrita, il tasso di mortalità infantile e la percentuale di bambini sottopeso. Nonostante a livello globale l’Indice sia sceso dai 19,8 punti del 1990 ai 14,7 del 2012, l’Africa Sub Sahariana e l’Asia Meridionale mantengono tuttora valori elevati con 22,5 e 20,7 punti. In tali aree del mondo si trovano i cosiddetti Paesi a Sviluppo Minimo (PSM), sui quali si concentrerà maggiormente l’attenzione del presente articolo. Lo scenario che emerge è quello di una sfida senza precedenti, sfamare una popolazione che nel 2050 dovrebbe raggiungere i 9 miliardi di persone, in realtà equivalenti a 12 miliardi (attuali) se si considera la progressiva adozione di stili di vita tipici dei Paesi industrializzati da parte dei

¹ I paragrafi 1 e 2 vanno attribuiti al prof. L. Di Comite, quelli successivi alla dott.ssa S. Giordano.

Paesi cosiddetti emergenti. A mettere a dura prova la sostenibilità dello sviluppo globale e la lotta alla fame concorrono vari fattori, dalla sempre maggiore diffusione della dieta occidentale, all'aumento dei prezzi dei beni alimentari, alla scarsità di terra e di acqua, fino agli evidenti cambiamenti climatici. L'articolo intende inoltre analizzare, a sessant'anni circa dalla pubblicazione dei volumi di J. De Castro, lo "stato" della fame nel mondo, nelle sue implicazioni quantitative e qualitative, nonché esplorare un ventaglio di possibili soluzioni o strategie volte a raggiungere una crescita economica sostenibile.

2. Nutrire 9 miliardi di persone: una sfida possibile?

Alla fine del XVIII secolo fu Malthus a diffondere, attraverso le sue opere, un concetto che risulta ancora oggi molto in auge, ovvero l'idea che il continuo e costante aumento della popolazione ponga in serio pericolo la sussistenza stessa del nostro pianeta. Dalla constatazione che la popolazione tenderebbe a raddoppiare ogni 25 anni, e che il medesimo tasso di crescita non può applicarsi alle superfici coltivabili, l'economista inglese giunse a proporre come rimedio il controllo rigido delle nascite. Due secoli più tardi, il Club di Roma e, a seguire, gli odierni neo-malthusiani riprendono tale impostazione; il primo sostenendo che l'incremento demografico a ritmi esponenziali conduce inevitabilmente al degrado ambientale e all'innalzamento dei prezzi alimentari, i secondi, interrogandosi sul numero di esseri umani che la Terra può ragionevolmente sopportare, giungendo alla conclusione che il controllo delle nascite e la riduzione dei consumi da parte dei Paesi industrializzati sono azioni necessarie a preservare l'ambiente e ad ovviare all'impossibilità oggettiva di realizzare un aumento della produzione di alimenti in linea con i fabbisogni della popolazione globale.

Può la soluzione risiedere nel cercare di produrre di più? Stando alle stime diffuse periodicamente dal World Population Prospects delle Nazioni Unite², nel 2050 la popolazione mondiale dovrebbe toccare la cifra di circa 9,2 miliardi di persone, il che comporterebbe la necessità di raddoppiare la produzione alimentare mondiale. Viene da chiedersi se ciò risulti realistico. Già negli anni Sessanta la cosiddetta Rivoluzione Verde aveva fornito una prima risposta a tale questione; l'impiego della genetica e della chimica, nonché una serie di politiche pubbliche ad

² Fonte : World Population Prospects, the 2012 Revision : <http://esa.un.org/wpp/>.

hoc, avevano consentito di raddoppiare nei trenta anni precedenti la produzione di molti alimenti.³

Ad oggi, tuttavia, è necessario muoversi in una prospettiva di lungo termine, prendendo in considerazione le tensioni evidenti che si stanno delineando sul mercato dei prodotti agricoli, sia dal lato della domanda che da quello dell'offerta.

Passando ad analizzare i fattori di tensione, risulta importante riprendere il discorso della "esplosione demografica".⁴ In effetti, volendo scomporre i fattori, è stato calcolato che fino al 2050 sarà necessario un incremento della produzione agricola mondiale del 76% al fine di pareggiare il dato demografico e del 2% al fine di fronteggiare l'evoluzione della composizione della popolazione (cfr. Collomb, 1999). A questo va aggiunta la pressione derivante dal costante modificarsi degli stili di vita alimentari; la nuova "classe media", in particolare in paesi quali Cina, India e nel Sud-Est asiatico, sta progressivamente aumentando il consumo di alimenti di origine animale (carne, pesce, latticini), la cui produzione richiede una considerevole quantità di risorse.⁵ Dal lato dell'offerta, non può essere trascurato il diffondersi, in molti Paesi, di politiche volte ad incentivare la produzione di bio-carburanti, con l'inevitabile effetto collaterale di sottrarre terre fertili alla produzione agricola. L'effetto congiunto dell'urbanizzazione, del riscaldamento globale e dell'inquinamento ha determinato una continua diminuzione della quantità di terra arabile disponibile per ogni 1000 abitanti, passata dai circa 410 ettari negli anni Sessanta ai circa 220 del 2005, fino ad una stima di appena 160 ettari per il 2050.⁶ E ciò che rende tale situazione ancora peggiore è il verificarsi di una sorta di neo-colonialismo: i paesi con maggiori risorse finanziarie acquistano o affittano terre coltivabili nei paesi più poveri a fini speculativi e con un ritorno esclusivamente "domestico", peggiorando la loro già difficile situazione (cfr. Brown, 2012, pp101-113).

3. I Paesi a sviluppo minimo (PSM)

Il gruppo di Paesi a Sviluppo Minimo preso in esame comprende trentuno Paesi localizzati esclusivamente in Africa e Asia, ad eccezione di Haiti (cfr. Stigliano,

³ Nei Paesi a sviluppo minimo (PSM), la produzione di cereali è passata da 60 milioni di tonnellate alla fine degli anni Sessanta a 120 milioni di tonnellate alla fine degli anni Novanta. Fonte FAO: <http://faostat.fao.org/site/567/default.aspx#ancor>.

⁴ Risulta immediato il legame fra adolescenza e aumentate richieste energetiche, al fine di accompagnare e favorire la crescita.

⁵ Si stima che il solo impiego di cereali necessari ad ottenere un chilogrammo di carne di pollo, maiale e vitello risulti rispettivamente pari a 4, 6 e 12 chilogrammi.

⁶ Fonte World Bank Indicators Database: <http://data.worldbank.org/topic/agriculture-and-rural-development>.

2012). Essi presentano un elevato grado di eterogeneità, quale emerge da una prima analisi dei dati riportati nella Tabella 1; è di immediata lettura il dato relativo al reddito pro-capite 2011⁷, il quale, salvo rari casi, non risulta superiore ai mille dollari, confermando la situazione di estrema difficoltà dei Paesi in questione.

I dati relativi alla mortalità nei primi cinque anni di vita sono certamente i maggiormente significativi nel quadro del presente articolo, e risulta immediata la drammaticità di un tasso medio superiore al 100‰, in particolare se confrontato con il dato UE (4,8‰) e quello USA (7,5‰). L'importanza di tali cifre risiede in due motivazioni: la prima è legata alle precarie condizioni igienico-sanitarie e alla sottanutrizione in cui le donne portano avanti il loro periodo di gravidanza, con inevitabili conseguenze sullo sviluppo del nascituro; la seconda attiene direttamente alle negatività legate alla stessa sottoalimentazione nel corso dei primi anni di vita. I neonati e i bambini, lì dove non subentra la morte, vanno incontro a serie conseguenze psico-fisiche, dal ritardo nello sviluppo fino a disturbi mentali e di apprendimento. Inoltre, per quanto appaia paradossale, gli stessi bambini sottanutriti risultano maggiormente soggetti a sviluppare problematiche quali obesità, diabete e colesterolo alto in età adulta, qualora abbiano l'opportunità di emigrare in un contesto urbano.

Un ulteriore dato certamente degno di nota è quello relativo alla percentuale di popolazione rurale, generalmente superiore al 70% a fronte del 26% dell'UE e del 17,6% degli USA. Risultano, quindi, evidenti le "riflessioni" sull'importanza fondamentale che l'agricoltura riveste in tali Paesi e sulla necessità di avviarsi e tutelare un concreto processo di sviluppo della stessa.

4. Lo sviluppo dell'agricoltura come strada prioritaria per la sicurezza alimentare

Il 2015 è alle porte; si avvicina, dunque, l'anno stabilito dalle Nazioni Unite per il raggiungimento dei cosiddetti "Obiettivi di sviluppo del Millennio", fra i quali quello di assicurare alla intera popolazione mondiale l'accesso ad un adeguato livello di nutrimento. Attuare quella che è stata definita come la Rivoluzione Doppia Verde (cfr. Griffon, 2002), ovvero produrre di più consumando meno, risulta l'unica soluzione al problema della fame; porre in essere una serie di politiche efficaci volte ad instaurare e sviluppare una agricoltura rispettosa dei singoli eco-sistemi. La strada verso il raggiungimento della sicurezza alimentare

⁷ Fonte : World Bank GNI per capita in US Dollars
<http://data.worldbank.org/indicator/NY.GNP.PCAP.CD>.

passa attraverso quattro fondamentali tappe: innovazione, mercati, persone e politica.

Tabella 1 – Indicatori socio-economici per alcuni Paesi a Sviluppo Minimo (PSM)

Paese	Reddito pro-capite 2011 (in dollari USA)	Popolazione rurale (in %)	Mortalità infantile (in ‰)
Angola	3.830	41	158
Benin	780	55	106
Burkina Faso	580	73	146
Burundi	250	89	139
Ciad	720	78	169
Congo	190	66	168
Eritrea	430	79	68
Etiopia	370	83	77
Guinea	430	65	126
Madagascar	430	67	62
Malawi	360	84	83
Mali	610	65	176
Mozambico	460	69	103
Niger	360	82	125
Ruanda	570	81	54
Senegal	1.070	57	65
Sierra Leone	460	61	185
Somalia	Low	62	180
Sudan	1.310	67	86
Tanzania	540	73	68
Togo	570	62	110
Uganda	510	84	90
Zambia	1.160	61	83
Afghanistan	470	76	101
Bangladesh	780	72	46
Cambogia	820	80	43
Laos	1.130	66	42
Myanmar	Low	67	62
Nepal	540	83	48
Yemen	1.070	68	77
Haiti	700	47	70
UE	34.053,6	26	4,8
USA	48.620	17,6	7,5
Mondo	9.514,2	48	51,4

Note: Fonte Dati World Bank Database Pil Reddito pro-capite è espresso in US\$. Il dato della produzione agricola e della produzione alimentare è calcolato per singolo Paese in base ad un indice calcolato usando come base quello del 2004-2006. Viene esclusa la produzione di mangime e quella di caffè e tè, in quanto privi di elementi nutritivi. La Popolazione rurale è stimata in base a dati Nazioni Unite; il Tasso di mortalità sotto i 5 anni, dati in base a stime Level & Trends in Child Mortality (UNICEF, WHO, World Bank, UN DESA, UNPD).

L'innovazione risulta necessaria a dotare l'agricoltura di un'ampia gamma di tecnologie, frutto di una attività di ricerca e sviluppo sia pubblica che privata; tecnologie che devono necessariamente essere economicamente accessibili ed eco-compatibili. Altresì fondamentali risultano la creazione e la gestione di una rete efficiente di mercati a supporto delle produzioni agricole e dei redditi delle

aziende, in particolare di quelle medio-piccole. Ne segue come corollario l'attenzione alle persone, ai piccoli contadini e alle donne, le quali garantiscono, nel loro molteplice ruolo di contadine, madri ed educatrici, un legame fondamentale tra produzione e consumo. Infine, la politica deve necessariamente essere lungimirante, favorendo gli investimenti in ricerca e sviluppo e onorando gli impegni presi all'interno dei trattati internazionali.

Illuminante risulta il Report della FAO dal titolo "The State of food and agriculture", il quale fornisce un quadro completo dei flussi di investimenti nei paesi a basso e medio reddito. I dati disponibili mostrano come siano gli agricoltori ad investire nelle proprie aziende circa quattro volte l'ammontare degli investimenti effettuati dai governi nel settore agricolo, superando anche gli apporti dei donatori internazionali privati e degli investitori stranieri. Ciò spiega perché siano gli agricoltori a dover essere al centro di qualsiasi strategia volta ad aumentare la quantità e l'efficacia degli investimenti in agricoltura. La creazione delle migliori condizioni per un favorevole clima di investimento dipende in buona parte dai mercati e dai governi; i primi generano incentivi di prezzo tali da segnalare ad agricoltori e altri privati imprenditori quando e dove esistono opportunità di reddito; i governi sono responsabili della creazione dei presupposti legali, politici e istituzionali che consentono agli investitori privati di operare in modo socialmente responsabile.

Gli elementi da tenere in considerazione risultano: buona governance, stabilità macroeconomica, politiche commerciali trasparenti, efficaci istituzioni di mercato e pieno rispetto per i diritti di proprietà. Importante risultano altresì il meccanismo degli incentivi per gli investimenti in agricoltura e la tassazione relativa allo stesso settore agricolo, nonché la incorporazione dei costi ambientali e la conseguente creazione di meccanismi volti ad agevolare il passaggio a sistemi di produzione sostenibili. Sono i piccoli agricoltori a dover affrontare difficoltà specifiche quali povertà estrema, deboli o inesistenti diritti di proprietà della terra, scarso accesso ai mercati e ai servizi finanziari ed estrema vulnerabilità agli shock internazionali. La creazione di organizzazioni capaci di associare i piccoli agricoltori potrebbe consentire il superamento dei suddetti vincoli e la nascita di una efficace rete di protezione tale da contrastare la carenza di risparmi privati, il difficile accesso al credito e la mancanza di assicurazione contro i rischi. Ulteriore problema è posto dall'aumento dei flussi internazionali di fondi diretti alle grandi acquisizioni di terreni da parte di società private; in assenza di un solido quadro normativo di riferimento, siffatte operazioni rischiano di produrre un negativo impatto a livello locale in termini ambientali, sociali ed economici e di compromettere la futura crescita delle comunità. La sfida è dunque quella di migliorare le capacità dei governi e delle comunità locali di negoziare contratti di investimento realmente vantaggiosi.

Ancora, gli investimenti pubblici in ricerca agricola e sviluppo, in istruzione e in infrastrutture rurali portano a risultati superiori a quelli ottenuti attraverso sussidi diretti, rafforzando il tessuto socio-economico delle aree rurali interessate, con un effetto di lungo periodo in termini di crescita del settore agricolo e conseguente riduzione della povertà e della diffusione della fame.

Infine, numerose risultano le soluzioni “originali” ed insolite; in primo luogo l’agroforestazione, ovvero la pratica che associa le colture ad attività di gestione e conservazione di foreste e boschi limitrofi ai campi, e che costituisce una importante fonte di sostentamento per circa 1,6 miliardi di persone nel mondo. Vengono lavorati e venduti frutta, oli, rimedi curativi a base di erbe, e, ovviamente, il legname. Inoltre i cibi selvatici delle foreste risultano notevolmente nutrienti, preziosi in caso di carestie o di distruzione delle colture a causa di siccità, alluvioni, parassiti o malattie. Altresì fondamentale è il ruolo delle foreste nel promuovere la biodiversità, proteggere la terra dall’erosione e dalle alluvioni e nel ridurre la quantità di gas serra nocivi nell’ambiente.

Ancora, le coltivazioni idroponiche con le quali si producono piante senza l’utilizzo di terra ma solo con una miscela di nutrienti disciolti nell’acqua, efficienti in quanto richiedono circa un settimo del terreno e un quarto dell’acqua normalmente utilizzati per produrre tradizionalmente gli stessi quantitativi. Esempi concreti sono le “zattere agricole” in Bangladesh, idonee a contrastare gli allagamenti dei campi durante la stagione dei monsoni; esse hanno consentito ad una vasta parte della popolazione di uscire dal tunnel della fame e della povertà, di nutrirsi autonomamente e di vendere i prodotti in eccesso sui mercati.

Riferimenti bibliografici

- AA.VV. 1999. Alimentation mondiale 2050; Bien nourrir les hommes sans dégrader la Planète. Parigi: L’Harmattan.
- AA.VV. 2012. Comment nourrir le Monde?. La Tour d’Aigues: Editions de l’Aube et Passion Céréales.
- AA.VV. 1999. Géopolitique de la faim. Parigi: Presses Universitaires de France.
- AA.VV. 2009. Nourrir les hommes. Un dictionnaire. Neuilly: Atlante.
- BAUDOIN F., PARLONGUE D. 2009. La faim dans le Monde: crises d’aujourd’hui et défis de demain. in Collection Libres Opinions. Paris: Presses des Mines.
- BROWN L.R. 2012. Full Planet, Empty Plates. The new Geopolitics of Food scarcity. New York: Earth Policy Institute – W.W. Norton&Company.
- BRUEGEL M. et al. 2009. Profusion et Pénurie. Les hommes face à leurs besoins alimentaires. in Collection Tables des Hommes. Rennes: Presses Universitaires de Rennes.
- BRUNEL S. 2002. Famines et Politique. Paris Presses de Sciences Politiques.

- CALAME M. 2008. La tourmente alimentaire. Pour une politique agricole mondiale. Paris Editions Charles Léopold Mayer.
- CHARVET J.P. 2010. Atlas de l'Agriculture. Comment pourra-t-on nourrir le Monde en 2050?. Paris: Editions Autrement.
- COLLOMB P. 1999. Une voie étroite pour la sécurité alimentaire d'ici à 2050. Paris: Economica FAO. <http://www.fao.org/docrep/003/x3002f/x3002f00.HTM>
- CRIBB J. 2010. The coming famine. The global food Crisis and what we can do to avoid it. Los Angeles: University of California Press.
- DE CASTRO J. 1946. Geografia da fome: a fome no Brasil. Rio de Janeiro: O Cruzeiro.
- DE CASTRO J. 1952. Geografia da fome. São Paulo: Edit. Brasiliense.
- FUMEY G. 2012. Géopolitique de l'alimentation. Auxerre: Sciences Humaines Editions.
- GHERSI G. et al. 2005. Nourrir 9.000.000.000 d'hommes. Parigi: Association pour la diffusion de la pensée française.
- GRIFFON M. 2002. Révolution Verte, Révolution Doublement Verte. Quelles technologies, institutions et recherche pour les agricultures de l'avenir?. in Mondes en développement, 1/2002 (n.117).
- STIGLIANO M. 2012. In tema di paesi a sviluppo minimo. in Rivista Italiana di Economia, Demografia e Statistica, Vol. LXVI n. 1.

SUMMARY

Geography of famine: sixty years after!

The issue regarding the contrast between the potential population growth and the potential increase in sustenance (alias, food supply) is as old as time. At the beginning of the Fifties a book entitled "The Geography of Hunger" was published; the author, Josué De Castro, took stock of the multiple questions regarding the issues deriving from hunger and malnutrition immediately after the Second World War. Since then, after about sixty years, despite reality has experienced remarkable changes, the issue of hunger is still a living matter. The following article will examine the "state" of World hunger as well as explore a range of feasible solutions and strategies aiming at reaching a sustainable economic growth in developing countries. The article will argue that the common denominator of all these solutions and strategies is investing in productive, innovative and sustainable agriculture.

Luigi DI COMITE, Università di Bari, luigi.dicomite@uniba.it
Simona GIORDANO, Università di Bari, simo_giordano@yahoo.it

TIPOLOGIE DI COLLEGAMENTO SU FERRO TRA CAPOLUOGHI DI PROVINCIA: IL CASO DI TRENITALIA

Valentina Ferri, Leonardo Palmisano

1. Introduzione

L'obiettivo del lavoro è di effettuare un'analisi delle soluzioni offerte al cliente che acquista un biglietto on-line dalla compagnia Trenitalia, principale società italiana del trasporto ferroviario ed S.p.A. partecipata al 100% da Ferrovie dello Stato Italiane. Le tratte considerate sono quelle che collegano i capoluoghi di regione. I dati sono stati raccolti dal sito ufficiale della società, in cui all'utente è consentito effettuare l'acquisto del biglietto on line. L'interfaccia che il sito offre, consentendo l'acquisto via internet, prevede nella ricerca la possibilità di avere soluzioni con meno cambi e la possibilità di avere tutte le soluzioni (con maggior numero di cambi). Nella ricerca, finalizzata alla raccolta dei dati, sono state utilizzate le opzioni "Tutti i treni" perché comprende anche treni diversi dalle frecce e "Tutte le soluzioni", così da poter avere non solo le migliori in termini di tempo di percorrenza ma tutte le soluzioni di viaggio. I dati riguardano esattamente la data del 30-5-2013, e si riferiscono agli spostamenti tra i capoluoghi di regione.

2. Metodologia

Al fine di raggruppare le situazioni simili dal sono state sperimentate diverse tecniche, si è esclusa la metodologia K-means in quanto i risultati non erano soddisfacenti. La cluster analysis gerarchica con metodo di Ward (1) (metodo che minimizza la varianza ed indica, per ogni tappa del processo, i due gruppi dalla cui fusione deriva il minimo incremento possibile della devianza "entro"), ha fornito alcune soluzioni su cui si sono individuati dei criteri che hanno portato alla scelta di utilizzare la ripartizione in otto clusters. Anzitutto dall'osservazione del dendrogramma, i clusters presentavano molti rami che si riunivano entro una ristretta gamma di distanze, in secondo luogo si è riscontrata la presenza di rami dell'albero tassonomico più lunghi (questi ultimi rappresentano le distanze rispetto agli altri clusters). Nello specifico, inoltre, gli otto clusters individuati per l'analisi

delle soluzioni offerte all'utente del sito di Trenitalia, sono risultati maggiormente rispondenti all'analisi esplorativa e più efficaci in quanto ben definiti.

$$DEV_T = \sum_{s=1}^p \sum_{i=1}^n (x_{is} - \bar{x}_s)^2 = \sum_{i=1}^n \sum_{s=1}^p (x_{is} - \bar{x}_s)^2 \quad (1)$$

Le variabili incluse nell'analisi riguardano innanzitutto la distribuzione dei treni durante la giornata, a tal proposito si sono individuate tre fasce: la prima dalle 6 alle 14; la seconda dalle 14 alle 22 e la terza dalle 22 alle 06 del mattino. Oltre questa differenziazione, si sono prese in considerazione le soluzioni dirette e quelle con cambio. Delle soluzioni con cambio si è effettuata un'analisi approfondita valutando se si tratta di un solo cambio o più di un cambio. Inoltre, quando le soluzioni offerte dalla compagnia sono dirette, si è individuata la quantità di frecciarossa, freccia Bianca, frecciargento, intercity/eurocity, intercity night/euronight night, bus e regionali. L'analisi è stata resa più approfondita da altri dati: è stata rilevata la quantità delle soluzioni con un cambio di cui almeno un treno fosse una "Freccia" o un intercity. Le variabili incluse nella cluster analysis riguardano l'intera giornata di Giovedì 30 maggio 2013 e sono: Soluzioni totali; Soluzioni dalle 6.00 alle 14.00; Soluzioni 14.00-22.00; Soluzioni 22.00-6.00; Soluzioni con più di un cambio; Soluzioni con un cambio di cui entrambi intercity/eurocity/regionali (anche notturni); Soluzioni con un cambio di cui almeno uno è Freccia Bianca/ Frecciargento/ Frecciarossa; Soluzioni dirette con Freccia Bianca; Soluzioni dirette con Frecciargento; Soluzioni dirette con Frecciarossa; Soluzioni dirette con InterCity /EuroCity; Soluzioni dirette con Intercity Night/Euronight; Soluzioni dir. con regionali e bus.

3. Risultati dell'analisi dei clusters

Dall'analisi emerge che il **primo cluster** comprende le tratte con un numero medio di 10,2 soluzioni nella giornata, con un campo di variazione da 4 a 22 soluzioni per tratta, fortemente concentrate nella mattinata. Osservando la distribuzione nella giornata, dalle 6 alle 14 sono disponibili 6 soluzioni in media per ognuno dei percorsi, dalle 14 alle 22 circa 3,3 e successivamente 1,1 ca. Non risultano presenti soluzioni dirette nella giornata e si rilevano in media 9,1 soluzioni per tratta in cui i passeggeri devono effettuare più di un cambio, affrontando così un viaggio meno confortevole (Tab. 1). Di questo cluster fanno parte 128 tratte, in maggior misura 29/128 sono soluzioni che collegano i

capoluoghi da Nord e arrivano al Sud¹ e quelle che partono da un capoluogo del Sud e arrivano a Nord 28/128 (Tab.2). Anche il **cluster 2** è molto numeroso e prevede una scarsa quantità di soluzioni durante la giornata: 9,3 in media, il minimo sono 2 soluzioni per tratta e il massimo 25. La distribuzione tra le fasce orarie prese in considerazione non è equa, la maggiore concentrazione di soluzioni si registra nelle ore del mattino. Nonostante un gran numero di soluzioni con cambio per il primo e il secondo cluster, si registra, in questo caso un minor numero di soluzioni con più di un cambio, aumenta così la qualità del viaggio in quanto il passeggero non è costretto ad effettuare molti spostamenti con bagagli da una vettura all'altra. Dunque, in questo raggruppamento sono presenti viaggi leggermente più confortevoli grazie alla maggior quantità di viaggi in cui è previsto un unico cambio. Inoltre, approfondendo l'analisi di tali soluzioni in cui è previsto un solo cambio, all'interno si rileva un numero significativo di soluzioni con almeno una "Freccia", ciò presuppone che una parte del viaggio risulterà abbreviata in termini di tempo e in una vettura abbastanza recente con una serie di comfort (Tab. 1). Ciò che purtroppo stupisce è che all'interno di questo cluster rientrano le tratte che collegano i capoluoghi del Mezzogiorno e che dunque, per la breve distanza che intercorre tra gli stessi, le soluzioni con cambio non dovrebbero rappresentare una quota significativa (Tab. 2).

Il terzo raggruppamento si può definire intermedio, esso comprende i collegamenti in cui per ogni percorso considerato si prevedono da 13 a 35 soluzioni, dunque un range decisamente più ampio rispetto ai precedenti. Si tratta, peraltro, di soluzioni ben distribuite tra la mattina e il pomeriggio, molto meno presenti risultano le soluzioni notturne. Nel raggruppamento sono contenuti quei collegamenti con maggiore diversificazione tra soluzioni dirette e con cambio: le prime sono mediamente 17,5 per tratta, raggiungendo un massimo di 34 soluzioni. I collegamenti diretti sono maggiormente FrecciaBianca, da considerare una caratteristica che pone in una situazione intermedia le tratte che ne fanno parte proprio perché non sono diretti ad alta velocità ma con maggiori comfort rispetto ai regionali o agli intercity (Tab. 1). Di questo cluster fanno parte 14 tratte di cui 4 sono collegamenti Nord-Nord.

Le tratte che fanno parte del **quarto cluster** prevedono in media 18,3 soluzioni nell'arco della giornata, il minimo è di 11 soluzioni, il massimo è di 32. Nonostante l'offerta al passeggero sia significativa, nel cluster si rileva un numero cospicuo di soluzioni con cambio, mediamente 16,2, rispetto ai 2,1 diretti. Trattasi di soluzioni prevalentemente con 1 cambio, uno dei quali nella maggior parte dei casi è una

¹ Sud e Mezzogiorno verranno utilizzate in maniera equivalente, con tale definizione si intendono Abruzzo, Molise, Basilicata, Campania, Calabria e Sicilia. Le restanti regioni sono classificate come Centro e Sud secondo la distribuzione prevista dall'Istat.

“Freccia”, a garanzia di una maggiore velocità/maggiore comfort (Tab. 1). Il cluster comprende principalmente le tratte Nord-Nord: 19/62, le tratte più omogenee che prevedono un elevato numero di cambi riguardano principalmente aree che per la distanza, spesso esigua, non dovrebbero mostrare tale situazione (Tab. 2).

Tabella 1 - Clusters: valore medio, minimo e massimo per soluzione proposta da Trenitalia e relative caratteristiche.

Cluster	Soluzioni Val.	Solh 6-14	Solh 14-22	Solh 22-6	Tot diret	Tot cambio	Tot con 1 cambio	Tot > 1 cambio	2 intercity/reg	2 di cui 1 freccia	Dir Frecciabianca	Dir Frecciargento	Dir Frecciarossa	Dir IC/EC	Dir ICN/En	Dir Reg Autobus	
		Max	Min	Med	Max	Min	Med	Max	Min	Med	Max	Min	Med	Max	Min	Med	Max
C1 u 1	Max	22	13	12	3	1	22	14	22	6	6	0	0	0	1	0	0
	Min	4	1	0	0	0	4	0	4	0	0	0	0	0	0	0	0
	Med	10,2	5,9	3,3	1,1	0,0	10,3	2,1	9,1	0,5	0,6	0,0	0,0	0,0	0,0	0,0	0,0
C1 u 2	Max	25	11	12	3	22	18	14	8	7	10	6	18	0	4	2	6
	Min	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Med	9,3	4,8	3,9	0,6	2,7	6,8	4,7	2,0	1,7	3,0	0,2	1,1	0,0	0,4	0,2	0,5
C1 u 3	Max	35	17	16	4	34	19	17	2	5	17	20	7	1	6	7	12
	Min	13	8	4	0	8	0	0	0	0	0	0	0	0	0	0	0
	Med	25,4	12,9	10,8	1,8	17,5	8,9	8,9	0,4	1,2	8,1	10,5	1,1	0,3	2,2	2,7	1,6
C1 u 4	Max	32	23	14	6	15	31	25	16	10	23	3	2	12	3	2	2
	Min	11	4	0	0	0	6	0	0	0	0	0	0	0	0	0	0
	Med	18,3	9,4	6,5	2,4	2,1	16,2	11,7	4,2	1,7	10,5	0,3	0,0	0,9	0,4	0,2	0,1
C1 u 5	Max	22	13	11	4	8	18	16	8	15	11	6	0	0	1	2	5
	Min	9	4	3	0	0	9	8	0	7	0	0	0	0	0	0	0
	Med	16,3	8,3	6,5	1,5	1,2	14,8	12,9	2,5	11,2	2,6	0,5	0,0	0,0	0,2	0,2	0,7
C1 u 6	Max	68	33	35	5	63	12	12	1	12	0	1	24	24	6	3	8
	Min	57	27	26	2	47	1	1	0	2	0	0	24	22	0	1	0
	Med	60,4	29,6	29,8	3,4	56,4	4,0	3,6	0,2	4,0	0,0	0,2	24,0	22,8	4,6	2,0	3,8
C1 u 7	Max	59	28	28	4	55	21	21	2	10	20	9	3	37	20	3	10
	Min	23	9	12	1	16	0	0	0	0	0	0	0	0	2	0	0
	Med	37,2	17,6	17,3	2,3	31,2	6,0	5,6	0,4	1,8	3,6	1,4	1,1	17,6	6,9	1,1	3,2
C1 u 8	Max	52	18	31	6	43	11	4	0	1	4	9	17	13	4	2	27
	Min	33	11	14	3	32	0	0	0	0	0	2	0	0	0	1	14
	Med	36,8	14,8	17,5	4,5	36,5	0,3	0,3	0,0	0,3	0,0	4,8	4,3	3,3	2,3	1,3	20,8

Fonte: Elaborazioni proprie su fonte RFI

I collegamenti tra capoluoghi che fanno parte del **quinto cluster** non hanno un gran numero di soluzioni in media nella giornata: 16,3. Per quanto riguarda le tratte con un solo cambio, la maggior parte sono caratterizzate dalla combinazione di due intercity/regionali: mediamente sono 11,2. Le soluzioni dirette sono in media 1,2

per tratta: un numero davvero modesto (Tab. 1). Di tale raggruppamento fanno parte maggiormente le tratte riguardanti cinque capoluoghi centraliani su 13 tratte in totale presenti nel cluster (Tab. 2).

Tabella 2 - Numero di tratte per zona del capoluogo di partenza e d'arrivo (v.a.)

	Partenze	Arrivi			Tot. Riga
		Nord	Centro	Sud	
Cluster 1	<i>Nord</i>	18	10	29	57
	<i>Centro</i>	9	0	11	20
	<i>Sud</i>	28	8	15	51
	<i>Tot. Colonna</i>	55	18	55	128
Cluster 2	<i>Nord</i>	2	7	14	23
	<i>Centro</i>	9	3	13	25
	<i>Sud</i>	11	11	25	47
	<i>Tot. Colonna</i>	22	21	52	95
Cluster 3	<i>Nord</i>	4	2	2	8
	<i>Centro</i>	2	1	1	4
	<i>Sud</i>	1	1	0	2
	<i>Tot. Colonna</i>	7	4	3	14
Cluster 4	<i>Nord</i>	19	8	9	36
	<i>Centro</i>	7	2	1	10
	<i>Sud</i>	12	4	0	16
	<i>Tot. Colonna</i>	38	14	10	62
Cluster 5	<i>Nord</i>	1	1	0	2
	<i>Centro</i>	1	4	0	5
	<i>Sud</i>	2	2	2	6
	<i>Tot. Colonna</i>	4	7	2	13
Cluster 6	<i>Nord</i>	0	2	0	2
	<i>Centro</i>	2	2	0	4
	<i>Sud</i>	0	0	0	0
	<i>Tot. Colonna</i>	2	4	0	6
Cluster 7	<i>Nord</i>	7	2	2	11
	<i>Centro</i>	2	0	2	4
	<i>Sud</i>	2	2	0	4
	<i>Tot. Colonna</i>	11	4	4	19
Cluster 8	<i>Nord</i>	5	0	0	5
	<i>Centro</i>	0	0	0	0
	<i>Sud</i>	0	0	0	0
	<i>Tot. Colonna</i>	5	0	0	5

Fonte: Elaborazioni proprie su fonte RFI

La più elevata quantità media di soluzioni si rintraccia nel **sesto cluster**, essa è anche ben distribuita durante tutte le fasce orarie. Il cluster è, inoltre, interessato dal numero più elevato di soluzioni dirette, di cui la metà sono Frecciarossa e Frecciargento. All'interno di questo raggruppamento sono comprese le 6 tratte che collegano Bologna, Roma, Firenze (Tab. 2). I collegamenti tra questi tre capoluoghi sono, dunque, tra i migliori offerti da Trenitalia, sia dal punto di vista quantitativo, sia da quello qualitativo (Tab. 1). Anche nel **Cluster 7** si rileva un

significativo numero di soluzioni dirette, tuttavia esse risultano molto eterogenee in quanto sono presenti anche numerose soluzioni con un cambio. Si registra, peraltro, nel medesimo gruppo una grande quantità di treni durante tutte le fasce orarie. La presenza di soluzioni con un cambio è caratterizzata da molti treni di cui almeno uno è una “Freccia” (Tab. 1). In questo cluster sono comprese 7/19 tratte, tutte riguardanti i collegamenti tra capoluoghi del Nord (Tab. 2). Nell’**ottavo cluster** sono presenti mediamente 36,8 soluzioni per tratta, esse risultano maggiormente diversificate e ben distribuite all’interno delle varie fasce orarie. La quantità totale di collegamenti è molto elevata (Tab. 1). Non ci sono soluzioni con più di un cambio e si rileva una quantità importante di frecce. In questo caso, come nel precedente, sono comprese tratte che collegano capoluoghi del Nord (Tab.2).

4. Collegamenti tra le macroaree italiane e soluzioni per l’utente

Le tre ripartizioni considerate (Nord/Sud/Centro) comprendono un numero diverso di regioni e dunque di capoluoghi, si è ritenuto opportuno valutare in che percentuale le tratte che collegano le macroaree fossero presenti nei vari clusters. E’ stato calcolato dunque il rapporto tra numero di tratte nel cluster e numero totale per ogni collegamento nelle e tra macroaree, il denominatore utilizzato è riportato in Tabella 3.

Tabella 3 – Numero di tratte totali per zona del capoluogo di partenza e d’arrivo (v.a.)

	Arrivi Nord	Arrivi Centro	Arrivi Sud	Tot. Riga
Partenze Nord	56	32	56	144
Partenze Centro	32	12	28	72
Partenze Sud	56	28	42	126
Tot. Col.	144	72	126	342

Fonte: Elaborazioni proprie su fonte RFI

Per ogni tipo di percorso si è rilevata (Tab. 4) la percentuale più rilevante e il cluster in cui la stessa si trova. Il 33,93% delle 56 tratte Nord-Nord è concentrato nel cluster 4, la maggior concentrazione delle tratte Nord-Centro si ha nello stesso cluster con il 25% delle 32 tratte considerate. Diversamente delle tratte Nord-Sud la maggiore concentrazione si rileva nel primo cluster, comprendendo quest’ultimo il 51,79% dei 56 percorsi. Per le partenze da un capoluogo centrale e gli arrivi in città settentrionali, si evince la maggiore concentrazione del 28,13 sia nel primo sia nel secondo cluster e, invece, per i collegamenti Centro-Centro, il 33% è nel quinto cluster. Ancora, i collegamenti Centro-Sud si trovano per il 46,4% nel secondo cluster. E’ evidente che i percorsi che hanno un capoluogo meridionale di partenza

o d'arrivo sono concentrati tra i primi due clusters, diversamente dalle restanti aree i cui collegamenti sono maggiormente concentrati nel quarto e nel quinto cluster.

Le tratte che collegano il Mezzogiorno alle tre macro-aree hanno invece per le destinazioni a Nord il 50% nel primo cluster, per quelle in Italia Centrale e Meridionale il 39,29% e il 59,52% nel secondo cluster.

Tabella 4 – Cluster con la maggiore concentrazione di ogni combinazione (%)

	Arrivi Nord	Arrivi Centro	Arrivi Sud
Partenze Nord	Cluster 4	Cluster 4	Cluster 1
	33,93%	25%	51,79%
Partenze Centro	Cluster 1 e 2	Cluster 5	Cluster 2
	28,13%	33%	46,43%
Partenze Sud	Cluster	Cluster 2	Cluster 2
	50%	39,29%	59,52%

Fonte: Elaborazioni proprie su fonte RFI

5. Conclusioni: l' "isola"/isolamento del Mezzogiorno.

L'analisi mette in luce l'elevata disomogeneità caratterizzante in termini quantitativi e qualitativi l'offerta a passeggeri italiani che viaggiano tra i capoluoghi di regione. Le soluzioni dirette che collegano i capoluoghi del Mezzogiorno ed il Mezzogiorno con altre aree italiane sono esigue nonché testimoni di una situazione preoccupante. Lo studio peraltro è esente dal considerare i collegamenti effettuati dalla compagnia Italo con cui si offre un maggior numero di soluzioni ad alta velocità al passeggero che si sposta da Salerno in su.

Come emerge dai più recenti rapporti di Banca d'Italia (2011-2012), le vocazioni del Sud sono fiaccate da una debolezza infrastrutturale aggravata dall'abbandono ferroviario. Per di più, in questi anni l'intervento pubblico nel settore infrastrutture e trasporti si è significativamente contratto. La quota degli investimenti al Sud sul totale nazionale è passata dal 37% del 1996 al 30% del 2012. A rendere il panorama più complesso interviene un altro fattore: un recente studio della Banca d'Italia, testimonia come nel settore delle infrastrutture l'intervento "aggiuntivo" a favore del Mezzogiorno ha spesso sostituito quello ordinario e «ad un aumento dei finanziamenti europei (compreso il cofinanziamento nazionale) ha corrisposto una diminuzione di circa il 20% delle altre fonti» (Cannari et al, 2009).

Concludendo, la scarsità di risorse è accompagnata da una serie di criticità che hanno caratterizzato la gestione degli investimenti nel Mezzogiorno del Paese e legati a tempi di realizzazione troppo lunghi e ad una scarsa visione sistemica

(ACAM, 2009). Gli investimenti in infrastrutture risultano volano di sviluppo e di crescita in quanto riducono i costi di trasporto delle imprese e ampliano i mercati di acquisizione dei fattori di sbocco delle produzioni, contribuendo ad aumentare l'accessibilità dei luoghi e dei mercati e la mobilità di cose e persone (Banister and Berechman, 2000). Tale volano è purtroppo ben lungi dal lambire l'“isola del Mezzogiorno”, il cui isolamento inizia poco più a sud di Napoli.

Riferimenti bibliografici

- ACAM 2009. Dario Gentile, *Un mezzogiorno da connettere*.
- BANISTER AND BERECHMAN 2000. *Transport Investment and Economic Development*, London, UK: UCL Press.
- CANNARI L., MAGNANI M., PELLEGRINI G. 2009. *Quali politiche per il Sud? Il ruolo delle politiche nazionali e regionali nell'ultimo decennio*, in “Questioni di Economia e Finanza” n. 50, luglio.
- RUSSO F. 2009. *Un Mediterraneo da integrare – L'Italia e i corridoi mediterranei del trasporto*
- VIESTI G. (2008), *Mezzogiorno a tradimento- il Nord, il Sud e la politica che non c'è*, Laterza.

SUMMARY

Train routes connecting the italian regional capitals: the case of Trenitalia.

The work ends with considerations about the itineraries in which passengers have more possibilities in terms of offer in the timeslots taken into account, in terms of time they reach the destination and comfort due to the greater or lesser number of transfer and the time that elapses between a train and the other where the journey is made up of several sections. The cluster analysis shows a lot of differences among the offer of South, Centre and North of Italy. Travels are very uncomfortable for people who want to move from South to Centre or North of Italy (going and return) and also between capital cities of southern region of Italy.

CHI SI STABILISCE IN ITALIA? TRANSIZIONE VERSO LA RESIDENZA DEGLI IMMIGRATI EXTRACOMUNITARI

Marco Fortini, Luca Mancini, Luigi Marcone,
Eleonora Mussino, Evelina Paluzzi

1. Introduzione

L'Italia è ormai diventato un paese d'immigrazione e questa tendenza sta sempre più influenzando sulle dimensioni e sul profilo demografico della popolazione. Negli ultimi decenni si è assistito a una fase di stabilizzazione della popolazione straniera la quale tende ad inserirsi in un progetto di vita sempre più condizionato dal contesto ospitante. L'aumento delle acquisizioni di cittadinanza, i ricongiungimenti familiari, l'incremento di minori e delle nascite sono solo alcuni degli indicatori che testimoniano questo processo di consolidamento e radicamento della presenza straniera in Italia. Al 1° gennaio 2012, con un incremento di circa 102 mila unità rispetto all'anno precedente, i cittadini stranieri non comunitari regolarmente presenti in Italia ammontano a 3.637.724 unità (Istat 2012), di cui il 21,2% ha ottenuto il permesso di soggiorno per motivi familiari. Dall'ultimo dato disponibile, al 9° ottobre 2011 risultano residenti in Italia 4.029.145 stranieri (comunitari e non) i quali rappresentano quasi il 7% della popolazione residente.. Nel 2011, inoltre, sono nati circa 79mila bambini stranieri che costituiscono il 14,5% del totale delle iscrizioni in anagrafe per nascita.

Il rilievo assunto negli ultimi decenni dall'immigrazione straniera in Italia ha richiamato l'attenzione delle istituzioni e della comunità scientifica, non soltanto relativamente alla dimensione del fenomeno ma anche alle sue caratteristiche. Di qui la necessità di disporre di dati sempre più aggiornati e affidabili sugli stranieri presenti nel paese. Il 15° Censimento della popolazione e delle abitazioni, attraverso innovazioni di processo, ha cercato di rispondere a questo fabbisogno informativo. Le novità introdotte, in particolare l'uso di registri amministrativi, hanno contribuito a migliorare la qualità e la copertura delle informazioni sulla popolazione straniera immigrata in Italia. Per la prima volta, infatti, il censimento si è avvalso di dati amministrativi: le liste anagrafiche comunali, come strumento principale per l'invio dei questionari ai rispondenti, e le Liste Integrative da Fonti Ausiliarie derivanti in parte dall'archivio dei permessi di soggiorno, come strumenti per il recupero mirato e sistematico della sottocopertura anagrafica. Ciò ha consentito di semplificare il complesso mosaico che compone il fenomeno

migratorio in due componenti: in primis, una componente stabile, costituita dai cittadini stranieri iscritti in anagrafe; ma ancor più importante ai fini dell'esaustività censuaria, una componente presumibilmente non stabile, costituita dai cittadini stranieri non iscritti in anagrafe ma in possesso di un titolo a soggiornare in Italia.

Obiettivo del presente lavoro è di partire dai risultati delle operazioni effettuate dalla Direzione dei censimenti dell'Istat (Fortini et al, 2011), finalizzate alla rilevazione degli stranieri non comunitari non iscritti in anagrafe (sottocopertura anagrafica), per cogliere eventuali legami tra caratteristiche socio-demografiche e migratorie del suddetto segmento di popolazione straniera e l'acquisizione di residenza. Da queste analisi ci aspettiamo di apprendere qualcosa in più sul processo di stabilizzazione degli stranieri arrivati in Italia con un regolare permesso di soggiorno. Precedenti studi sull'integrazione degli stranieri hanno già mostrato come l'acquisizione di residenza sia strettamente collegata con l'integrazione strutturale e politica degli immigrati (Blangiardo et al. 2006) e che la mancata residenza sia associata a un ritardo nell'accesso al sistema sanitario nazionale (Lariccia et al. 2013a), a esiti negativi al parto (Lariccia et al. 2013b), a un maggiore sfruttamento lavorativo (Ferrara et al. 2012) e a una minore integrazione sociale ed economica nel contesto di arrivo (Strozza and Mussino 2012, Mussino et al 2013). Riteniamo, pertanto, opportuno studiare le caratteristiche individuali associate all'iscrizione anagrafica la quale non soltanto porta vantaggi agli iscritti (iscrizione al sistema sanitario nazionale, carta d'identità, diritti politici e così via), ma rappresenta un valore aggiunto per l'intero sistema statistico nazionale. Come ha mostrato il censimento e visto il sempre maggiore utilizzo delle fonti amministrative a fini censuari, i registri comunali forniscono una fonte statistica fondamentale. L'accuratezza di tali informazioni è un bene pubblico, che informa gran parte della azione di governo sia a livello nazionale che locale.

Questo lavoro parte quindi dagli sforzi fatti nelle fasi di programmazione e attuazione del censimento che, attraverso tecniche di *record linkage*, ha abbinato i singoli individui presenti nell'archivio dei permessi di soggiorno al 1° gennaio 2011 con la popolazione iscritta al registro comunale alla stessa data. Nonostante l'archivio dei permessi di soggiorno dal 2007 sia riferito quasi esclusivamente ai cittadini non comunitari, rimane tuttora la fonte principale per lo studio socio-demografico della presenza straniera in Italia. L'approccio multi-fonte e l'impiego di diversi archivi amministrativi consentono di disporre di informazioni aggiuntive attraverso le quali è possibile studiare non soltanto la struttura del fenomeno migratorio e le sue connotazioni territoriali ma anche il processo di stabilizzazione della componente straniera della popolazione che risiede in Italia. L'obiettivo è quindi di evidenziarne i legami con alcune caratteristiche e comportamenti socio-demografici attraverso l'uso di modelli logistici e tecniche di *event history*.

2. Dati e Metodi

2.1. Fonti e RL

Il dataset utilizzato per questo lavoro è dato dall'integrazione di fonti amministrative e di dati censuari. I dati amministrativi utilizzati sono a loro volta il risultato di operazioni di integrazione (*record linkage*) tra una serie di archivi che hanno portato alla determinazione di due liste di individui: le liste anagrafiche comunali delle famiglie e delle convivenze (LAC) e le liste integrative da fonti ausiliarie (LIFA). Le LAC rappresentano il fulcro del nuovo censimento assistito da lista, consentendo l'invio di questionari a mezzo posta, mediante l'impiego degli indirizzi anagrafici, oltre 20 milioni di famiglie residenti. In assenza di un'anagrafe nazionale, l'utilizzo delle LAC comunali ha comportato un ingente lavoro di controllo e normalizzazione da parte dell'ISTAT su oltre 61 milioni di record trasmessi elettronicamente dai singoli comuni.

Sebbene le LAC dovrebbero fornire la „fotografia“ costantemente aggiornata della popolazione effettivamente residente in un certo territorio, errori di copertura, dovuti a mancate cancellazioni di individui iscritti in anagrafe e non più residenti e a mancate iscrizioni di individui abitualmente dimoranti, sono piuttosto comuni. Questi disallineamenti tra popolazione ufficialmente ed effettivamente residente sono più marcati per la popolazione straniera abitualmente dimorante in Italia. Per questo motivo, in occasione del censimento sono state utilizzate le liste ausiliarie di individui costruite a partire da archivi amministrativi alternativi alle LAC, contenenti segnali di presenza complementari alle liste anagrafiche.

Le LIFA nascono dalla sintesi di due archivi principali: l'Anagrafe Tributaria (AT) per i cittadini italiani e comunitari e i Permessi di Soggiorno (PS) per i cittadini stranieri non comunitari (Fortini et al., 2011, Casale et al., 2012). L'archivio dei PS conteneva originariamente circa 4 milioni di record provenienti da 11 registri dedicati a diverse tipologie di permesso – lavoro, ricongiungimento familiare, rinnovi, richieste sotto esame, ecc. – tenuti dai Dipartimenti di Pubblica Sicurezza e delle Libertà Civili del Ministero dell'Interno. La presenza di sovrapposizioni tra gli archivi ha reso necessario un passaggio preliminare di de-duplicazione dei record individuali che ha permesso di individuare ed eliminare circa 700 mila duplicati.

Nel record linkage tra PS de-duplicato (PS+) e LAC, così come nella fase precedente di ricerca delle duplicazioni, sono state impiegate strategie probabilistiche per sopperire all'assenza di chiavi primarie di aggancio (quali, ad esempio, il codice fiscale) con standard adeguati di qualità e completezza. Dei 3,3 milioni di record presenti in PS+ circa 1,86 milioni sono individui con regolare permesso di soggiorno iscritti in anagrafe allo stesso indirizzo specificato nella

domanda di permesso, 660,000 sono iscritti in anagrafe ma con indirizzi LAC e PS non coincidenti (stesso comune o altro comune), mentre i restanti 780,000 record si riferiscono a cittadini stranieri abitualmente dimoranti con regolare permesso a soggiornare in Italia ma non iscritti in alcuna anagrafe. L'integrazione con i dati censuari ha consentito, in una seconda fase, di rimuovere dal dataset gli individui non censiti poiché essi hanno presumibilmente lasciato il paese. Per i censiti, oltre alle informazioni anagrafiche, ivi inclusa la data di iscrizione in anagrafe, il dataset contiene anche le principali caratteristiche del permesso di soggiorno, inclusa la data ufficiale di ingresso in Italia, necessaria per costruire una variabile di durata tra ingresso nel paese ed iscrizione in anagrafe.

Tabella 1 – Distribuzione dei censiti e degli iscritti in anagrafe

Primo permesso		Censiti		Totale
	No	Si		
Iscritti	No	31.2	8.9	
	Si	11.0	48.9	
Totale				887,297

Per problemi di qualità del dato e di disponibilità delle informazioni negli archivi, per queste elaborazioni ci siamo concentrati soltanto sull'archivio principale dei permessi in corso di validità, l'archivio delle nuove domande di permesso in fase di pre-processamento e l'archivio delle domande in fase di scrutinio/istruttoria.¹ Inoltre, poiché spesso quando si rinnova il permesso le informazioni sono sovrascritte, per non incorrere nel rischio di *anticipatory analysis*, abbiamo lavorato soltanto sugli individui che alla data del censimento fossero in possesso del loro primo permesso. L'ultima selezione effettuata per una conformità del dato è stato quello di concentrare la nostra attenzione solo sui cittadini non comunitari (tabella 1).

2.2. Metodi

Per includere nell'analisi le determinanti della iscrizione anagrafica si è deciso di fare ricorso in un primo momento alla regressione logistica multivariata per stimare gli *odds-ratios* (OR). Ipotizzando che la propensione ad acquisire la cittadinanza sia legata al tempo trascorso in Italia, in un secondo momento si è provveduto ad analizzare i *rischi relativi* (RR) attraverso una regressione esponenziale costante a tratti.

¹ Restano esclusi, tra gli altri, le domande respinte ed i permessi scaduti.

3. Risultati

Le analisi proposte in questo articolo riguardano il sottoinsieme di nuovi immigrati regolarmente presenti in Italia al 1° gennaio 2010 (ovvero con il primo permesso di soggiorno valido alla data di riferimento del censimento). Considerando i diversi modelli migratori delle comunità presenti in Italia, ipotizziamo che ci sia forte eterogeneità di comportamento tra le diverse cittadinanze e che queste differenze persistono quando controlliamo per le caratteristiche socio-migratorie degli individui. Per questo motivo abbiamo proceduto con un approccio *stepwise* (tabella 1).

Tabella 2 – *Odd ratio e relative risk ratio di iscriversi in anagrafe per la variabile cittadinanza*

	Regressione Logistica						Regressione esponenziale			
	Modello 0		Modello 1		Modello 2		Modello 3		Modello 4	
	OR	sign	OR	sign	OR	sign	RR	sign	RR	sign
Albania	1.67	***	1.67	***	1.82	***	1.45	***	1.30	***
Ucraina	0.80	***	0.76	***	1.16	***	1.08	***	0.98	*
Moldova	0.92	***	0.81	***	1.36	***	1.11	***	0.98	**
Cina	0.38	***	0.36	***	0.37	***	0.55	***	0.65	***
Filippine	0.82	***	0.80	***	0.98		0.94	***	0.85	***
India	0.46	***	0.50	***	0.78	***	0.91	***	1.02	*
Egitto	0.45	***	0.46	***	0.59	***	0.73	***	0.89	***
Marocco	1		1		1		1		1	
Tunisia	0.77	***	0.83	***	0.88	***	0.91	***	1.02	*
Perù	1.08	***	1.05	**	1.45	***	1.20	***	1.03	**
Paesi a sviluppo avanzato	0.08	***	0.07	***	0.08	***	0.16	***	0.71	***
Altri Paesi a forte pressione migratoria	0.68	***	0.68	***	0.98	*	0.98	**	1.00	

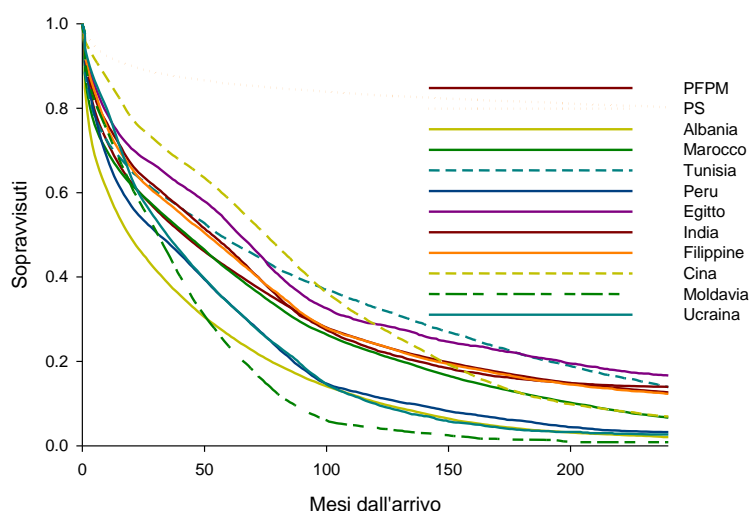
Modello 0: Cittadinanza; Modello 1: Modello 0 + Sesso, Stato civile, ripartizione geografica; Modello 2: Modello 1 + età all'arrivo, motivo del permesso, tipologia di archivio, coorte di arrivo; Modello 3: Modello 2 + durata ; Modello 4 solo censiti

I risultati nella tabella 1 mostrano una certa variabilità tra le diverse cittadinanze relativamente alla propensione ad iscriversi in anagrafe. Nel Modello nullo (modello 0), dove controlliamo solo per la variabile cittadinanza, rispetto ai marocchini soltanto gli albanesi sembravano avere degli OR significativamente più alti. Una volta controllato per le variabili demografiche (Modello 1), non si rilevano significative variazioni. Le cose cambiano quando controlliamo per le caratteristiche migratorie, come il motivo e l'anno di ingresso (Modello 2). Infatti, si evidenzia come ci sia un modello dei cittadini provenienti dall'est Europa nella scelta di acquisire la residenza e si rafforzi la propensione dei peruviani.

Nei modelli di *Event History* (3 e 4) le differenze, anche se persistono, sono meno forti. Effettivamente andando a vedere la rappresentazione grafica delle

curve di sopravvivenza (figura 1) non è tanto la densità quanto la cadenza che distingue i comportamenti per cittadinanza.

Figura 1 - Sopravvissuti per cittadinanza



Un ulteriore sforzo è stato fatto prendendo in considerazione solo gli immigrati censiti. L'ipotesi infatti è che i modelli precedenti sovrastimino i non iscritti, visto che gli immigrati possono aver lasciato il paese durante il periodo di osservazione (modello 4). D'altro canto, è anche vero che esiste un effetto selezione sulla variabile indipendente dal momento che, se un individuo è in LAC, è sensibilmente "più a rischio" di essere censito essendo il censimento italiano della popolazione 2011 assistito dai registri anagrafici. Le differenze tra il rischio relativo di iscrizione sono tipicamente più contenute rispetto al modello 3. Come prevedibile il motivo della presenza segna una eterogeneità nella propensione ad iscriversi (tabella 3): se gli stranieri arrivati in Italia per motivi di famiglia hanno una propensione maggiore ad iscriversi, quelli arrivati per motivi di studio o di salute difficilmente tendono a iscriversi. Tra gli altri risultati non mostrati in tabella per ragioni di spazio (Modello 3 e 4) ma degni di nota² emerge che gli stranieri dimoranti nelle regioni del Nord Est hanno una maggiore propensione all'iscrizione rispetto alle altre ripartizioni geografiche italiane, così come le donne e gli

² I coefficienti commentati ma omessi dalle tabelle così come tutte le stime non commentate sono disponibili su richiesta.

individui coniugati. Infine gli individui di età inferiore a 17 anni hanno un rischio relativo maggiore rispetto alle altre fasce di età di iscriversi in anagrafe. Nel modello 4 questo comportamento si estende anche agli over-65 i quali sono circa il 25% più „a rischio“ di iscrizione rispetto alle fasce di età intermedie.

Tabella 3 – *Odd ratio e relative risk ratio di iscriversi in anagrafe per la variabile Motivo del permesso*

Motivo del permesso	Regressione esponenziale			
	Modello 3		Modello 4	
	RR	sign	RR	sign
Adozione e minori	1.726	***	1.935	***
Rifugiati	0.483	***	0.861	***
Famiglia	2.006	***	2.189	***
Religiosi	0.053	***	0.304	***
Lavoro	<i>1</i>		<i>1</i>	
Salute	0.491	***	0.761	***
Studio e Ricerca	0.149	***	0.598	***
Lungo periodo (ce)	2.410	***	2.373	***
Altro	2.080	***	1.879	***

4. Conclusioni

La disponibilità di dati provenienti dagli sforzi fatti per il censimento ha permesso di eseguire analisi innovative che hanno mostrato come non si possa prescindere dall'effettuare considerazioni differenziate per cittadinanza e per storia migratoria nello studio dell'acquisizione di residenza. Nasce quindi l'esigenza di adeguare il vigente sistema di raccolta dei dati anche ai fini del censimento permanente.

RIFERIMENTI BIBLIOGRAFICI

- BLANGIARDO M., STROZZA S., TERZERA L. 2006. Indicatori di integrazione degli immigrati in Italia. In BLANGIARDO G.C., FARINA P. (Eds), *Il Mezzogiorno dopo la grande regolarizzazione. Immagini e problematiche dell'immigrazione*, Volume terzo, Franco Angeli, Milano: 153-189
- CASALE D., MANCINI L., MARCONE L., PELLIZZARO F. e PEZONE A. 2012. *The Use of Population Registers in the 15th Italian Census: Challenges and Preliminary Evidence*. UNECE Working paper n.9, UNECE-Eurostat Expert Group Meeting on Censuses Using Registers, Conference of European Statisticians, Geneva, 22-23 May 2012.
- FERRARA R., MUSSINO E., STROZZA S. 2012. Lo sfruttamento lavorativo: dimensione del fenomeno e caratteristiche degli immigrati coinvolti. In: *Vivere da*

- immigrati nel casertano. Profili variabili, condizioni difficili e relazioni in divenire*. Collana ISMU Iniziative e Studi sulla Multietnicità, Franco Angeli.
- FORTINI M., GALLO G., BENASSI F., MANCINI L., MARCONE L. 2011. *Counting in non-Italian residents: the use of the "Permit to Stay" archive in the next Population Census*. Proceedings of the Second ITACOSM Conference, Edizioni Plus – Pisa University Press.
- ISTAT 2012. Cittadini non comunitari regolarmente soggiornanti.
- LARICCIA F., MUSSINO E., PINNELLI A., PRATI S. 2013a. *Antenatal care in Italy: Differences between Italian and Foreign Women*. GENUS N. 2, 2013 (Futhercoming).
- LARICCIA F., MUSSINO E., PINNELLI A., PRATI S. 2013b. *Differences in perinatal outcomes: the role of citizenship*. MULINO (Futhercoming).
- MUSSINO E., STROZZA S., TERZERA L. 2013. Dimensioni e determinanti dell'integrazione degli immigrati. Il ruolo delle provenienze e delle realtà d'insediamento, Giornate di Studio sulla popolazione Bressanone 6-8 February.
- STROZZA S., MUSSINO E. 2012. Le condizioni di integrazione (Integration conditions), In: *Vivere da immigrati nel casertano*. Profili variabili, condizioni difficili e relazioni in divenire. Collana ISMU Iniziative e Studi sulla Multietnicità. Franco Angeli.

SUMMARY

Who settles down in Italy? Transition to residency of non-EU migrants

In the last decades Italy has become an immigration country. The importance of studying this phenomenon and the characteristics of the population involved are pressuring the official statistics on the need of reliable data. In 2011 thanks to the innovation of the 15th population and housing census and the use of the auxiliary list (LIFA) we were also able to exploit and combine this information with additional sources to provide a better overview of the demographic and behavioral characteristics of the foreign immigrant population in Italy. The objective of this work is to present the results of the record linkage procedure between different sources and to explore the potential of this information for research. Using event history techniques we will highlight the links between socio-demographic characteristics and the mobility and the stabilization of migrants.

Marco FORTINI, Istituto Nazionale di Statistica, fortini@istat.it.

Luigi MARCONE, Istituto Nazionale di Statistica, lmarcone@istat.it.

Luca MANCINI, Istituto Nazionale di Statistica, lmancini@istat.it.

Eleonora MUSSINO, Istituto Nazionale di Statistica, mussino@istat.it.

Evelina PALUZZI, Istituto Nazionale di Statistica, paluzzi@istat.it.

LA VULNERABILITÀ POST-DIVORZIO: SPAGNA E ITALIA A CONFRONTO¹

Maria Carella, Thaís García Pereiro,

1. Introduzione

In letteratura, da circa un decennio, il fenomeno della rottura coniugale viene inteso come un processo che genera situazioni di vulnerabilità socio-economica per i soggetti coinvolti. Diversi studi sostengono che l'epoca successiva alla separazione identifica un periodo contraddistinto dal rischio e dall'incertezza in cui gli individui sono costretti a riconsiderare i loro "progetti di vita" tenendo conto delle nuove e, spesso ridotte, risorse disponibili (Vignoli e Ferro, 2009; Zanatta, 2008).

Nel presente articolo si propone lo sviluppo empirico di un approccio teorico sul post-divorzio (García Pereiro e Solsona, 2011) che, trovando il suo fondamento concettuale nel contesto della vulnerabilità, esamina gli effetti prodotti dal scioglimento del matrimonio sugli ambiti di vita che riguardano gli aspetti economici, la salute e lo stato di benessere generale degli ex-coniugi.

La vulnerabilità post-divorzio segue un modello che replica le relazioni di genere costruite all'interno del matrimonio. Tali relazioni possono assumere forme diverse in base ad alcune caratteristiche socio-demografiche come l'anno di formazione dell'unione, l'età dei consorti e la loro appartenenza ad un gruppo sociale (Solsona, 2011).

In questo lavoro, facendo ricorso ai dati provenienti dall'Indagine sulle Condizioni di Vita (EU-SILC) dell'anno 2009, ci si propone di analizzare gli stili di vita delle categorie divorziati/separati, suddivisi per genere, in Italia e Spagna. Tale analisi sarà condotta in relazione a gruppi di indicatori che afferiscono a tre sfere: economica, della salute e sociale.

In questo senso, si cercherà di perseguire tre obiettivi: 1) comprendere come e in che misura lo stato di vulnerabilità, solitamente riscontrabile nel processo post-divorzio, trasforma ed incide sulle dimensioni di vita oggetto di studio; 2)

¹ Lavoro svolto nell'ambito del progetto di ricerca "Post-divorcio y vulnerabilidad social en España. Experiencias de hombres y mujeres en al ámbito económico y de la salud" (Ref. CSO2009-09891) diretto da Montserrat Solsona e finanziato dal Plan Nacional de I+D+I del Ministerio de Economía e Competitividad. L'esecuzione complessiva del lavoro va intesa svolta dagli autori (citati in ordine alfabetico) in stretta collaborazione: tuttavia, si attribuisce il par.1 a M. Carella, il par.2 a T. G. Pereiro, il par.3 e 4 vanno attribuiti ai due autori.

osservare le differenze di genere che riguardano separati/divorziati; 3) valutare similarità e divergenze nelle conseguenze prodotte dal divorzio in Italia e in Spagna.

2. Dati e metodologia

Numerosi studi hanno dimostrato che lo scioglimento del vincolo matrimoniale modifica radicalmente situazioni e status di vita degli ex-coniugi. Nella sfera economica, la vulnerabilità generata dal divorzio comporta un'immediata riduzione delle risorse monetarie di entrambi gli individui coinvolti (Andreb e Bröckel, 2007; Aassve et al., 2007), mentre in riferimento allo stato di salute, tale vulnerabilità si manifesta nel peggioramento (oggettivo e percepito) della condizione sia fisica che mentale o emozionale (Andreb e Bröckel, 2007; Colman et al., 2010). Alcune ricerche hanno riscontrato, inoltre, che lo stato di fragilità successivo alla rottura coniugale incide anche sulla vita sociale dei separati. Dopo il divorzio si trasformano le reti interpersonali, cambiano tipologie e modalità di frequentazioni e variano le attività dedicate al tempo libero (Kalmijn e Uunk, 2007; Symoens e Bracke, 2010; Solsona, 2011).

Per mancanza di dati di tipo longitudinale che consentirebbero il monitoraggio delle biografie pre e post-divorzio, in questa sede si è scelto di confrontare le popolazioni di separati e divorziati in entrambi i Paesi, considerando tale raffronto come una proxy adatta alla lettura dei segni della vulnerabilità (Simò et al., 2013). A questo proposito, per neutralizzare gli effetti *protettivi* del matrimonio e *selettivi* del divorzio, le analisi sono state condotte osservando solo la popolazione che mantiene nel tempo invariato lo status di separato/divorziato. In altri termini, sono stati esclusi i divorziati e separati, uomini e donne, che al momento dell'Indagine, pur dichiarando tale stato civile, avevano formato un'altra unione (conviventi). Pertanto, il campione oggetto di studio è composto da 2.456 individui separati/divorziati di cui 1.061 di nazionalità spagnola e 1.395 italiani.

In questo lavoro sono stati utilizzati i micro-dati trasversali dell'Indagine sulle Condizioni di Vita (EU-SILC) dell'anno 2009 che contiene numerose informazioni sulle tre dimensioni della vulnerabilità post-divorzio esaminate in questo studio. Questa fonte ci ha consentito di osservare contemporaneamente undici indicatori: quattro afferenti lo stato di salute degli individui, tre riguardanti la vita sociale ed i restanti quattro relativi alla sfera economica (Tabella 1). L'EUSILC include anche diverse variabili socio-demografiche di particolare interesse negli studi sul post-divorzio che possono attenuare o rafforzare le situazioni di vulnerabilità. A titolo esemplificativo, livello d'istruzione, età degli ex-coniugi, presenza di figli a carico sono senza dubbio caratteristiche importanti che modificano ulteriormente le condizioni di vita della coppia spezzata. Per queste ragioni tali caratteristiche sono state inserite come variabili di controllo nelle analisi di seguito condotte.

Tabella 1- Indicatori EU-SILC 2009 utilizzati per misurare la vulnerabilità post-divorzio in Spagna e in Italia.

SFERA	VARIABILE	DESCRIZIONE	CATEGORIE
Salute (4)	Salute (soggettiva)	Come va in generale la sua salute?	Bene Nè bene nè male Male
	Limitazioni nelle attività abituali	A causa di problemi di salute, Lei ha delle limitazioni che durano da almeno 6 mesi nelle attività abitualmente svolte?	Sì No
	Visite mediche	Numero di visite mediche negli ultimi 12 mesi	Meno di 5 Più di 5
	Malattia cronica	Lei è affetto da malattie croniche o problemi di salute di lunga durata? (da almeno 6 mesi)	Sì No
Sociale (3)	Tempo ad amici e parenti	Lei dedica regolarmente tempo ad amici e parenti?	Sì No
	Si dedica al tempo libero	Si dedica al tempo libero? Svolge qualunque attività per ricrearsi, divertirsi, ecc.	Sì No
	Spende soldi in consumi personali	Lei spende soldi per i suoi consumi personali relativi al tempo libero?	Sì No
Economica (4)	In grado di far fronte a spese impreviste	Sarebbe in grado di far fronte a spese impreviste di un ammontare approssimativo di 800 euro con risorse proprie?	Sì No
	Reddito annuo	Reddito annuo medio dichiarato	Media (euro)
	Indicatore di povertà	Costruito da EU-SILC (definita in funzione della linea di povertà relativa, 60% del valore mediano della distribuzione del reddito a livello nazionale)	Non a rischio A rischio
	Ritorno nucleo di origine	Proxy costruita in relazione alla posizione della madre o del padre nel nucleo familiare	Sì No

3. Vulnerabilità post-divorzio in Spagna e in Italia: misure di genere

Nella Figura 1 vengono riportati gli indicatori relativi alla vulnerabilità post-divorzio nei contesti descritti (salute, sociale ed economico). La distribuzione percentuale di ciascuna delle variabili è stata calcolata in funzione di tutte le possibili risposte consentite all'intervistato. Tuttavia si è deciso di mostrare graficamente solo le frequenze percentuali delle risposte che esprimono chiaramente la vulnerabilità per ciascun ambito osservato, al fine di perseguire un duplice obiettivo: 1) rendere immediato il confronto tra Spagna e Italia per rilevare e misurare le diversità tra i due paesi; 2) cogliere le differenze di genere riscontrabili tra le due popolazioni di divorziati/separati.

Partendo dalla sfera che interessa la salute degli indagati, si osservano differenze di rilievo nelle risposte che mostrano un aggravamento oggettivo dello stato degli individui (presenza di una malattia cronica e frequenza delle visite mediche). In Spagna la condizione fisica degli individui divorziati e separati risulta decisamente peggiore rispetto a quella dichiarata in Italia: uomini e donne di nazionalità spagnola colpiti da malattia cronica supererebbero rispettivamente di

circa 9 e 14 punti percentuali uomini e donne italiane con stesse problematiche. In Spagna appaiono ancora più interessanti le differenze di genere: lo stato di salute delle donne sembrerebbe più precario rispetto a quello degli uomini. Di fatto, la percentuale di donne separate/divorziate che dichiarano di essere affette da malattia cronica e di frequentare il medico più di 5 volte all'anno (39,7% e 36,2%) è in entrambi i casi superiore a quella degli uomini (34,8%; 26,4%).

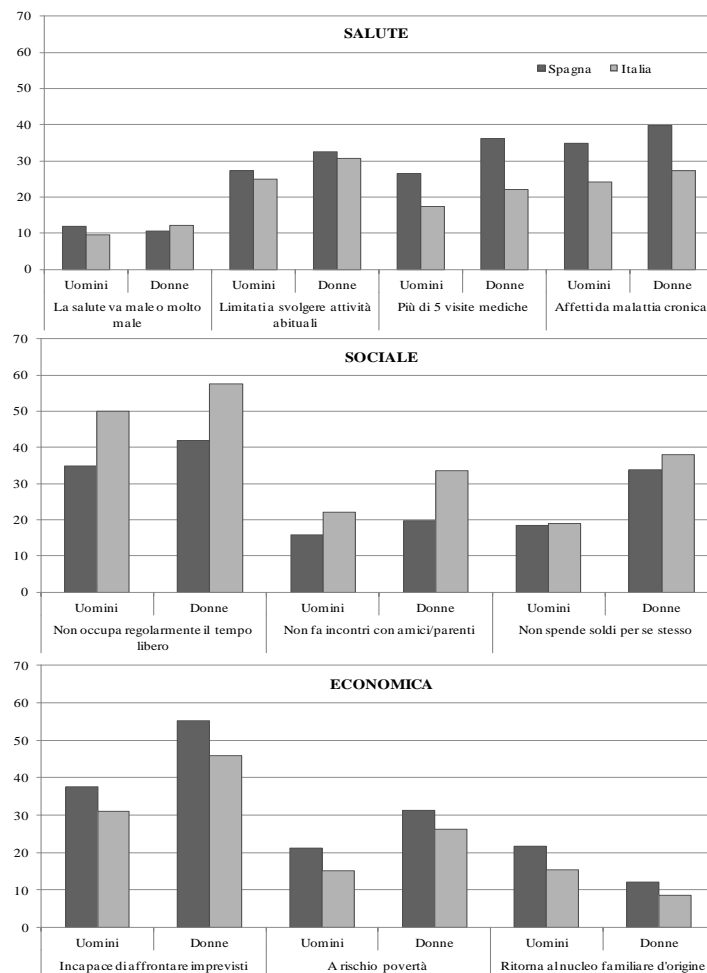
Per quanto riguarda la vulnerabilità in ambito sociale, i comportamenti differenziali tra separati/divorziati si rinvengono osservando modalità e tempi utilizzati per il tempo libero. Il 57,6% delle donne separate/divorziate in Italia dichiarano di non impiegare il proprio tempo in attività ricreative (contro il 41,9% delle spagnole) e il 33,6% non frequenta amici e parenti (contro il 19,7% in Spagna): in generale con la rottura del matrimonio si riscontra una maggiore tendenza alla socializzazione da parte dei divorziati di nazionalità spagnola e una certa ritrosia ad intessere nuove amicizie o a sviluppare rapporti già esistenti soprattutto da parte delle donne italiane. Nel contesto sociale le differenze tra i due sessi sono ancor più evidenti osservando il caso italiano e riguardano non solo lo spazio temporale dedicato ai rapporti inter-personali, ma anche il denaro utilizzato per il proprio benessere fisico. Invero, la percentuale di donne separate che dichiarano di non spendere soldi per attività finalizzate alla loro persona (38%) è molto più elevata di quella degli uomini (19%).

L'analisi delle variabili in ambito economico mostra uno stato di vulnerabilità post-divorzio sostanzialmente omogeneo nei due paesi: sia per gli italiani che per gli spagnoli divorziati si riducono le disponibilità monetarie e peggiorano le condizioni economiche. Tuttavia gli individui divorziati/separati in Spagna sembrerebbero più penalizzati: la misura media del loro reddito appare inferiore a quella degli italiani, sebbene in Italia si riscontrino significative differenze fra il reddito post-divorzio femminile e quello maschile (circa 12.000 euro). A riguardo in entrambi i paesi si osserva un forte squilibrio tra i sessi: le donne dichiarano una maggiore difficoltà a soddisfare tutti i bisogni e a sostenere spese impreviste dopo il divorzio e sembrerebbero essere esposte ad un maggior rischio di povertà. Si osserva comunque una maggiore propensione da parte degli uomini divorziati a ricongiungersi con il nucleo familiare d'origine: la percentuale di uomini che dopo la rottura coniugale ritorna a vivere con genitori/parenti si aggira intorno al 21,8% in Spagna e al 15,5% in Italia.

Per procedere ad uno studio più critico del fenomeno, le variabili sono state testate in un contesto multivariato al fine di individuare i soggetti e gli ambiti maggiormente colpiti dalla vulnerabilità post-divorzio.

A tal proposito si è fatto ricorso alla regressione logistica binaria mediante la quale sono stati stimati due modelli distinti per ciascun paese.

Figura 1- Distribuzione percentuale degli indicatori della vulnerabilità post-divorzio (salute, sociale ed economica). Spagna ed Italia 2009.



Fonte: Elaborazione propria. EU-SILC 2009.

Il primo è stato costruito con le variabili esplicative che si riferiscono alle tre dimensioni già descritte, il secondo è stato stimato introducendo alcune variabili che definiscono il profilo socio-demografico dei soggetti indagati. In entrambi i modelli, il sesso (componente femminile del campione) è funzione delle variabili indipendenti di cui sono riportati nella Tabella 2 gli odds ratio per ciascuna sfera della vulnerabilità post-divorzio.

Tabella 2 - Risultati del modello di regressione relativo alle differenze di genere negli ambiti della vulnerabilità post-divorzio. Spagna e Italia, 2009.

		Spagna		Italia		
		Mod. 1	Mod. 2	Mod. 1	Mod. 2	
Salute	(Bene)					
	Salute (soggettiva)	Nè bene nè male	0,77	0,92	1,17	1,29
		Male	0,43 **	0,57 **	0,99	1,28
	Limitazioni nelle attività abituali	Si	1,01	1,03	0,93	1,06
		(No)				
	Visite mediche	(Meno di 5)				
	Più di 5	1,71 **	1,74 **	1,17	1,23	
Malattia cronica/problemi lunga durata	Si	1,09	1,23	0,96	0,96	
	(No)					
Sociale	Dedica tempo per condividere con amici e parenti	(Si)				
		No	0,82	0,83	1,35 *	1,50 **
	Si dedica al tempo libero (attività)	No				
		(Si)	0,85	0,92	0,75 *	0,73 *
Spende soldi in consumi personali	(Si)					
	No	2,16 **	1,97 **	2,13 **	1,81 **	
Vive coi genitori	Si	0,51 **	0,51 **	0,55 **	0,63 **	
	(No)					
Economica	In grado di far fronte a spese impreviste	(Si)				
		No	1,75 **	1,64 **	1,21	1,15 *
	Reddito annuo	Log	1,06	0,84	0,89	0,68 **
Indicatore di povertà	(Non a rischio)					
	A rischio		1,51 *	1,13 *	1,25	1,1 *
Ha più di 40 anni	Si		1,21		0,96	
	(No)					
Titolo di studio	(Universitario)				0,68 *	
	Non universitario		0,91			
Controlli	Condizione nella professione	(Impiegato/a)				
		Non impiegato/a		0,67 *		-
Condizione abitativa	Non proprietario			0,69 *	0,86	
	(Proprietario)					
Presenza di figli a carico	(Senza f. dipendenti)					
	Figli dipendenti		5,32 **		7,68 **	
		N	1.061	1.061	1.395	1.395
		Log-likelihood (-2)	1.255	1.139	1.678	1.486
		Pseudo R2	11	24	10	26

*p<=0,05 **p<=0,01

Categorie di riferimento in parentesi

Fonte: Elaborazione propria. EU-SILC 2009

Due evidenze appaiono interessanti e sembrerebbero contraddistinguere i due paesi: in Spagna la probabilità che le donne divorziate/separate si rechino dal medico più di cinque volte all'anno è 1,7 volte superiore rispetto a quella degli uomini. Tuttavia la probabilità che le stesse dichiarino uno stato di salute precario (la salute va male) è inferiore a quella degli uomini. Detto altrimenti, con il divorzio la condizione fisica della componente femminile spagnola sembrerebbe

peggiore oggettivamente, mentre la percezione di malessere è più avvertita dal genere maschile.

In Italia, invece, è interessante notare come i modelli confortino le analisi descrittive che interessano la dimensione sociale. Rispetto agli uomini le donne manifestano non solo una minor attitudine a frequentare regolarmente familiari e amici e a svolgere attività di svago, ma altresì si mostrano meno propense ad impiegare denaro in attività finalizzate alla propria persona (quest'ultimo aspetto interessa anche le donne spagnole). Per quanto concerne gli aspetti economici tutte le stime sembrano confermare l'assunto che dopo il divorzio le donne appaiono i soggetti più svantaggiati in entrambi i Paesi. Dai modelli si desume che la rottura del matrimonio genera una forte riduzione delle disponibilità monetarie femminili: difatti, la probabilità che una donna dichiari di non poter affrontare spese impreviste con risorse proprie è più alta 1,15 volte in Italia e 1,64 volte in Spagna rispetto a quella degli uomini. La variabile che stima il rischio di povertà rafforza quanto detto: per la popolazione femminile i valori sono sempre più elevati di quelli riguardanti la categoria maschile. Contrariamente, la probabilità di ritorno al nucleo familiare d'origine dopo il divorzio è maggiore per gli uomini: si può supporre pertanto che questa propensione tipicamente maschile a ricongiungersi con i propri genitori derivi anch'essa da un disagio economico il quale costringe ad ammortizzare le spese utilizzando tutti gli espedienti possibili.

4. Considerazioni conclusive

Da tempo gli studiosi sono concordi nel ritenere che la rottura coniugale rappresenta un evento critico che influisce su tutti gli aspetti di vita degli ex-coniugi e che dopo il divorzio le donne siano più esposte al rischio di impoverimento economico e declassamento sociale rispetto agli uomini (Zanatta, 2008). Le analisi condotte relativamente ai separati/divorziati in Italia e Spagna avvalorano le evidenze comprovate dalla letteratura, confermando la dicotomia esistente nell'Europa Occidentale tra nord e sud in tema di protezione sociale a sostegno dei soggetti coinvolti dal scioglimento di un nucleo familiare. La mancanza di misure di welfare adeguate nei due paesi esaminati trasformano le famiglie d'origine degli ex-coniugi in ammortizzatori sociali, al punto che i genitori italiani e spagnoli consentono ai figli separati, più che altrove in Europa, la possibilità di ritrovare alloggio e conforto nella loro abitazione d'origine.

Riferimenti bibliografici

ANDREB H-J., BROCKEL, M. 2007. "Income and life satisfaction after marital disruption in Germany". *Journal of Marriage and Family*, Vol.69 N.2, 500-512.

- AASSVE A., BETTI G. et al. 2007. *Marital disruption and economic well-being: a comparative analysis*. Paper presentato nel Fifth Meeting of the European Network for the Sociological and Demographic Study of Divorce, Londres:UK.
- COLMAN E., SYMOENS S., BRACKE P. 2010. *Differences in health care use among the divorced and the married: are they need-based?*. Paper presentato nell'Eighth Meeting of the European Network for the Sociological and Demographic Study of Divorce, Valencia: España.
- GARCÍA PEREIRO T., SOLSONA M. 2011. El divorcio como nudo biográfico. Una revisión de la literatura reciente desde la perspectiva de la vulnerabilidad post-divorcio. *Documents d'Anàlisi Geogràfica*, Vol. 57, N.1, pp 105-126.
- KALMIJN M., UUNK. W. 2007. Regional value differences in Europe and the social consequences of divorce: A test of the stigmatization hypothesis. *Social Science Research*, 36, 447-468.
- SIMÓ C., GONZÁLEZ ME., HERNÁNDEZ A.; MUÑOZ D. (forthcoming). "La salud de la población separada/divorciada. Una exploración con la Encuesta Europea de Salud de 2009". *REIS*.
- SOLSONA M. 2011. Biografías de divorcio: Cambios de estado borrosos y trayectorias familiares complejas. *Notas de Población*, 93, 77-106.
- VIGNOLI D., FERRO I. 2009. Rising marital disruption in Italy and its correlates.. *Demographic Research*, Vol. 20, N.4, pp.11-36.
- ZANATTA, A. L. 2008. *Le nuove famiglie. Felicità e rischi delle nuove scelte di vita*. Bologna: Il Mulino.

SUMMARY

Post-divorce vulnerability: the case of Spain and Italy

This paper analyses post-divorce vulnerability in Spain and Italy; the focus of the study regards cross-country comparisons: Being divorced in Italy or Spain has different impacts on post-divorce lives? Is more difficult to be divorced for men or women in Italy or in Spain? This research study the consequences that marital disruption has for its protagonists in the economic, health and social ambits comparing the outcomes obtained for both Italy and Spain and the identifying gender gaps. The results seem to confirm a post-divorce situation that is clearly difficult for women in both countries.

Thaís GARCÍA PEREIRO, Università degli Studi di Bari "Aldo Moro",
t.garcia.pereiro@uniba.it

Maria CARELLA, Università degli Studi di Bari "Aldo Moro",
maria.carella1@uniba.it

ANALYSIS OF FOOD CONSUMPTION IN EUROPE VIA TIME SERIES CLUSTERING

Margherita Gerolimetto, Christine Mauracher

1. Introduction

The second half of the twentieth century has been the theatre of an astonishing rebirth after WWII, which brought to unprecedented generalized socio-economic and wealthy levels, especially for Western countries. Such great economic changes that took place in this period profoundly alter not only the economies of Western countries but many aspects of their social life allowing people to improve their lifestyles and therefore their consumption.

In the last thirty years, the evolution of food consumption in European countries has been analysed by several authors and some general conclusions can be derived from those works. Some focused on the general picture (Blandford, 1984; Gil, et al. 1995; Gracia and Albisu, 2001), others analysed more specific regions (Elsener and Hartmann, 1998; Pingali, 2006) and finally other studies were conducted on the evolution of the food categories (Grigg, 1996; Schneider 2002).

The central aim of this work is to analyse how food consumption patterns have changed from 1961 to 2009 in Europe through the elaboration of FAO data via a statistical tool that, to our knowledge, had never been employed in this kind of studies, *i.e.* the cluster analysis of time series. This statistical tool allows analysing the trend followed by the thirteen food categories in each of the 39 European nations and to group the countries characterized by a similar evolutionary path.

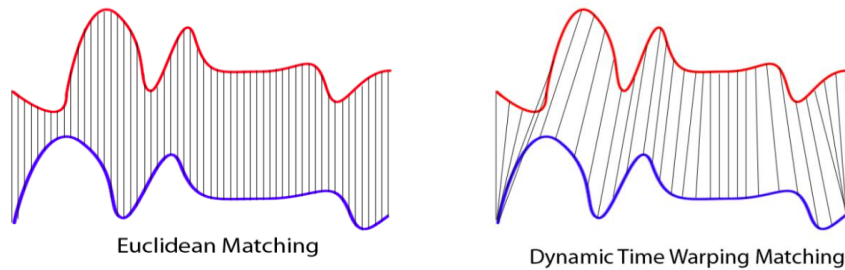
The paper is structured as follows: the second section illustrates clustering time series, the third section presents the data set and the results. The fourth section concludes.

2. Clustering time series with the Dynamic Time Warping (DTW)

To cluster time series there are two approaches (Liao, 2005). In the first approach, existing methods for clustering static data are modified in such a way

that time series data can be handled. In the second approach, time series data are converted into static data so that the traditional algorithm for clustering can be directly used. In this work the former approach has been chosen, also known as raw-data-based approach. It works directly with raw time series data and its major peculiarity, compared to static data cluster analysis, is in that the usual similarity/distance measures for static data are replaced with other measures, specific for time series. In particular, in order to overcome the flaws of the Euclidean Distance, in this paper we employ the Dynamic Time Warping (DTW) distance that, instead, allows an elastic shifting in the time axis, to adjust time series sequences that are similar but out phase, as illustrated in Figure 1.

Figure 1: *Euclidean vs DTW distance*



The principle of the DTW algorithm is to stretch and squeeze locally the two time series, as shown in Figure 1, so that the correct distance between the two time series must be calculated by summing the elements that have been aligned, i.e. between points in time that are in correspondence. Searching for the best match is the central point of the DTW. At the beginning this was done with forms of approximate indexing (among the others, Park et al. 2001), but more recently Keogh and Ratanamahatama (2004) introduce a new method for exact indexing in DTW.

At the core of the DTW technique is the concept of the warping curve $\phi(k) = (\phi_x(k), \phi_y(k))$, $k=1, \dots, n$, where $\phi_x(k) \in \{1, \dots, N\}$ and $\phi_y(k) \in \{1, \dots, N\}$

The warping function maps the time index of time series x_t and y_t and the average cumulated distortions between the two warped time series is

$$d_\phi(x, y) = \sum_{k=1}^T d(\phi_x(k), \phi_y(k)) \frac{m_\phi(k)}{M_\phi}$$

Where $m_\phi(k)$ is a weight and M_ϕ is the normalization constant which guarantees

that comparisons along different time path of the cumulated distortions. The optimal path is the one such that

$$D(x, y) = \min_{\phi} d_{\phi}(x, y)$$

Among all possible warps, only the reasonable ones are of interest, i.e. those such that the deformation of the time axes of x_t and y_t makes them resemble to each other as much as possible. With this in mind, some constraints are usually imposed, for instance monotonicity, to preserve the time order.

3. Empirical analysis with time series clustering

This study examines the 39 European countries, whose territories belong to Europe and with the boundaries and names used internationally in 2009.¹

The dataset used in this work contains information on the average daily per capita calorie (Kcal) consumption of 14 food aggregates² annually detected from 1961 to 2009. Data come from the Food Balance Sheets calculated by FAO³ and the food aggregates are: alcoholic beverages, animal fats, cereals, eggs, fish and seafood, fruits and tree nuts, meat and offals, milk, miscellaneous, starchy roots and pulses, stimulants and spices, sugar and sweeteners, vegetable oils and oilcrops, vegetables.

The use of caloric equivalents facilitates the analysis, since they are assumed to be common across countries for a given food product and it simplifies the cross country comparisons. However, the use of caloric weights implies greater emphasis on food products which are high in calories per unit e.g. fats, oils, cereal and starchy foods (Blandford, 1984).

We now move to the results of the analysis, using the hierarchical agglomerative clustering algorithm and the DTW distance, implemented, in particular, with the complete linkage method. Among all food categories, here for

¹ They are: Albania, Austria, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Moldova, Montenegro, Netherlands, Norway, Poland, Portugal, Romania, Russia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, The Former Yugoslav Republic of Macedonia (the FYROM), Ukraine and United Kingdom.

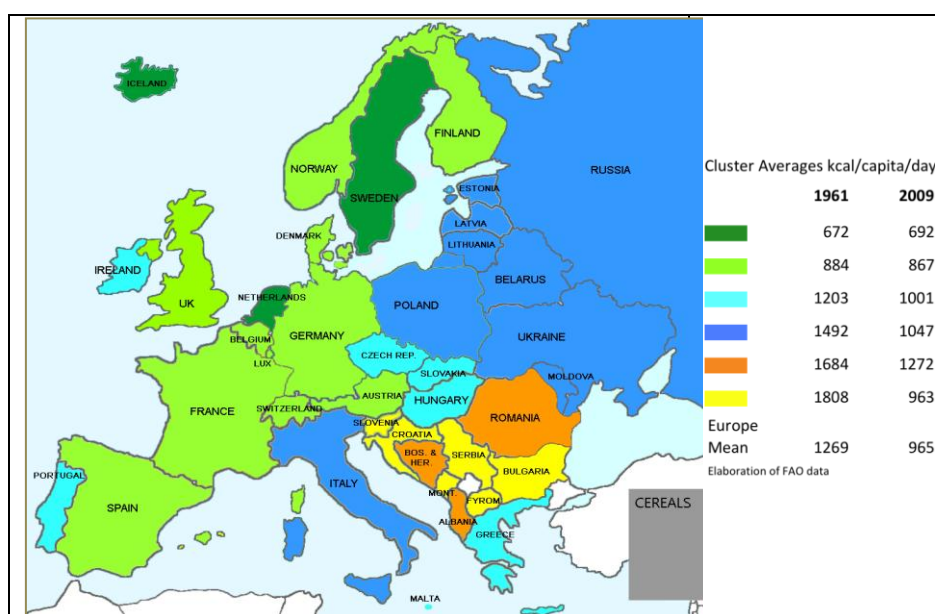
² The sum of these fourteen categories gives the 100 per cent of the daily total caloric intake per capita of every country analysed in this work.

³ It is important to clarify that these are not consumption data, but estimates of the available supplies, calculated from national statistics of production with the exclusion of exports, food put into stores and used for seed, feed and industrial purposes. A further reduction of 10% is made for waste estimation between farms and retail outlets. The final figures are food supplies available at a retail level in calories per capita.

space reasons, we focus only on cereals and meat, i.e. two important categories in the European diet, but other results relatively to the other food categories are available upon request from authors.

As for cereals consumption, the dendrogram has been cut into six groups and to each of them it is assigned a colour in order to obtain the European map of Cereal consumption (Figure 2).

Figure 2: Map of Cereal Consumption in Europe (kcal/capita/day)

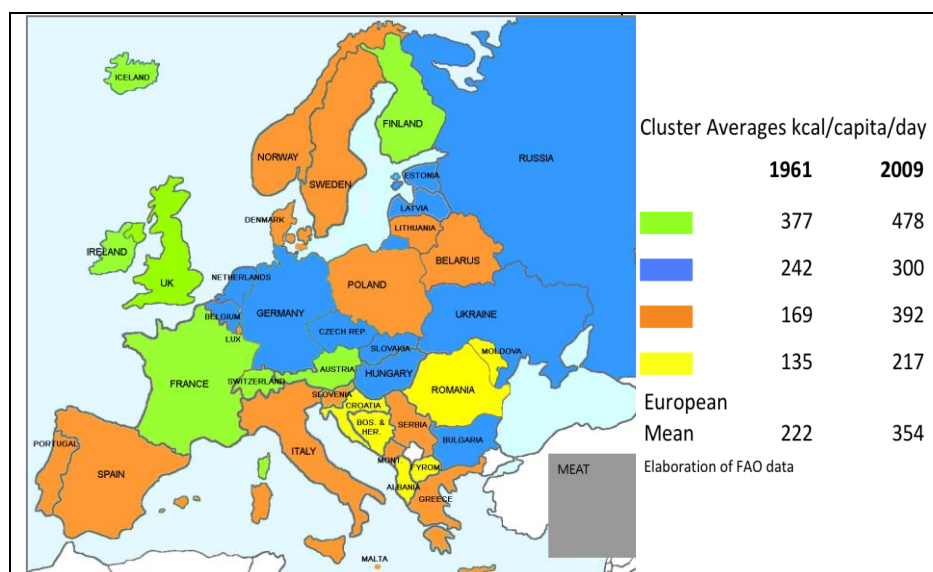


The countries painted in blue share almost the same pattern: they start with a high consumption, above the European mean, and the values decrease constantly over the period stabilizing in the last years around 1100 kcal/capita/day, still larger than the European mean. It is worth noticing that the majority of the members of this group are countries formerly belonging the Soviet Bloc and located in the North East of Europe. Unexpectedly, Italy also belongs to this cluster although it follows a slightly different path, with a sharp drop in early 1980s and from that point until nowadays the values are almost stable. Countries belonging to the light blue group follow a decreasing but less steep path. The light green group includes countries following a U-shape curve evolution. In particular, in the early years the cereal intake was lower than the average, however, with the amazing growth of the last decades, the cereal consumption increased and got closer to the European

mean.⁴

Moving to meat consumption, via the clustering analysis, yet again hierarchical agglomerative clustering algorithm, implemented, in particular, with the complete linkage method and the DTW distance, four clusters have been identified (Fig. 3).

Figure 3: Map of meat Consumption in Europe (kcal/capita/day)



Countries in this cluster mostly belong to Western and Northern Europe; Spain has a fairly different path since its consumption declined for the whole period. The dark green group is characterized by a pretty stable trend, with consumption values generally lower than the European mean ranging around 600 and 700 calories per capita per day. Finally, the yellow group shares a pattern which is slightly decreasing from 1961 to 1991 and shows a dramatic fall in consumption in 1992 and from that shock all the data stabilized to a level that is just above the European mean, except for Bulgaria where cereals consumption decreased over the period. The orange group is composed by countries whose pattern is peculiar but all of them share a high consumption of cereals. It is worth noting that in the 1960s, the diet in the Balkans is highly based on cereals consumption, which accounts for more than 50 per cent of the daily caloric per capita intake. This characteristic still holds afterwards for Albania, in which the share of cereals in the diet is even above

⁴ Such increase could be due to an increasing concern in health and cereals are very good against fibre-deficiency diseases (Grigg, 1996)

60 per cent until the early 1990s.

The green-painted group is characterised by a high consumption of meat, always higher than the European mean. On the opposite side, there is the yellow group including countries whose consumption of meat is always under the mean line. For the first three decades the trend is generally growing for most of the countries, but in the early 1990s there was a sharp collapse in consumption, probably this is linked to the Balkan wars and the following changes in political regimes. Between these two extreme groups there are other two clusters characterised by middle consumption levels: the orange-painted group, including countries whose initial meat consumption was quite low and that reached the European mean and eventually even crossed it over. The blue group includes almost all the countries linked to the Soviet Bloc. They started from middle consumption levels, which were generally a little higher than the orange cluster, and grew until 1990. Then the consumed meat amount dropped in the early 1990s but in the last fifteen years consumption has increased again.

It needs to be pointed out that the increasing pattern is common for almost all the members of the group, but with different rates. It is not by chance that the collapse in consumption happened just after the fall of the Berlin Wall. In most of these countries the consequent transition phase from a planned economy to a market one involved several changes in consumers' purchasing powers and so their food habits. In particular, the consumption of meat is one of the categories which were affected the most since it is a luxury good, due to its very high cost.

5. Conclusions

The main findings of this paper are the similarities and peculiarities in food consumption patterns between 1961 and 2009 in Europe. Food consumption is variably affected by different factors including food availability, food accessibility and food choice, which in turn may be influenced by geography, demography, disposable income, urbanization, globalization, marketing, religion, culture and consumer attitudes (Kearney, 2010). Moreover changes in the distribution system have had an important effect on food consumption patterns. Also trade liberalization is another relevant factor that has led to changes in food intake. In the considered period several events marked the history of Europe and deeply influenced the eating habits of the European population, e.g. two oil crises, the fall of the Berlin wall, the Balkan wars.

It is important to underline that the data used were not enough to perform disintegrated analyses which would give in further research more clarifying results.

References

- BLANDFORD D. 1984. Changes in food consumption patterns in the OECD area, *Review of Agricultural Economics*, Vol. 11, No. 1, pp. 43-64.
- ELSNER K., HARTMANN M. 1998. *Convergence of Food Consumption Patterns Between Eastern and Western Europe*. IAMO Discussion paper No. 13-1998, (Institut für Agrarentwicklung in Mittel und Osteuropa).
- GIL J. M., GRACIA A., PÉREZ y PÉREZ L., 1995, Food Consumption and Economic Development in the European Union, *European Review of Agricultural Economics*, Vol. 22 , No. 3, pp. 385-399.
- GRACIA A., ALBISU L. M. 2001. Food Consumption in the European Union: Main Determinants and Country Differences, *Agribusiness*, Vol. 17 , No. 4, pp. 469-488.
- GRIGG D. 1996. The Starchy Staples in Food Consumption, *Annals of the association of American Geographers*, Vol. 86, No. 3, pp. 412-431.
- KEARNEY J. 2010. Food consumption trends and drivers, *Philosophical Transactions of the Royal Society B*, Vol. 365, No. 1554, pp. 2793–2807.
- KEOGH E.J, RATANAMAHATAMA C.A. 2004. Exact indexing of Dynamic Time Warping. *Knowledge and Information Systems*, Vol. 7, pp. 358-386.
- LIAO T. W. 2005. Clustering of Time Series Data - a survey, *Pattern Recognition*, Vol. 38, pp. 1857-1874.
- PARK S., KIM S., CHU W. 2001. Segment-based approach for subsequence searches in sequence databases. In: *Proceedings of the 16th ACM symposium on applied computing*, Las Vegas, NV, pp 248–252.
- PINGALI P. 2006. Westernization of Asian Diets and the Transformation of Food Systems: Implications for Research and Policy, *Food Policy*, Vol. 32, No. 3, pp. 281-298.
- SCHNEIDER A. 2002. Overview of the market and consumption of pulses in Europe. *British Journal of Nutrition*, Vol. 88, No. 3, pp. 243-250.

SUMMARY

The purpose of the paper is to investigate the evolution of food consumption in the European countries. In particular the attention will be focused on the food consumption patterns and the major differences occurred in the period between 1961 and 2009. The statistical technique, that has been used, is the Cluster Analysis of Time Series, with the measure of the Dynamic Time Warping (DTW).

Margherita GEROLIMETTO, Lecturer in Economics Statistics, University of Venice

Christine MAURACHER, Professor in Agricultural Economics, University of Venice

A SAMPLING ESTIMATOR OF THE BONFERRONI INEQUALITY INDEX

Giovanni Maria Giorgi, Alessio Guandalini

1. Introduction

Bonferroni (1930) proposed the inequality index (B) as an alternative to the Gini index (R). For about half a century B remained almost forgotten because it was ostracized by Corrado Gini and his followers, who tried to prevent any measures of inequality from overshadowing the concentration ratio (R).

Recently, B has been revalued and studied for its features and for its interesting applications in social and economic contexts.

Since the '90s, thanks to an increasing availability of sample data, some researchers started to study also the sampling properties of B .

The main aim of this work is to propose a sampling estimator of B and to show its properties.

After a brief introduction on B and a comparison with R (section 2), in section 3 the sampling estimator is derived. In section 4 the estimator is compared through a simulation study to the percentile standard estimator of Bonferroni, \hat{B}_{100} , and to the estimator based on weighted observations, \hat{B}_r . Conclusions and future prospects of research are discussed in section 5.

2. The Bonferroni inequality index

The expression of B proposed by Bonferroni (1930, p. 55 and p. 85) is a function of partial means:

$$B = \frac{1}{N-1} \sum_{i=1}^{N-1} \frac{(\mu - \mu_i)}{\mu} = 1 - \frac{1}{N-1} \sum_{i=1}^{N-1} \frac{\mu_i}{\mu}, \quad (1)$$

where $0 \leq B \leq 1$, and

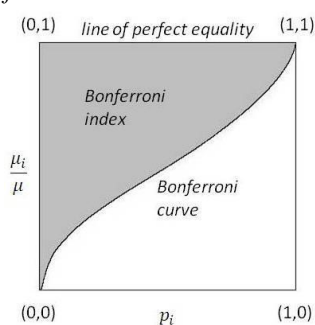
$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad \mu_i = \frac{1}{N} \sum_{j=1}^i x_j \quad i = 1, 2, \dots, N$$

are the general and the partial means for units sorted in non-decreasing order with respect to the X variable.

The B index satisfies the axiomatic properties required for inequality indices [i.e. the principle of transfer, of proportional addition to incomes, of proportional addition to persons, of symmetry, of normalization and of operationality] (Giorgi, 1998, p. 142) and $B \geq R$ holds, because it gives bigger weights to units with lower values in the X ranking (see, e.g., De Vergottini, 1940 pp. 318-319 and Pizzetti, 1951 p. 302). Therefore, it is more sensitive to lower levels of the income distribution (see, e.g., Giorgi & Mondani, 1995).

The Bonferroni index is linked to the Bonferroni curve (Figure 1) which is obtained by plotting the cumulative proportion of recipients ($p_i = i/N$), arranged in non-decreasing values of X , versus the corresponding ratio between partial mean and total mean (μ_i/μ).

Figure 1 – An example of Bonferroni curve.



The polygonal line joining the points $(p_i, \mu_i/\mu)$ is the Bonferroni curve. If all the recipients in the population have the same quantity of X (i.e. equal to μ) the Bonferroni curve coincides with the line of perfect equality that joins the coordinate points $(0,0)$, $(0,1)$, $(1,1)$.

The area between the Bonferroni curve and the line of perfect equality is the concentration area, which is equal to the value of B (Giorgi & Crescenzi, 2001, p. 572-573).

3. Sampling estimators of B

In the sampling framework we have a finite population $U = \{1, 2, \dots, N\}$ from which a sample s of size n is drawn according to a sample design $p(\cdot)$. The inclusion probability $\pi_k = \Pr\{k \in s\}$ is assumed to be known and strictly positive for all the units in the population. The inverse of inclusion probability is referred to as the sampling weight or expansion coefficient and denoted by $w_k = 1/\pi_k$.

In some applications (see e.g. Barcena-Martin and Silber, 2011 and 2012) the percentile estimator, \hat{B}_{100} , is often used to derive the value of the Bonferroni inequality index from sample data. This method consists in determining the percentiles of the distribution. Then the mean (\bar{x}_i) of X within the range defined by the $(i - 1)$ -th and the i -th percentile are computed, and B is derived by

$$\hat{B}_{100} = \frac{1}{100 - 1} \sum_{i=1}^{100-1} \frac{(\bar{x} - \bar{x}_i)}{\bar{x}}$$

using only the 100 points derived as previously explained.

For a better use of sample size and sample design, the weighted observations are usually considered

$$\hat{B}_r = 1 - \sum_{k=1}^{n-1} \frac{w_k \bar{x}_k}{N \bar{x}} \quad (2)$$

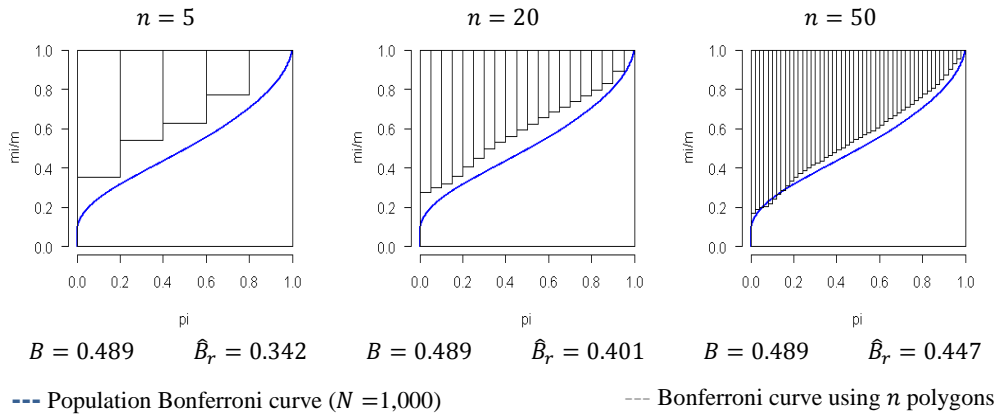
In practice, all the sample units are considered and the Bonferroni area is splitted in n rectangles, where $\bar{x} = \sum_{k=0}^n x_k w_k / \sum_{k=1}^n w_k$ is the sampling mean and $\bar{x}_k = \sum_{j=0}^k x_j w_j / \sum_{j=0}^n w_j$ the partial mean until the k -th unit, where, obviously, $w_0 = 0$ and $x_0 = 0$.

We want to point out that, unlike \hat{B}_{100} , where the area is decomposed in 100 interval of equal size (1/100), in (2) the size of the $n + 1$ intervals depends on the sampling weights and it is exactly equal to w_k / N .

In simple random sampling designs there are no differences among the size of the intervals, because the sampling weights are equal for each unit (i.e. $w_k / N = (N/n)(1/N) = 1/n$). Instead, differences in interval size arise when considering more complex sample designs that assign different sampling weights to the units.

However, in this case the concentration area is underestimated (see Gastwirth, 1972 p. 309, for the Gini index), especially for small samples. In fact, at least when the sample observed is unbiased, the piecewise line connecting the points generally lies above the curve (see Figure 2) and, moreover, the contribute to inequality of the population unit represented by the k -th unit in the sample is underestimated.

Figure 2 – Bonferroni area estimated by rectangles area (\hat{B}_r).



Better estimates can be derived considering a linear trend within the classes. In practice, the estimator

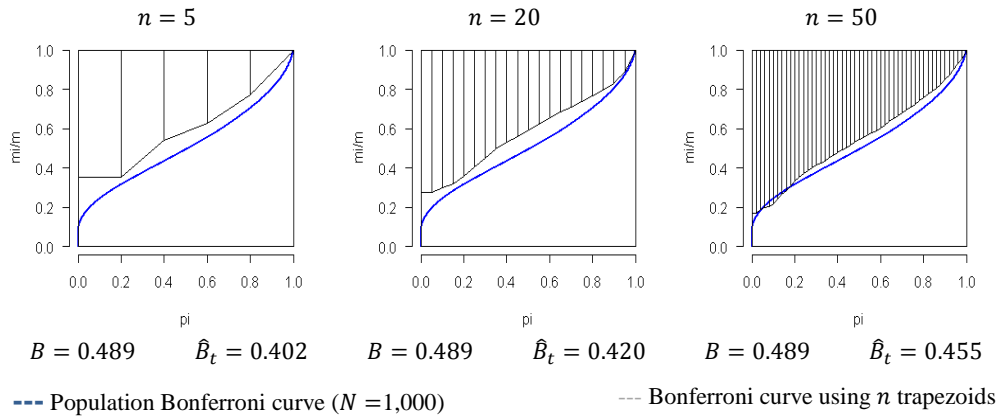
$$\hat{B}_t = 1 - \frac{1}{2n} \frac{1}{\bar{x}} \sum_{k=1}^n (\bar{x}_k - \bar{x}_{k-1}), \tag{3}$$

is computed summing the area of trapezoids the Bonferroni area can be decomposed in (see Figure 3).

With respect to \hat{B}_{100} , \hat{B}_t makes best use of the information capacity of the sample using all its size and not only 100 points as a synthesis of the whole sampling distribution. Moreover, with respect to \hat{B}_r , it recovers the inequality contribution of the units in the population represented by the generic sample unit k , while, when using the rectangles, it is considered only the inequality contribution of the sample unit k . This represent the most remarkable property of the proposed estimator and the reason of its advantage whit respect to \hat{B}_r , especially for small samples, as it is shown in the following section.

In some way, this estimator follows what is done by Eurostat (2004, p. 39) to derive the Gini ratio sampling estimator on Eu-Silc data.

Figure 3 – Bonferroni area estimated by trapezoids area (\hat{B}_t).



4. Simulation study

In this section, we compare the three estimators shown in section 3, that is \hat{B}_{100} , \hat{B}_r and \hat{B}_t .

To evaluate the performance of the estimators, we consider a superpopulation of 1,000,000 units described by a log-normal distribution. From this population we have drawn, with a simple random sampling with replacement, 1,000 populations of 100,000 units each and computed B . From each population, still with a simple random sampling with replacement, 100 samples are drawn and then used to derive B by the estimators seen in the previous paragraph.

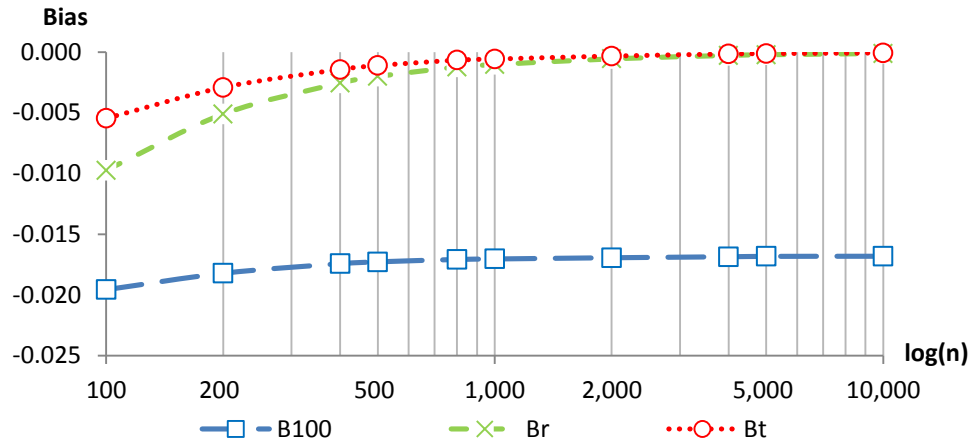
This procedure has been carried out for several sample sizes, respectively $n = 100, 200, 400, 500, 800, 1,000, 2,000, 4,000, 5,000$ and 10,000 units.

Figure 4 shows that all three estimators are affected by negative bias, that is, they underestimate the value of the index. The bias of \hat{B}_{100} is bigger than those of the other two that seem to be asymptotically unbiased.

Considering the accuracy of the estimators through the relative root mean squared error ($rRMSE$) it is possible to see, in Figure 5, that \hat{B}_t has better performance, in particular for small samples.

\hat{B}_t and \hat{B}_r are asymptotically equivalent and make better use both of the whole sample size and the sample design (the size of the intervals considered depends on the sampling weights). However, \hat{B}_t , considering trapezoids area, recovers information that \hat{B}_r is not able to catch.

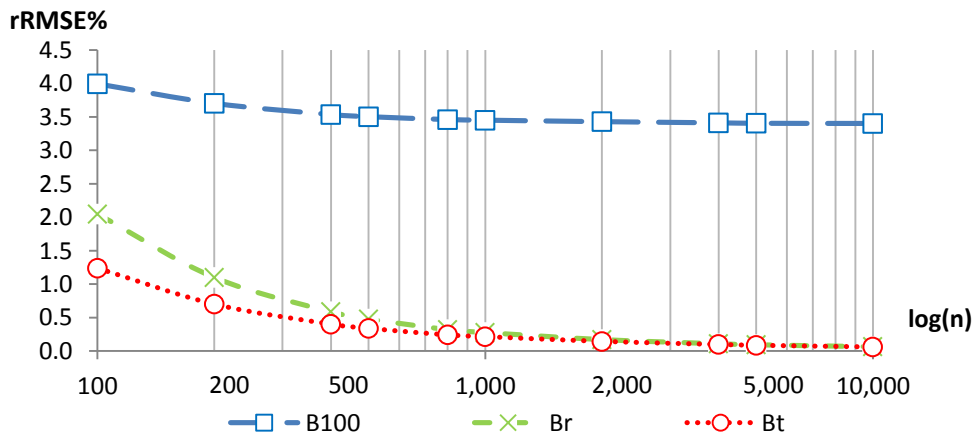
Figure 4 – Bias of \hat{B}_{100} , \hat{B}_r and \hat{B}_t for increasing n .



Note: n is expressed on logarithmic scale.

Using \hat{B}_{100} we have an advantage in computational burden, but also a loss of information. Even for large n , only 100 points are considered to derive the Bonferroni area. Moreover, the synthesis of the whole distribution that is obtained by the average of X between two consecutive percentiles adds a further approximation in the estimation of B . In fact, the error made by using \hat{B}_{100} stabilizes around 3.5% of the value of the index for n bigger than 500 and, even further increasing sample size, there are no significant improvements in estimates.

Figure 5 – $rRMSE\%$ of \hat{B}_{100} , \hat{B}_r and \hat{B}_t for increasing n .



Note: n is expressed on logarithmic scale.

5. Conclusions

In this paper a sampling estimator of the Bonferroni index based on the corresponding curve is proposed and compared with the percentile estimator and the estimator based on the weighted observations, often used in empirical applications. These estimators have been firstly illustrated and then tested for simple random sampling designs. In this case, the results are all in favor of \hat{B}_t .

The \hat{B}_t estimator decomposes the area between the Bonferroni curve and the line of perfect equality in trapezoids, as it is done by Eurostat to derive the Gini index from sample data.

The estimator \hat{B}_r is asymptotically equal to \hat{B}_t . However, \hat{B}_t is more convenient especially for small samples. The percentile estimator, \hat{B}_{100} , considering only 100 points, has a small computational advantage but its performances are always worst when compared to the other two.

However, to provide more general results, we think that further investigations are needed to test performances of the considered estimators for more complex sample designs.

Acknowledgements

The present work has been realized within the grant for the project “Indici classici di disuguaglianza e variabilità: nuove prospettive di ricerca” (Sapienza 2013).

References

- BARCENA-MARTIN E., SILBER J. 2011. On the concepts of Bonferroni segregation index curve. *Rivista Italiana di Economia, Demografia e Statistica*, Vol. 62, No. 2, pp. 57-74.
- BARCENA-MARTIN E., SILBER J. 2013. On the generalization and decomposition of the Bonferroni index. *Social Choice and Welfare*, Vol. 41, No. 4, pp. 763-787.
- BONFERRONI C.E. 1930. *Elementi di statistica generale*. Firenze: Libreria Seber.
- DALTON H.D. 1925. *Some aspects of the inequality of incomes in modern communities*. London: Routledge.
- DE VERGOTTINI M. 1940. Sul significato di alcuni indici di concentrazione. *Giornale degli Economisti e Annali di Economia*, Anno II (nuova serie), No. 5-6, pp. 317-347.

- EUROSTAT 2004. Common cross-sectional EU indicators based on EU-SILC; the gender gap. *Working Group on Statistics on Income and Living Conditions (EU-SILC)*, 29-30 March, Luxemburg, pp. 1-42.
- GASTWIRTH, J.L. 1972. The estimation of the Lorenz Curve and Gini Index. *The Review of Economics and Statistics*, Vol. 54, No. 3, pp. 306-316.
- GINI C. 1914. Sulla misura di concentrazione e della variabilità dei caratteri. *Atti del R. Istituto Veneto di SS.LL.AA.*, Vol. 73, parte II, pp. 1203-1248 (English traslation in *Metron*, 2005, Vol. 63, No. 1, pp. 3-38).
- GIORGI G.M. 1998. Concentration index, Bonferroni. In Kotz S. *et. al* (Eds.) *Encyclopedia of Statistical Sciences*, Update 2. New York: Wiley-Intersciences, pp. 141-146.
- GIORGI G.M., CRESCENZI M. 2001. A look at the Bonferroni inequality measure in a reliability framework. *Statistica*, Vol. 61, No. 4, pp. 571-583.
- GIORGI G.M., MONDANI R. 1995. Sampling distribution of Bonferroni inequality index from an exponential population. *Sankhya*, Series B, Vol. 57, No. 1, pp. 10-18.
- PIZZETTI E. 1951. Relazioni fra indici di concetrazione. *Statistica*, Anno XI, No. 3-4, pp. 294-316.

SUMMARY

A sampling estimator (\hat{B}_t) of the Bonferroni inequality index based on the Bonferroni curve, is presented. \hat{B}_t is compared through a simulation study with the percentile estimator and \hat{B}_r , the estimator that uses weighted observations. It is shown that the \hat{B}_t estimator, that decomposes the area between the Bonferroni curve and the line of perfect equality in trapezoids, has better performance than the other two, especially for small samples.

Giovanni Maria GIORGI, Department of Statistical Sciences, "Sapienza" University of Rome, giovanni.giorgi@uniroma1.it.
Alessio GUANDALINI, Department of Statistical Sciences, "Sapienza" University of Rome, alessio.guandalini@uniroma1.it

IL MERCATO DEL LAVORO IN ITALIA DAL 1977 AD OGGI

Cinzia Graziani, Silvia Loriga, Michele Antonio Salvatore, Andrea Spizzichino

1. Introduzione

La recente diffusione, da parte dell'Istituto Nazionale di Statistica, di nuovi dati ricostruiti sui principali indicatori del mercato del lavoro, offre nuove possibilità per analisi di lungo periodo a partire dal 1977¹.

Le serie trimestrali ricostruite, coerenti tra loro a tutti i livelli di aggregazione (per ripartizione geografica, genere, grandi classi di età, condizione lavorativa, posizione nella professione e settore di attività economica), consentono di superare i limiti dei dati della Rilevazione sulle Forze di Lavoro caratterizzati dai break generati in occasione dell'introduzione di innovazioni nell'indagine².

In questo lavoro sono stati evidenziati i principali cambiamenti che hanno caratterizzato l'occupazione e la disoccupazione negli ultimi 35 anni, sia nel complesso sia rispetto alle principali variabili demografiche e socio-economiche.

L'analisi, condotta sui dati di media annua, si è basata sull'utilizzo di tassi e incidenze.

Nel secondo paragrafo si illustra l'evoluzione del tasso di occupazione, con riferimento anche al dettaglio per genere, ripartizione geografica e settore di attività economica. Inoltre sono state analizzate le dinamiche del tasso di occupazione congiuntamente a quelle del prodotto interno lordo (Pil).

Alla descrizione dell'andamento del tasso di disoccupazione negli ultimi 35 anni è invece dedicato il terzo paragrafo: particolare attenzione è stata rivolta, oltre che all'analisi per genere e ripartizione, alle dinamiche che hanno caratterizzato il tasso di disoccupazione nella fascia di popolazione dei 15-24enni.

Nelle conclusioni, infine, si dà cenno degli sviluppi futuri del lavoro.

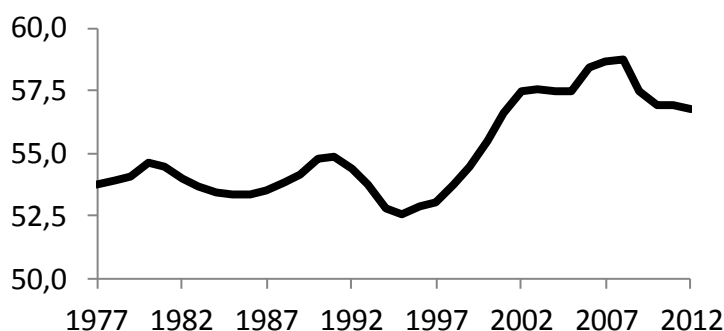
¹ I dati delle serie ricostruite sono disponibili sul datawarehouse I.Stat (dati.istat.it). Per maggiori informazioni sui risultati della ricostruzione si consulti: ISTAT. 2013. Occupati e disoccupati: dati ricostruiti dal 1977. Statistiche Report, Roma, 24 aprile 2013.

² Per maggiori informazioni sulla metodologia utilizzata si consulti Graziani et al. 201.

2. Crescita della partecipazione femminile al mercato del lavoro e sviluppo del settore terziario

Tra il 1977 e il 2012 il tasso di occupazione dei 15-64enni è passato dal 53,8% al 56,8% (Figura 1). L'incremento è attribuibile esclusivamente alla crescita della partecipazione delle donne al mercato del lavoro: il tasso di occupazione femminile è infatti aumentato di 13,6 punti percentuali (dal 33,5% al 47,1%), più che compensando la diminuzione di 8,1 punti (dal 74,6% al 66,5%) registrata per quello maschile. L'incidenza delle donne sul totale degli occupati è dunque salita dal 31,5% al 41,3%; tuttavia il tasso maschile rimane sostanzialmente superiore a quello femminile e, pur riducendosi il divario, sono ancora lontani gli obiettivi di Europa 2020 che vogliono, per il nostro Paese, il tasso maschile al 73% e quello femminile al 52%.

Figura 1 – Tasso di occupazione 15-64 anni. Medie annue 1977-2012, valori percentuali.



Anche a livello territoriale si notano dinamiche differenti: nel Nord e nel Centro l'aumento registrato tra le donne è superiore alla riduzione tra gli uomini evidenziando nel complesso incrementi nel tasso di occupazione molto simili (rispettivamente dal 57,5% al 65,0% nel Nord e dal 53,4% al 61,0% nel Centro); nel Mezzogiorno invece l'incremento femminile non ha compensato la forte riduzione registrata tra gli uomini e il tasso, che già partiva da un valore più contenuto, è ulteriormente sceso dal 48,9% al 43,8%.

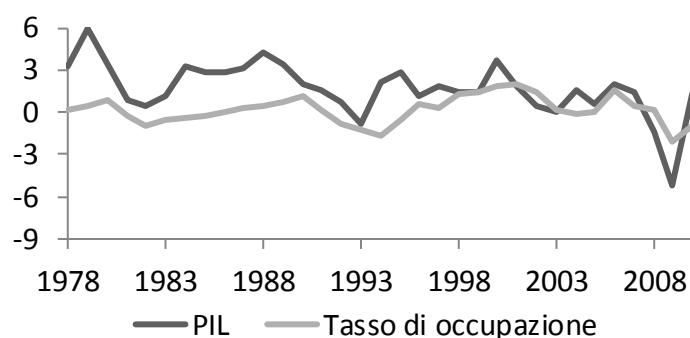
Tra il 1977 e il 1995 il range di variazione del tasso di occupazione è stato di soli 2,4 punti percentuali: il minimo, pari al 52,5%, si è registrato nel 1995 e il massimo, pari al 54,9%, nel 1991. Dal 1996 al 2008 si è registrata una crescita di 6,2 punti percentuali e successivamente un calo di 1,9 punti che ha portato il tasso al 56,8% nel 2012. È interessante analizzare come tale dinamica si sia

accompagnata a quella dell'output di produzione, espresso in termini di prodotto interno lordo (Pil) (Figura 2).

Tra la fine degli anni Settanta e la metà degli anni Novanta l'andamento del tasso di occupazione ricalca quello del Pil (calcolato dalla Contabilità nazionale). Confrontando le serie storiche dei rispettivi tassi di variazione annui, è evidente come in questo periodo, caratterizzato peraltro da crescita economica, i rallentamenti e gli incrementi nella crescita del Pil anticipino le riduzioni e gli aumenti del tasso di variazione dell'occupazione.

Dalla metà degli anni Novanta al 2007 sia il tasso di occupazione sia il Pil presentano dinamiche tendenzialmente positive, seppur fortemente variabili, e vicine tra loro, con la crescita del Pil nel periodo 2000-2003 inferiore alla crescita dell'occupazione. Mentre le fasi di espansione del Pil registrate nel periodo precedente si sono accompagnate a incrementi decisamente modesti dell'occupazione, dalla metà degli anni Novanta al 2007 si osservano incrementi decisamente più forti del tasso di occupazione in corrispondenza di variazioni del Pil più contenute.

Figura 2 – Tassi di variazione annui del tasso di occupazione 15-64 anni e del PIL. Periodo 1977-2010, valori in punti percentuali.



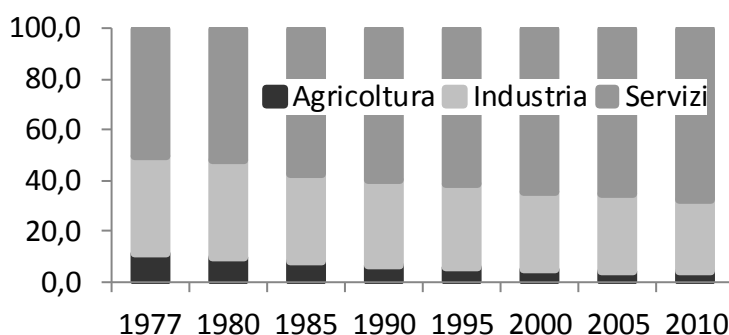
La crescita più forte dell'occupazione si può ricondurre a due fattori principali: il processo di terziarizzazione e la diffusione di forme di lavoro atipiche.

L'accelerazione che ha caratterizzato il processo di terziarizzazione nei primi anni del Duemila ha mostrato un rallentamento della produttività e una contestuale crescita dell'occupazione. La liberalizzazione delle forme di lavoro atipiche ha contribuito notevolmente all'espansione dell'occupazione, soprattutto tra le donne, maggiormente disponibili ad accettare forme di lavoro con orario ridotto e/o a termine data la necessità di conciliare il lavoro con la cura della famiglia. La fase di recessione economica, del 2008 e del 2009, anticipa la fase di calo che fino ad oggi

ha interessato l'occupazione portandola agli attuali record negativi (Destefanis 2012).

Il processo di terziarizzazione ha caratterizzato l'intero periodo considerato portando la quota di occupati nei servizi dal 50,6% del 1977 al 67,6% del 2010³ trasformando profondamente il quadro occupazionale riguardo alla composizione settoriale (Figura 3). L'incremento dell'occupazione nei servizi risulta rilevante in tutte le ripartizioni con quote che aumentano dai 12,2 punti percentuali nel Centro (da 59,3% a 71,5%) ai 18,5 punti del Nord (da 45,3% a 63,8%). L'incidenza degli occupati in agricoltura sul totale risulta più che dimezzata passando dall'11,0% del 1977 al 3,9% nel 2010 (il calo più consistente è stato rilevato nel Mezzogiorno, dal 19,3% al 6,7%); anche la quota di occupati nell'industria ha registrato una forte diminuzione, pari a quasi 10 punti percentuali, passando dal 38,4% dell'occupazione totale nel 1977 al 28,5% nel 2010 (la maggiore diminuzione al Nord, dal 47,6% al 33,2%).

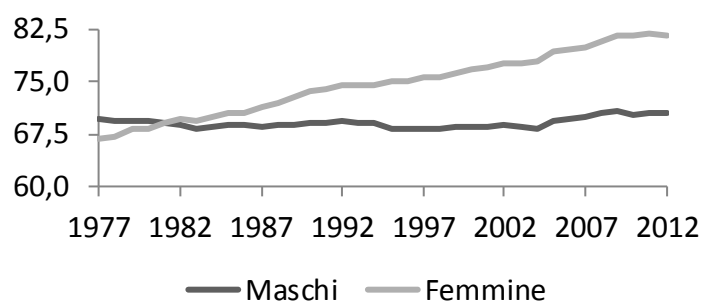
Figura 3 – Quota di occupati per settore di attività economica. Medie annue 1977-2010, valori percentuali.



Anche la distribuzione per posizione professionale ha subito forti modifiche con la percentuale di lavoratori dipendenti che è salita dal 68,8% al 75,2%. I cambiamenti vanno attribuiti, quasi del tutto, alla componente femminile per la quale la quota di occupazione dipendente è cresciuta di quasi 15 punti percentuali, passando dal 66,9% del 1977 all'81,7% del 2012 (Figura 4). Tra gli uomini invece l'aumento di lavoratori alle dipendenze è stato inferiore a un punto percentuale, passando dal 69,7% al 70,5%.

³ I dati per settore d'attività economica risultano confrontabili solo fino al 2010 in quanto a partire dal 2011 è stata adottata la nuova classificazione delle attività economiche ATECO2007 (Nace Rev. 2) che ha sostituito la precedente classificazione ATECO2002.

Figura 4 – Quota di occupati dipendenti secondo il genere. Periodo 1977-2010, valori percentuali.



L'aumento della quota di lavoro dipendente, soprattutto per le donne, si accompagna al processo di terziarizzazione. Nei servizi infatti aumenta di 10,0 punti percentuali passando dal 64,7% del 1977 al 74,7% nel 2010; in particolare per le donne sale dal 34,6% al 54,1%.

Diverse sono le dinamiche negli altri settori economici: in agricoltura si conferma la crescita del lavoro alle dipendenze che passa dal 42,7% del totale occupati nel 1977 al 48,1% nel 2010; tuttavia la percentuale di donne dipendenti cala dal 34,9% al 30,2%. Diverso ancora è il caso dell'industria, in cui la quota di lavoratori dipendenti scende dall'81,6% al 78,8%, con la componente femminile che passa dal 25,8% al 22,4%.

3. L'evoluzione nel tempo della disoccupazione

Per quanto riguarda la disoccupazione, il tasso è passato dal 6,4% del 1977 al 10,7% del 2012. Osservando la dinamica, che ha visto l'alternarsi di fasi di crescita e di calo, è interessante evidenziare la crescita di 4,6 punti percentuali registrata dal 2007 al 2012: una variazione di tale entità non è mai stata rilevata in passato e ha annullato in metà del tempo la diminuzione di 5,2 punti avvenuta tra il 1998 e il 2007 (Figura 5).

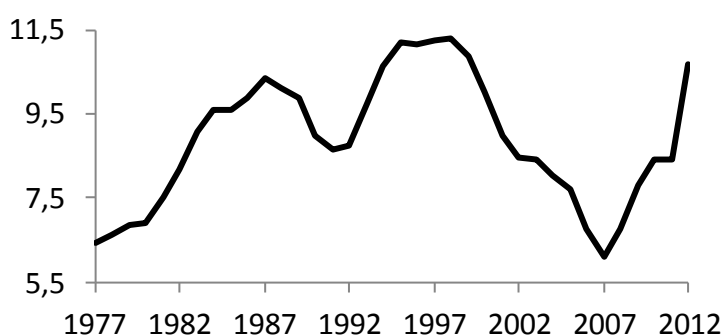
Dal 1977 al 2012 l'incremento è stato di 5,5 punti percentuali (dal 4,3% al 9,9%) per la componente maschile, mentre per la componente femminile la variazione è risultata più contenuta (dal 10,7% all'11,9%).

A livello territoriale il tasso di disoccupazione presenta andamenti simili nelle tre macro aree, seppur con intensità differenti. Il Mezzogiorno mostra l'incremento

più accentuato (dall'8,0% del 1977 al 17,2% del 2012) a fronte di quello più moderato del Centro (dal 5,5% al 9,5%) e soprattutto del Nord (dal 5,8% al 7,4%).

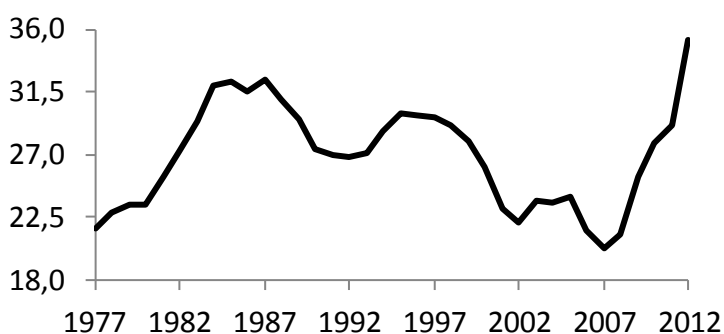
Nel Nord e nel Mezzogiorno l'aumento deriva da una crescita del tasso di disoccupazione maschile, superiore rispetto al calo registrato per le donne. Nel Centro invece si è registrato un aumento sia per gli uomini sia per le donne, con intensità maggiore per i primi.

Figura 5 – Tasso di disoccupazione. Medie annue 1977-2012, valori percentuali



Tra quanti cercano lavoro, i giovani hanno sempre mostrato maggiori criticità. Il tasso di disoccupazione dei 15-24enni è salito dal 21,7% del 1977 al 35,3% del 2012 (Figura 6). L'aumento ha coinvolto sia gli uomini sia le donne: per i primi, il tasso è cresciuto dal 18,1% al 33,7%, per le seconde dal 25,9% al 37,5%.

Figura 6 – Tasso di disoccupazione 15-24 anni. Medie annue 1977-2012, valori percentuali



Le difficoltà di questa componente della forza lavoro emergono ampiamente confrontando gli indicatori specifici giovanili con quelli complessivi. Il divario tra

il tasso di disoccupazione dei 15-24enni e quello complessivo si è progressivamente allargato dai 15,3 punti del 1977 ai 24,6 del 2012, da notare che dal 1985 al 2002 si era invece ridotto fino a 13,6 punti e che, in soli 5 anni, dal 2008, è cresciuto di 11 punti percentuali.

A livello territoriale la crescita della disoccupazione giovanile presenta intensità e dinamiche differenti. Nel Mezzogiorno il tasso ha registrato l'incremento maggiore passando dal 28,3% al 46,9%, nel Nord il tasso è salito dal 17,5% all'attuale 26,6%, mentre nel Centro dal 22,4% al 34,7%. In particolare nel Centro-Nord il tasso più basso si è registrato nel 2002, pari a 9,9% nel Nord e a 16,9% nel Centro.

4. Conclusioni

La ricostruzione delle serie storiche diffusa recentemente dall'Istat consente di superare i break presenti nei dati della Rilevazione sulle Forze di Lavoro rendendo così possibile la conduzione di analisi di lungo periodo. La disponibilità di questi nuovi dati consente di valutare le dinamiche che hanno caratterizzato il mercato del lavoro nel periodo precedente la crisi economica che ha colpito l'Italia nei primi anni Novanta.

Per il futuro sarebbe auspicabile raggiungere un maggior livello di dettaglio, così da integrare la serie precedentemente ricostruita a partire dal quarto trimestre del 1992 (Spizzichino e Gatto 2011). La ricostruzione dal quarto trimestre 1992 garantisce, infatti, anche il dettaglio per regione, classi d'età quinquennale, carattere, posizione e tipologia di orario dell'occupazione, nonché per la durata della disoccupazione.

Riferimenti bibliografici

- CONTINI B., TRIVELLATO U. 2005. *Eppur si muove: dinamiche e persistenze nel mercato del lavoro italiano*. Bologna: Il Mulino.
- DESTEFANIS S. 2012. Alcune considerazioni sul mercato del lavoro italiano alla luce della ricostruzione delle serie storiche territoriali per il mercato del lavoro, 1861-2011. *SVIMEZ, Nord e Sud a 150 anni dall'Unità d'Italia*.
- GRAZIANI C., LORIGA S., MARTINI A., SPIZZICHINO A. 2011. La ricostruzione delle serie storiche sul mercato del lavoro dal 1977. *Rivista Italiana di Economia, Demografia e Statistica*, Vol. LXV, No. 3/4, Luglio-Dicembre 2011.
- ISTAT 2013. Occupati e disoccupati: dati ricostruiti dal 1977. *Statistiche Report*, Roma, 24 aprile 2013.

- ISTAT 2011. I tempi del lavoro. *Argomenti n.40*, Roma.
- SPIZZICHINO A., GATTO R. 2011. Serie storiche sul mercato del lavoro: riconciliazione con la popolazione ufficiale intercensuaria. *Collana Istat Working Paper, n°13*.
- ISTAT 2006. La Rilevazione sulle Forze di Lavoro: contenuti, metodologie, organizzazione. *Metodi e Norme n. 32*, Roma.

SUMMARY

The Italian Labour Market from 1977 up today

Istat recently disseminated reconstructed data on the main labour market indicators, providing further opportunities for long-term analysis since 1977.

Reconstructed time series are consistent with data currently released and allow to overcome the limitations due to the breaks generated by innovations introduced in the Labour Force Survey over time.

In this paper we analyze the main changes occurred on employment and unemployment in the last 35 years according to the main demographic and socio-economic variables. The analysis shows an increase in employment due to the female growth and to the male stability. Female employment rate rises from 33.5% in 1977 to 47.1% in 2012, while the male one decreases from 74.6% to 66.5%.

The share of employment in the services sector rises from 50.6% to 67.6%.

The growth in employment among women and in the services sector is associated with the growth of employees, whose share rises from 68.8% to 75.2%.

The unemployment rate increases for both men (from 4.3% to 9.9%) and women (from 10.7% to 11, 9%). Among the unemployed, young people always shows critical issues: the unemployment rate of people aged 15-24 rises from 21.7% in 1977 to 35.3% in 2012.

In the future it's our intention to achieve a greater level of detail, as done for the series reconstructed from the 4th quarter of 1992.

Cinzia GRAZIANI, Istituto Nazionale di Statistica, cingraziani@istat.it

Silvia LORIGA, Istituto Nazionale di Statistica, siloriga@istat.it

Michele Antonio SALVATORE, Istituto Nazionale di Statistica, salvatore@istat.it

Andrea SPIZZICHINO, Istituto Nazionale di Statistica, spizzich@istat.it

MAJOR, INITIAL FAMILY-LIFE TRANSITIONS AMONGST YOUNG ITALIANS

Sara Grubanov-Bošković, Karra Greenberg

1. Introduction

Leaving the parental home together with completing formal education, entering the labor market, and forming a stable romantic partnership represent key events in transitioning into the independent, adult social status in Italy. The main route for leaving the home has traditionally been marriage but, in the last several years, other living arrangements such as cohabitation and living alone are gaining more importance across Europe (Corijin, Manting, 2000). Our paper aims to better understand the likelihood with which these alternative forms of family and living arrangements are pursued and whether there are particular groups of young Italians that are more likely to engage in some forms of major family-life transition as opposed to others.

2. Theoretical Background of the Italian Transition to Adulthood

The features of the Italian transition to adulthood have previously been described as the so called “Mediterranean” or the “latest-late” pattern of transition according to which a delayed marriage is not compensated by the rise of non-marital youth cohabitation and non-family households, but by a longer permanence in the parental home (Mencarini, Solera, 2011; Aassve *et al.*, 2002; Billari *et al.*, 1998). Other studies have proved that the young adults in Italy show the most relevant delay in terms of age in the choice of leaving the parental home with respect to the other advanced economies (Sironi *et al.*, 2012). This prolongation of the transition has taken place within traditional models of family formation – i.e. by marriage – and therefore the incidence of cohabitation or living alone is relatively small, though increasing (see e.g. Naldini, Jurado, 2010). The longer period of formal studying and the consequent delay of entrance into the labor market have certainly contributed to postponing the transition to a new and independent living arrangement. The economic difficulties have been the main reason for not leaving the parental home, mostly for women due to the persisting

presence of marked gender asymmetries on the labor market (Mencarini, Solera, 2011; Istat, 2009; Pupo, Licursi, 2003)¹. In this scenario, the low share of cohabitation as an alternative way of gaining residential autonomy is highly influenced by its low level of social acceptance which is officially evident at the legislative level (see e.g. Ongaro, 2004). Cohabitation is thus observed to be a phase preceding marriage and it terminates with the arrival of children which transitions the choice of cohabitation into the choice of marriage (see e.g. Scabini, Rossi, 2006). The social context was also reported to be important for the choice of living alone: it is a phenomenon more common for men and more frequent in big urban areas of North Italy where living as a single young adult is more socially acceptable (see e.g. Benassi, Novello, 2009).

3. Hypotheses

We hypothesize that higher educated men and women under the age of 40 are less likely to marry, cohabit, or live alone than they are to live at home. Increased time in school may translate into less time to invest in the search for a long-term, romantic partner plus less money with which to live alone. Of those men and women who have the resources to leave the parental home, and a desire to engage in a co-residential romantic partnership, we hypothesize that they may engage in cohabitation as a substitute for marriage because they may still not have the financial resources fully necessary to engage in marriage. We hypothesize that young, working Italian men and women may be only slightly more likely to marry than to cohabit relative to living at home with family. For women, we hypothesize that the most economically independent women will have the greatest likelihood of living alone as opposed to any other type of living arrangement. For men with resources we hypothesize that living alone is a viable alternative to marrying, although marrying will be more likely than living alone.

4. Data, Method and Descriptive Output

In this article, we propose an alternative approach for studying the young adults' initial life transitions by directly testing all family-life living options against each other — options that young Italians are naturally faced with upon choosing to leave the parental home. In the absence of longitudinal data, 2012 data from the

¹Despite achieving economic independence, some young adults decide to remain in the home of origin to take care of their parents, mostly in the case of Italian women, or because they already have large degrees of autonomy and independence while living in the *liberal* family of origin (Istat, 2009).

Istat survey “Indagine Multiscopo sulle Famiglie” is used. Pulling men and women ages 19-40 we generated a subsample of 11.403 observations related to the state which was restricted to individuals who had never married or who were currently in their first marriage². The subsample is representative both at the household and individual level for all of Italy. Considering the family of origin as well as the variety of routes for leaving the parental house and starting independent living, we create a multinomial dependent variable “Living arrangement” composed of the following mutually exclusive categories:

- 1) Never married Italians living with their parents or other relatives –*Parent*–;
- 2) Never married Italians living alone –*SPH*–;
- 3) Never married Italians that cohabit with their romantic partner –*Cohab*–;
- 4) Italians in their first marriage – *First Marriage*.

By applying a multinomial logistic regression, where “Living arrangement” is set to be the dependent variable, it was possible to evaluate the effect of different socio-demographic characteristics on the likelihood of not marrying and living with parents or relatives, of living alone, of cohabiting, or of being married for the first time. Since important differences exist between women and men, such as the timing of major life transitions in Italy, separate models were run for both. Also, considering that the risk of leaving home fluctuates strongly with age, we created 4 age groups: 19-24; 25-29; 30-35; 36-40. Other independent variables included are educational level³ and professional status⁴ since extended time in school and/or difficulties in finding a job emerged to be correlated with remaining in the parental home (Istat, 2009). Regional differences were controlled.

Tables 1 corroborate the idea of a delayed transition into adulthood of young Italians since 61,0% of men ages 25 to 40 and 54.1% of women ages 25 to 40 still live with their parents. It also confirms that the marriage is still a dominant transitional route although with substantial gender differences. On the other hand, the alternative living arrangement such as living alone and cohabiting appears not to be marginal among never married young Italians (with a total share of 12,8% for women and 15,2% for men). The mean ages and the share of each age group in every considered living arrangement add support to the hypothesis that

²The data do not permit us to determine if these young Italians have lived alone or cohabited prior to their current living arrangement. What we know, and underline as a crucial assumption for our first life-transition analysis, is that they have not been married before, or, in case they are married, it is their first marriage experience. In this way we eliminate the effect of divorce, separation, and widowhood on the choice of living arrangement.

³The education was summarized in 4 categories: postgraduate (that includes PhD, four-year university degree – old cycle – and Master degree); higher education (Bachelor, college, academy degrees); high-school; less than high school (incorporating primary education and lower secondary education).

⁴Within the professional status variable, the category “unemployed” incorporates individuals seeking a new job, seeking the first job as well as those who are classified in the dataset as unable to work.

cohabitation and living alone are transitional phases for some — after the age of 35 marital unions are slightly more common. While living alone is more common for men and cohabitation for women, the common feature for both living arrangements is that they are more prevalent in Northern Italy. However, a statistical analysis is necessary to determine if there are significant differences between these living arrangements and those that engage in them.

Table 1 – Descriptive Statistics

	Female				Male			
	SPH	Cohab	First Marriage	Parent	SPH	Cohab	First Marriage	Parent
Total	5,8	7,0	41,4	45,7	9,5	5,7	27,5	57,4
Mean age	32,4	31,6	34,4	26,2	32,9	32,3	35,4	26,9
Age Group								
19-24	8,1	13,3	2,7	45,8	6,3	11,2	0,8	42,0
25-29	21,1	22,3	11,6	28,1	18,8	18,1	7,0	26,3
30-35	38,5	33,8	39,3	17,4	38,0	36,6	36,2	20,6
36-40	32,3	30,7	46,4	8,7	36,9	34,1	55,9	11,1
Postgraduate	25,8	17,1	15,1	13,5	14,8	9,4	10,6	8,2
Higher educ.	12,7	8,2	4,4	9,4	6,5	3,9	3,7	6,2
High-School	45,3	46,3	45,7	58,8	51,7	47,1	48,0	57,6
Less than High-School	16,1	28,4	34,8	18,4	26,9	39,6	37,7	28,0
Housewife	0,0	10,0	32,4	2,6	-	-	-	-
Student	5,9	6,6	0,9	32,8	3,1	4,5	0,1	23,0
Employed	78,3	68,5	51,3	36,9	85,0	84,0	87,6	48,0
Unemployed	15,8	14,8	15,4	27,6	11,9	11,5	12,3	28,9
North-West	24,5	29,7	20,0	16,3	23,0	33,2	21,1	17,4
North-East	26,1	35,5	20,9	17,4	23,7	33,5	21,3	18,9
Center	19,6	19,7	16,3	17,5	19,7	16,6	16,6	16,6
South & Island	29,8	15,1	42,8	48,8	33,6	16,6	41,0	47,0

Source: own elaboration of Istat data (2012).

5. The Demographics of Who is Likely to Live in a Particular Arrangement

As expected, as women age, their likelihood of living outside of their parental home dramatically increases for every type of alternative living arrangement (marriage, living alone, and cohabitation). Of all these forms of major life transition from the parental home, marriage – when compared to living alone or cohabiting – is *not* the most likely for individuals aged 19-29⁵. In focusing on the

⁵ All cited differences in odds between categories of the dependent variable have been tested and are statistically significant.

⁶ The difference between odds were calculated by using relative risk ratio ratios (RRRR).

transition from a younger adult to an older adult, i.e. for women 30-35 compared to women 25-29 the odds of entering cohabitation or living alone increases approximately 125,0%-160,0%, respectively; whereas in the case of first marriage there is a 413,0%⁶ increase in the odds between these two age groups. Interestingly, living alone is *more* likely than marriage for women between the ages of 30-35. This begins to provide evidence that living alone and cohabiting might be viable options to marriage, even at older ages, for some young Italians making their first major life transition. Moreover, as a woman's level of educational attainment increases, the odds of getting married are reduced, which suggests that marriage, for highly educated women, may be increasingly unattractive. In terms of professional status, compared to women who are unemployed, employed women are 130,0% more likely to live alone than live at home, 86,0% more likely to cohabit and 63,0% more likely to marry. Although it is clear that women who identify as housewives have a higher probability of being married (or vice versa), this relationship is less intuitive for never married cohabiters: employed women are still 86,0% more likely to cohabit than unemployed women, but it is not near the 480,0% higher likelihood of being a housewife. These findings suggest that there may be two different types of female cohabiters in contemporary Italy. The first type are the working women who stay attached to work because they enjoy financial empowerment and may not want to engage in a traditional marriage with strict gender-role delineations which would separate them from the labor market.

The other type of female cohabiters is the more probable: they are less attached to the labor force and engage in domestic responsibilities. Such a finding is backed up by odds ratios that report that women with the lowest education have the greatest odds of cohabiting or marrying and suggests that cohabitation in Italy may be far more similar to marriage in its character than it is in other countries. Thus, in these cases, cohabitation may be an adaptive family form for those with poor financial prospects. These findings are very important because they speak to an emerging dichotomy amongst Italian women: those with the least attachment to the workforce are the most likely to be married or cohabit and those with strong attachment to the workforce are the most likely to live alone.

It cannot be ignored that the causal direction is not clear – women may have left the work force after marriage and not beforehand – but the final relationship is nonetheless clear: Italian women have increased odds of being attached to the labor force, relative to being unemployed, when they are *not* involved in a marital or cohabiting union. Relative to living at home, all alternative forms of partnership were more likely for women in the North than in the South of Italy. While not surprising for living alone or cohabitation, for marriage this may be a commentary on greater economic opportunities in the North which enable greater probabilities of young couples feeling financially prepared to marry. Similar to women, as men

age, their likelihood of living outside of their family of origin dramatically increases for every type of alternative living arrangement, with the highest odds ratios for the oldest age group. Consistent with women, they are also more likely to live alone rather than marry between the ages of 30 and 35. At all levels of education for high school completion and above, young Italian men are more likely to live in their parental home than to cohabit or marry.

Table 2 – Multinomial regression outcomes (odds ratios)

Covariates	Female			Male		
	SPH	Cohab	First Marriage	SPH	Cohab	First Marriage
Age						
19-24	,087***	,120***	,016***	,073***	,119***	,007***
25-29	,253***	,272***	,091***	,254***	,274***	,064***
30-35	,650***	,611***	,463***	,573***	,618***	,362***
36-40	<i>Reference category</i>					
Educational attainment						
Postgraduate	,852	,406***	,448***	,966	,405***	,450***
Higher educ.	1,318	,555**	,505***	1,151	,441**	,532***
High-School	,858	,550***	,721***	,997	,573***	,715***
Less than High-School	<i>Reference category</i>					
Professional Status						
Housewife	1,77e-01	5,807***	19,632***	-	-	-
Student	,601*	,627**	,186***	,790	1,015	,097***
Employed	2,306***	1,700**	1,628**	3,152***	2,735***	3,593***
Unemployed	<i>Reference category</i>					
Geographical Partition						
North-West	1,826***	5,387***	1,565***	1,380**	4,413***	1,012
North-East	1,777***	6,024***	1,499***	1,330*	4,163***	,966
Center	1,484*	3,587***	1,218*	1,287	2,429***	,883
South and Island	<i>Reference category</i>					

Note: Reference category. Never married persons living with their parents or other relatives

Source: own elaboration of 2012 Istat data.

The similarity in likelihood between marrying and cohabiting for this demographic again speaks to the idea that cohabitation and marriage are very similar in many regards in Italy. Additionally, there appears to be little distinction between socioeconomic groups and the choice to cohabit or marry. The nature of this similarity must be investigated more. Employed men are more likely to engage in every form of living arrangement relative to living at home: they are 259,0% more likely to marry, 215,0% more likely to live alone, and 173,0% more likely to cohabit. This raises more questions about the nature of single-living for young men in Italy since the likelihood of marrying when employed, relative to living at home, was only slightly greater than the likelihood of living alone.

5. Conclusions

Relative to living at home, both Italian men and women under the age of 40 are less likely to marry, cohabit, or live alone with the more education that they have. This is likely due to the fact that an increased time in school translates into less time to invest in the search for a long-term, romantic partner plus less money with which to live alone. However, when employment status is explored the likelihoods of cohabitation or marriage appear similarly common amongst working men and working women. Yet, for working women, living alone is the most likely living arrangement while for men, marriage is most likely one, followed by living alone. The findings indicate that young Italian women may sense the incompatibility between work and marriage in Italy and/or perhaps a growing enjoyment of single life for both women and men with independent finances in Italy. Additionally, less educated young Italian men and women, or working men and women, engage in cohabitation and marriage with roughly equal likelihoods relative to living in the parental house, or relative to being unemployed. More so, being a housewife is the most likely employment status for both female cohabiters and females that are married. Such similarities and a lack of demographic contrasts in these associations speak to the idea that cohabitation and marriage are similar in character in Italy and the choice to cohabit may have far more to do with economic pressures as opposed to ideological reasons.

References

- AASSVE *et al.* 2002. Leaving Home: A Comparative Analysis of ECHP Data, *Journal of European Social Policy*, Vol. 12, No.4, pp. 259–275.
- BILLARI F., ONGARO F. 1998. The transition to adulthood in Italy. Evidence from cross sectional surveys, *Espace, populations, sociétés*, Vol. 2, pp. 165-179.
- CORIJIN M., MANTING D. 2000. The Choice of Living Arrangement After Leaving the Parental Home. In DE BEER, DEVEN (Eds.) *Diversity in Family Formation*, Vol. 8. Kluwer Academic Publishers, pp. 33-58.
- ISTAT 2009. *Le difficoltà nella transizione dei giovani allo stato adulto e le criticità nei percorsi di vita femminili*, Roma: Istat.
- MENCARINI L., SOLERA C. 2011. Changing Paths to Adulthood in Italy. Men and Women Entering Stable Work and Family Careers, *Carlo Alberto Notebooks*, No. 219.
- PUPO S., LICURSI R. 2003. Famiglia e dintorni. In FANTOZZI (Ed.) *Giovani in Calabria*, Soveria Mannelli: Rubettino, pp. 81-102.

SIRONI E, ROSINA A. 2012. The Problematic Transition to Adulthood in Italy: Comparison Before and After the Beginning of the Global Crisis. Conference paper: *The Transition to Adulthood After the Great Recession*, Milano.

SUMMARY

The aim of this paper is to assess the effects of individual socio-demographic features in determining the likelihood of Italian young adults (19-40) to leave their parental home and acquire their residential autonomy through living alone, cohabitation, or first marriage. The analysis uses the 2012 Istat data from the Multipurpose Household Survey.

Sara GRUBANOV-BOŠKOVIĆ, Università di Bari – Hebrew University of Jerusalem (The Avraham Harman Institute of Contemporary Jewry),
saragrubanov@gmail.com
Karra GREENBERG, University of California, Los Angeles (UCLA),
karra.greenberg@gmail.com

POVERTÀ E DEPRIVAZIONE IN ITALIA: UN'ANALISI MULTILIVELLO

Mario Mastrangelo

1. Introduzione

Le tecniche di analisi tradizionali della povertà si basano fondamentalmente su una rigida dicotomizzazione della popolazione in poveri e non poveri e sulla unidimensionalità della variabile adottata per tale dicotomizzazione, usualmente il reddito o la spesa per consumi. A partire dalla seconda metà del secolo scorso, a questa lettura se ne è affiancata un'altra, incentrata sulla natura multidimensionale della povertà, che contesta l'assunzione implicita di equivalenza tra risorse economiche disponibili e livello di benessere propria della prima: il reddito diventa solo una delle numerose dimensioni della povertà, tra le quali, ad esempio, la possibilità di conseguire standard adeguati in settori quali l'educazione e la salute, e l'attenzione è concentrata sulla qualità della vita (Atkinson 1987; Ringen 1988).

Diviene centrale il tema dell'esclusione sociale, ovvero il processo prodotto dall'accumulazione ed interazione tra loro di diversi fattori di rischio sociale e ambientale, che indichiamo come eventi di deprivazione, i quali spingono gli individui che vi sono esposti verso una condizione di povertà (Whelan e Maitre 2012a, 2012b, 2011). La conferma di questa crescente importanza giunge dalla inclusione degli indicatori di deprivazione nell'EU 2020 Poverty Target.

In questa nota, si mettono a confronto tre indicatori distinti per investigare il fenomeno della povertà, il primo di natura monetaria e gli altri due basati su forme di deprivazione, e, mediante il ricorso ad opportuni modelli multilivello, si adottano contemporaneamente un punto di vista micro, in base alle caratteristiche delle unità elementari di analisi (nel caso specifico le famiglie) e un punto di vista macro, in base alle caratteristiche delle unità di secondo livello, le regioni italiane.

2. Dati e metodo

L'analisi è stata svolta utilizzando i dati dell'indagine campionaria sulle famiglie "Reddito e condizioni di vita" del 2011, condotta dall'ISTAT nell'ambito del progetto EU-SILC (European Statistics on Income and Living Conditions). Il

campione italiano è composto da 47842 individui e 19399 famiglie. Si è scelto di utilizzare la famiglia come unità di analisi, dal momento che gli indicatori non monetari sono registrati a quel livello. Per gli indicatori monetari sono stati utilizzati i dati relativi all' Household Reference Person (HRP), ovvero l'individuo responsabile per l'abitazione (nel caso in cui più persone condividono questa responsabilità è stata scelta la più anziana). Questa scelta, che implica l'assunzione che tutti i membri della famiglia condividano lo stesso (eventuale) livello di deprivazione, è ampiamente utilizzata in letteratura (Guio 2009).

I primi due indicatori utilizzati sono già compresi nel data set EU-SILC (variabili *HX080* e *RX060*), il terzo è un indicatore composito appositamente costruito. In particolare, il primo indicatore (I1) assume valore 1 se la famiglia è a rischio di povertà, ovvero dispone di un reddito equivalente¹ al di sotto del 60% del reddito equivalente mediano nazionale, e 0 altrimenti. Come noto il rischio di povertà non misura la povertà, ma un reddito basso rispetto a quello degli altri residenti, che non implica necessariamente uno standard di vita basso.

Il secondo indicatore (I2) si basa sul concetto di deprivazione severa, e assume valore 1 se la famiglia presenta almeno quattro eventi di deprivazione (tra i seguenti 9: a) Non poter sostenere spese impreviste di 800 euro; b) Non potersi permettere una settimana di ferie all'anno lontano da casa; c) Avere arretrati per il mutuo, l'affitto, le bollette, etc; d) Non potersi permettere un pasto adeguato ogni due giorni; e) Non poter riscaldare adeguatamente l'abitazione; f) Non potersi permettere una lavatrice; g) Non potersi permettere un televisore a colori; h) Non potersi permettere un telefono; i) Non potersi permettere un'automobile.) e 0 altrimenti. In questo caso le deprivazioni vengono assunte di uguale importanza.

Il terzo indicatore (I3) è sempre basato su misure di deprivazione, ma è stato ideato assegnando loro diversa importanza in base alla rispettiva capacità di discriminazione: è stato in altre parole costruito un opportuno sistema di pesi in modo da privilegiare le forme di deprivazione più discriminanti, ovvero quelle meno frequenti (prevalence weighting approach). Gli item di EU-SILC relativi ad eventi di deprivazione sono stati sottoposti ad una preliminare analisi fattoriale esplorativa che ha portato all'individuazione di sei dimensioni principali (Deprivazione di base, Deprivazioni a livello di consumi, Deprivazioni sui consumi domestici, Deprivazioni sulle utenze, Deprivazione a livello di qualità del vicinato, Necessità di ricorrere all'aiuto di persone non coabitanti) che sono state rese attraverso la scelta di 23 indicatori semplici. Per verificare la coerenza interna della scala multi-item risultante è stato utilizzato il coefficiente alfa di Cronbach, il quale

¹ Si intende per reddito equivalente il reddito totale disponibile di una famiglia diviso per il numero dei suoi membri dopo averli equivalenti, ovvero opportunamente pesati in base all'età utilizzando la scala di equivalenza OECD modificata.

ha fornito una misura pari a 0,737, superiore alla soglia comunemente adottata pari a 0,7. L'indicatore I3 per ciascuna famiglia i è così definito:

$$I3_i = \sum_{k=1}^{23} w_k I_k \quad (1)$$

e il peso dell'indicatore semplice k è pari a:

$$w_k = 1 - f_k(1) \quad (2)$$

L'indicatore normalizzato varia tra 0 (l'unità di analisi non presenta alcuna forma di deprivazione) e 1 (l'unità di analisi le presenta tutte) e in questo modo permette non solo di stabilire se una famiglia è in uno stato di deprivazione o meno, ma di misurarne l'intensità (Barcéna-Martin *et al.* 2013). Sebbene il Regolamento EU-SILC richieda solamente la produzione di indicatori a livello nazionale, in Italia l'ISTAT fornisce stime affidabili anche a livello regionale. In questo modo, i dati presentano una struttura gerarchica che permette l'uso di modelli di regressione multilivello (logistica per i primi due indicatori e lineare per il terzo) aventi le famiglie come unità di primo livello e le regioni come unità di secondo livello, particolarmente interessanti perché consentono di valutare l'effetto esercitato dalle covariate di primo e di secondo livello sulla variabile dipendente e l'effetto residuo esercitato dalle singole regioni. Le variabili esplicative di primo livello sono il sesso (riferimento: uomini), il livello di istruzione (*istr_*, codificato in quattro livelli, riferimento: licenza elementare o nessun titolo di studio conseguito), la tipologia di famiglia (*fam_*, codificata in 8 modalità, riferimento: single), il titolo di possesso dell'abitazione (*tenure_*, codificata in 4 modalità, riferimento: proprietario) la condizione lavorativa (*cond_*, codificata in 7 modalità, riferimento: dipendente a tempo indeterminato), l'età dell'HRP (*age_*, codificata in 3 modalità, riferimento: fascia di età 30-64) e il reddito familiare equivalente (*Redeq*, escluso come covariata per l'indicatore I1). Le variabili esplicative di secondo livello sono il tasso di disoccupazione (*T_dis*), il tasso di irregolarità del lavoro (*T_irr*), il livello di istruzione della popolazione adulta (*T_istr*), la percentuale di popolazione di età superiore o uguale ai 65 anni (*P65+*), l'indice di Gini relativo ai redditi netti (*IGRed*) e la spesa pro capite per interventi e servizi sociali (*Sp_ss*).

3. Discussione

Per tutti e tre gli indicatori scelti come variabili dipendenti si registra un Coefficiente di Partizione della Varianza (VPC) significativo: 0,1062 per I1, 0,1588 per I2 e 0,0862 per I3. Dunque, con riferimento ad esempio all'indicatore

basato sul rischio di povertà, quasi l'11% della varianza complessiva è attribuibile a differenze a livello regionale. E' possibile pertanto applicare modelli multilivello. La tabella 1 mostra le stime dei parametri per i tre indicatori; è indicato con MA il modello contenente le sole variabili esplicative relative al primo livello, e con MB il modello completo contenente anche le variabili esplicative del secondo livello.

Tabella 1 – Multilevel random intercept model for basic deprivation: HRP and macro predictors.
(* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.)

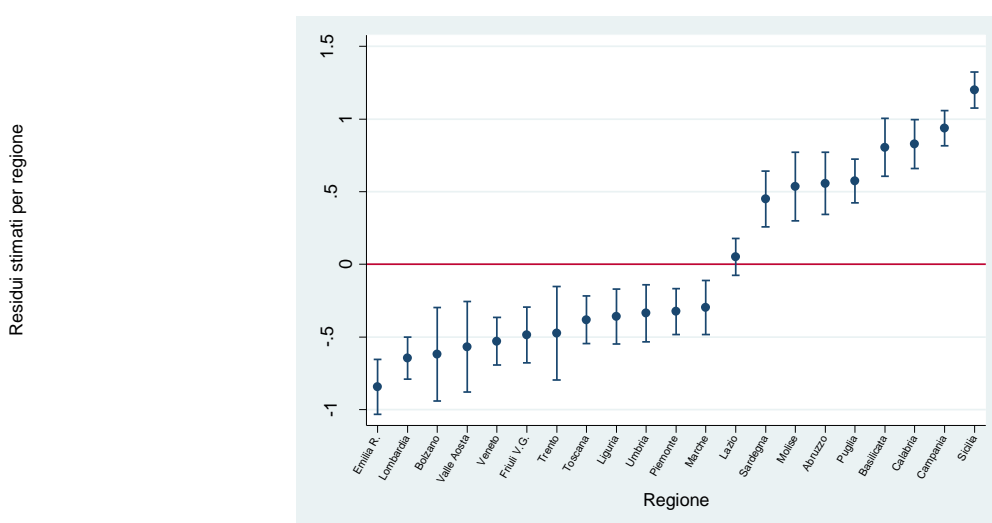
	I1 MA	I1 MB	I2 MA	I2 MB	I3 MA	I3 MB
Sex	0,0318	0,0305	0,1685*	0,1700**	0,0092***	0,0093***
Age_2	0,9637***	0,9687***	-0,1023	-0,1010	0,0075*	0,0075*
Age_3	-0,4354***	-0,4353***	-0,2114*	-0,2133*	-0,0077***	-0,0077***
Istr_2	-0,5472***	-0,5451***	-0,4436***	-0,4418***	-0,0151***	-0,0150***
Istr_3	-1,2791***	-1,2793***	-1,1309***	-1,1309***	-0,0423***	-0,0423***
Istr_4	-1,7959***	-1,7935***	-1,6513***	-1,6514***	-0,0559***	-0,0558***
Family_2	-0,9339***	-0,9397***	-0,4298***	-0,4347***	-0,0201***	-0,0201***
Family_3	-0,8044***	-0,8079***	-0,4471***	-0,4485***	-0,0161***	-0,0161***
Family_4	-1,2224***	-1,2308***	-0,1854*	-0,1920*	-0,0161***	-0,0161***
Family_5	0,8322***	0,8324***	-0,0823	-0,0800	0,0124***	0,0125***
Family_6	-0,3335***	-0,3397***	-0,4884***	-0,4956***	-0,0173***	-0,0174***
Family_7	0,1152	0,1110	-0,6039***	-0,6053***	-0,0197***	-0,0198***
Family_8	-0,0791	-0,0847	-0,2257	-0,2302	-0,0075*	-0,0076*
Cond_2	1,2851***	1,2865***	0,5044***	0,5046***	0,0292***	0,0293***
Cond_3	0,9965***	0,9964***	-0,2456*	-0,2484*	-0,0068***	-0,0068***
Cond_4	2,0202***	2,0148***	0,9968***	0,9946***	0,0714***	0,0713***
Cond_5	0,8165***	0,8172***	-0,0515	-0,0502	0,0005	0,0006
Cond_6	1,1115***	1,1063***	0,7397***	0,7493***	0,0494***	0,0494***
Cond_7	1,5617***	1,5590***	0,1904	0,1795	0,0179***	0,0177***
Tenure_2	0,4786***	0,4785***	0,6807***	0,6869***	0,0321***	0,0323***
Tenure_3	0,6170***	0,6115***	1,2521***	1,2397***	0,0639***	0,0639***
Tenure_4	0,6808***	0,6803***	1,1274***	1,1289***	0,0733***	0,0734***
Redeq			-6e-05***	-6e-05***	-6e-07***	-6e-07***
T_dis		0,1095***		0,0578		0,0040**
T_irr		0,0229*		-0,0033		-0,0013*
T_istr		-0,0148		0,0299*		0,0005
P65+		-0,0409*		0,1002***		0,0004
IGRed		-0,1209		8,1739*		0,1975
Sp_ss		-0,0014		-0,0051***		-0,0002***
Const	-1,5258***	-1,0430	-1,3978***	-7,0313***	0,1217***	0,0398
VPC	0,0949***	0,0059*	0,1258***	0,0085*	0,0844***	0,0121**

3.1. Indicatore I1

La stima dell'effetto casuale in assenza di covariate permette di valutare l'effetto complessivo delle unità di secondo livello. La figura 1 riporta il caterpillar plot del modello nullo, ovvero senza covariate di alcun livello (non riportato nella

tabella 1), che mostra i residui ordinati delle regioni, con i relativi intervalli di confidenza al 95%. Tali residui sono gli scostamenti di ciascuna regione dal log-odds medio generale, rappresentato dalla linea orizzontale a zero; le regioni per quali gli intervalli di confidenza non intersecano tale linea si discostano pertanto dal log-odds medio in maniera statisticamente significativa, per eccesso o per difetto.

Figura 1 – Caterpillar plot showing region residuals and 95% confidence intervals for II.



Il grafico mostra che questo avviene per tutte le regioni ad esclusione del Lazio: in particolare nella parte sinistra del grafico si trova un gruppo di regioni che presentano un log-odds inferiore a quello medio (comprendente tutte le regioni del Nord Italia e alcune del Centro), il contrario avviene nella parte destra (dove si trovano tutte le regioni del Sud Italia e alcune del Centro).

L'introduzione delle covariate di primo livello non cambia molto il VPC, che passa da 0,106 a 0,095. Limitandosi ai parametri significativi, l'essere a rischio di povertà è più probabile per le famiglie il cui HRP possiede un livello di istruzione basso, di età inferiore ai trent'anni e per le famiglie costituite da monogenitori con uno o più figli. Infine, tutte le condizioni lavorative diverse da quella di riferimento- in particolare, come è ovvio, lo stato di disoccupazione- accrescono la probabilità di essere a rischio di povertà così come lo stato di affittuario e mutuatario. Nel modello B l'introduzione delle covariate di secondo livello annulla il VPC; le uniche variabili significative sono il tasso di disoccupazione e la

percentuale di popolazione sopra i 65 anni, con effetti opposti in linea con le risultanze delle covariate di primo livello.

3.2. *Indicatore I2*

Lo studio delle stime dei residui del modello nullo mostra 3 gruppi di regioni: il primo, costituito dalla gran parte delle regioni del Nord Italia più Lazio e Umbria, che si discosta dal logg-odds medio per difetto; il secondo costituito da Toscana, Liguria, Piemonte, Molise e Sardegna, i cui intervalli di confidenza intersecano la linea del log-odds medio e infine il terzo gruppo, costituito dalle regioni del Sud Italia più Marche e Abruzzo, che se ne discostano per eccesso. Come nel caso precedente, l'introduzione delle covariate di primo livello non cambia molto il VPC, che passa da 0,158 a 0,126.

Guardando alle stime dei parametri, si possono osservare differenze rispetto all'indicatore I1. Infatti, la probabilità di essere soggetti a deprivazione severa è maggiore per le famiglie il cui HRP è single, donna, appartenente alla fascia di età centrale; rispetto all'indicatore I1 la condizione di lavoratore autonomo o libero professionista riduce tale probabilità rispetto alla condizione di riferimento, ovvero dipendente a tempo indeterminato. Il coefficiente del reddito equivalente ha ovviamente segno negativo, si conferma il ruolo incentivante del basso livello di istruzione e dello stato di affittuario e di mutuatario. Introducendo le covariate di secondo livello nel modello B il VPC diventa assolutamente trascurabile. Le variabili significative sono la percentuale di popolazione sopra i 65 anni e l'indice di Gini sui redditi netti, che presentano coefficienti positivi, e la spesa pro capite per interventi e servizi sociali, che presenta un coefficiente negativo.

3.3. *Indicatore I3*

Lo studio delle stime dei residui del modello nullo produce risultati molto simili a quelli dell'indicatore I2, evidenziando un primo gruppo di regioni costituito dalla gran parte delle regioni del Nord Italia più Umbria, Toscana ed Emilia Romagna, che presentano un valore medio di I3 inferiore a quello generale; un gruppo centrale costituito da Liguria, Molise e Sardegna (tutte e tre presenti nel gruppo centrale anche nel caso dell'indicatore I1) in cui la media non si discosta da quella generale e un gruppo di regioni a destra, costituito dalle regioni del Sud Italia più Lazio, Marche e Abruzzo, che se ne discostano per eccesso. L'introduzione delle covariate di primo livello non cambia molto il VPC, che passa da 0,086 a 0,084. I coefficienti stimati hanno segni e significatività molto simili a quelli stimati per l'indicatore I2, con alcune differenze. Per quanto riguarda l'età, I3 si comporta come I1: l'appartenenza alla fascia di età sotto i 30 accresce il valore

dell'indicatore rispetto alla fascia centrale, il contrario avviene per la fascia di età oltre i 65 anni. Come nel caso di I1, risulta positivo il coefficiente relativo alle famiglie costituite da monogenitori con uno o più figli.

Anche in questo caso l'introduzione delle covariate di secondo livello rende trascurabile il VPC; le variabili significative sono il tasso di disoccupazione, il tasso di irregolarità del lavoro e la spesa pro capite per interventi e servizi sociali.

Riferimenti bibliografici

- ATKINSON A.B. 1987. On the measurement of poverty, *Econometrica*, Vol. 55, No. 4, pp. 749-764.
- BARCENA-MARTIN E., LACOMBA B., MORO-EGIDO A.I., PEREZ MORENO S., 2013. Country Differences in Material Deprivation in Europe, *Review of Income and Wealth*, doi:10.1111/roiw.12030.
- GUIO, A.C., 2009. A European Union Approach to Material Deprivation using EU-SILC and Eurobarometer data, *IRISS Working Paper series*, Vol. 19.
- RINGEN S. 1988. Direct and indirect measures of poverty, *Journal of Social policy*, Vol. 17, No. 3, pp. 351-366.
- SNIJDERS T.A.B., BOSKER R.J., 1999. *Multilevel Analysis*. London: Sage.
- WHELAN C.T., MAITRE B., 2012. Understanding material deprivation: a comparative European analysis, *Research in Social Stratification and Mobility*, Vol. 30, pp. 489-503.
- WHELAN C.T., MAITRE B., 2012. Material Deprivation, Economic Stress and Reference Groups in Europe: An Analysis of EU-SILC 2009. Amsterdam: AIAS, Gini Discussion Paper 36.
- WHELAN C.T., MAITRE B., 2011. Measuring Material Deprivation with EU-SILC: Lessons from the Irish Survey, *European Societies*, Vol. 9, No. 2, pp. 147-173.

SUMMARY

Deprivation and poverty in Italy: a multilevel analysis

The aim of this paper is comparing three different indicators in order to analyze poverty; first indicator is a traditional monetary measure based on risk of poverty, and the other two, descending from the concept of multidimensional poverty, are based on material deprivation levels. We make use of the European Union Survey on Income and Living Conditions 2011, using Italian sample, which also guarantee reliable estimates at regional level. To take into account the resulting hierarchical structure of the data, we apply multilevel logistic and linear regression models, considering households as level 1 units clustered within regions, that are the level 2 units. In this way we adopt simultaneously a micro-level and a macro-level perspective.

Our results show relevant differences among the three indicators, and how region specific factors seem to be much more relevant than individual effects to explain region differences in material deprivation.

AN INTEGRATED ARCHIVE OF THE LIFESTYLES OF FAMILIES¹

Silvestro Montrone, Antonella Massari, Paola Perchinunno, Stefania Girone

1. Introduction

More and more often statistical data are collected with the preconceived objective of combining the information obtained by different investigations, whether sample surveys or censuses, both to extend the extent of knowledge and to guarantee a greater quality of data. The growing potential of computer science makes possible the acquisition, organization and creation of bigger and bigger data archives of ever greater quality.

The methodologies with a statistical basis utilized for the integration of data from a number of sources are: the techniques of Record Linkage (or exact matching) and of Statistical Matching (or synthetical Matching). The techniques of Record Linkage use specific algorithms that are adapted to identify pairs of records (related to the same statistical unit) in two different databases. The purpose of the methodology is therefore to integrate and match information which, though contained in different archives, can be attributed to the same statistical unit that is not identifiable by means of a single code that does not contain errors. The objective of the techniques of Statistical Matching is to identify records relating to similar units and to estimate the unified distribution of a number of variables observed in different data archives and to merge informative records. Both matching techniques make it possible to obtain integrated archives, adopting reasonable and statistically verifiable assumptions.

The present work will describe a model of data integration through a methodology of Statistical Matching (hot deck distance) for the integration of two surveys (EuSilc-Istat and Lifestyle Survey-University of Bari). The construction of an integrated database on the basis of these two surveys may be useful for the study of consumer behavior in relation to specific groups of commodities, in order to analyze the decisions taken by families with regard to saving, to examine economic and social inequality, and to study the impact of public policies by means of

¹ This contribution is the result of joint reflections by the authors, with the following contributions attributed to: S. Montrone (chapter 1,5), to A. Massari (chapter 4), to P. Perchinunno (chapter 2), to S. Girone (chapter 3).

simulations. The coexistence of multiple and differentiated objectives triggers the need to obtain a very general and versatile integrated file, which provides ongoing detailed information on the different types of spending, on levels of saving, on the distribution of incomes, on the occupational conditions of the members of the family unit, etc.

2. Techniques in data integration

2.1 Introduction

The statistical matching techniques are being used in order to link information that come from two or more datasets whose units contain similarities regarding a set of variables previously defined. The basic assumption is that the sources which should be integrated contain information on a series of common variables as well as information on different variables that are never been jointly observed and it would be interesting to relate them to each other.

The idea of using the integration techniques on data that come from multiple sources in order to combine and enrich the information gathered in several surveys is not a recent one: the first methodological examples date back to the '60s. The first important applied examples are in Okner (1972 e 1974) and Ruggles N. N. and Ruggles R. (1974). In recent years, there has been a renewed interest for these techniques after a broader recognition of their application potential, in particular in the area of national statistical institutes' activities (Schoier, Torelli, Zacchigna, Egidi, Sabbadini, 2006).

The methodologies, with statistical basis, that are used to integrate the data deriving from multiple sources can be grouped in two types:

1. *record linkage* or exact matching;
2. *statistical matching* or synthetic matching.

The techniques of exact matching (*record linkage*) aim at identifying the pairs of records, related to the *same* statistical unit, belonging to different datasets by using specific algorithms.

The techniques of statistical matching aim at identifying the *similar units*, estimating the joint distribution of multiple observed variables in different datasets and fusing the information records. Both matching techniques, the exact and the statistical one, provide integrated archives by adopting rational and statistically controllable assumptions (D'Orazio, Di Zio, Scanu, 2002).

2.2 The statistical matching

In the case of statistical matching the initial assumption is to have *two different*

archives that contain information records on two groups of units selected from the same population. Therefore, the starting point of the issue is the existence of two archives:

1. containing information on common variables (the socio-demographic type of variables) as well as on different variables that are never been jointly observed;
2. the sources' units to be integrated are separated, or rather the sample surveys are independent among them (Montrone, Perchinunno, 2005).

Hence, in the statistical matching the individual records deriving from two or more sources are linked, on the basis of their *similarities*, through a set of characteristics measured in each source. We consider here two datasets containing two files: the file A and the file B. In order to make the statistical matching of these files, it is necessary that the common information regarding the units is available in each file. Let X_A be the set of variables measured in the file A, and let X_B be the set of variables measured in the file B, it is assumed that these two sets of variables can be transformed in one set with common characteristics. We can indicate the individual's characteristics measured in the both datasets as the vector $\mathbf{X} = (X_I, \dots, X_P)$. The remaining variables in each file, that are not overlapping, are indicated as $\mathbf{Y} = (Y_I, \dots, Y_Q)$ in the file A and as $\mathbf{Z} = (Z_I, \dots, Z_R)$ in the file B.

The common variables X , defined from now on as the *integration variables*, are used to identify the units to be linked, while the non-common variables (Y e Z), defined as the *descriptive variables*, represent the information that is the object of the matching procedures. Therefore, it determines a situation in which the information that is simultaneously gathered on the same units for Y and Z is missing.

The aim of the statistical matching is to create a file, the *file C (the synthetic file)* in which each record contains all variables X , Y and Z . For each unit in the file A, a similar unit in the file B is identified, whereas the similarity is evaluated in terms of a function of variables X . The variables Z in the file B (defined as *the donor*) are then attributed to the matching record in the file A, creating thus a *record with full data* (X, Y, Z).

Considering the two archives A and B (which in this study case are the archives of family consumption and income), we will match the common variables in order to link the descriptive variables.

Some of the methods used for statistical matching are the following:

- *hot deck random*;
- *hot deck rank*;
- *hot deck distance*.

The method applied in this work is the *hot deck distance* which consists of matching the record a from A with the record b^* from B that is the “closest”

considering the variety of common variables. Hence, we will have that:

$$d_{a,b^*} = |x_a^A - x_{b^*}^B| = \min_{1 \leq b \leq n_B} |x_a^A - x_b^B| \quad (1)$$

Each unit of the base dataset (i.e. archive defined as the receiving one) is operatively associated to another unit of the second dataset (i.e. the archive defined as the donor) by applying the *distance function* $d(x_i, x_j)$ which is computed on the integration variables and assumes much lower values as the individuals are more similar among themselves (*nearest neighbor match*).

3. Construction of the integrated Eu-Silc and Lifestyles archive

3.1 Data sources

This report evaluates the possibility of utilizing two different sample sources to construct an integrated database with information relating to the economic situation of families. The two surveys used are *Eu-Silc* (ISTAT) and *Lifestyles* (University of Bari "A. Moro").

The *Eu-Silc survey* conducted by Istat collects detailed information on incomes and on family expenditure for the purchase of goods and services for final consumption, on family typology (in terms of composition and characteristics of the family head), on living conditions and spending habits, and relevant information about individual incomes, on the employment of savings and on family wealth.

On the other hand, the *Lifestyles survey* collects information on income, on spending behavior, and on the recourse to loans by families with children.

It is therefore interesting to verify the possibility of creating integration between the two archives, though always bearing in mind that the available data in both cases concern different samples from the same population: in addition, neither of the surveys covers a sample such as to allow the construction of a single database containing information relating to both income and consumption.

3.2 Harmonization of information from two archives

The first phase that is necessary to the accomplishment of statistical matching is the so called harmonization phase of the data from the two archives.

This is an essential process to verify the real comparability of the surveys.

The separate steps involved in this harmonization process can be summarized as follows:

- Analysis of the input sources. The input sources from the separate archives are analyzed in terms of the principle data acquired completeness and validity.
- Selection of stratification variables that characterize the families (number of family members, working condition).
- Selection of common variables (sex, age groups, marital status, tax-and-bills-payment difficulties, school and clothing subsidies, possession of dishwashers, washing machines and refrigerator).
- Selection of integration variables from the Eu-Silc (ability to make ends meet each month; possibility to spend a one-week vacation out of home; saving ability and loan application) or from the Lifestyle (perception on the progress of the family's economic situation; potentially required monthly income to live without economic problems).
- Harmonization of the common variables. The selected common variables are standardized in terms of classification. The codes of one of the two sources are then inverted.

At the end of this procedure the first part of the two harmonized archives will contain the information that is common to both, and the second part the data of each source.

3.3 Integrated archive results

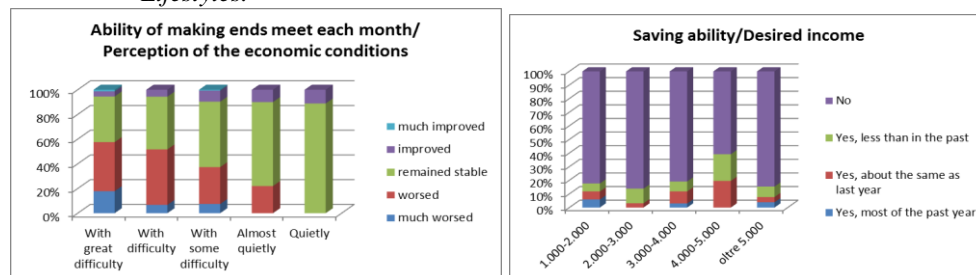
The study case concerns two simulations where the definition of the threshold value for the *distance function*, in the first case is set to be very restrictive, equal to zero (i.e. considering 5 matching variables only the matches that had a distance equal to zero were selected) while in the second case the value is set to be equal to 0,1 (i.e. considering 10 matching variables only those matches with a distance less or equal to one were selected).

The final result of elaborations was the integrated archive that derives from a combination of two surveys and contains *311 records in the first case and 365 matched records in the second one*.

The cross examination of results between the integration variables, that are the variables observed only in the Eu-Silc or only in the Life Styles, appears to be particularly interesting. In particular, comparing the data related to the "perception of the economic conditions" from the Life Styles with the one related to the "ability of making ends meet each month" from the Eu-Silc, it follows that there is a certain percentage of families which, although it makes ends meet each month without difficulties, perceives that its economic condition is deteriorated over the past year.

In addition, comparing the data on the "saving ability" Eu-Silc with the "desired income" Life Style, it follows that the saving ability is common only to those families that have a desired income above 4.000 euro per month.

Figure 1 – Results from the combination of matching variables between the Eu-Silc and Lifestyles.



4. Construction of hardship profiles

4.1 The cluster analysis

The next step in the integrated archive analysis was to consider a clustering procedure that will identify some poverty profiles, not previously defined, to which each family with socio-economic attitudes derived from the matching of two archives can be assigned. The cluster analysis proves to be very convenient since it provides “quite different” clusters (i.e. heterogeneous) among themselves where each one of them is composed of units (families) with a high grade of “natural association”

The variety of approaches to the cluster analysis have though a common necessity of defining one dissimilarity or distance matrix between the n observation couples, that represents a point from which every algorithm is being generated. The cluster analysis technique chosen is the one defined as the Two-Step. It is an extension of distance measures used by Banfield and Raftery (1993) based on the model and introduced for data with continuous attributes. The Two-Step algorithm has two advantages: it deals with mixed type of variables and automatically determines the optimal number of clusters, although it allows us to set a desired number of clusters.

The Two-Step procedure, very efficient for large datasets, is a scalar cluster analysis algorithm and is able to treat simultaneously continuous and categorical variables or attributes. It is being solved in two steps: in the first step, defined pre-clustering, the records are pre-classified in many small sub-cluster; in the second step the sub-clusters (generated in the first step) are grouped in the number of clusters that optimizes the BIC (Bayesian Information Criterion) defined as:

$$BIC_K = -2l_k + r_k \log n \quad (2)$$

where r_k is the number of independent parameters and:

$$l_k = \sum_{v=1}^k \xi_v \quad (3)$$

is the function of *log-likelihood*, for the step with k clusters, which can be interpreted as the dispersion within the cluster. It also represents the entropy within the k clusters in the case in which only the categorical variables are considered.

4.2 Identified profiles

The cluster analysis allowed us to identify several profiles of families derived from the integrated archive Eu-Silc with the Life Style.

By running the cluster analysis on the integrated archive, 5 clusters to which different profiles of surveyed families are associated have emerged. The most important variable in the definition of the profiles resulted to be the perception of the economic situation progress. In particular:

- Cluster 1: families with more than 5 members, which claim that their situation has remained stable over the time or has improved, have no kind of debts nor difficulties paying taxes, bills, school subsidies;
- Cluster 2: families that perceive a deterioration of their economic situation over the past year, have difficulties paying taxes, bills, school subsidies and have other kinds of debts;
- Cluster 3: families with less than 5 members, which claim that their situation has remained stable over the time or has improved, have no kind of debts nor difficulties paying taxes, bills, school subsidies;
- Cluster 4: families that perceive a deterioration of their economic situation over the past year, but have no kind of debts nor difficulties paying taxes, bills, school subsidies;
- Cluster 5: families that perceive a deterioration of their economic situation over the past year, have difficulties paying taxes, bills, school subsidies, but do not have other kinds of debts.

5. Concluding remarks

Through our analysis we have attempted to quantify the influence of income and of family typology (number of members) in order to understand how the lifestyles of the families evolve. The estimates of the risk of poverty based on “objective” indicators, such as income or recourse to going into debt, are completely independent of the state of awareness of those directly involved. It is, however, also useful to observe the “subjective” perception of Italian people in relation to their standard of living and to the recurring causes of economic and social hardship.

It is hoped that the variations regarding the new family profiles that emerge in general from analyses carried out with different criteria provide important suggestions not only for describing and understanding the phenomenon of economic hardship better, but also to obtain indications for the social policies to contrast poverty.

References

- BANFIELD J.D. RAFTERY A.E. 1993. Model-based Gaussian and non-Gaussian clustering. *Biometrics*, 49, p. 803-821.
- OKNER B.A. (1972), "Constructing a new data base from existing micro data sets: the 1966 merge file", *Annals of Economic and Social Measurement*, 1.
- D'ORAZIO M., DI ZIO M., SCANU M. 2002. *Statistical Matching and Official Statistics*, Quaderni di Ricerca ISTAT, 1.
- MONTRONE S., PERCHINUNNO P. 2005. La stima della povertà basata su modelli, *Annali del Dipartimento di Scienze Statistiche, Facoltà di Economia, Università degli Studi di Bari*, n. 4.
- RUGGLES N. N., RUGGLES R., 1974. A strategy for merging and matching micro data sets, *Annals of Economic and Social Measurement*, 3.
- SCHOIER G., TORELLI N., ZACCHIGNA A., EGIDI V., SABBADINI L.L. 2006. L'abbinamento statistico di dati dal sistema di indagini multiscopo: prime proposte e evidenze empiriche, in "Metodi statistici per l'integrazione di dati da fonti diverse", a cura di Liseo, Montanari, Torelli, Ed. Franco Angeli.

SUMMARY

The present work will describe a model of data integration through a methodology of Statistical Matching (hot deck distance) for the integration of two surveys (EuSilc-Istat and Lifestyle Survey-University of Bari). The construction of an integrated database on the basis of these two surveys may be useful for the study of consumer behavior in relation to specific groups of commodities, in order to analyze the decisions taken by families with regard to saving, to examine economic and social inequality, and to study the impact of public policies by means of simulations.

Silvestro MONTRONE, Università degli Studi di Bari silvestro.montrone@uniba.it

Antonella MASSARI, Università degli Studi di Bari antonella.massari@uniba.it

Paola PERCHINUNNO, Università degli Studi di Bari

paola.perchinunno@uniba.it

Stefania GIRONE, Università degli Studi di Bari stefaniagirone@libero.it

A PROPOSAL FOR A SEMIPARAMETRIC CLASSIFICATION METHOD WITH PRIOR INFORMATION

Luciano Nieddu, Cecilia Vitiello

1. Introduction

The aim of discriminant analysis is to determine a function (classifier) that, on the basis of a set of covariates \underline{x} , best predicts a categorical variable y labeling the class c a unit belongs to.

If the data at hand have been previously classified by an expert the problem is known as supervised classification (Watanabe, 1985) and can be further classified as classification with perfect supervisor and with imperfect supervisor (Katre and Krishnan, 1989). Otherwise the problem is referred to as unsupervised classification.

A plethora of methods have been suggested in order to determine the best classifier. According to Wolpert (1996), if the interest is on the generalization performance of a classifier without making any prior assumptions on the data, then no classification algorithm is inherently superior to any other or even to random guessing.

A general taxonomy of classification rules divides the methods into parametric and non-parametric. In between those two categories the semi-parametric approach via finite mixture models is widely used in unsupervised classification but has found its way to supervised classification (Hastie and Tibshirani, 1996). The method we suggest follows the same approach but differs in some relevant aspects, namely: when finite mixtures of multinormal distribution are involved, no constraints are imposed on the covariance structure of each component (resulting in a more flexible method). Moreover we have adopted a weighted likelihood approach where weights express the information given by the expert. This can easily be extended to handle classification with imperfect supervisor.

The outline of the paper is as follows: in Section 2 the proposed method will be introduced and the parameter estimators will be derived. In Section 3 the experimental results will be presented and in Section 4 conclusions will be drawn.

2. Proposal

Let $\underline{y} = \{y_1 \dots y_G\}$ be an observable multinomial random variable with $y_{ig}=1$ iff unit i ($i=1, \dots, n$) belongs to group g . We introduce an unobservable multinomial random variable $\underline{z} = \{z_1 \dots z_K\}$ and model $f(\underline{x}_i | y_{ig} = 1)$ as a K components finite mixture of multinormal $\phi(\underline{x}_i; \underline{\mu}_{gk}, \Sigma_{gk})$ with masses $p(z_{ik} = 1), k = 1, \dots, K$.

$$f(\underline{x}_i | y_{ig} = 1) = \sum_k^K p(z_{ik} = 1) \phi(\underline{x}_i; \underline{\mu}_{gk}, \Sigma_{gk}).$$

We then derive the likelihood in a weighted form as follows:

$$L(\underline{\theta}, \underline{x}, \underline{y}) = \prod_i^n \sum_k^K \sum_g^G y_{ig} p(z_{ik} = 1) p(y_{ig} = 1 | z_{ik} = 1) \phi(\underline{x}_i; \underline{\mu}_{gk}, \Sigma_{gk})$$

In order to simplify notation let's write

- $(\underline{\mu}_{kg}, \Sigma_{kg}) = \theta_{kg}$
- $\phi(\underline{x}_i, \underline{\mu}_{kg}, \Sigma_{kg}) = f_i(\theta_{kg})$
- $p(y_{ig} = 1 | z_{ik} = 1) = \pi_{g|k}$
- $p(z_{ik} = 1) = \pi_k$

Maximum likelihood estimates (MLE) are obtained iteratively and involve three blocks of parameters.

The first refers to the location and scale parameter θ_{kg} , for fixed π_k and $\pi_{g|k}$ which can be obtained as standard MLE solving the following:

$$\begin{cases} \sum_i^n w_{igk} \frac{\delta \log f_i(\theta_{kg})}{\delta \theta_{kg}} = 0 & g = 1, \dots, G; k = 1, \dots, K \\ w_{igk} = \frac{y_{ig} \pi_{g|k} \pi_k f_i(\theta_{kg})}{\sum_k \sum_g y_{ig} \pi_{g|k} \pi_k f_i(\theta_{kg})}, & g = 1, \dots, G; k = 1, \dots, K; i = 1, \dots, n \end{cases}$$

The other two blocks refer respectively to probabilities π_k and $\pi_{g|k}$ and are both constrained problems.

$$\begin{cases} \sum_i^n \frac{y_{ig} \pi_{g|k} f_i(\theta_{kg})}{\sum_k \sum_g y_{ig} \pi_{g|k} \pi_k f_i(\theta_{kg})} + \lambda = 0, & k = 1, \dots, K \\ \sum_k \pi_k = 1, \end{cases}$$

$$\hat{\pi}_k = \sum_i^n \frac{1}{n} \frac{\sum_g y_{ig} \pi_{g|k} \pi_k f_i(\theta_{kg})}{\sum_k \sum_g y_{ig} \pi_{g|k} \pi_k f_i(\theta_{kg})} \quad k = 1, \dots, K$$

And finally, for each $k=1, \dots, K$, further constrained MLE with $\sum_g \pi_{k|g} = 1$.

$$\left\{ \begin{array}{l} \sum_i^n \frac{y_{ig} \pi_k f_i(\theta_{kg})}{\sum_k \sum_g y_{ig} \pi_{g|k} \pi_k f_i(\theta_{kg})} + \lambda^k = 0 \quad g = 1, \dots, G; \\ \sum_g \pi_{k|g} = 1 \end{array} \right.$$

$$\hat{\pi}_{g|k} = \sum_i^n \frac{1}{nk} \frac{y_{ig} \pi_{g|k} \pi_k f_i(\theta_{kg})}{\sum_k \sum_g y_{ig} \pi_{g|k} \pi_k f_i(\theta_{kg})}$$

which are straightforward extensions of standard mixture model results.

3. Experimental Results

In this section the results of a simulation study will be presented to evaluate the performance of the proposed method when compared with well-known parametric and non-parametric methods. Then the application to a benchmark dataset from the UCI¹ repository will be shown.

3.1. Simulation study

At this stage, to easily visualize the results, only two class data in \mathbb{R}^2 have been considered, generating points around centroids whose locations have been randomly selected on planar curves.

Varying number of centroids have been used. For each centroid $n \in \{20; 30; 50\}$ points have been generated adding MVN noise with $\Sigma = \begin{bmatrix} 2 & 0.8 \\ 0.8 & 2 \end{bmatrix}$. Three types of curves have been used: lines, parabolas and cubic curves. For each class all the three configurations have been used, yielding 9 possible configurations.

Table 1 displays the choices for the parameters of the curves bearing the centroids and the range for the x variable for class 0 and class 1. The corresponding values for the y s of the centroids have been computed from the equations of each curve. The ranges for the x variable for the two classes have been chosen as only partially overlapping to guarantee a discriminant power for the variable itself.

¹ <http://archive.ics.uci.edu/ml>

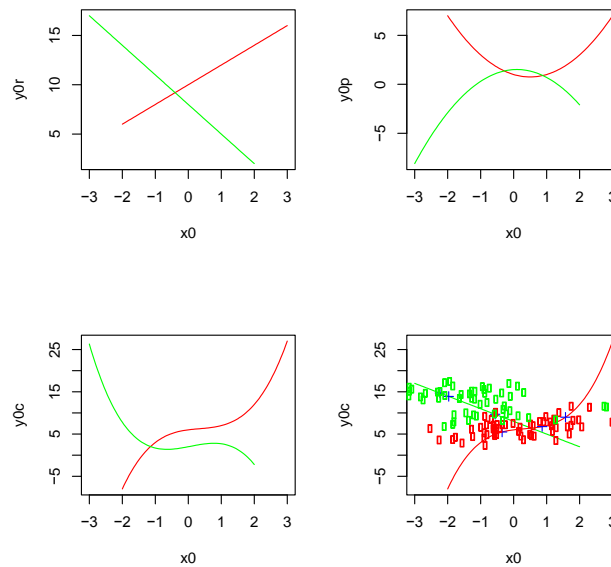
Table 1 – *Simulation setup*

	Class 0	Class 1
Range for x	$[-2; +3]$	$[-3; +2]$
Line	$y = 2x + 10$	$y = -3x + 8$
Parabola	$y = x^2 - x + 1$	$y = x^2 + 0.2x + 1.5$
Cubic	$y = x^3 - x^2 + x + 6$	$y = -x^3 + 0.2x^2 + 1.5x + 2$

In Figure 1 some configurations of curves for class 0 and class 1 and an example of a dataset have been displayed.

The proposed method has been compared with two well-known parametric methods, namely linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA), and two non-parametric methods such as k -nearest neighbor ($k=1$) (1-NN) and Classification trees (TREE) (Hastie and Tibshirani, 2003).

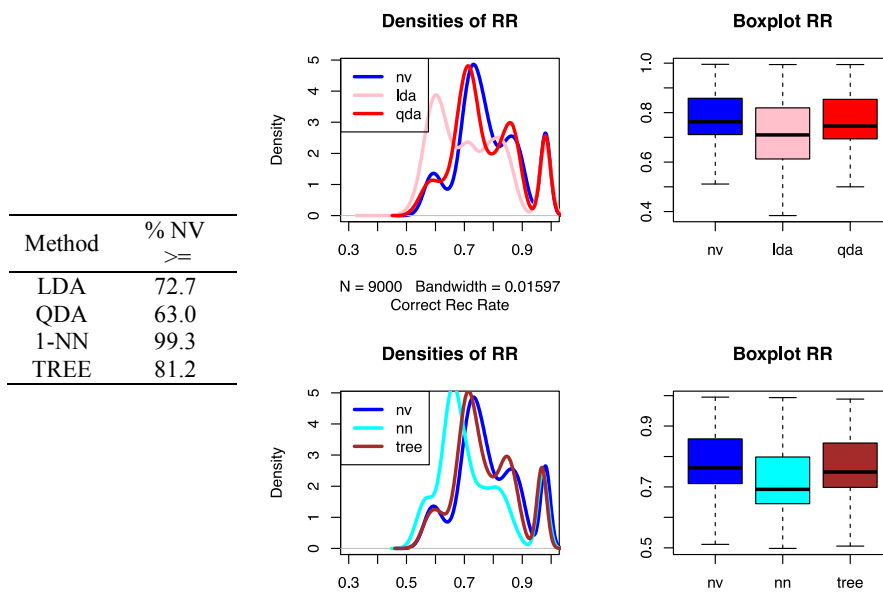
Figure 1 – *Curves bearing the centroids for each class: 3 of the possible 9 configurations. In the lower right corner an example of dataset with cubic and line*



For each configuration of the curves bearing the centroids 100 trials have been carried out and, in each trial, the assessment of the performances of the various techniques have been determined using leave one out cross validation. The simulation has been carried out with the R package using the functions `rpart()`, `lda()`, `qda()` and `knn.cv()` for classification trees, linear discriminant analysis, quadratic discriminant analysis and 1-nearest neighbor respectively.

In a first phase of the simulation study, the number of centroids (nm) and the number of points (n) for each centroid have been allowed to vary. For each configuration of nm and n we have computed the number of times that the proposed method (NV in the following) has bested the others or reached an equal performance. In Figure 2 the total percentages of times that the proposed method has at least performed as well as the others have been displayed (table on the left part of the figure). On the right hand side, the density estimates of the correct recognition rates (RR) for the proposed method and the other classification algorithms, over the whole simulation, have been displayed together with the boxplots of the distributions of the correct recognition rates.

Figure 2 – Distribution of the percentage of times that the proposed methods performed at least as best as the others (left) and distributions of the correct recognition rates (RR) (densities and boxplots)(right).



The top part of the figure show the comparison with parametric techniques: LDA seems to be the method with the worst average performance. The performance of NV and QDA are comparable but the distribution of the proposed method is shifted to the right hand side, implying a better average performance (as also shown by the boxplots). This was to be expected since quadratic discriminant analysis has a better adaptive capacity to the data than linear discriminant analysis, which assumes equal covariance matrices for all the classes. The performance of

the proposed method, when compared with non-parametric methods, is displayed on the bottom part of Figure 2. NV works much better than classification trees and almost always bests or performs equally as well as 1-NN (99.3% of cases). This is clearly shown by the density estimate of the 1-NN recognition rate which is the leftmost density in the picture, immediately followed by the classification tree density estimate. Therefore on average the proposed method works at least as well as the other classification methods that have been considered in the simulation. It must be stressed that this results is an overall result over all the 9 configurations of bearing-centroid curves and over various number of centroids ($nm=10,25,35$) and points per centroid ($n=20,30,50$).

To determine if there is an effect of the varying number of centroids, a simulation with $n=30$ points per centroids and varying number of centroids (nm) has been undertaken. Once again the number of times that the proposed method has performed better or as well as the other methods has been considered. In Table 2 such percentages have been displayed together with the number of centroids for each classification algorithm:

Table 2 – *Percentage of the times the proposed method has performed at least as well as the other classification methods over varying number of centroids.*

Method	Number of centroids											
	2	4	6	8	10	12	14	16	18	20	25	35
LDA	83.0	83.4	81.9	81.0	79.0	77.6	76.4	78.9	75.7	75.5	74.4	72.6
QDA	80.4	79.6	74.2	71.8	71.0	69.9	68.2	66.8	67.5	66.5	64.9	62.9
1-NN	93.3	96.8	97.9	98.4	98.2	98.5	98.6	99.4	99.2	99.2	99.4	99.5
TREE	87.1	92.1	87.1	86.2	84.2	81.6	80.3	83.4	80.8	80.8	80.7	81.2

There seems to be a tendency of the proposed method to work at least as well as 1-nn when the number of centroids increases. Such a tendency is not shared by the others methods although the percentage of times the proposed method works as well as classification trees seems to stabilize around 80%. Such percentage seems to show a decreasing trend with the increasing number of centroids for LDA and QDA. This can be ascribed to the fact that with an increasing number of centroids with fixed number of points per centroid and considering the narrow range of the values that the x can assume, the scatter plots for the two classes seem to show a multinormal distribution which is the case where LDA and QDA work best.

3.2. Real data application

In order to test the performance of the proposed method on real data and to compare it to some widely used classification methods, a credit scoring benchmark dataset (Australian credit scoring) has been used (Murphy and Aha, 2001). The

data is composed of 690 records: 307 instances are creditworthy applicants and 383 instances are from not creditworthy applicants. Each instance contains 6 nominal, 8 numeric attributes, and 1 class attribute (accepted or rejected). To protect the confidentiality of data, the attributes names and values have been changed to meaningless symbolic data.

Considering the multinormality hypothesis of the proposed method only the 3 continuous variables out of the 6 numerical have been considered in the study. The performance of the proposed method has been once again compared to nearest-neighbor classifier (1-NN), linear and quadratic discriminant analysis (LDA, QDA) and classification trees (TREE). Leave one out has been used to assess the correct recognition rate. In Table 3 the correct recognition rates have been displayed

Table 3 – *Correct Recognition Rate for the Australian Credit Scoring data: comparing the new proposal to standard methods*

Method	Correct Recognition Rate
Proposed method (NV)	0.716
LDA	0.658
QDA	0.641
1-NN	0.601
TREE	0.674

The proposed method shows a correct recognition rate of 71.6% based only on 3 continuous attributes. The other methods show a recognition rate which is clearly lower, with 1-NN having the worst performance.

4. Conclusions

In this paper an adaptive classification method based on finite mixtures and a-priori information has been presented. The proposed methodology has been tested on a simulation study and on a real benchmark dataset. The proposed approach more than holds its own when compared with parametric and non-parametric well established methods. The proposed methodology assumes the structure of subclasses on the data at hand and the presence of prior information on the classification of the data. The method has been tested in the case of crisp prior information but can be easily extended to the case where there is fuzzy information on the prior classification of the data. Further studies are necessary to test the actual performance on the method also w.r.t. sensitivity and specificity.

Acknowledgements

The authors would like to thank Marco Alfò for the time spent with them discussing alternatives to obvious solutions.

References

- WATANABE S. 1985. *Pattern recognition: human and mechanical*, New York: John Wiley & Sons, ISBN 0-471-80815-6
- KATRE U. A., KRISHNAN T. 1989. Pattern recognition with an imperfect supervisor. *Pattern Recognition*. 22, 4 (August 1989), 423-431.
- MURPHY P.M., AHA D.W. 2001. UCI repository of machine learning databases. Department of Information and Computer Science, University of California Irvine, CA. Available from <http://www.ics.uci.edu/mllearn/MLRepository.html>
- QUINLAN J.R. 1986. Induction of decision trees. *Machine Learning*, 1(1), 81–106.
- WOLPERT D.H. 1996. The Lack of A Priori Distinctions between Learning Algorithms, *Neural Computation*, 8, 1341-1390.
- HASTIE T., TIBSHIRANI R. 1996. Discriminant analysis by Gaussian mixtures, *Journal of the Royal Statistical Society series B*, Vol. 58 pp. 158-176
- HASTIE T., TIBSHIRANI R., FRIEDMAN J.H. 2003. *The Elements of Statistical Learning*. Springer, corrected edition.

SUMMARY

Classification methods are usually grouped into three main categories, ranging from unsupervised classification to supervised classification, passing through classification with imperfect supervisor. Our proposal tries to span a bridge between these two banks. The method proposed has been tested on a simulation study yielding very interesting results.

Luciano NIEDDU, Università degli Studi Internazionali di Roma LUSPIO,
l.nieddu@unint.eu

Cecilia VITIELLO, “Sapienza” - Università di Roma - Dipartimento di Statistica,
cecilia.vitiello@uniroma1.it

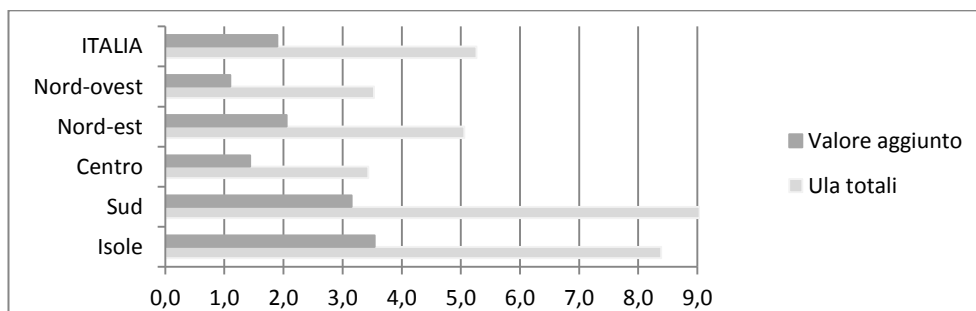
RECENTI TRASFORMAZIONI DEL LAVORO NEL SETTORE AGRICOLO

Paola Naddeo

1. Premessa

L'agricoltura storicamente ha rappresentato per l'Italia un serbatoio di forza lavoro, poco qualificata, a cui attingere per promuovere lo sviluppo di altri settori produttivi. Come testimoniano numerosi studi, nel settore agricolo prevalgono le micro-imprese di tipo familiare, con un ridotto ricorso a manodopera esterna, spesso utilizzata con forme di lavoro cosiddetto atipico, se non irregolare: poche giornate all'anno, con una distribuzione degli orari di lavoro che riflette le caratteristiche peculiari del settore (Idda e Pulina, 2011; Isfol, 2012). Solo negli anni più recenti sembrerebbe che qualche cosa possa essere cambiato.

Figura 1 - Valore aggiunto a prezzi di mercato e Ula totali nel 2010 (quota del settore agricolo su totale economia)



Fonte: elaborazioni su dati Istat, Conti economici regionali

La crisi economica e lo sviluppo di nuovi servizi (Giovannini, 2012) ha permesso, specie nel Nord, l'affermarsi di situazioni in cui l'organizzazione agricola ha assunto le caratteristiche tipiche di produzione capitalistica; al contempo, specie in diverse aree del Sud, continuano a registrarsi forme di lavoro "arcaiche", con il ricorso al caporalato, reso possibile anche da un eccesso di offerta rappresentato dal lavoro irregolare immigrato (Pugliese, 2012).

I dati del Censimento dell'agricoltura rappresentano dunque un osservatorio privilegiato per conoscere le modifiche che si sono prodotte nel settore agricolo nell'ultimo decennio, ma presentano l'inconveniente di non poter fornire indicazioni di cosa è successo nel corso dei diversi anni del decennio.

L'agricoltura rappresenta meno del 2% del valore aggiunto nazionale, con forti differenze per aree territoriali: si passa da poco più dell'1% nell'Italia nord-occidentale a oltre il 3,5% delle Isole (Figura 1). In termini di unità di lavoro totali le quote risultano sensibilmente più elevate, come riflesso della bassa produttività che caratterizza il settore agricolo, con oscillazioni tra il 3,5% circa dell'Italia nord-occidentale e il 9% circa del Sud, con una media nazionale pari a circa il 5%.

2. Obiettivo del lavoro e metodologia

Nel presente lavoro si analizza come le recenti modifiche nella domanda di servizi attinenti il settore agricolo, quali ad esempio l'agriturismo, le produzioni biologiche, quelle DOP (Denominazione di Origine Protetta) e IGP (Indicazione Geografica Protetta), abbiano influenzato l'occupazione del settore, con particolare riferimento ai giovani (intesi come coloro che hanno un'età inferiore a 30 anni¹) e alle donne. La fonte dei dati utilizzata è rappresentata dal Censimento dell'agricoltura nel 2010.

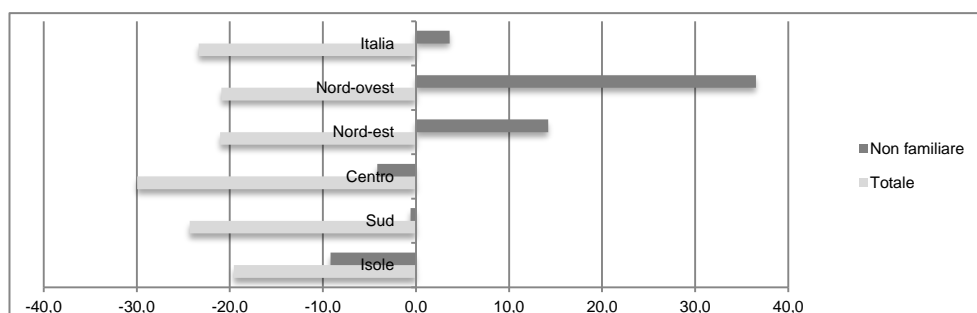
Dopo un'analisi descrittiva delle principali caratteristiche dell'occupazione agricola per ripartizioni territoriali, si procede ad un'analisi di cluster a livello provinciale, al fine di verificare come le diverse aree territoriali abbiano finora reagito alle nuove opportunità offerte dal settore. Questo tipo di analisi è spesso utilizzata, come nel caso presente, quando si vogliono effettuare analisi a livello territoriale basate su tassonomie. La cluster analysis è, infatti, una delle tecniche di analisi multivariata utilizzata per fare classificazioni; si presenta più facile da un punto di vista matematico e non richiede la presenza di un modello a monte o la verifica di ipotesi statistiche. Non per questo, però, la cluster analysis può essere considerata un metodo inferiore ad altri (Romesburg, 2004). L'analisi si differenzia, ad esempio, da quella discriminante, perché non richiede una predefinita dei gruppi, ma questi vengono formati sulla base di qualche indicatore di similarità (Jain, Murty e Flynn, 1999). Com'è noto, esistono numerosi modelli di cluster analysis e i relativi risultati possono differire sensibilmente, ad esempio perché si sono utilizzati diversi indicatori di similarità o dissimilarità.

¹ La scelta di considerare una classe di età più ampia rispetto a quella tradizionalmente considerata (fino a 24 anni) per i giovani, in linea con recenti studi in ambito comunitario, è motivata dalla recente evoluzione socio-economica (si veda al riguardo European Commission, 2008).

3. L'occupazione agricola secondo i dati del Censimento

Nel periodo tra i due censimenti, 2000 e 2010, l'occupazione complessiva, misurata in termini di numero totale di giornate di lavoro, diminuisce in media di oltre il 20% (Figura 2). In realtà il calo è interamente da addebitare alla contrazione delle ore lavorate per la manodopera familiare, mentre aumenta l'impiego di quella non familiare, ciò vale soprattutto per le regioni del Nord Italia, mentre nelle altre aree si registra una contrazione, seppure più contenuta di quella per l'insieme della manodopera. L'effetto complessivo è un aumento dell'incidenza dell'occupazione non familiare sul totale che si attesta nel 2010 attorno al 20% come media nazionale, oscillando tra il 15% circa del Nord-ovest e il 25% circa del Sud.

Figura 2 - Giornate di lavoro per la manodopera totale e per quella non familiare (variazioni percentuali 2010/2000)



Fonte: elaborazioni su dati Istat, 5° e 6° Censimento generale dell'Agricoltura

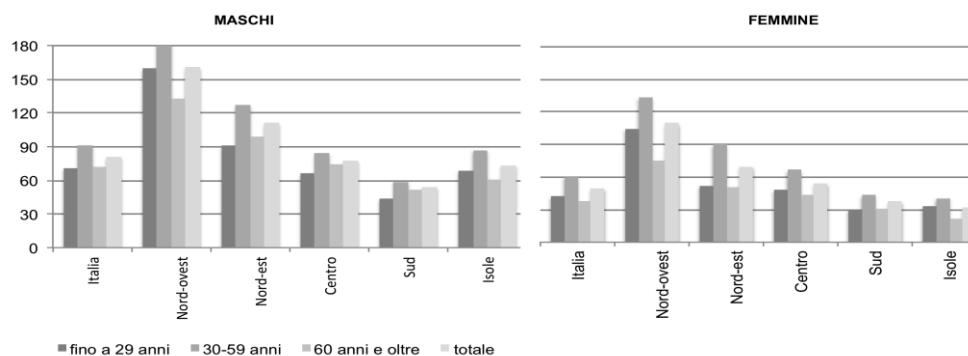
L'occupazione familiare rappresenta ancora la quota maggioritaria in agricoltura, settore che, proprio a causa della netta prevalenza di manodopera familiare, risulta ancora polverizzato in numerose microimprese. Anche la realtà dell'occupazione familiare si presenta con differenze marcate per area geografica.

In media i maschi della manodopera familiare risultano lavorare 80 giornate l'anno, con forti oscillazioni, passando da 54 giornate nel Sud a 160 nel Nord-ovest (figura 3). La distribuzione per classi di età mostra un andamento a campana, con un massimo che si raggiunge nelle classi di età centrali, ovvero in corrispondenza della classe di età 35-39 anni (nella figura 3 per ragioni espositive sono riportate solo alcune classi di età aggregate).

Andamenti analoghi, anche se per valori inferiori si riscontrano per la manodopera familiare femminile. In media le donne dichiarano di essere occupate per 50 giornate l'anno, con valori che oscillano tra le 31 giornate delle Isole e le 109 giornate del Nord-Ovest. Sembrerebbe, quindi, potersi affermare l'esistenza di una segmentazione per quanto concerne le caratteristiche della manodopera

familiare tra Nord e Sud Italia, dove nella prima una quota rilevante della manodopera familiare, anche giovanile, appare svolgere un'attività su base continuativa, mentre nel Sud e Isole (specie per le donne) il lavoro agricolo per la manodopera familiare è molto più discontinuo.

Figura 3 - Giornate di lavoro pro-capite per età della manodopera familiare, 2010



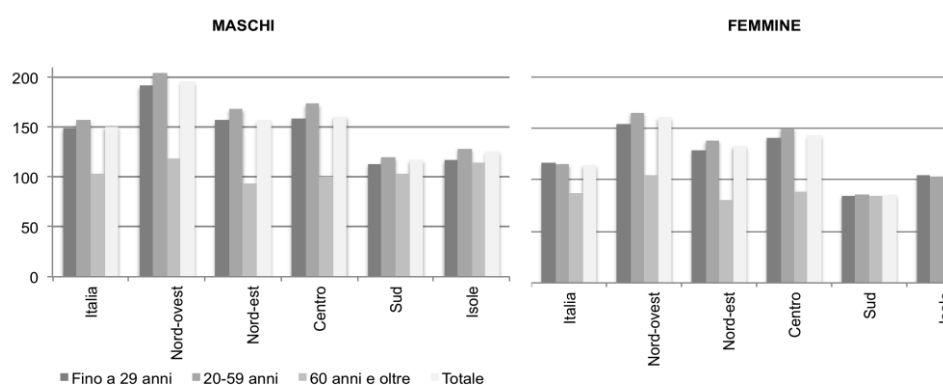
Fonte: elaborazioni su dati Istat, 6° Censimento generale dell'Agricoltura

Un'ulteriore conferma di tale divergenza può ricavarsi osservando i dati relativi alla condizione professionale per la manodopera familiare. In generale circa il 58% dei giovani maschi e il 38% delle giovani femmine si dichiara come occupato; si tratta di valori che risultano superiori a quelli medi della manodopera familiare (rispettivamente il 50% e il 33% circa). Ciò che differenzia i giovani dal resto della popolazione è l'alta incidenza dei disoccupati e di coloro che sono alla ricerca della prima occupazione: complessivamente circa il 16% dei giovani (maschi e femmine), contro il 4% in media. Se si considerano gli studenti, circa il 24% dei giovani maschi e il 35% delle giovani donne si dichiara in tale condizione, contro il 2% dell'intera popolazione.

Importanti differenze si registrano a livello territoriale. Per quanto concerne i maschi la quota di coloro che si dichiarano occupati oscilla tra circa l'80% del Nord-ovest e circa il 50% di Sud e Isole. In queste due aree territoriali circa il 15% dei giovani si dichiara disoccupato e circa l'8% alla ricerca di una prima occupazione (nel Nord-Ovest rispettivamente il 2,2% e l'1,8%). Per le giovani femmine si registrano differenze analoghe, anche se la quota di occupati risulta più contenuta, a vantaggio di quella degli studenti. Nel Nord-ovest il 65% della manodopera familiare si dichiara come occupato, mentre nel Sud e nelle Isole la percentuale è pari a circa il 30%. In queste due aree la quota di femmine disoccupate o in cerca di prima occupazione è pari a circa il 22%, contro il 6% del Nord-ovest.

Le differenze territoriali osservate per la manodopera familiare si registrano anche per la manodopera non familiare. A fronte di una media nazionale pari, per i maschi, a circa 150 giornate pro-capite, nel Nord-ovest si ha un valore prossimo a 200, mentre nel Sud e nelle Isole attorno a 120 giornate. Per i giovani si oscilla tra un valore di circa 113 giornate pro-capite del Sud a oltre 191 del Nord-ovest, contro una media nazionale pari a 149 giornate (figura 4).

Figura 4 - Giornate pro-capite per età della manodopera non familiare, 2010



Fonte: elaborazioni su dati Istat, 6° Censimento generale dell'Agricoltura

Andamenti analoghi si hanno per quanto concerne la componente femminile: 85 giornate pro-capite nel Sud, oltre 160 nel Nord-ovest. Come accade per gli uomini, il dato dei giovani non si discosta molto da quello medio, anzi in questo caso la media nazionale per i giovani (116 giornate) è leggermente superiore alla media nazionale complessiva (114).

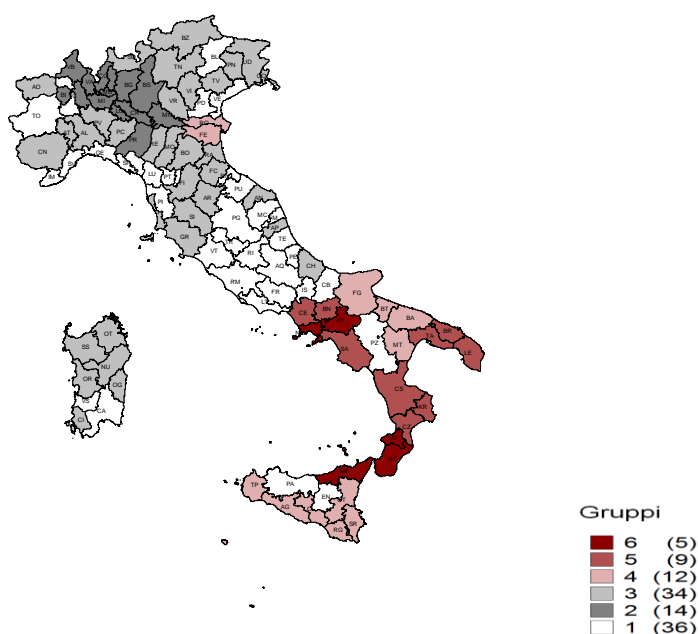
I dati fin qui presentati mostrano come, nonostante la maggiore incidenza dell'occupazione e del valore aggiunto dell'agricoltura nel Mezzogiorno, nelle regioni settentrionali, ed in particolare nel Nord-ovest si registra maggiore produttività, maggiore stabilità occupazionale, ovvero un modello più "capitalistico" e meno "familistico-solidaristico".

Il diverso carattere della produzione agricola tra macro-aree territoriali si ricava anche osservando i valori relativi all'incidenza delle attività remunerate connesse all'agricoltura (ad esempio l'agriturismo), ovvero di quelle attività che, legate comunque al settore primario, tendono ad incrementare il valore dei servizi resi, consentendo anche occasioni per una occupazione più stabile. A fronte di un'incidenza sul monte ore totale lavorato (il Censimento non registra dati di fatturato) in media nazionale pari al 3,7%, si ha un'oscillazione tra l'1,4% del Sud e il 6,5% del Nord-ovest.

4. L'analisi di cluster

I dati precedentemente analizzati per le macro-aree territoriali sono stati utilizzati per condurre una cluster analysis a livello provinciale. Si è utilizzato il metodo di Ward, si tratta di un metodo gerarchico agglomerativo, per il quale si inizia con una partizione formata da n gruppi (o grappoli), pari alle unità di analisi, per giungere, attraverso riallocazioni successive al numero ritenuto ottimale di grappoli. Nel presente lavoro tale numero è stato posto pari a 6, ritenuto quello che permette di giungere ad una certa omogeneità dei gruppi.

Figura 1 — Ripartizione delle province italiane per caratteristica del settore agricolo – 2010



Il metodo di Ward è diretto alla minimizzazione della varianza all'interno dei gruppi. Una partizione è ritenuta migliore quanto più la varianza è minore all'interno dei gruppi e maggiore tra i gruppi (si ricorda che la varianza totale può essere espressa come somma della varianza interna ed esterna). Le aggregazioni successive nel metodo di Ward avvengono unendo quei grappoli che determinano il minor incremento della varianza interna (ad ogni aggregazione, ovviamente, aumenta la varianza interna e diminuisce quella esterna). Come ogni metodo gerarchico, il metodo di Ward presenta l'inconveniente di non permettere la

riallocazione tra i gruppi, per cui non è detto che l'allocazione finale sia quella ottimale.

Il primo gruppo, formato da 36 province, rappresentato nella mappa con il colore bianco, è composto da province che hanno iniziato un percorso deciso verso attività più innovative nel settore agricolo, anche se non hanno ancora raggiunto i livelli del secondo gruppo. In queste province la manodopera è prevalentemente di tipo familiare.

Il secondo gruppo (14 province) si caratterizza per la presenza di produzioni ad elevato valore aggiunto che hanno creato occasioni di lavoro per i giovani maschi, con un forte aumento della manodopera non familiare.

Il terzo gruppo (34 province) è rappresentato da province che hanno trovato il loro punto di forza nella valorizzazione dei prodotti tipici e di qualità, nelle quali il lavoro rimane soprattutto di tipo familiare con ridotti spazi per le donne.

Nel quarto gruppo (12 province) sono presenti province con elevata incidenza del valore aggiunto in agricoltura), produzioni di tipo tradizionale con bassa incidenza di giovani e donne.

Il quinto gruppo (9 province) si differenzia dal successivo per una più forte attenzione alle produzioni Dop e per un maggior numero di giornate pro-capite per la manodopera non familiare.

Il sesto gruppo (5 province) è composto da province nelle quali l'impiego agricolo sembra ancora svolgere un ruolo "difensivo": il lavoro familiare rappresenta per una parte dei giovani e donne l'unica alternativa alla carenza di occasioni all'esterno; anche il lavoro non familiare presenta elevati elementi di discontinuità.

5. Conclusioni

L'analisi condotta ha mostrato l'esistenza di nuovi dualismi territoriali nel mercato del lavoro del settore agricolo. Da un lato vi sono diverse province, sostanzialmente quelle settentrionali, caratterizzate da un peso ridotto del numero di giornate di lavoro nel settore agricolo, ma nelle quali lavorazioni di tipo più innovativo sono riuscite negli ultimi anni ad attrarre giovani, soprattutto maschi.

Dall'altro vi sono numerose province, in particolare nel Sud Italia, nelle quali l'occupazione agricola, specie di carattere familiare, funziona ancora come cuscinetto per una manodopera, giovanile e femminile, che stenta a trovare al di fuori dell'azienda agricola occasioni di lavoro. La manodopera non familiare, che registra valori elevati, è ancora legata a lavorazioni più tradizionali, con basso valore aggiunto tranne poche eccezioni.

Riferimenti bibliografici

- EUROPEAN COMMISSION 2008. YOUTH: Young in Occupations and Unemployment: Thinking of their better integration in the labour market - Final Report.
- GIOVANNINI E. 2012. Il volto dell'agricoltura tra complessità e cambiamento. *AgriRegioniEuropa*, Anno 8, No. 31, <http://www.agriregionieuropa.univpm.it>.
- IDDA L., PULINA P. (Eds.) 2011. *Impresa agricola familiare, capitale umano e mercato del lavoro*, Milano: Franco Angeli.
- ISFOL 2012. *Dimensioni e caratteristiche del lavoro sommerso/irregolare in agricoltura*. Roma: I libri del Fondo Sociale Europeo.
- JAIN A.K., MURTY M.N., FLYNN P.J. 1999. Data clustering: a review. *ACM Computing Survey*, Vol. 31, No. 3, pp. 264-323.
- PUGLIESE E. 2012. Il lavoro agricolo immigrato nel Mezzogiorno e il caso di Rosarno. *Mondi Migranti*, No. 3, pp. 7-28.
- ROMESBURG H.C. 2004. *Cluster Analysis for researcher*. North Caroline: Lulu Press.

SUMMARY

Recent trends in the agricultural labour market

The Agricultural Census data allow us to verify several changes that the agricultural sector has been experiencing in Italy. It would seem that the development of demand for certain services related to agriculture has helped to attract a number of young workers. This is particularly true for some areas of Northern Italy. However, several areas of Southern Italy are still characterized by traditional type of production. After a descriptive analysis of Census data, I proceed to a cluster analysis in order to verify how the Italian provinces are positioned with respect to new products, having particular regard to young people and women. The results show the existence of new forms of dualism between North and South Italy.

AGE STRUCTURE EVOLUTION IN SOME SUB-SAHARAN COUNTRIES: THE ADVANTAGE OF AGEING¹

Michela C. Pellicani, Valeria Moro

1. Introduction

Demography, resorting to the formal rigor of mathematics of populations, builds scenarios for the future that contribute to feed collective fears. For more than fifty years - from the fundamental works of Frank W. Notestein (Notenstein F.W., 1945) and Kingsley Davis (Davis K., 1945) on the demographic transition - the fear indicated by the projections has been the overpopulation, in other words, the pressure that the population growth is supposed to exercise on the means of production, in particular on material resources and labour.

The fear of an overly rapid population growth has focused particularly on the Southern populations that came later in time (than those of the North) in the phase of the population explosion.

Later, however, the evident dramatic declines in fertility in different countries of the South, the United Nations has had to revise their projections assuming that the world population could reach a maximum amount significantly lower than previously expected.

As argued by the economist Walt Whitman Rostow, theorist of "The Stages of Economic Growth", in the next few decades, the planet will not be threatened by overpopulation, but, on the contrary, by the reduction of growth and, therefore, by the perspective of demographic stagnation (Rostow W.W., 1998).

This finding reveals a reverse fear which could affect, once again, in an insidious manner, the South more than the North since the South manifests a marked retardation in the awareness of the on-going slowdown of its growth and of the onset of a progressive ageing process. It is not, in fact, only a challenge for industrialized countries, which have already become familiar with the idea of ageing, but also, and perhaps above all, for the least developed countries that are not prepared for such a change.

¹ This paper is the result of the collaboration between the two Authors, in particular, the paragraphs 1 and 3 are attributable to M.C. Pellicani; the paragraphs 2 and 5 to V. Moro and the paragraph 4 to both Authors.

2. A macro-region with strong elements of heterogeneity

In our analysis, in the broader context of sub-Saharan Africa, we decided to consider the age structure evolution of the countries that compose, according to the UN, the “Middle Africa”.

The countries of this area were mainly affected by the French domination, with the exception of Cameroon (partly under British control) and Angola (Portuguese enclave) the last country to reach the Independence (1975).

Besides the historical and political affinities, practiced religions are also very homogeneous: Christianity is prevalent (mainly Catholic) followed by Animism² (Istituto Geografico De Agostini, 2011).

The region is, however, characterized by a deep heterogeneity in demographic, socio-cultural and economic terms. From a demographical point of view, with a total of almost 130 million people, the area includes countries with an amount of population ranging from 1.5 million of Gabon to nearly 70 million of the Democratic Republic of the Congo, passing by about 20 million of Angola and Cameroon³ (UN, 2011).

The infant mortality rate takes very different values: from 131.0 ‰ of Chad to 51.0 ‰ of Gabon. As a consequence, life expectancy at birth, still very low, ranges from 44.5 years (males) and 47.3 years (females) of the Central African Republic to 60.2 years (males) and 62.4 years (females) of Gabon. Finally, we find these two countries at the extremes also with regard to the total fertility rate, ranging from 6.2 children per woman in Chad to 3.4 in Gabon (UN, 2011).

In socio-cultural terms, illiteracy is still widespread⁴, especially in the peripheries of the cities and in the countryside, where the population lives in great poverty, often in improvised settlements.

Such situation of decay is predominant in countries like Chad and Angola, torn by many years of civil war that determined millions of displaced persons, or Democratic Republic of the Congo, which, besides the civil war, is also affected by the scourge of malnutrition, and also Cameroon, never affected by *coups d'état* but tormented by the epidemics concerning the whole area⁵. Despite the notable sanitary interventions of prevention, the HIV/AIDS continues to dominate the

² Only in Chad prevails the Islamic religion (53.0%) (Istituto Geografico De Agostini, 2011).

³ In this analysis we didn't consider other two countries of the region - Equatorial Guinea and São Tomé and Príncipe - because of their limited demographic dimension.

⁴ The literacy rate in 2010 ranges from 35.0% of Chad to 88.0% of Gabon (CIA, 2013).

⁵ Middle Africa is currently considered the area of the greatest risk of international transmission of polio, because of the limited impact of measures of control and the historical spread of this disease between Angola and Democratic Republic of the Congo (WHO, 2010).

scene, whose incidence among the young adult/adult population (15-49 years old) in 2011 ranges from 2.1% of Angola⁶ to 5.0% of Gabon (UNAIDS, 2012).

3. Methodology and data

Among the numerous questions concerning the African continent in terms of its future prospects, certainly there is the question related to the consequences of the changes of demographic behaviour on the age structure of its populations.

In this occasion we choose an age, 25 years, that concentrates as much the fears as the hopes. That is the age of the end of the studies (when the case), of the entry in the labour market (if not yet the case), of the leave of the parental *foyer* to start (later) its own family and finally, the age of the stronger geographic mobility (Fargues Ph. - Pellicani M.C., 2000).

To try to approach as much as possible the “lived reality”, we simulated the individual perspective rather than the population one.

In generational terms and at the moment of the entry in the productive (25 years) and reproductive (30 years) life, we tried to answer to some questions that the individual will have to face: what are the charges that parents do represent for a single individual and what are the charges represented by the children that the individual will have; with which handicaps and advantages, therefore, the individual coming from the South starts, nowadays, its adult life.

On this purpose we constructed two indicators that allow us to illustrate the assessment that the average person (in a statistical sense) can do about its situation (Fargues Ph. - Pellicani M.C., 2000):

$$- \text{ ascending charge index (a.c.i.)} = 2 * (I_{55}^{g-30} / I_{30}^{g-30}) / (ISF^g * I_{25}^g) \quad (1)$$

$$- \text{ descending charge index (d.c.i.)} = ISF^{g+30} / 2 \quad (2)$$

To the past evolution we added a prospect analysis (on the base of the 2010 revision of the United Nations population prospects, medium variant) considering the generations 1960-2030.

4. Generations to the test

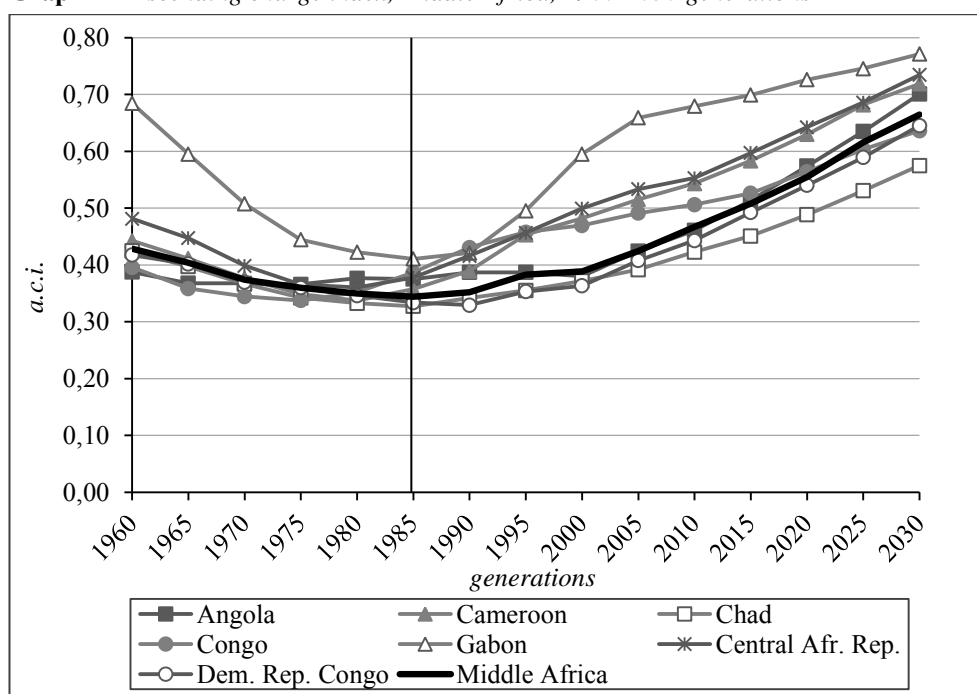
Always bearing in mind the general situation of the region, characterized by a late and generally modest increase in life expectancy and by a still high fertility

⁶ In Angola, the epidemic of HIV/AIDS is more recent than the rest of Africa; it doesn't have, therefore, yet reached its peak (UNAIDS, 2012).

rate, the ascending charge has had a decreasing trend at least till the 1985 generation. In 2010, there are, on average in the considered area, 0.34 surviving parents for each 25-year-old individual (Graph 1).

The combined effect of the reduction in parents fertility and of the increase in their life expectancy will begin to be noted from the 1995 generation, which will be 25 years old in 2020. In the considered period, the index value will increase from 0.43 for the 1960 generation to 0.66 for the 2030 generation. In other words, the ascending charge will become heavier of 53.5%.

Graph 1 – Ascending charge index, Middle Africa, 1960-2030 generations



Source: our processing on UN data (UN, 2011)

In opposition to what we observed for the ascending charge, with regard to the descending charge, there is a clear trend to convergence of the values towards the fateful level of generations replacement⁷.

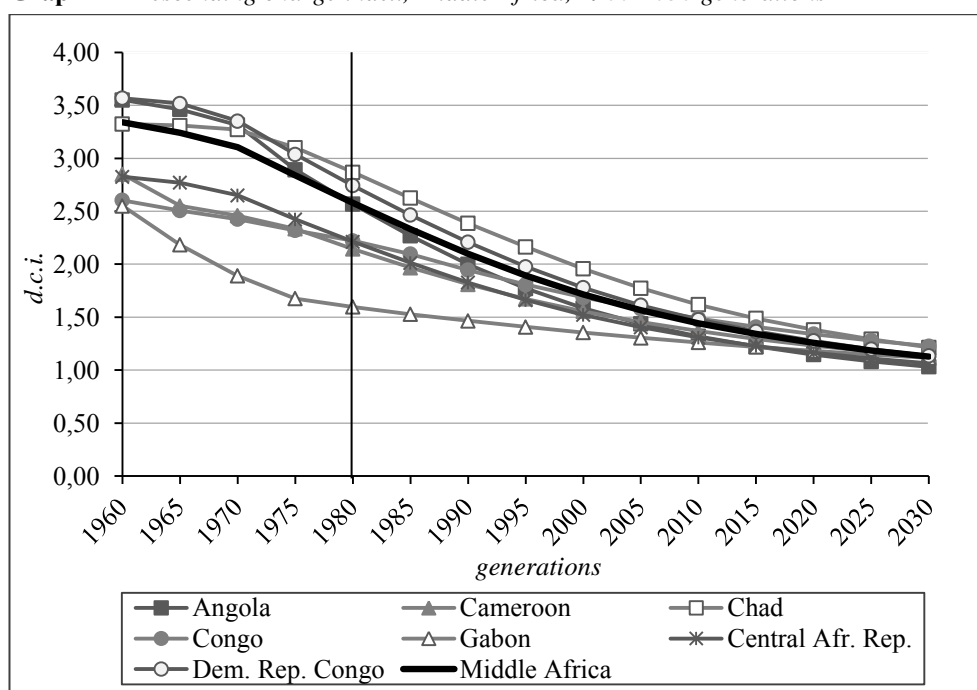
Compared to 3.34 descendants for the 1960 generation, the 1980 generation (the one that is 30 years old in 2010) is exposed to a (still) high charge even if in progressive decline (2.58) to reach the minimum value of 1.13 for the 2030

⁷ The hypothesis, underlying the projections of the United Nations, of the convergence of fertility levels to the threshold value (2.1 children per woman) is, however, yet to be demonstrated.

generation (Graph 2). The latest generations, therefore, will find themselves in a very different context compared to the generations that preceded them.

Even more than the general evolution, the variations by country result very significant, both in terms of intensity and timing of decrease. It is interesting, in this regard, to underline the behaviour of Gabon that, before the other countries of the area, has had a spectacular decrease of mortality and that, starting from the lowest fertility rate⁸ will reach 1.12 descendants for the 2030 generation.

Graph 2 – Descending charge index, Middle Africa, 1960-2030 generations



Source: our processing on UN data (UN, 2011)

Finally considering both indicators we can conclude that the generations 1990-2005 will find themselves in a period of exceptional demographic opportunity in comparison to the generations that immediately preceded them, profiting from a further fertility decrease and from a not (yet) excessive charge represented by the care of their parents.

⁸ In 1960 the total fertility rate of Gabon was equal to 4.3 children per woman compared to 7.4 children per woman of Angola (UN, 2011).

5. Conclusions

As we have had occasion to show, between the overpopulation of yesterday and the stagnation of tomorrow, there is, perhaps, a brief moment during which demography can be neutral, or even, represent an opportunity rather than a burden.

The generations that in the next few years will become adult will have the opportunity, in fact, to invest in their work as a source of personal well-being as well as of increase of the life quality of the future generations, for example in terms larger and better access to health care and education.

Based on the considerations formulated so far, none of considered generations in these countries has experienced or will experience a burden towards the generations that preceded them as heavy as that suffered by populations of other regions, especially the North ones, at the same time there is no doubt that Middle Africa countries must become aware of the implications of their fertility decline.

This phase of major demographic transformations will have as a consequence a substantial change in the amount and, even more relevant, in the population structure, leading to an increasingly long coexistence and inequality between the different generations.

Countries such as those we analysed will, therefore, have to carefully evaluate the significant effects produced by structural changes in order to calibrate the possible interventions. At the same time it will be fundamental to keep in mind that, in the increased demand, for example in terms of education (see *theory of quantity for quality trade-off*), social progress and democratization will play an even more important role than that played by the demographic evolutions.

References

- BEEGLE K., DE WEERDT J., DERCON S. 2009. The intergenerational impact of the African orphans crisis: a cohort study from an HIV/AIDS affected area, *International Journal of Epidemiology*, Vol. 38, No. 2, pp. 561-568.
- CALDWELL J.C., ORUBULOYE I. O., CALDWELL P. 1999. Fertility Decline in Africa: A New Type of Transition?, *Population and Development Review*, Vol. 18, No. 2, pp. 211-242.
- CALDWELL J.C. 2000. Rethinking the African AIDS epidemic, *Population and Development Review*, Vol. 26, No. 1, pp. 117-135.
- CIA - CENTRAL INTELLIGENCE AGENCY 2013. *The World Factbook 2013-14*, Washington, DC.
- DAVIS K. 1945. The World Demographic Transition, *The ANNALS of the American Academy of Political and Social Science*, Vol. 237, pp. 1-11.

- EASTWOOD R., LIPTON M. 2011. Demographic transition in sub-Saharan Africa: How big will the economic dividend be?, *Population Studies*, Vol. 65, No. 1, pp. 9-35.
- EL-BADRY M.A. 1992. Emerging Population Issues in Africa, *International Statistical Review*, Vol. 60, No. 2, pp. 119-127.
- FARGUES PH., PELLICANI M.C. 2000. La possibile influenza del contesto demografico sulla mobilità del fattore umano: un'analisi per generazioni, *Rivista Italiana di Economia Demografia e Statistica*, Vol. LIV, No. 3, pp. 131-159.
- HATTON J.T., WILLIAMSON J.G. 2003. Demographic and Economic Pressure on Emigration out of Africa, *The Scandinavian Journal of Economics*, Vol. 105, No. 3, pp. 465-486.
- ISTITUTO GEOGRAFICO DE AGOSTINI 2011. *Calendario Atlante De Agostini 2012*, Novara: Istituto Geografico De Agostini.
- KALIPENI E. 1995. The Fertility Transition in Africa, *Geographical Review*, Vol. 85, No. 3, pp. 286-300.
- KRAVDAL Ø. 2002. Education and Fertility in Sub-Saharan Africa: Individual and Community Effects, *Demography*, Vol. 39, No. 2, pp. 233-250.
- LEGRAND T.K., BARBIERI M. 2002. The Possible Effects of Child Survival on Women's Ages at First Union and Childbirth in Sub-Saharan Africa, *European Journal of Population*, Vol. 18, No. 4, pp. 361-386.
- NEUBERT D. 2005. Researching Africa South of the Sahara: A Sociologist's Perspective, *Africa Spectrum*, Vol. 40, No. 3, pp. 429-444.
- NOTESTEIN F.W. 1945. Population, the Long View. In SCHULTZ E. (Ed) *Food for the World*, Chicago III: University of Chicago Press, pp. 36-57.
- PELLICANI M.C. 2012. Rivoluzione della struttura per età e "primavera araba": esigenza di nuove forme di solidarietà intergenerazionale nel quadro del processo di democratizzazione. In VALLERI M. et al. (Eds), *Il Mediterraneo: uno studio e una passione. Scritti in onore di Luigi Di Comite*, Bari: Cacucci Editore, pp. 301-313.
- ROSTOW W.W. 1998. *The Great Population Spike and After. Reflections on the 21st Century*. New York: Oxford University Press.
- TABUTIN D., SCHOUMAKER B. 2002. The Demography of Sub-Saharan Africa from the 1950s to the 2000s. A Survey of Changes and a Statistical Assessment, *Population*, Vol. 59, No. 3/4, pp. 457-556.
- UNAIDS 2012. *Global report: UNAIDS report on the global AIDS epidemic 2012*, Ginevra: WHO Press,
http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2012/gr2012/20121120_UNAIDS_Global_Report_2012_en.pdf.
- UN - UNITED NATIONS 2011. *World Population Prospects. The 2010 Revision*, New York: United Nations.

- UN - UNITED NATIONS 2002. *Hiv/Aids and fertility in Sub-Saharan Africa: a review of the research literature*, New York: United Nations, <http://www.un.org/esa/population/publications/fertilitysection/HIVAIDSPaperFertSect.pdf>.
- WORLD BANK 2013. *Africa Development Indicators 2012/13*, Washington, DC: World Bank, <https://openknowledge.worldbank.org/handle/10986/13504> License: CC BY 3.0 Unported.
- WHO - WORLD HEALTH ORGANIZATION 2010. *Global Polio Eradication Initiative. Strategic Plan 2010-2012*, Geneva: WHO Press, http://www.polioeradication.org/Portals/0/Document/StrategicPlan/StratPlan2010_2012_ENG.pdf.
- ZIMMER Z., DAYTON J. 2005. Older Adults in Sub-Saharan Africa Living with Children and Grandchildren, *Population Studies*, Vol. 59, No. 3, pp. 295-312.

SUMMARY

The planet will be threatened, says Rostow, not by overpopulation but by the perspective of the demographic stagnation that constitutes the true challenge not only for the industrialized countries but equally, and perhaps above all, for the countries less developed (that are less prepared to such a change).

This is the case also of the sub-Saharan countries, that have not been sufficiently analyzed under this interesting point of view.

Aware that it is not possible to have accurate data, we used, for our calculations, the data provided by the United Nations in order to have an homogeneity and reliability (as much as possible) for all considered countries.

Our main aim, therefore, is to outline the frame of reference in terms of past and future structural changes of the populations of some sub-Saharan countries through the construction of two indicators that allow us to illustrate the assessment that the average person (in a statistical sense) can do about its situation. We tried to answer some questions that the individual, entering in the productive and reproductive life, will have to face: what are the charges that parents do represent for a single individual and what are the charges represented by the children that the individual will have.

As a result of our analysis, we notice that the considered countries will have to carefully evaluate the significant effects produced by structural changes in order to calibrate the possible interventions. At the same time it will be fundamental to keep in mind that, in the increased demand, for example in terms of education, social progress and democratization will play an even more important role than that played by the demographic evolutions.

Michela C. PELLICANI, Università degli Studi di Bari "Aldo Moro",
michelacamilla.pellicani@uniba.it

Valeria MORO, Università degli Studi di Bari "Aldo Moro", vale.moro@yahoo.it

SODDISFAZIONE, ASPETTATIVE E PROSPETTIVE DEGLI STUDENTI STRANIERI DELL'UNIVERSITÀ DELLA CALABRIA

Roberta Saladino

1. Introduzione

Un aspetto molto importante del fenomeno migratorio sono gli studenti internazionali, da considerare una delle espressioni più significative del fenomeno della mobilità, non da ultimo per il potenziale impatto sulla concorrenza e sulla costruzione di società interculturali. La Commissione Europea, cosciente di questa posta in gioco, ha proposto uno studio specifico "International Student Survey" ai 27 Stati membri, che fanno parte dell'European Migration Network (EMN), un programma istituzionale che in Italia fa capo al Ministero dell'Interno, Direzione Centrale delle Politiche per l'Immigrazione e per l'Asilo. L'indagine in Italia è stata condotta dall'Istituto di Ricerche sulla Popolazione e le Politiche Sociali del Consiglio Nazionale delle Ricerche e dal Centro Studi e Ricerche IDOS, con la collaborazione di diverse Università tra cui l'Università della Calabria (UniCal).

Lo strumento d'indagine è stato un questionario di 60 domande in forma chiusa elaborato da un gruppo di lavoro congiunto IDOS e IRPPS-CNR, al fine di rilevare le condizioni di vita e di studio, il grado di soddisfazione, le aspettative e le prospettive per l'inserimento nel mondo del lavoro degli studenti¹ non-comunitari che hanno scelto di svolgere il proprio percorso formativo negli Atenei italiani. Il presente contributo analizza i risultati dell'indagine realizzata presso l'UniCal.

2. L'indagine in UniCal

L'indagine presso l'UniCal è stata avviata nel mese di novembre 2012 e portata a termine nel mese di aprile 2013, la somministrazione del questionario realizzata con tecnica PAPI, è avvenuta in momenti diversi:

¹ Studenti il cui ingresso e soggiorno in Italia è condizionato alle previsioni del Testo Unico sull'immigrazione

1) nei corsi di lingua italiana, erogati dall'UniCal presso il Dipartimento di Lingue e Scienze dell'Educazione;

2) durante l'evento culturale "Ciao Indonesia", organizzato dall'Associazione ASIC² il 15 dicembre 2012. L'evento è stato strutturato con diverse attività (seminari, mostre, concerti musicali e danze), rappresentando una vera finestra aperta sull'Indonesia, a tutto ciò ha partecipato l'ambasciatore della Repubblica Indonesiana;

3) nella pausa pranzo e/o cena (presso il Centro Residenziale UniCal).

Al campione sono stati ammessi, oltre agli studenti non comunitari iscritti ai corsi di laurea triennale e di laurea magistrale, anche gli studenti Erasmus Mundus e quelli di Master, di Dottorato e di Corso di Specializzazione.

Il questionario è stato somministrato a 302 studenti stranieri non comunitari, si tratta di una indagine campionaria la cui dimensione del campione è abbastanza rilevante, se pensiamo infatti, che gli studenti stranieri non comunitari per l'a.a. 2012-2013 sono in tutto 505. All'interno del campione sono rappresentate 31 nazionalità (su 51). Tra queste, le più numerose, che complessivamente assommano a più della metà degli intervistati, sono gli studenti cinesi (35%), indonesiani (10%), ecuadoriani (9%), turchi (7%) e russi (6%). La nazionalità più numerosa in UniCal è la cinese con il 27% degli studenti, tutto ciò non deve sorprenderci poiché la Cina è il Paese con il maggior numero di studenti outgoing: nel 2012 la quota cinese sul mercato degli studenti internazionali era pari al 16%, la più grande in assoluto, seguita dall'India con il 7%.

Anche la comunità degli indonesiani è numerosa: il 30% degli studenti indonesiani presenti in Italia è iscritto all'UniCal. Tutto ciò è dovuto al grande impegno dell'Ateneo a partecipare alle Fiere Indonesiane, al fine di far conoscere le sue offerte formative e i servizi che offre agli studenti stranieri. Questa sua costante partecipazione alle Fiere ha permesso all'Università di farsi conoscere ed essere apprezzata dagli studenti indonesiani, che hanno nel corso degli anni svolto un "passaparola" determinando un ulteriore incremento della loro presenza. La notevole presenza degli studenti ha comportato il coinvolgimento anche del governo indonesiano, che annualmente partecipa attraverso il suo ambasciatore alle iniziative che l'ASIC organizza.

Le caratteristiche socio-demografiche del campione sono le seguenti:

- 1) l'età media è pari a 25 anni;
- 2) il 51% degli intervistati sono donne;
- 3) solo il 3% è coniugato;

² "Associazione degli studenti indonesiani in Calabria"

4) il 7% dichiara di vivere in Italia con la sua famiglia, ciò indica che le immigrazioni per altre ragioni in questo caso per studio non è più un fenomeno trascurabile;

5) il 63% ha almeno un genitore con un titolo universitario, questo dato tende a mostrare che gli intervistati provengono in larga maggioranza da situazioni familiari con livello di istruzione alto o medio-alto.;

6) solo il 28% degli intervistati è uno studente lavoratore.

3. Metodologia applicata

I dati dei questionari sono stati studiati attraverso una strategia di analisi integrata quindi mediante un concatenamento di metodi fattoriali che è tipico della Scuola Francese di Analisi dei dati. Avendo un questionario che contiene per lo più variabili qualitative ho scelto un'analisi delle corrispondenze multiple, dopo di che ho fatto una cluster analysis sui risultati dell'analisi fattoriale. Nella tradizione della Scuola Francese dell'Analyse des Données (che ha come capostipite lo statistico Benzecri) questa viene anche chiamata Tandem Analysis perché c'è proprio questa idea di classificare gli individui non sulla base delle variabili originarie ma sulla base delle relazioni di associazioni tra le variabili considerate nell'indagine.

Nel set di variabili osservate ho valutato quali dovevano giocare il ruolo di attive e quelle invece che dovevano essere impiegate come supplementari. Le variabili considerate attive sono in tutte 13: 1) Come valuti il tuo rendimento nello studio?, 2) Come valuti la tua esperienza nell'Università?, 3) Come valuti il tuo rapporto con il personale docente?, 4) Consideri la segreteria didattica della tua Università efficace ed adeguata alle tue esigenze?, 5) Perché sei venuto a studiare in Italia?, 6) Come valuti il tuo livello di integrazione in Italia?, 7) Qual è l'atteggiamento degli studenti italiani verso gli studenti stranieri?, 8) Secondo te, come sono percepiti gli studenti internazionali nella società italiana?, 9) Vorresti lavorare in Italia dopo la laurea? 10.a) Se sì, indicare solo la motivazione principale, 10.b) Se no, indicare solo la motivazione principale, 11) Ritieni che la preparazione fornita dall'Università sia adeguata per accedere al mondo del lavoro?, 12) Ripeteresti l'esperienza?, 13) Come ti sei trovato in Italia?.

La scelta di considerare le variabili appena menzionate come attive è dovuto al fatto che, esse ci consentono di descrivere meglio il fenomeno (analizzare il grado di soddisfazione, di integrazione e le prospettive degli studenti).

4. Risultati dell'indagine

Nella rappresentazione ottenuta considerando il primo e il secondo asse fattoriale, abbiamo una inerzia quindi una percentuale di informazione originaria rappresentata sul piano bidimensionale pari al 16%. In realtà questo è un problema noto dell'analisi delle corrispondenze multiple, poiché stiamo rappresentando sostanzialmente delle relazioni ortogonali tra modalità che appartengono alle stesse variabili, in maniera artificiale si introduce una certa sfericità nei dati, questo vuol dire che l'inerzia spiegata che vediamo rappresentata in termini percentuali sulla mappa fattoriale (Vedi fig. 1) è una misura pessimistica del potere esplicativo dell'analisi fatta. Allora, possiamo rappresentare in termini di inerzia spiegata/aggiustata con una correzione proposta da Benzecri, tenendo conto quindi del numero di modalità complessivo che appartengono alle variabili attive, un valore che invece è del 40%, questa correzione consente di avere una miglior lettura della inerzia spiegata e quindi della variabilità del fenomeno tenendo conto che in questa analisi particolare c'è proprio questo problema (distorsione).

Figura 1 – Mappa fattoriale: livello di integrazione



Fonte: "International Student Survey" UniCal

Andando a visualizzare sulla mappa fattoriale le modalità che ci consentono di spiegare cosa abbiamo sul primo e il secondo asse fattoriale (si dice battezzare gli assi), osserviamo come nel primo asse fattoriale leggiamo delle modalità che per una accezione positiva esprimono un ottimo livello di integrazione, quindi vediamo che per esempio gli studenti intervistati hanno risposto che hanno avuto un'ottima esperienza, ritengono che la società italiana abbia un atteggiamento nei loro confronti positivo, valutano il loro rendimento nello studio eccellente, ottimo il

loro rapporto con il personale docente, ripeterebbero la loro esperienza, valutano il loro livello di integrazione eccellente e vorrebbero lavorare in Italia.

Questo buon livello di integrazione presente sul primo asse è contrapposto a un livello di segno opposto di integrazione nel secondo asse, in questa parte della mappa fattoriale sono presenti gli studenti che ritengono di aver avuto un'integrazione insufficiente, valutano la loro esperienza sia all'interno dell'UniCal che nella società italiana negativa.

La particolare forma triangolare della nube dei punti (Vedi fig.1) è spiegabile con il fatto che in questo particolare caso, il secondo asse ha una configurazione unipolare, nel senso che, assume valori elevati tanto in positivo che in negativo soltanto da un lato del grafico, nella parte che abbiamo individuato come di minor integrazione. Tutto ciò può essere spiegato con il fatto che c'è una differenziazione, nella parte bassa del grafico sono presenti le modalità che possono essere individuate come conflitti sociali, infatti i rispondenti hanno nel loro profilo di risposte indicato come molto negativa la loro esperienza, una rivalità da parte degli studenti italiani e un'ostilità da parte della società italiana. Nella parte superiore invece c'è un grande accento verso quelli che possono essere considerati dei conflitti culturali, infatti questi studenti non hanno una soddisfacente esperienza universitaria, ritengono che la preparazione fornita dall'Università non sia adeguata per accedere al mondo del lavoro quindi non ripeterebbero l'esperienza.

Chiaramente nel primo asse fattoriale vi è uno schiacciamento della nube dei punti perché fondamentalmente chi si è integrato bene ha avuto un'inclusione sia sociale che culturale.

Sui risultati di questa analisi fattoriale è possibile andare a costruire una classificazione automatica, classificazione automatica che è tale perché fondamentalmente non si hanno a priori delle categorie di soggetti, ma i risultati sono ottenuti sulla base dei profili di risposta ai questionari e quindi possiamo riconoscere in percentuale 5 categorie (Vedi fig. 2) che sono state battezzate nel seguente modo: soddisfatti (20%), opportunisti (31%), disillusi (7%), insicuri (25%) e insoddisfatti (17%).

Le modalità più caratterizzanti per la categoria dei soddisfatti sono:

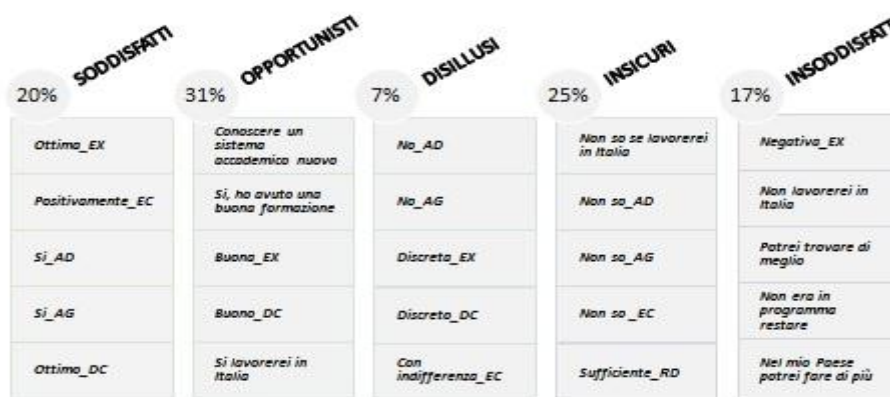
- 1) Ottima l'esperienza universitaria,
- 2) Positiva la percezione che la società italiana ha nei confronti degli studenti internazionali;
- 3) Adeguata la preparazione fornita dall'Università per accedere al mondo del lavoro;
- 4) Ripeterei l'esperienza universitaria;
- 5) Ottimo il rapporto con il personale docente.

Le modalità più caratterizzanti per la categoria degli opportunisti sono:

- 1) Venuta/o a studiare in Italia per conoscere un sistema accademico nuovo;

- 2) Buona la formazione ricevuta;
- 3) Buona l'esperienza universitaria;
- 4) Buono il rapporto con il personale docente;
- 5) Sì, lavorerei in Italia.

Figura 2 – Cinque categorie di studenti intervistati



Fonte: "International Student Survey" UniCal

Le modalità più caratterizzanti per la categoria dei disillusi sono:

- 1) Inadeguata la preparazione fornita dall'Università per accedere al mondo del lavoro;
- 2) Non ripeterei l'esperienza;
- 3) Discreta l'esperienza universitaria;
- 4) Discreto il rapporto con il personale docente;
- 5) Con indifferenza la società italiana percepisce gli studenti internazionali.

Le modalità più caratterizzanti per la categoria degli insicuri sono:

- 1) Non so se lavorerei in Italia;
- 2) Non so se la preparazione fornita dall'Università è adeguata per accedere al mondo del lavoro;
- 3) Non so se ripeterei l'esperienza;
- 4) Non so valutare come sono percepiti gli studenti internazionali nella società italiana;
- 5) Sufficiente il rendimento nello studio.

Le modalità più caratterizzanti per la categoria degli insoddisfatti sono:

- 1) Non lavorerei in Italia;
- 2) Negativa l'esperienza universitaria;
- 3) Troverei un lavoro migliore in un altro Paese;
- 4) Non era nel mio programma restare;

5) Con la mia laurea nel mio Paese posso fare di più.

Il fatto che queste modalità siano le più caratterizzanti per le 5 categorie lo si può considerare attraverso un test statistico. Test statistico che sostanzialmente è un test costruito sull'ipotesi che il modello probabilistico sottostante sia di tipo ipergeometrico, quindi la statistica test che viene costruita tiene conto delle differenze tra la proporzione di rispondenti di quella modalità nell'intero collettivo e la proporzione di rispondenti di quella modalità nella categoria che stiamo considerando, per cui sulla base dei risultati di questi test di significatività che sono stati condotti per assicurare validità ai risultati che vediamo riportati nella figura 2, sono in effetti queste le modalità di risposta che hanno presentato una differenza significativa nelle cinque categorie.

5. Conclusioni

Dall'analisi eseguita si evince che la maggioranza degli studenti non comunitari iscritti all'UniCal viene specificatamente per motivi di studio.

Dalla cluster analysis effettuata si evince che il 20% degli studenti intervistati si considera soddisfatto dell'esperienza universitaria, ritenendo adeguata la preparazione fornita dall'UniCal per accedere al mondo del lavoro. Il 31% lavorerebbe in Italia dopo la laurea mentre il 17% no, poiché troverebbero di meglio in un altro Paese. C'è da dire inoltre che, probabilmente è proprio il fatto che il giudizio sulla società italiana in generale, e in particolare sulla sua disposizione verso gli immigrati, non sia totalmente positivo per la maggior parte degli studenti intervistati a far sì che il 25% di loro non abbia ancora deciso se rimanere o no in Italia dopo la Laurea.

L'indagine empirica ha permesso anche di capire che un buon livello di integrazione è riconducibile a due fattori: 1) fattori culturali, relativi quindi a tutti gli aspetti qualitativi del sistema universitario (esempio la presenza delle English facilities nei Corsi di Studio); e a fattori sociali, relativi ovviamente alla società (esempio l'atteggiamento che la società italiana ha nei confronti degli studenti internazionali). Gli studenti intervistati che hanno infatti una buona integrazione nel "mondo universitario" hanno anche un buon livello di integrazione nella società. Sebbene le variabili supplementari nella mappa fattoriale (Vedi figura 1) non sono state rappresentate, possiamo dire che è emersa una forte associazione tra il livello di integrazione e il livello di conoscenza della lingua italiana, infatti coloro che hanno una buona conoscenza dell'italiano ha anche una buona integrazione. La lingua è uno strumento di integrazione sicuramente importante, un livello elevato di conoscenza della lingua chiaramente comporta un ottima esperienza di vita universitaria e nella società in generale.

Dall'analisi effettuata, si evince infine, come la presenza degli studenti stranieri non è soltanto influenzata dalla qualità del sistema universitario italiano ma anche al fatto che gli sbocchi professionali per i laureati in Italia sono limitati perché, generalmente, chi sceglie di frequentare gli studi universitari all'estero lo fa spesso prevedendo poi di restare a lavorare nel Paese dove consegnerà il titolo.

Riferimenti Bibliografici

- BRANDI M.C. 2008. Emigrazione di élite: neolaureati e ricercatori italiani all'estero, in *Fondazione Migrantes, Rapporto Italiani nel Mondo, Edizioni IDOS*, Roma, pp. 156-162.
- Di FRANCO G. 2011. *Tecniche e Modelli di Analisi Multivariata*. Milano: FrancoAngeli.
- GoLini A. 2006. *L'immigrazione straniera: indicatori e misure di integrazione*. Bologna: Il Mulino
- SALADINO R. 2013. *Università della Calabria: la dimensione internazionale in un Ateneo del Meridione*. Rapporto EMN Italia. Vol. VI, pp. 179-181.
- WALLACE J.B. 1999. The case for student as customer. *Quality Progress*, Vol. 32, n°2, pp. 47-51.

SUMMARY

Satisfaction, expectations and prospects of foreign students of Calabria

This paper analyzes the results of the "International Students: presence and impact," made at the UniCal. The survey instrument was a questionnaire of 60 questions, developed by a joint working group IDOS and IRPPS-CNR, has been structured in order to detect the conditions of life and study, the degree of satisfaction, expectations and prospects for entering into employment of non-EU students who have chosen to do their training in the Italian universities.

DAL BENESSERE ALL'ANALISI DELL'EFFICIENZA DI UN TERRITORIO*

Domenico Summo, Tommaso Pepe

1. Introduzione

Negli ultimi anni sono state sviluppate diverse metodologie statistiche con la finalità di restituire misure sintetiche sul livello di benessere economico, sociale ed ambientale percepito dalla popolazione di un territorio, allo scopo di superare le parzialità e le distorsioni derivanti dal PIL; se quest'ultimo in passato era considerato sia come indicatore di riferimento per misurare lo sviluppo economico di un territorio sia per misurare il benessere, per la supposizione dell'esistenza di un rapporto univoco e diretto tra il grado di sviluppo e la relativa ricchezza prodotta, oggi tale misura risulta essere condizionata non solo dagli aspetti economici, ma anche dai fattori sociali ed ambientali.

L'obiettivo del presente lavoro ha riguardato innanzitutto l'implementazione di una metodologia di analisi in grado di misurare il benessere a livello complessivo. Posta l'esigenza di voler sviluppare una metodologia in grado di misurare il benessere sulla base di una batteria di indicatori elementari distribuiti in diverse dimensioni e, per ognuna di esse, in diversi domini latenti e coscienti delle difficoltà riscontrate nella formulazione di un punteggio sintetico, nel presente lavoro si è partiti dalla supposizione che le diverse dimensioni si pongano in combinazione lineare tra loro e che le problematiche connesse al peso che ognuna di esse assume nella determinazione della percezione complessiva del benessere possano essere superate ricorrendo a misure di dissimilarità.

Passando dall'analisi dei bisogni di un individuo a quella condotta su una intera unità territoriale, si transita verso modelli teorici di efficienza; questo ha portato ad evidenziare che la complessità del fenomeno non è solo legata al numero di dimensioni che lo condizionano, ma anche al punto di vista da cui lo stesso viene osservato e studiato. Se nella misura del benessere si deve legare l'osservazione più al territorio nel quale un individuo risiede, implementando di conseguenza modelli diversi per ogni ripartizione territoriale, nell'esaminare la performance di un

* Il lavoro è frutto congiunto dell'impegno degli stessi autori.

territorio si deve partire dall'ipotesi che ogni unità territoriale dovrebbe raggiungere gli stessi risultati e dovrebbe posizionarsi su un'unica frontiera di efficienza. Dal punto di vista più propriamente metodologico, è stato sviluppato un modello teorico basato sulla *Deterministic Frontier Analysis* (DFA).

2. L'indicatore composito proposto per una misura sintetica del benessere

In termini generali, supposta l'esistenza di una interdipendenza tra le misure di benessere implementate¹ di tipo multi-direzionale e considerato che nello spazio i

¹Alla formulazione dell'indicatore composito si giunge attraverso la determinazione di misure relative alla percezione che gli individui hanno del benessere a livello dimensionale. Pertanto, richiamando quanto già proposto dagli stessi autori (Summo D., Pepe T., 2011), in cui si riteneva che la misura del benessere è un fenomeno che non può essere studiato considerando solo pochi indicatori di carattere socio-economici ed ambientali, ma che ha bisogno di essere osservato nella sua totalità attraverso una adeguata metodologia di analisi statistica multivariata ed un'ampia batteria di indicatori elementari e ricorrendo all'analisi fattoriale in componenti principali con lo scopo di ridefinire nuovi domini latenti ed assegnare un peso diverso agli indicatori elementari ed ai relativi nuovi domini, è stata formulata prima di tutto una misura del benessere a livello dimensionale sulla base di due distinte componenti, quali: la misura del grado di benessere conseguito nell'*s*-esima dimensione da ciascuna unità territoriale della *g*-esima ripartizione e la misura della distanza dei risultati di ciascuna unità territoriali rispetto ai valori ritenuti ottimali. Distribuite, allora, le *N* unità territoriali in *g* ripartizioni, individuate le *s* dimensioni del fenomeno, per ognuna di esse, individuati gli *m* domini che le spiegano, selezionate le diverse variabili attraverso le quali – dominio per dominio – spiegare la più alta percentuale del fenomeno, isolati gli aspetti più significativi del fenomeno e rideterminato i diversi domini sulla base dell'analisi fattoriale in componenti principali, sono state determinate prima di tutto le due componenti $\varphi(X)$ e $\psi(X)$. Per ottenere $\varphi(X)$ è stata calcolata la misura sintetica di ogni singolo indicatore esplicativo di ogni fattore, ponderando per il suo peso fattoriale e dividendo per la somma dei pesi del fattore che la contiene. In simboli:

$$\begin{aligned} \varphi_{sgi}(X) &= \varphi_{sgi}(x_{s,g,1,i}, \dots, x_{s,g,n,i}) = \\ &= \left[\sum_j^m \left(\sum_i^{N_j} p_{sgij} x_{sgij} / \sum_i^{N_j} p_{sgij} \right) b_{sgj} \right] / \sum_j^m b_{sgj} = \sum_j^m \beta_{sgj} b_{sgj} / \sum_j^m b_{sgj} \end{aligned} \quad (1)$$

dove x_{sgij} è l'*i*-esimo indicatore elementare esplicativo inserito nel *j*-esimo fattore appartenente alla *s*-esima dimensione ed alla *g*-esima ripartizione territoriale, mentre p_{sgij} è il peso fattoriale ad essa associato. Con lo stesso procedimento è stata determinata la successiva trasformata $\psi_{sgi}(X)$, ottenuta come funzione degli scarti tra il risultato conseguito da ciascuna unità osservata ed il risultato ottimale che ciascuna potrebbe conseguire per raggiungere livelli ottimali di benessere. Per giungere alla sua definizione, è stato calcolato il valore medio (λ_{sgij}), di quei valori non anomali di ogni indicatore elementare situati tra il terzo quartile e gli outliers più grandi – per quegli indicatori con andamento crescente della propria performance – oppure tra gli outliers più piccoli ed il primo quartile – per quegli indicatori con andamento decrescente della propria performance. Intesa come difficoltà incontrata da ciascuna unità osservata nel raggiungere livelli ottimali, risulta determinata da:

$$\begin{aligned} \psi_{sgi}(X) &= \psi_{sgi}(x_{s,g,1,i}, \dots, x_{s,g,n,i}) = \\ &= \left\{ \sum_j^m \left[\sum_i^{N_j} k p_{sgij} (\lambda_{sgij} - x_{sgij}) / \sum_i^{N_j} p_{sgij} \right] b_{sgj} \right\} / \sum_j^m b_{sgj} = \sum_j^m \delta_{sgj} b_{sgj} / \sum_j^m b_{sgj} \end{aligned} \quad (2)$$

risultati rilevati in una generica unità territoriale possono dipendere anche da quelli rilevati nelle unità vicine, si pone il problema di quale criterio adottare per sintetizzare tali relazioni; partendo allora dalle s dimensioni del benessere, si è ritenuto utile ricorrere alle misure di distanza per porle in relazione lineare e giungere, di conseguenza, ad una misura sintetica del benessere complessivamente percepito dalla popolazione dell' i .esima unità territoriale nella g .esima ripartizione. Supposto che le s dimensioni del fenomeno influenzino secondo pesi differenti la sensazione complessiva di benessere, si ha:

$$B_{ig}(X) = \sum_s \alpha_{sgi} f_{sgi}(B_{sgi}(X)) \quad (1)$$

dove α_{sgi} rappresenta il peso da assegnare all' s .esima dimensione del benessere – indicata con $B_{sgi}(X)$ – nell' i .esima unità territoriale e limitatamente alla g .esima ripartizione. Si pone inoltre la necessità di inserire un sistema di pesi proprio per il differente contributo dato da ciascuna dimensione alla percezione complessiva del fenomeno; per questo motivo, proprio per la sua adattabilità a cogliere la diversità tra i profili a livello multidimensionale si è fatto ricorso alla distanza di Minkowski come elemento di ponderazione nella costruzione dell'indicatore sintetico proposto. Poi, nell'ambito dell'analisi spaziale dei dati, la costruzione di una matrice delle dissimilarità consente, di fatto, di rappresentare le relazioni territoriali che sussistono tra le unità oggetto di studio; questo permette, quindi, di evidenziare le differenze che la i .esima unità territoriale presenta con le altre. Nonostante siano una semplificazione della realtà, misurare le dissimilarità attraverso gli indicatori elementari di partenza ha permesso di ottenere una misura che differenzia la i .esima unità territoriale dalla h .esima unità (Zaccomer G. P., 1995) (Arbia G., 2006); quindi, la logica sottostante che ha guidato nel modello, parte essenzialmente dal presupposto che tra le diverse unità territoriali, sia pure non contigue tra di loro, possano sussistere delle interrelazioni che consentano di individuare elementi di omogeneità o di eterogeneità.

Volendo, quindi, confrontare il risultato raggiunto dalla i .esima unità territoriale nella s .esima dimensione del benessere con quello conseguito dalla h .esima unità

La costante k assume valore pari a $+1$, in presenza di un indicatore elementare con andamento crescente in termini di performance, e valore pari a -1 , in presenza di indicatori con andamento decrescente della performance. Determinate entrambe le trasformate si passa alla definizione della forma funzionale del modello di benessere per l' s .esima dimensione, data da:

$$B_{sgi}(X) = F(X_{sgi} | \varphi_{sgi}, \psi_{sgi}) = \frac{\exp[\varphi_{sgi}(X) - \psi_{sgi}(X)]}{1 + \exp[\varphi_{sgi}(X) - \psi_{sgi}(X)]} \quad (3)$$

La (3) costituisce un indicatore composito del fenomeno osservato con range di riferimento compreso tra 0 ed 1.

che ha presentato il risultato peggiore e rapportando tale differenza al relativo range di variazione (Del Colle E., 2006), ne è derivato un indicatore composito, così definito:

$$B_{ig}(X) = \sum_s \left(\frac{\sum_i w_{sgh}}{\sum_s \sum_i w_{sgh}} \sqrt{\frac{[B_{sgi}(X) - B_{sg\min}(X)]^2}{[B_{sg\max}(X) - B_{sg\min}(X)]^2}} \right) \quad (2)$$

dove w_{sgh} esprime il coefficiente di dissimilarità relativo alla *i*.esima unità territoriale rispetto alla *h*.esima unità nella *s*.esima dimensione del benessere, limitatamente alla *g*.esima ripartizione, mentre $B_{sg,\min}(X)$ e $B_{sg,\max}(X)$ rappresentano rispettivamente il valore minimo ottenuto nella *s*.esima dimensione del benessere e quello massimo. Il punteggio conseguito da una unità territoriale è stato sintetizzato considerando la distanza euclidea tra il valore registrato ed il valore più basso, rapportato alla distanza intercorrente tra il miglior valore riscontrato ed il peggiore, ottenuti nella stessa dimensione. Tale risultato è stato, poi, ponderato rapportando la sommatoria dei coefficienti di dissimilarità della *i*.esima unità territoriale esaminata con tutte le altre e la sommatoria complessiva dei coefficienti di dissimilarità.

Oltre ad ottenere una linearizzazione delle *s* dimensioni del benessere, in questo modo si sono ottenuti valori compresi tra zero ed uno sia per ogni dimensione del benessere sia per *i* pesi assegnati; moltiplicando tra loro *i* pesi per il valore ottenuto da $B_{sgi}(X)$, si ottiene ancora un valore compreso tra zero ed uno, ancora più piccolo dei singoli fattori di partenza. Questo non deve sorprendere in quanto insito nella moltiplicazione di fattori positivi minori dell'unità; ricordando che la somma di tutti *i* pesi attribuiti all'*i*.esima unità territoriale è uguale ad uno e che anche $B_{sgi}(X)$ assume in genere valori compresi tra zero ed uno, si evidenzia che questo ultimo potrebbe assumere valori anche superiori di poco all'unità solo in condizioni particolari (in presenza di livelli di dissimilarità prossimi allo zero per *s*-1 dimensioni ed un alto valore sia in termini di dissimilarità sia di benessere nella *s*.esima dimensione).

3. Dal concetto di benessere a quello di performance di un territorio

Nel momento in cui dai bisogni dei cittadini si passa ad esaminare una unità territoriale nella sua interezza, si transita inevitabilmente dal tema della qualità della vita a quello della performance di un territorio ed all'analisi del suo grado di competitività; anche nel conseguimento degli obiettivi di coesione socio-economica di un territorio diventa importante implementare misure legate al grado

di efficienza raggiunto dalla i -esima unità territoriale. Si passa, allora, da una analisi condotta per ripartizioni, al confronto di tutte le N unità territoriali osservate con un'unica frontiera ideale.

Spostando, quindi, l'attenzione dall'individuo all'intera i -esima unità territoriale, bisogna necessariamente introdurre il concetto di frontiera; in effetti, nel conseguire obiettivi di coesione e di maggiore omogeneità nei comportamenti a livello territoriale, è necessario individuare dei punti di riferimento, da intendersi come elementi di ottimalità. In tale contesto, nella frontiera deterministica la componente di disturbo casuale agisce solo nella direzione di ridurre l'output osservato rispetto a quello ottimale, per effetto del maggiore o minore livello di performance che caratterizza la singola unità territoriale considerata; ne segue che nessuna unità dovrebbe trovarsi al disopra della linea di frontiera. Tuttavia, l'output, oltre a dipendere dagli input considerati e dal rispettivo livello di performance, è in funzione di disturbi aleatori che caratterizzano le determinazioni empiriche collegate alla influenza di altri fattori non controllabili, oltre che da errori di misurazione.

Pertanto, si propone di seguito una metodologia di analisi per la misura del livello di adattamento di ciascuna unità territoriale esaminata rispetto a quelli che possono essere definiti standard di ottimalità e, in tal caso, individuare possibili elementi di rigidità che impediscono alla stessa di conseguire migliori risultati. Trovandosi quindi in presenza di unità omogenee, è possibile definire una frontiera di ottimalità che descrive come i vari input $B_{si}(X)$ possano combinarsi per ottenere gli output $B_i(X)$.

Al fine di stimare il livello di ottimalità dell' i -esima unità territoriale considerata, è stata definita una funzione $f(B_{si}(X); z_i)$ dipendente sia dal livello degli input primari sia da altri possibili indicatori elementari z_i che possono spiegare l'eventuale livello di performance (Lothgren M., 1997) (Richmond J., 1974). Sulla base di un modello di regressione lineare i cui parametri sono stimati utilizzando il metodo dei minimi quadrati corretto, la relazione tra input ed output per la i -esima unità territoriale risulta essere del tipo:

$$B_i(X) = \beta_0 + \sum_s \beta_s B_{si}(X) + u_i \quad (3)$$

dove $B_i(X)$ è il valore osservato dell'output dell' i -esima unità territoriale, $B_{si}(X)$ esprime il benessere nella s -esima dimensione e relativo alla i -esima unità territoriale considerata, β_0 e β_s sono i parametri da stimare ed u_i è una variabile casuale associata al livello di performance, paragonabile sostanzialmente alla difficoltà riscontrata dalla i -esima unità nel conseguire un livello di performance

più vicino a quello di ottimalità. Poiché quest'ultima non ha media zero, ai fini della stima con il metodo dei minimi quadrati corretto, la (3) va a sua volta trasformata, sommando e sottraendo la media della variabile u_i all'equazione precedente, in modo da definire una nuova variabile casuale a media nulla $\varepsilon_i = (u_i - \mu)$ ed una nuova intercetta $\beta_0^* = (\beta_0 + \mu)$; il modello diventa pertanto:

$$B_i(X) = (\beta_0 + \mu) + \sum_s \beta_s B_{si}(X) + (u_i - \mu) \quad (4)$$

$$B_i(X) = \beta_0^* + \sum_s \beta_s B_{si}(X) + \varepsilon_i \quad (5)$$

La variabile ε_i , espressa in forma vettoriale, è allora formata da una componente deterministica (z_i) e da una componente aleatoria, associata agli usuali disturbi aleatori ed assunta indipendente ed identicamente distribuita secondo una normale con media zero e varianza costante². Pertanto, stimati i parametri della (5), il passo successivo è quello di andare a stimare il massimo valore ottimale in modo da poter proporre in questo contesto anche una misura sia del livello di efficienza raggiunto dall' i -esima unità territoriale sia dei relativi elementi di rigidità; per determinare la frontiera di performance cui ciascuna unità territoriale dovrebbe tendere, a livello generale, per ogni indicatore elementare considerato viene individuato un valore ottimale a livello unidimensionale. L'insieme di tutti tali punti ottimali vanno a rappresentare la relativa frontiera cui ogni unità territoriale dovrebbe tendere a livello di ogni dimensione; poi, per costruire la relativa frontiera generale, è stato applicato il modello di regressione stimato anche a tali valori ottimali. Questo ha permesso di passare da una frontiera multidimensionale di punti nello spazio alla relativa sintesi. Si è ottenuta, in tal modo, la seguente funzione ottimale:

$$B_i^f(X) = \beta_0^* + \sum_s \beta_s B_s^f(X) \quad (6)$$

La (6) spiega il massimo livello di output ottenibile se l' i -esima unità territoriale conseguisse il valore ottimale di $B_s^f(X)$ (ovvero qualora avesse conseguito

² Il vettore:

$$\varepsilon_i = (u_i - \mu) = \sum_s \gamma_s z_{si} + V_i$$

è stato ottenuto ponendo in combinazione lineare i residui con il divario territoriale tra ed entro le unità in grado di spiegare lo stato di difficoltà incontrato per conseguire livelli di performance più vicini a quelli ritenuti ottimali. Per la misura di tali divari territoriali è possibile ricorrere implementazione di un indicatore di entropia, come ad esempio lo stesso indice di Theil (Bracalente B., 1991).

nell'esima dimensione i valori ottimali in tutti gli indicatori elementari inseriti nell'analisi). Il rapporto tra la funzione (5) e la funzione ottimale (6), compreso tra zero ed uno, fornisce una misura della performance di un territorio, espressa in termini di efficienza, mentre il relativo complemento algebrico rappresenta la relativa componente di rigidità (Rizzi D., 1999).

4. Alcune considerazioni

Nel presente lavoro è stata proposta una misura sintetica della qualità della vita, sviluppata secondo un approccio metodologico; confrontare tra loro unità territoriali appartenenti ad una stessa ripartizione permette sia di differenziarsi da quanto già proposto in letteratura sia di determinare la misura del benessere sulla base di elementi di maggiore omogeneità. A tale scelta si è arrivati dopo aver riflettuto sul fatto che sostanzialmente la percezione del benessere da parte di individui appartenenti ad una determinata unità territoriale potrebbe essere influenzata (in termini di pesi) da indicatori elementari probabilmente diversi da quelli che influenzerebbero, invece, un individuo di un'altra unità territoriale.

Passando dal tema del benessere all'analisi della performance di una unità territoriale, la funzione frontiera è stata determinata partendo dall'ipotesi che tutte le unità osservate in teoria dovrebbero conseguire gli stessi risultati di ottimalità.

Nonostante l'elevato interesse verso le misure sulla qualità della vita e sulla performance di un territorio, bisogna evidenziare che esse non contribuiscono affatto a rendere migliore la vita di un individuo se non vengono applicate in maniera sistematica e se i loro risultati non vengono interpretati in maniera corretta, al fine di individuare i punti di forza e gli aspetti critici sui quali intervenire.

Riferimenti Bibliografici

- ARBIA G. 2006. *Spatial econometrics: statistical foundations and applications to regional convergence*. Berlino: Springer Verlag.
- BRACALENTE B. 1991. Analisi di dati spaziali. In MARBACH G. (Ed.), *Statistica economica*, Torino: Utet.
- CICCARELLI A. 1998. *L'articolazione della competitività a livello territoriale*. Torino: Giapichelli.
- DEL COLLE E. 2006. *Tecnopoli. L'articolazione territoriale della competitività in Italia*. Milano: Franco Angeli.

- LOTHGREN M. 1997. *A Multiple Output Stochastic Ray Frontier Production Model*, Stockholm School of Economics, W. P. Series in Economics and Finance, n. 158, February.
- RICHMOND J. 1974. *Estimating the Efficiency of Production*. International Economic Review.
- RIZZI D. 1999. *L'efficienza dei dipartimenti dell'Università Ca' Foscari di Venezia via DEA e DFA*, Università Ca' Foscari, Nota di Lavoro n. 09/99, Venezia.
- SUMMO D., PEPE T. 2011. Divari economici tra le province dell'Italia Meridionale. *Rivista della Società Italiana di Economia, Demografia e Statistica*, volume LXV N. 3/4.
- ZACCOMER G. P. 2008. *Economia, Statistica e Territorio. Informazione e metodologia statistica per la conoscenza dell'economia del Friuli Venezia Giulia*. Forum, Udine.

SUMMARY

From welfare to analysis of a territory

This paper will be to develop a statistics methodology in order to measure the total well-being of an individual. This total perception of well-being is determined placing in linear combination the various dimensions and assigning them a various weight on the base of heterogeneity measures. By the analysis of the well-being of an individual is occurred to the analysis of the entire territorial unit passing accordingly to efficiency measures.

Domenico SUMMO, Professore Associato, Università degli Studi di Bari "Aldo Moro", dosummo@dss.uniba.it.

Tommaso PEPE, Dottore di Ricerca, Università degli Studi di Bari "Aldo Moro", pepetommaso@libero.it.

MISURARE L'EUROPA CHE CAMBIA: DALLE STRATEGIE DI LISBONA AL PIANO 2020

Valentina Talucci

1. La definizione del problema

L'interesse della Comunità Scientifica intorno a temi quali lo sviluppo umano, la qualità della vita o il benessere è piuttosto recente; come è noto, i primi autorevoli studi si possono ricondurre alla seconda metà degli anni '80 rappresentati dal *capability approach* (Sen A., 1985). Un tema così "giovane" è affrontato sotto diverse prospettive disciplinari ed è ancora fortemente legato ad aspetti diversi tra loro come quelli etico-normativi, metodologici e di misura nonché di politica pubblica.

Gli obiettivi di questo lavoro sono sia teorici che metodologici e sembrano in linea con i recenti obiettivi internazionali di misurazione del progresso e dei progressi degli stati membri. Si tenterà da un lato di seguire concettualmente il corso degli studi legati alla comprensione e, se possibile, alla misurazione delle trasformazioni sociali ed economiche in Europa, dall'altro di effettuare una sintesi degli indicatori elementari previsti dal Piano Europeo 2020 che porti verso una misura comparativa spaziale e temporale.

L'approccio seniano risulta il più coerente al fine di analizzare le misure della qualità della vita e la sostenibilità nei processi di sviluppo umano, anche in contesti socio-economici avanzati.

Il benessere individuale viene concettualizzato non come una condizione statica e materialista, definita dal possesso ad un certo istante di tempo di un dato ammontare di risorse materiali, ma come un processo all'interno del quale i mezzi e le risorse acquisibili rappresentano uno strumento, essenziale ed irrinunciabile, per ottenere benessere, ma non ne costituiscono di per sé una metrica adeguata a misurare il benessere effettivo delle persone o la qualità della vita che esse riescono a realizzare con i mezzi a loro disposizione (Chiappero Martinetti E., 2009).

L'approccio delle capacità non si limita ad estendere l'attenzione al di fuori della sola dimensione monetaria, ma intende richiamare l'attenzione sulla pluralità dei fattori personali e familiari e sulle molteplicità di contesti: sociali, ambientali, economici, istituzionali, culturali, che agiscono nella determinazione del processo di benessere individuale.

Ci si è mossi dunque in un ambito interdisciplinare cercando di mantenere un equilibrio e una sinergia tra il piano teorico e quello empirico, tra aspetti metodologici e sostantivi. Sono stati affrontati, infatti, sia aspetti socio-economici, sia aspetti statistico-metodologici tesi alla misurazione del fenomeno. Nel quadro teorico di riferimento, una delle parole chiave è la multidimensionalità, intesa come differenziazione e integrazione dei fenomeni indagati.

Sen ha dimostrato una certa preferenza verso un principio che può essere così riassunto: “Ottimizzare le condizioni di qualsiasi soggetto a condizione di non intaccare quelle di qualche altro soggetto che sta peggio di lui”. Ciò rappresenta la forma più radicale che l’egualitarismo può assumere senza arrecare danno agli stessi soggetti che intendono prioritariamente favorire (Somaini, 2002).

Il framework teorico di questo articolo parte dallo studio delle deprivazioni materiali legate ai beni primari (Teoria della giustizia di Rawls, 1971), ma le estende in un’ottica multidimensionale ad altri aspetti che influiscono sulle condizioni di vita delle persone (Teoria dell’uguaglianza delle capacità fondamentali di Sen). Gli interventi di welfare hanno il dovere di intervenire su sacche di popolazione vulnerabili, deprivate, a forte rischio di esclusione, con il preciso intento di programmare azioni o piani di intervento ugualitari, dove ogni individuo deve essere messo in condizione di attuare a pieno le proprie aspirazioni (Teoria del diritto di Dworkin).

Un nuovo aspetto che sta assumendo sempre maggiore centralità è la sostenibilità che nel processo di sviluppo umano ha una valenza molto importante poiché implica l’ampliamento delle scelte individuali. In particolare, significa soddisfare i bisogni delle generazioni presenti senza però compromettere le capacità e le opportunità di quelle future, ossia equità e solidarietà intergenerazionale.

Qualunque principio di giustizia sociale mette la persona al centro di un processo di sviluppo in modo tale da assicurare *empowerment*, partecipazione, uguaglianza delle opportunità tra uomini e donne, equità nella distribuzione del reddito, contenimento della povertà e uno sviluppo sostenibile duraturo nel tempo.

2. Dalla teoria alla misurazione

Nel corso degli ultimi anni la discussione in merito all’uso del PIL (GDP) come misura del benessere ha occupato molto spazio all’interno delle sedi istituzionali (OCSE, Commissione Europea, Istituti nazionali di statistica) e dei media. Da un lato si sente la necessità di tenere traccia delle misure che per anni sono state usate – e che non sono del tutto abbandonate - nella misurazione del fenomeno e dall’altro ricercare, all’interno di nuovi studi e nuovi approcci, elementi che

affrontino il problema in un'ottica multidimensionale; l'obiettivo è quello di allargare la visuale a dimensioni che devono accostarsi, con eguale importanza, a quella del reddito e della ricchezza in generale.

In seguito al mandato del Consiglio Europeo di Lisbona (2000), gli Stati membri e la Commissione hanno cercato di sviluppare strategie comuni e un certo grado di compatibilità per quanto concerne gli indicatori. Il lavoro è stato svolto dal Comitato per la protezione sociale e dal suo sottogruppo tecnico in particolare; il sottogruppo era interessato al miglioramento degli indicatori nel campo della povertà e dell'esclusione sociale. Tale sfera d'attività deriva dall'accordo politico raggiunto al Consiglio Europeo di Nizza, durante il quale sono stati definiti gli obiettivi appropriati nella lotta contro la povertà e l'esclusione sociale e gli Stati membri e la Commissione sono stati invitati a sviluppare indicatori comuni.

Al Consiglio Europeo di Stoccolma del marzo 2001, i capi di Stato e di governo hanno conferito mandato al Consiglio di adottare, entro la fine dell'anno, un gruppo di indicatori in materia di integrazione sociale accettati comunemente. Tali indicatori dovrebbero consentire agli Stati membri e alla Commissione di valutare i progressi compiuti nel conseguimento dell'obiettivo definito dal Consiglio europeo di Lisbona, vale a dire quello di ottenere un impatto decisivo sull'eliminazione della povertà entro il 2010, migliorare la comprensione dei fenomeni della povertà e dell'esclusione sociale nel contesto europeo e di individuare e scambiare buone pratiche. Già nel dicembre 2001 il Consiglio Europeo di Laeken adottò una lista di indicatori volti a definire la situazione di povertà ed esclusione sociale.

Nel Novembre del 2007 la Commissione Europea organizzava a Bruxelles una conferenza internazionale dall'eloquente titolo "*Beyond GDP*"; la conferenza si chiudeva con la dichiarazione del presidente Barroso: "è ormai tempo di andare oltre il PIL". La Commissione Europea (20 Agosto 2009) si è così espressa al Consiglio e al Parlamento Europeo "Non solo PIL. Misurare il progresso in un mondo in cambiamento"; assumendosi l'impegno di orientare le proprie politiche in virtù di un set di indicatori più ampio dichiarando "Esistono validi motivi per completare il PIL con statistiche che riprendano gli altri aspetti economici, sociali ed ambientali dai quali dipende fortemente il benessere dei cittadini" (www.ilsole24ore.com; Dossier Italia, 7 Novembre 2009).

Il 2010 è stato dichiarato "Anno europeo della lotta alla povertà e all'esclusione sociale". Oggi vi è un complesso insieme di indicatori comuni, suddivisi per liste tematiche, al fine di effettuare il monitoraggio degli effetti delle politiche di welfare nei paesi UE.

L'adozione della nuova Strategia EU2020 che segue quella adottata a Lisbona 10 anni, punta a rilanciare l'economia dell'UE nel prossimo decennio, dedicando particolare attenzione alle politiche economiche, occupazionali e sociali. In un mondo che cambia l'UE si propone di diventare un'economia intelligente,

sostenibile e solidale. Queste tre priorità che si rafforzano a vicenda intendono aiutare l'UE e gli Stati membri a conseguire elevati livelli di occupazione, produttività e coesione sociale. In pratica, l'Unione si è posta cinque ambiziosi obiettivi – in materia di occupazione, innovazione, istruzione, integrazione sociale e clima/energia – da raggiungere entro il 2020. Ogni Stato membro ha adottato per ciascuno di questi settori i propri obiettivi nazionali. Interventi concreti a livello europeo e nazionale vanno a consolidare la strategia.

3. La scelta degli indicatori

I 5 obiettivi che si è posta l'Unione Europea fanno capo a 5 dimensioni tematiche per ciascuna delle quali è stato scelto un set di indicatori il cui numero risulta variabile; ciascun indicatore ha un verso concorde o discorde con il fenomeno studiato. Il problema del verso dell'indicatore implica una necessaria operazione di trasformazione che dipende dalla natura delle variabile e che può condizionare i risultati di sintesi.

Un aspetto importante nella valutazione dell'utilizzo degli indicatori risiede nel modo in cui questi servono a programmare e raggiungere degli obiettivi mirati.

1) Dimensione del Lavoro: innalzamento al 75% del tasso di occupazione (per la fascia di età compresa tra i 20 e i 64 anni e genere).

2) Dimensione Economica R&S: aumento degli investimenti in ricerca e sviluppo al 3% del PIL dell'UE. Questo indicatore recepisce gli aspetti delle *capabilities* prendendo in considerazione lavori creativi intrapresi in modo sistematico al fine di accrescere l'insieme delle conoscenze, compresa la conoscenza dell'uomo, della cultura e della società e l'uso di tale bagaglio di conoscenze per concepire nuove applicazioni;

3) Dimensione Ambientale: Cambiamenti climatici e sostenibilità energetica 2020:

3.1) riduzione delle emissioni di gas serra del 20% (o persino del 30%, se le condizioni lo permettono) rispetto al 1990. Questo indicatore evidenzia le tendenze in emissioni totali di origine antropica del „paniere di Kyoto“ di gas ad effetto serra. Le emissioni di gas serra aggregate sono espressi in unità di CO₂ equivalenti. L'indicatore non comprende le emissioni e gli assorbimenti relativi all'uso del suolo, cambiamenti di uso del suolo e alla silvicoltura, né comprendono le emissioni prodotte dal trasporto aereo internazionale e del trasporto marittimo internazionale;

3.2) Energia rinnovabile, espresso come quota di energie rinnovabili nel consumo finale lordo di energia (20%). Questo indicatore è calcolato sulla base di statistiche dell'energia di cui al regolamento statistiche sull'energia. Può essere considerato una stima dell'indicatore descritto nella direttiva 2009/28/CE, in quanto

il sistema delle statistiche per alcune tecnologie di energia rinnovabile non è ancora pienamente sviluppato per soddisfare i requisiti della presente direttiva, dunque il contributo di queste tecnologie è piuttosto marginale per il momento;

3.3) Efficienza energetica. Si intende il consumo interno lordo escludendo ogni utilizzo non energetico dei vettori energetici (es. gas naturale utilizzato per la combustione, ma non per la produzione di prodotti chimici). Questa quantità è rilevante per misurare il vero consumo energetico e per confrontarlo con l'Europa. La "percentuale di risparmio" è calcolata utilizzando i valori del 2005 e le sue previsioni per il 2020; l'obiettivo di Europa 2020 è raggiunto quando questo valore raggiunge il livello del 20%;

4) Dimensione Istruzione:

4.1) Tasso di abbandono scolastico precoce. Questo indicatore mostra la percentuale di studenti in età compresa tra 18 e 24 anni che hanno abbandonato la scuola primaria, inferiore o superiore, e che hanno dichiarato di non aver ricevuto alcuna istruzione o formazione nelle quattro settimane precedenti l'indagine. L'obiettivo dell'UE per il 2020 è quello di diminuire la percentuale di abbandono scolastico a meno del 10%;

4.2) Tasso di istruzione superiore esprime la quota della popolazione di età 30-34 anni che ha completato con successo l'università (terzo livello). Questo indicatore misura l'obiettivo chiave della strategia Europa 2020 per aumentare la quota di 30-34 anni che ha completato l'istruzione terziaria o equivalente almeno al 40% nel 2020;

5) Dimensione Povertà: tasso di povertà, popolazione a rischio di povertà o di esclusione espresso come % sul totale della popolazione. A rischio di povertà sono le persone con un reddito disponibile equivalente al di sotto della soglia di rischio di povertà, fissata al 60% del reddito mediano nazionale equivalente monouso reddito (dopo i trasferimenti sociali). La "deprivazione materiale" comprende indicatori relativi al possesso di beni durevoli, alle caratteristiche dell'alloggio e la qualità dell'ambiente della dimora. Persone gravemente deprivate in senso materiale non possono permettersi di i) per pagare affitto o bollette, ii) a tenere a casa adeguatamente caldo, iii) affrontare spese impreviste, iv) mangiare carne, pesce o una proteina equivalente ogni due giorni, v) una settimana di vacanza lontano da casa, vi) un'automobile, vii) una lavatrice, viii) un televisore a colori, o ix) un telefono. Persone di età compresa tra 0-59 che vivono in famiglie in cui gli adulti lavorano meno del 20% del loro potenziale di lavoro totale durante l'anno passato. L'obiettivo alla lotta contro la povertà e all'emarginazione è ridurre di almeno 20 milioni le persone a rischio o in situazione di povertà ed emarginazione.

4. Una rappresentazione sintetica dei paesi

Il metodo delle componenti principali (ACP) consente di eliminare eventuali ridondanze nel complesso delle informazioni disponibili sulle unità di rilevazione, sostituendo alle variabili originali un minor numero di variabili che sono funzioni lineari delle prime e che “assorbono” una quota sufficiente dell’informazione complessiva. Data una matrice X_{np} , l’analisi in componenti principali risolve il problema delle informazioni ridondanti sostituendo alle p variabili iniziali, k variabili nuove di varianza massima e di importanza decrescente.

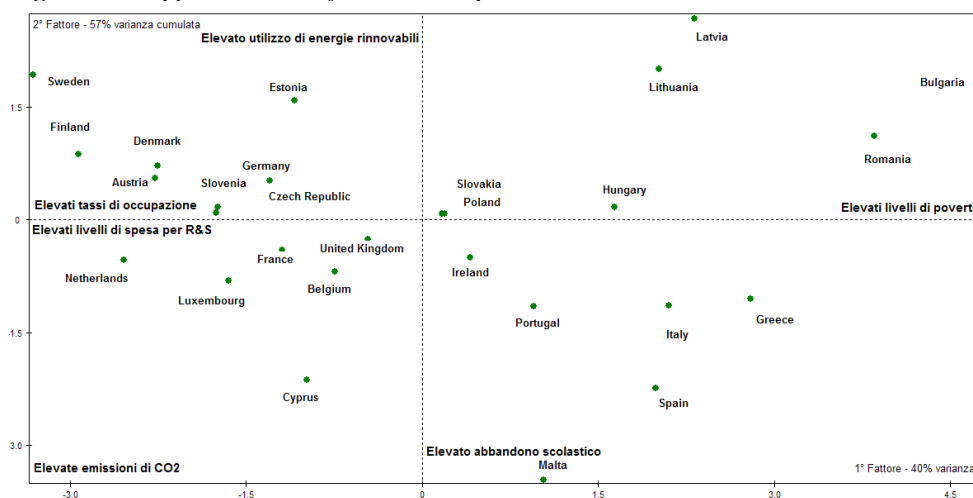
Si può meglio affermare che l’ACP permette di individuare in un insieme di variabili correlate, nuove e poche variabili sintetiche – dette componenti – tra loro non correlate, ottenute come combinazione lineare delle variabili di partenza. Le componenti offrono una sintesi il più possibile fedele ai dati originari, esse possono essere interpretate come una retta che meglio si adatta alla nube degli n punti nello spazio p -dimensionale.

L’obiettivo di questo lavoro è studiare e rappresentare il fenomeno attraverso un approccio multidimensionale, in cui possano emergere eventuali differenze territoriali. Gli indicatori scelti in ambito EU2020 sembrano ben cogliere le criticità connesse ai mutamenti socio-economici che stanno investendo l’Europa. Le evoluzioni della società moderna suggeriscono, infatti, che la dimensione unica del reddito non è più sufficiente a rappresentare il fenomeno della disuguaglianza sociale in cui agiscono molteplici fattori legati a dimensioni quali il lavoro, l’istruzione, l’ambiente e gli interventi dei governi volti al riequilibrio degli scompensi. Si è altresì coscienti che tale ricerca e tale risultato possano fornire un valido contributo, per studiosi del settore e *policy makers*, al disegno e all’individuazione di sacche di popolazione la cui condizione sociale è decisamente superiore o inferiore rispetto alla media nazionale. In particolare come si può vedere dalla figura 1 che segue si individuano chiaramente due fattori, il primo che intercetta la dimensione della povertà nel semiasse positivo e quella degli investimenti in ricerca e sviluppo in quello negativo; il secondo fattore invece intercetta la dimensione ambientale nel semiasse positivo e l’istruzione in quello negativo. Dall’intersezione dei due fattori è possibile individuare un piano fattoriale entro cui disegnare una geografia dell’Europa che risulta piuttosto coerente con i livelli di sviluppo delle nazioni europee. È possibile apprezzare per i paesi del nord Europa e le social-democrazie (Germania, Danimarca, Austria, Finlandia, Svezia) un maggior investimento in ricerca e sviluppo e una maggiore sensibilità ambientale che si sostanzia in utilizzo di energie rinnovabili, dimostrando così un maggiore livello di sostenibilità rispetto a paesi come la Francia, l’Olanda il Belgio che pur posizionandosi nel semipiano dei paesi lodevoli

ancora non investono in politiche energetiche volte alla riduzione delle emissioni di CO₂.

I restanti paesi delle aree del mediterraneo e dell'est europeo si caratterizzano invece per elevati tassi di abbandono scolastico (Portogallo, Spagna, Italia, Grecia); alti tassi di povertà (Romania, Bulgaria, Ungheria).

Figura 1 – Rappresentazione fattoriale dei paesi EU



Fonte: Nostra elaborazione su dati European Commission/Eurostat/Europe_2020_indicators/headline_indicators

La natura del rapporto tra le disuguaglianze e le prestazioni economiche è di centrale importanza per la politica di tutta l'UE, dato che il raggiungimento di un elevato livello di protezione sociale e la garanzia di una maggiore coesione sociale sono i principali obiettivi dell'Unione europea (insieme con il raggiungimento di una crescita economica sostenuta dal mantenimento e il rafforzamento della competitività). D'altro canto, una più equa distribuzione del reddito non è compatibile solo con i miglioramenti nella performance economica, ma potrebbe anche contribuire a conseguire obiettivi sociali che possono svolgere un duplice ruolo nel ridurre le disuguaglianze e rafforzare la competitività.

Concludendo perché la strategia Europa 2020 sia efficace e raggiunga realmente gli obiettivi sperati per ogni paese, è stato istituito un sistema di governo dell'economia per coordinare le azioni a livello UE e a livello nazionale. Esso si basa su tre pilastri: un'agenda economica rafforzata, con una maggiore sorveglianza da parte dell'UE; interventi per salvaguardare la stabilità dell'area dell'euro; misure per rimediare ai problemi del settore finanziario.

Riferimenti bibliografici

- CHIAPPERO MARTINETTI E. (ED.) 2009. *Sviluppo Umano sostenibile e qualità della vita*. Roma: Carocci.
- DWORKIN R. 1978. *Taking rights seriously*, Cambridge: Mass Belknap Harvard University Press.
- EUROPEAN COMMISSION EMPLOYMENT 2006. Social Affairs and Equal Opportunities DG Social protection and social integration Social and demography analysis. Portfolio Of Overarching Indicators And Streamlined Social Inclusion, Pensions, And Health Portfolios. Brussels.
- EUROPEAN COMMISSION 2003. Laeken indicators – Detailed Calculation methodology, working group statistics on Income, poverty & social exclusion, Luxemburg.
- RAWLS J. 1971. *A Theory Of Justice*, Cambridge: Mass Belknap Harvard University Press.
- SEN A.K. 1985. *Commodities and Capabilities*. Amsterdam: North Holland.
- SOMAINI E. 2002. *Uguaglianza: teorie, politiche, problemi*. Roma: Donzelli Editore.
- SAVOVA I. 2012. Europe 2020 Strategy – towards a smarter, greener and more inclusive EU economy? General and regional statistics, *Eurostat statistics in focus*.

SUMMARY

Measuring changing in Europe: from Lisbon Strategy to EU2020

In the recent years a large literature on indicators of wellbeing and quality of life has proliferated. The commonality is the departure from a strict economic view of these phenomena. Gross domestic product or gross national income – as also indicated by Simon Kuznets - are not good thermometers of the wellbeing levels of a society, but only rough measures of the value of commodities produced in a country in one year. In this paper we intend to review this international literature, drawing examples of indicators proposed in European Strategy 2020. Europe 2020, a strategy for jobs and smart, sustainable and inclusive growth, is based on five EU headline targets which are currently measured by eight headline indicators. While all these indicators sign an important step for the recognition of wellbeing as a multidimensional phenomenon, they are often rooted in very different approaches – when we can identify a relevant “theoretical” framework standing behind an indicator -, such as basic needs approach, primary goods approach, happiness approach, or capability approach.

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

- 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).
- XLIX Mobilità e sviluppo: il ruolo del turismo. (San Benedetto del Tronto, 24-26 maggio 2012).
- 50esima Trasformazioni economiche e sociali agli inizi del terzo millennio: analisi e prospettive (Università Europea di Roma, 29-31 maggio 2013).

B) GIORNATE DI STUDIO

- 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

