

DIPARTIMENTO DI
SCIENZE ECONOMICHE E METODI MATEMATICI

METODI E ANALISI STATISTICHE

2016



UNIVERSITÀ
DEGLI STUDI DI BARI
ALDO MORO

DIPARTIMENTO DI
SCIENZE ECONOMICHE E METODI MATEMATICI

METODI E ANALISI STATISTICHE

2016



Tutti i diritti di traduzione, riproduzione e adattamento,
totale o parziale, con qualsiasi mezzo (comprese
le copie fotostatiche e i microfilm) sono riservati

© Copyright 2016 by Università degli Studi di Bari Aldo Moro
www.uniba.it

Prima edizione: dicembre 2016

ISBN 978-88-88793-99-3

Coordinamento scientifico:

Ernesto Toma
Francesco D. d'Ovidio

Editing finale: F. D. d'Ovidio

Sommario

Ernesto Toma Presentazione	pag. 5
Giovanni Girone, Lucia Mongelli L'utilizzo delle terapie non convenzionali in Italia	« 7
Giuseppina Sacco, Pietro Sacco Il panorama dell'invecchiamento demografico nell'attuale situazione italiana	« 19
Lucia Mongelli Il ricorso ai servizi sanitari in Italia e in Puglia	« 35
Stefania Taralli, Monica Carbonara Il benessere equo e sostenibile dei territori pugliesi tra multidimensionalità e sintesi	« 53
Vito Ricci Antroponimia a Ruvo di Puglia tra XII e XIV secolo: alcune analisi statistiche	« 77
Najada Firza, Rossana Mancarella Degenze ospedaliere per malattie cardiovascolari in Puglia	« 91
Agata Maria Madia Carucci, Giovanni Vannella Sull'integrazione tra fonti amministrative e fonti statistiche per le imprese	« 105
Angela Maria D'Ugento Misurare la Salute: un'applicazione dell'indice di Mazziotta-Pareto	« 127
Caterina Marini, Vittorio Nicolardi L'aggiornamento di strutture di Contabilità Nazionale disaggregate	« 141
Laura Antonucci, Corrado Crocetta Un indicatore sintetico per il bilancio di genere dell'Università di Foggia	« 157
Rossana Mancarella Analisi dell'efficienza organizzativa nelle strutture sanitarie pugliesi tramite alberi di classificazione	« 167
Domenico Summo Total Factor Productivity Index	« 193

Presentazione

Il Dipartimento di Scienze Economiche e Metodi Matematici cura da molti anni una collana di “working paper” su temi economici, SERIES (*Southern Europe Research in Economic Studies*), a cui si sono affiancate nel corso degli anni altre collane di pubblicazioni edite a cura delle diverse sezioni che incarnano le “anime” di tale multidisciplinare Istituzione: la collana *Geografia Economica*, i *Saggi di Storia Economica* e infine gli *Studi di Diritto Pubblico*. Di tali collane si possono trovare sul portale del Dipartimento (<http://www.uniba.it/ricerca/dipartimenti/dse/ricerca>) i principali riferimenti bibliografici, e per molti l'intero testo in libera consultazione in file PDF.

Da un quinquennio al Dipartimento si è aggregata anche la sezione statistica, costola del precedente Dipartimento di Scienze Statistiche “Carlo Cecchi”; a parere degli studiosi afferenti a tale branca scientifica sembra ormai ora di riprendere la pubblicazione degli *Annali* che tale Dipartimento ha curato per un decennio, in prosecuzione della precedente collana degli *Annali dell'Istituto di Statistica*, creata nel 1927 e ben rinomata nel settore. Il nuovo nome di tale collana (che avrà cadenza non necessariamente annuale) è rappresentato in questo volume: *Metodi e Applicazioni Statistiche*. Anche se tale collana ha, al momento, soprattutto lo scopo di racchiudere i primi risultati di ricerche in corso, anche di studiosi esterni al Dipartimento, essa potrà svilupparsi in futuro verso ruoli più centrali nella ricerca scientifica.

In questo volume sono presentati in ordine sparso: articoli in tema di statistiche sanitarie, che talvolta sfiorano nella statistica medica (i lavori di Girone & Mongelli, Mongelli, Firza & Mancarella, D'Uggento, Mancarella), di statistica economica (Carucci & Vannella, Marini & Nicolardi, ma anche il lungo saggio di Summo), di statistiche sociali e assimilate (Taralli & Carbonara, Antonucci & Crocetta), di demografia (Sacco & Sacco), e persino su temi quasi storico-etimologici (la ricerca di Ricci).

Alcuni di questi articoli rappresentano già un lavoro compiuto, altri sono “work in progress”, ma tutti affrontano con competenza temi di interesse. Sperando che il volume incontri il favore dei lettori, colgo l'occasione per ringraziare tutti coloro che vi hanno contribuito, a partire dagli Autori ma senza dimenticare i *referees* anonimi che hanno consentito una adeguata revisione degli articoli qui proposti.

Bari, 29/12/2016

Il Direttore del Dipartimento di
Scienze Economiche e Metodi Matematici
Ernesto Toma

L'utilizzo delle terapie non convenzionali in Italia

Giovanni Girone¹, Lucia Mongelli^{2*}

¹ Università "Aldo Moro" di Bari (Professore emerito),

² Istat sede territoriale per la Puglia e Università "Aldo Moro" di Bari

Riassunto: In questa nota vengono analizzati i dati sul ricorso alle terapie non convenzionali in Italia in relazione ai caratteri anagrafici, nonché le motivazioni della loro adozione o alla loro rinuncia.

Keywords: Terapie non convenzionali; agopuntura; omeopatia; fitoterapia; trattamenti manuali; altre terapie non convenzionali.

1. Premessa

La disponibilità di dati aggiornati rappresenta una condizione indispensabile per l'analisi dello stato di salute di una popolazione. L'Istat, con il sostegno del Ministero della Salute e delle Regioni, con la Regione Piemonte come capofila, ha realizzato l'indagine multiscopo "Condizioni di salute e ricorso ai servizi sanitari", condotta tra il 2012 e il 2013. L'indagine permette di comprendere le condizioni di salute, le modalità di accesso ai servizi sanitari, i comportamenti di tutela della salute e di produrre stime a livello regionale.

L'indagine è campionaria e i dati sono stati raccolti direttamente presso gli individui. La numerosità campionaria è di circa 120 mila intervistati.

In questa nota ci limiteremo a considerare un tema sinora poco analizzato, il ricorso degli intervistati alle terapie non convenzionali (agopuntura, omeopatia, fitoterapia, trattamenti manuali, altre terapie non convenzionali). I numeri di coloro

* Corresponding author: mongelli@istat.it

Il presente articolo è frutto del lavoro comune degli Autori, ma è attribuibile a G. Girone il par. 1 e il par 10, a L. Mongelli i par. 2-9.

che hanno fatto uso di terapie non convenzionali negli ultimi tre anni (12.659) o negli ultimi 12 mesi (10.840) assicurano una buona affidabilità delle stime.

2. L'incidenza dell'uso delle terapie non convenzionali in Italia

Nella seguente Tab. 1 sono riportate le incidenze per mille intervistati di coloro che in Italia negli ultimi 3 anni e negli ultimi 12 mesi hanno fatto uso di terapie non convenzionali. D'ora in avanti, nel caso di presenza di risultati riferiti ai due periodi, limiteremo il nostro commento a quelli con riferimento al periodo più ampio perché sono più stabili.

Come può vedersi, la terapia non convenzionale più diffusa è l'omeopatia 40,1% degli intervistati, seguono i trattamenti manuali 34,6%, al terzo posto troviamo la fitoterapia 19,0%, al quarto l'agopuntura 10,6%, all'ultimo posto le altre terapie solo il 2,0%. Non si tratta di cifre trascurabili atteso che una persona ogni 25, ossia 2,4 milioni di abitanti utilizza l'omeopatia in Italia, all'altro estremo una persona ogni 100, ossia 600 mila fanno uso dell'agopuntura.

Tabella 1. *Soggetti che in Italia hanno fatto uso delle terapie non convenzionali ogni 1.000 intervistati*

Uso negli ultimi tre anni				
Agopuntura	Omeopatia	Fitoterapia	Trattamenti manuali	Altre terapie
10,6	40,1	19,0	34,6	2,0
Uso negli ultimi 12 mesi				
Agopuntura	Omeopatia	Fitoterapia	Trattamenti manuali	Altre terapie
7,7	36,3	16,7	28,6	1,7

3. L'incidenza dell'utilizzo delle terapie non convenzionali in Italia per sesso

Nella seguente Tab. 2 sono riportate le incidenze per mille intervistati di coloro che in Italia negli ultimi 3 anni e negli ultimi 12 mesi hanno fatto uso di terapie non convenzionali distintamente per maschi e femmine. Come può vedersi agevolmente, l'uso di tutte le terapie non convenzionali è più diffuso tra le donne che tra gli uomini.

Il divario è particolarmente marcato per l'utilizzo dell'omeopatia 50,0% per le donne contro il 29,5% per gli uomini, ossia il 69,5% in più per le prime, per la fitoterapia il 23,8% per le donne contro il 13,9% per gli uomini, ossia il 71,2% in più

per le prime, per l'agopuntura utilizzata dal 13,2‰ delle donne e dal 7,7‰ degli uomini, ossia il 71,4% in più dalle prime rispetto ai secondi.

Più contenuto il divario per le altre due terapie non convenzionali: il 38,7‰ delle donne contro il 30,1‰ degli uomini per l'utilizzo dei trattamenti manuali, ossia il 28,8% in più per le prime ed infine il 2,4‰ per le donne contro l'1,6‰ degli uomini per le altre terapie non convenzionali, ossia il 50% in più per le donne rispetto agli uomini.

Tabella 2. *Soggetti che in Italia hanno fatto uso delle terapie non convenzionali ogni 1.000 intervistati, per sesso*

Sesso	Agopuntura	Omeopatia	Fitoterapia	Trattamenti manuali	Altre terapie
	Usò negli ultimi tre anni				
Maschi	7,7	29,5	13,9	30,1	1,6
Femmine	13,2	50,0	23,8	38,7	2,4
Totale	10,6	40,1	19,0	34,6	2,0
	Usò negli ultimi 12 mesi				
	Maschi	5,3	26,7	11,8	24,6
Femmine	10,0	45,2	21,4	32,3	2,1
Totale	7,7	36,3	16,7	28,6	1,7

4. L'incidenza dell'utilizzo delle terapie non convenzionali in Italia per fasce di età

Nella seguente Tab. 3 sono riportate le incidenze per mille intervistati di coloro che in Italia negli ultimi 3 anni e negli ultimi 12 mesi hanno fatto uso di terapie non convenzionali distintamente per fasce di età.

L'incidenza dell'utilizzo delle diverse terapie non convenzionali negli ultimi 3 anni in Italia presenta 3 tipi di andamento: il primo dapprima crescente e successivamente decrescente con l'età (fitoterapia, trattamenti manuali), il secondo con andamento sempre crescente con l'età (agopuntura e altre terapie), il terzo con andamento alterno con l'età.

L'incidenza della fitoterapia parte dal 12,5‰ della prima fascia di età (fino 14 anni), sale fino a quasi raddoppiarsi 24,0‰ nella classe di età 30-49 anni, per calare al 19,4‰ nell'ultima classe di età. L'incidenza dei trattamenti manuali parte dal 6,9‰ della classe di età fino a 14 anni, sale fino al 44,0‰ della classe 30-49 anni e scende, sia pur di poco, al 41,1‰ dell'ultima classe di età.

L'utilizzo dell'agopuntura parte da un livello modesto (0,6‰) nella prima fascia di età e cresce sistematicamente sino al 14,8‰ dell'ultima fascia di età. Dinamica sempre crescente con l'età, sia pure su livelli modesti, anche per le altre terapie: dallo 0,6‰ della classe fino a 14 anni si sale sino al 2,7‰ degli ultracinquantenni.

Per l'omeopatia l'incidenza più elevata (59,5‰) si ha per la classe iniziale, si scende poi al 34,4‰ nella classe 15-29 anni, si risale al 51,2‰ nella classe 30-49 anni, per calare al 28,8‰ per la classe delle età più elevate.

Tabella 3. *Soggetti che in Italia hanno fatto uso delle terapie non convenzionali ogni 1.000 intervistati, per classi di età*

Classi di età	Agopuntura	Omeopatia	Fitoterapia	Trattamenti manuali	Altre terapie
	Uso negli ultimi tre anni				
Fino a 14	0,6	59,5	12,5	6,9	0,6
15-29	4,4	34,4	14,4	22,6	0,9
30-49	12,0	51,2	24,0	44,0	2,2
50 e più	14,8	28,8	19,4	41,1	2,7
Totale	10,6	40,1	19,0	34,6	2,0
	Uso negli ultimi 12 mesi				
Fino a 14	0,2	55,7	11,2	5,9	0,6
15-30	3,2	31,7	12,5	19,1	0,8
30-50	8,8	46,4	21,7	36,9	1,8
50 e più	11,0	25,1	16,6	33,4	2,4
Totale	7,7	36,3	16,7	28,6	1,7

5. L'incidenza dell'utilizzo delle terapie non convenzionali in Italia, per regioni

Nella seguente Tab. 4 sono riportate le incidenze per mille intervistati di coloro che in Italia negli ultimi 3 anni e negli ultimi 12 mesi hanno fatto uso di terapie non convenzionali, per regioni.

Le incidenze di quanti ne hanno fatto uso negli ultimi tre anni sono anche rappresentate nei 5 cartogrammi delle Fig. 1-5. È subito evidente per tutte le terapie non convenzionali una marcata differenziazione territoriale, con le incidenze più marcate nelle regioni del Centro-Nord e quelle molto più contenute nel Mezzogiorno. Per l'agopuntura le incidenze più elevate si hanno in Valle d'Aosta (40,0‰), in provincia di Bolzano (28,2‰) e nel Friuli Venezia Giulia (17,4‰), quelle più basse nel Molise (2,1‰), in Puglia (3,7‰) e in Sicilia (3,9‰).

Tabella 4. *Soggetti che in Italia hanno fatto uso delle terapie non convenzionali ogni 1.000 intervistati, per regione*

Regioni	Agopuntura	Omeopatia	Fitoterapia	Trattamenti manuali	Altre terapie
	Usò negli ultimi tre anni				
Piemonte	17,3	55,3	23,9	48,0	2,5
Valle d'Aosta	40,0	84,9	27,3	79,4	1,8
Lombardia	10,1	44,6	19,0	46,4	3,1
Bolzano	28,2	175,6	42,8	60,0	2,6
Trento	14,6	80,6	37,8	51,4	7,6
Veneto	11,7	59,7	28,3	49,8	3,9
Friuli Venezia Giulia	17,4	70,2	39,9	59,3	3,1
Liguria	14,6	46,2	19,4	29,3	0,9
Emilia Romagna	15,3	64,4	27,1	51,5	3,2
Toscana	15,1	48,1	24,4	38,2	3,1
Umbria	11,0	56,1	17,2	39,9	1,7
Marche	8,3	39,2	24,8	53,4	2,6
Lazio	10,2	44,6	20,5	31,7	1,1
Abruzzo	8,0	18,7	17,5	29,0	1,7
Molise	2,1	19,8	8,0	17,3	0,4
Campania	5,5	9,7	7,8	12,9	0,8
Puglia	3,7	8,1	8,9	16,8	1,4
Basilicata	6,8	6,8	5,9	11,1	0,3
Calabria	4,1	6,1	7,1	16,8	0,6
Sicilia	3,9	15,6	10,8	17,1	0,3
Sardegna	9,6	34,8	15,1	19,5	1,8
Italia	10,6	40,1	19,0	34,6	2,0
	Usò negli ultimi 12 mesi				
Piemonte	10,7	48,5	19,4	36,9	1,9
Valle d'Aosta	29,1	73,4	24,3	64,3	1,2
Lombardia	7,1	40,7	16,7	38,7	2,9
Bolzano	20,7	164,6	38,8	51,6	2,6
Trento	10,1	74,0	35,2	40,3	6,5
Veneto	8,4	55,2	26,4	41,7	3,2
Friuli Venezia Giulia	12,3	64,6	37,0	50,8	2,9
Liguria	11,7	43,0	17,6	24,9	0,5
Emilia Romagna	12,2	59,4	25,5	43,3	3,0
Toscana	11,5	44,9	21,0	32,2	3,1
Umbria	9,3	49,6	15,8	32,7	1,4
Marche	6,4	35,2	23,6	42,8	1,7
Lazio	7,8	39,0	17,6	26,5	0,8
Abruzzo	5,5	16,2	14,5	25,7	1,7
Molise	1,7	16,8	7,2	16,0	0,4
Campania	4,2	8,9	6,0	10,2	0,5
Puglia	2,8	7,1	7,2	13,8	1,4
Basilicata	6,2	5,9	4,9	9,8	0,3
Calabria	2,4	4,1	5,5	12,6	0,4
Sicilia	2,9	12,6	8,5	14,0	0,2
Sardegna	7,2	31,1	13,5	16,3	1,4
Italia	7,7	36,3	16,7	28,6	1,7

Figura 1. Incidenze % di quanti in Italia negli ultimi 3 anni hanno fatto uso di agopuntura.

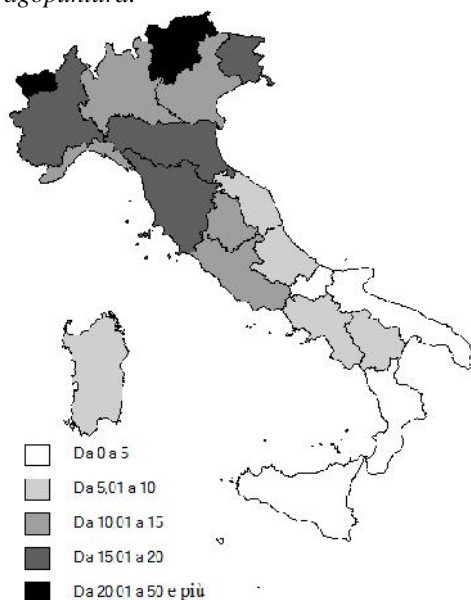


Figura 2. Incidenze % di quanti in Italia negli ultimi 3 anni hanno fatto uso di omeopatia.

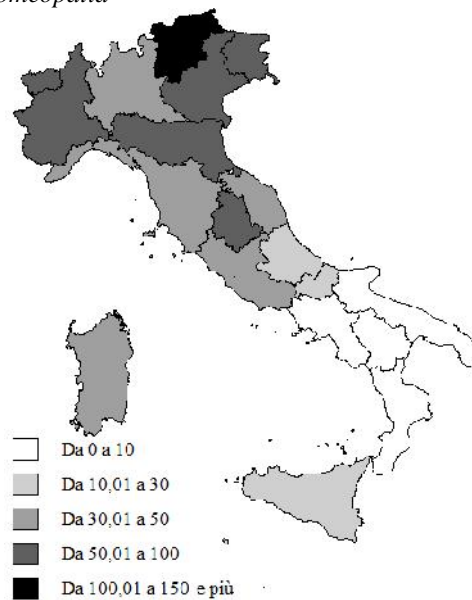


Figura 3. Incidenze % di quanti in Italia negli ultimi 3 anni hanno fatto uso di fitoterapia.

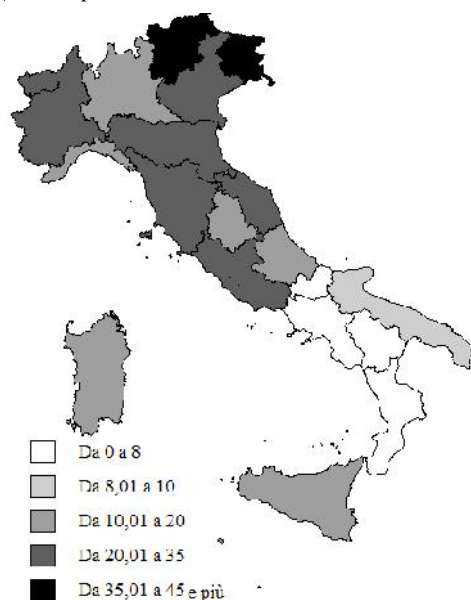


Figura 4. Incidenze % di quanti in Italia negli ultimi 3 anni hanno fatto uso di trattamenti manuali.

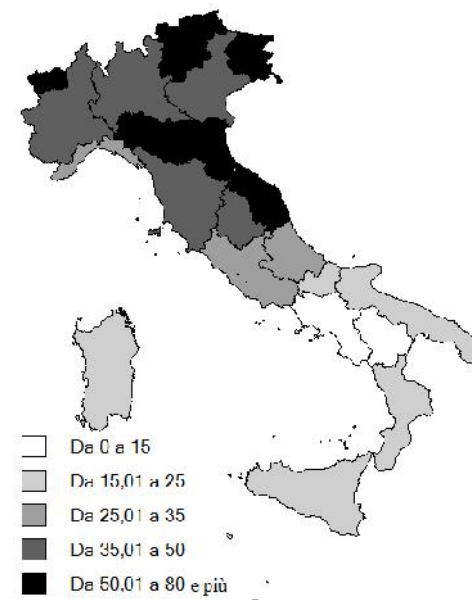
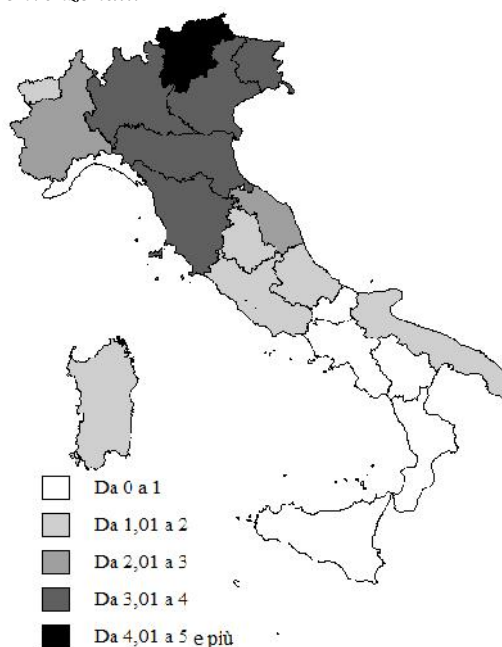


Figura 5. Incidenze % di quanti in Italia negli ultimi 3 anni hanno fatto uso di altre terapie non convenzionali.



Per l'omeopatia le incidenze più elevate si registrano in provincia di Bolzano (175,6‰), in Valle d'Aosta (84,9‰) e in provincia di Trento (80,0‰), quelle più contenute in Calabria (6,1‰), in Basilicata (6,8‰) e in Puglia (8,1‰).

Per la fitoterapia le incidenze più elevate si hanno in provincia di Bolzano (42,8‰), nel Friuli Venezia Giulia (39,9‰) e in provincia di Trento (37,8‰), quelle più contenute in Basilicata (5,9‰), in Calabria (7,1‰) e in Campania (8,0‰).

Le incidenze dei trattamenti manuali toccano i valori più elevati in Valle d'Aosta (79,4‰), in provincia di Bolzano (60,0‰) e nel Friuli Venezia Giulia (59,3‰), quelli più bassi in Basilicata (11,1‰), nella Campania (12,9‰) e in Puglia e in Calabria (16,8‰).

Infine, le altre terapie non convenzionali fanno registrare le incidenze più alte in provincia di Trento (7,6‰), nel Veneto (3,9‰) e in Liguria (3,2‰) e quelle più basse in Basilicata ed in Sicilia (0,3‰) e in Abruzzo (0,4‰).

La presenza sistematica dei valori più elevati solo nelle regioni settentrionali e dei valori più bassi solo nelle regioni meridionali è una conferma di una precisa stratificazione dell'utilizzo delle terapie non convenzionali.

7. Le motivazioni del continuare o sospendere l'utilizzo delle terapie non convenzionali in Italia

Nella seguente Tab.7 sono riportate le incidenze delle motivazioni che inducono gli intervistati che in Italia negli ultimi 12 mesi hanno fatto uso delle terapie non convenzionali a continuarne o a sospenderne l'utilizzo.

La motivazione più frequente a continuare ad utilizzare le terapie non convenzionali è costituita dai benefici avuti; ciò vale per tutte le terapie non convenzionali. Particolarmente per le altre terapie non convenzionali (84,4%), per i trattamenti manuali (83,2%) e per l'agopuntura (79,1%). Sempre prevalente, anche se su livelli meno elevati, la prevalenza della motivazione dei benefici ottenuti per la fitoterapia (58,0%) e per l'omeopatia (56,4%). Segue la motivazione della preferenza per i rimedi naturali con incidenze ancora elevate (4 ogni 10 di coloro che continueranno a utilizzare l'omeopatia e la fitoterapia). Più contenute le incidenze di coloro che intendono continuare l'utilizzo delle terapie non convenzionali perché ritenute rimedi più naturali per l'agopuntura (17,8%), per i trattamenti manuali (12,0%) e per le altre terapie (11,4%).

Nella stessa Tab.7 sono anche riportate le incidenze delle motivazioni che inducono gli intervistati a sospendere l'utilizzo delle terapie non convenzionali. L'incidenza della motivazione di assenza di benefici è prevalente per l'omeopatia (51,5%) e per l'agopuntura (47,1%), terapie per le quali le incidenze della motivazione dei costi elevati sono le più contenute (rispettivamente 16,5% e 14,2%). Le incidenze della motivazione per altri motivi sono prevalenti e particolarmente rilevanti per le altre tre terapie non convenzionali: fitoterapia (48,2%), trattamenti manuali (54,0%) e altre terapie non convenzionali (60,0%).

Tabella 7. % di coloro che hanno fatto uso di terapie non convenzionali negli ultimi 12 mesi e che hanno o non hanno intenzione di continuare a farne uso secondo la motivazione

	Agopuntura	Omeopatia	Fitoterapia	Trattamenti manuali	Altre terapie
Sì					
Perché ha avuto benefici	79,1	56,4	58,0	83,2	84,4
Perché preferiscono rimedi naturali	17,8	40,4	38,0	12,0	11,4
Per altri motivi	3,1	3,3	4,0	4,8	4,2
No					
Perché non ne ha avuto benefici	47,1	51,5	27,7	23,4	20,0
Perché costa troppo	14,2	16,5	24,1	22,5	20,0
Per altri motivi	38,7	32,1	48,2	54,0	60,0

8. Le motivazioni del continuare o sospendere l'utilizzo delle terapie non convenzionali in Italia, per sesso

Nella Tab. 8 sono riportate, distintamente per maschi e femmine, le incidenze degli intervistati che hanno fatto uso di terapie non convenzionali negli ultimi 12 mesi e che intendono o meno continuare a farne uso, nonché le incidenze delle relative motivazioni.

Per tutte le terapie la motivazione principale di continuare connessa ai benefici ottenuti prevale, sia pure di poco, negli uomini: per l'agopuntura 80,6% per i maschi contro il 78,5% delle femmine, per l'omeopatia 59,2% contro 54,8%, per la fitoterapia 60,0% contro il 57,0%, per i trattamenti manuali 86,6% contro 80,8% e per le altre terapie non convenzionali 92,2% contro 79,6%. Al secondo posto, la motivazione della preferenza dei rimedi naturali prevale tra le donne rispetto agli uomini: agopuntura 17,0% per gli uomini contro 18,1% delle donne, omeopatia 38,3% contro 41,5%, fitoterapia 36,3% contro 38,8%, trattamenti manuali 9,2% contro 13,9%, ed infine le altre terapie non convenzionali 6,3% contro 14,6%. Ancora prevalenza delle incidenze femminili per le altre motivazioni, cioè su cifre più contenute.

Tabella 8. % di coloro che in Italia hanno fatto uso di terapie non convenzionali negli ultimi 12 mesi e che hanno o non hanno intenzione di continuare a farne uso, secondo la motivazione ed il sesso

	Agopuntura		Omeopatia		Fitoterapia		Trattamenti manuali		Altre terapie	
	Maschi	Femmine	Maschi	Femmine	Maschi	Femmine	Maschi	Femmine	Maschi	Femmine
Si										
Perché ha avuto benefici	80,6	78,5	59,2	54,8	60,0	57,0	86,6	80,8	92,2	79,6
Perché preferisce rimedi naturali	17,0	18,1	38,3	41,5	36,3	38,8	9,2	13,9	6,3	14,6
Per altri motivi	2,4	3,4	2,4	3,7	3,6	4,2	4,2	5,3	1,6	5,8
No										
Perché non ha avuto benefici	41,4	50,3	54,3	49,7	37,3	21,2	19,6	26,6	26,7	16,0
Perché costa troppo	13,1	14,9	12,0	19,3	22,4	25,3	16,3	27,6	-	32,0
Per altri motivi	45,5	34,9	33,7	31,0	40,3	53,5	64,1	45,8	73,3	52,0

Veniamo, infine, all'analisi per genere delle incidenze delle motivazioni di non continuare ad utilizzare le terapie non convenzionali utilizzate negli ultimi 12 mesi.

L'assenza di benefici è indicata maggiormente dagli uomini che dalle donne per l'omeopatia (54,3% contro 49,7%), per la fitoterapia (37,3% contro 21,2%) e per le altre terapie non convenzionali (26,7% contro 16,0%). L'assenza di benefici è maggiormente indicata dalle donne per l'agopuntura (50,3% rispetto al 41,4% dei maschi) e per i trattamenti manuali (26,6% contro 19,6% dei maschi). La rinuncia a continuare l'utilizzo di terapie non convenzionali motivata dai costi della loro adozione è indicata maggiormente dalle donne: 14,9% contro il 13,1% per l'agopuntura, 19,3% contro 12,0% per l'omeopatia, 25,3% contro 22,4% per la fitoterapia, 27,6% contro 16,3% per i trattamenti manuali e infine 27,6% contro nessuna indicazione per le altre terapie non convenzionali. Elevate anche le incidenze degli altri motivi di rinuncia non specificati.

9. Incidenze di intervistati che hanno fatto uso di terapie non convenzionali negli ultimi 12 mesi in Italia che intendono continuare a farne uso, secondo classi di età

Nella seguente Tab. 9 sono riportate le incidenze di coloro che hanno fatto uso di terapie non convenzionali negli ultimi 12 mesi in Italia e che intendono continuare o meno a farne uso, secondo classi di età.

Emergono tre diversi andamenti: per l'agopuntura si registrano incidenze decrescenti con l'età, dalla totalità di continuare della fascia di età fino a 14 anni al 67,5% della fascia di età da 50 anni in su. Dinamiche alterne con tendenze incerte e di scarsa entità per le restanti terapie. Per l'omeopatia le incidenze di coloro che intendono continuare oscillano tra il 92,8% ed il 97,1%. Per la fitoterapia si registrano valori tra l'89,7% ed il 97,8%. Per i trattamenti manuali ancora andamento oscillante e valori più contenuti. Andamento analogo per le altre terapie non convenzionali

Tabella 9. % di coloro che in Italia hanno fatto uso di terapie non convenzionali negli ultimi 12 mesi e che hanno o non hanno intenzione di continuare a farne uso, per classi di età

	Agopuntura		Omeopatia		Fitoterapia		Trattamenti manuali		Altre	
	sì	no	sì	no	sì	no	sì	no	sì	no
Fino a 14	100,0	0,0	97,1	2,9	97,8	2,2	87,4	12,6	77,8	22,2
15-29	76,8	23,2	92,8	7,2	90,8	9,2	83,0	17,0	92,9	7,1
30-49	74,0	26,0	93,8	6,2	92,8	7,2	85,7	14,3	78,7	21,3
da 50 in poi	67,5	32,5	94,3	5,7	89,7	10,3	82,7	17,3	80,5	19,5
Totale	70,2	29,8	94,5	5,5	91,7	8,3	83,9	16,1	80,7	19,3

10. Conclusioni

Le incidenze di coloro che in Italia hanno fatto uso di terapie non convenzionali sono tutt'altro che trascurabili: si va dal 40,1% per l'omeopatia al 2,0% per le altre terapie non convenzionali; valori intermedi si registrano per i trattamenti manuali, per la fitoterapia e per l'agopuntura. L'utilizzo di terapie non convenzionali è di gran lunga più diffuso tra le donne che tra gli uomini. L'andamento in relazione all'età è abbastanza diversificato per le varie terapie. La diffusione delle terapie non convenzionali è largamente diffuso al Nord e scarsamente diffuso nel Mezzogiorno. La gran parte di quanti hanno fatto uso di terapie non convenzionali continuerà ad utilizzarle innanzitutto perché ne ha tratto benefici, in secondo luogo perché ritiene che trattasi di rimedi naturali. I pochi che rinunciano motivano la decisione innanzitutto per l'assenza di benefici e in subordine per i costi elevati.

Riferimenti bibliografici

- Federazione Nazionale degli Ordini dei Medici Chirurghi e Odontoiatri, Commissione per le pratiche alternative (1999). *Dal primato della medicina scientifica al confronto con le pratiche alternative*. Rivista SIMG, n. 2.
- Federazione Nazionale degli Ordini dei Medici Chirurghi e Odontoiatri (2002). Le linee guida sulle medicine e pratiche non convenzionali. *La Professione* 5/6:5.
- Fontanarosa P.B., Lundberg G.D., (1998). *Alternative medicine meets science*. JAMA 280:1618–1619.
- Giarelli G, Roberti di Sarsina P, Silvestrini B. (2007). *Le medicine non convenzionali in Italia*. Franco Angeli, Milano.
- Klingemann H.-G. (2000). *Terapie non convenzionali nella cura del cancro. L'efficacia dei trattamenti alternativi. Il metodo globale*. Hermes Edizioni, Roma.
- Pucci E., Cartechini E., Taus C., Giuliani G. (2004). Why physicians need to look more closely at the use of complementary and alternative medicine by multiple sclerosis patients. *European Journal of Neurology*, 11:263–267.
- Schmidt K., Jacobs P.A., Barton A. (2002). Cross-cultural differences in GP's attitudes towards complementary and alternative medicine: a survey comparing regions of the UK and Germany. *Complementary Therapy Medicine* 10:141–147.
- Società Italiana di Medicina Interna (2000). *Le medicine alternative*, Riv. Med. Lab. – JLM, vol. 1, N 4.

Fonte

Istat – Condizioni di salute e ricorso ai servizi sanitari. Anno 2013. File standard - elaborazione dei dati: 15/01/2016

Il panorama dell'invecchiamento demografico nell'attuale situazione italiana

Giuseppina Sacco, Pietro Sacco

*Università degli Studi di Bari Aldo Moro
Dipartimento di Scienze Economiche e Metodi Matematici*

Riassunto: Il fenomeno di maggior rilievo che ha caratterizzato la struttura della popolazione italiana del XX secolo ed ancor più contraddistinguerà il XXI secolo è l'invecchiamento della popolazione. Per "invecchiamento demografico" si intende l'aumento dell'incidenza della popolazione anziana sulla popolazione totale. Detto fenomeno per i paesi a sviluppo avanzato, in particolare per l'Italia degli ultimi 50 anni, viene assumendo valori sempre più elevati. Ciò pur essendo "un trionfo dello sviluppo e quindi uno dei grandi successi dell'umanità" tuttavia desta non poche preoccupazioni poiché implica conseguenze di ampia portata per tutti i settori della società prospettando un gran numero di sfide da dover affrontare. Invero, le sfide riguarderanno il sistema economico e produttivo ma, al tempo stesso, ciò potrebbe implicare anche opportunità, da tenere in debito conto, derivanti da un nuovo equilibrio numerico tra le generazioni. In questo saggio intendiamo focalizzare l'attenzione sul fenomeno invecchiamento relativo al periodo che va dal censimento del 1951 al 1° gennaio 2015. Solitamente gli indici che si costruiscono per la misura del grado di invecchiamento di una popolazione sono l'indice di vecchiaia dato dal rapporto tra gli ultrasessantacinquenni e la popolazione giovanile e l'indice di invecchiamento dato dal rapporto tra gli ultrasessantacinquenni e la popolazione totale. Nella presente nota utilizzeremo quest'ultimo indicatore, ritenuto adeguatamente sufficiente a stabilire se una popolazione è "vecchia".

Keywords: Invecchiamento demografico; Italia; incremento.

1. Evoluzione del grado di invecchiamento della popolazione italiana

L'evoluzione del grado di invecchiamento della popolazione italiana considerata nella presente nota, è stata osservata in due momenti temporali: dal passato al presente mediante l'analisi della "velocità" con cui il fenomeno si è manifestato nel

periodo che ha compreso il censimento del 1951 ed il 1° gennaio 2015; dal passato al futuro attraverso la stima di come si evolverebbero gli indici di invecchiamento mediante lo studio del trend osservato nel passato, sotto l'ipotesi che le condizioni che hanno determinato l'evoluzione del fenomeno osservato nel tempo passato, permangano anche nel futuro. Prima di entrare nel dettaglio delle misure dell'evoluzione temporale del fenomeno in osservazione, è opportuno procedere ad un confronto spaziale necessario per capire il livello raggiunto attualmente dal fenomeno "invecchiamento" in Italia confrontando lo stesso con gli indici riguardanti l'invecchiamento della popolazione di alcuni paesi sviluppati al 2014 (cfr. Tab.1).

Stimato che una popolazione può considerarsi "vecchia" se il suo grado d'invecchiamento supera il valore del 10%, si evince, dall'analisi tabellare, che tutti i paesi indicati hanno ormai raggiunto un livello d'invecchiamento critico. In particolare l'Italia, con un indicatore pari al 22,0%, è, subito dopo il Giappone, tra i paesi quello col più alto livello di invecchiamento demografico al mondo.

Tabella 1. *Indici di invecchiamento in alcuni paesi del mondo - 2014*

Paesi	Popolazione (milioni di abitanti)	$I_v = \frac{P_{65-\xi}}{P_{0-\xi}} \cdot 100$
Stati Uniti	319	14,4
Giappone	127	25,7
Federazione russa	144	13,2
Regno Unito	64	17,5
Francia	66	18,7
Germania	81	21,1
Italia	61	22,0

Fonte: Banca Mondiale

1.1 "Velocità" del grado di invecchiamento

Per studiare la "velocità" di variazione che hanno manifestato gli indici di invecchiamento nel periodo di tempo in osservazione, abbiamo calcolato il ritmo di variazione, inteso come una "velocità media", attraverso il seguente modello di crescita geometrico:

$$I_{v,t} = I_{v,0}(1+r)^t$$

ove $I_{v,0}$ e $I_{v,t}$ sono gli indici di invecchiamento della popolazione ai tempi 0 e t mentre r rappresenta la variazione media annua dell'indice di invecchiamento da un anno all'altro. Dalla formula precedente si deduce:

$$r = \sqrt[t]{\frac{I_{v,t}}{I_{v,0}}} - 1 .$$

Per avere una visione completa del grado di invecchiamento nel nostro paese, abbiamo calcolato gli indici d'invecchiamento e la relativa velocità media, per entrambi i sessi e per tutte le regioni italiane, nel periodo compreso tra il censimento del 4/11/1951 e il 1° gennaio 2015. Le Tabb. 2 e 3 racchiudono gli indici di invecchiamento della popolazione italiana in complesso (MF) e, distintamente per tutte le regioni italiane, quelli maschili (M) e femminili (F) nonché la “velocità” r vale a dire il tasso con cui è cresciuto il livello d'invecchiamento dal 1951 all'1.1.2015.

Osservando i dati della Tab.2 si nota che l'indice di invecchiamento della popolazione italiana in complesso (MF) nel 1951 era dell'8,2%, mentre all'1.1.2015 è passato al 21,7%: un aumento notevole che posiziona l'Italia tra i paesi più “vecchi” al mondo. Parallelamente il livello d'invecchiamento della popolazione maschile (M) è passato dal 7,6% al 19,2%, mentre quello della compagine femminile (F), che era nel 1951 pari all'8,8%, nel 2015 ha raggiunto ben il 24,1%.

Inoltre, la geografia dell'invecchiamento marca la Liguria (28,0%) come la regione con il più alto livello di invecchiamento per entrambi i sessi, mentre la regione con il più basso indice di invecchiamento si rivela essere la Campania (17,6%).

Tabella 2. *Indici di invecchiamento della popolazione delle regioni italiane*

Regioni	M		F		MF	
	1951	1.1.2015	1951	1.1.2015	1951	1.1.2015
Piemonte	10,5	21,7	12,2	27,0	11,4	24,5
Valle d'Aosta	8,7	20,0	9,5	24,9	9,1	22,5
Liguria	10,0	24,5	11,3	31,1	10,6	28,0
Lombardia	7,4	18,8	8,7	24,3	8,0	21,6
Trentino A.A.	7,4	17,4	8,9	22,2	8,2	19,8
Veneto	7,0	19,0	8,3	24,2	7,7	21,7
Friuli Venezia Giulia	7,9	22,0	10,4	28,1	9,2	25,1
Emilia-Romagna	8,1	20,8	9,3	25,9	8,7	23,4
Toscana	9,2	22,1	10,4	27,2	9,8	24,8
Umbria	8,0	22,0	8,4	26,9	8,2	24,6
Marche	7,4	21,1	8,8	26,2	8,1	23,7
Lazio	6,3	18,3	7,6	22,9	7,0	20,7
Abruzzo	7,7	20,2	8,8	24,9	8,3	22,6
Molise	7,7	20,7	8,8	25,9	8,2	23,4
Campania	5,8	15,6	7,2	19,4	6,5	17,6
Puglia	6,8	18,4	7,4	22,4	7,1	20,5
Basilicata	6,2	19,3	7,0	23,9	6,6	21,6
Calabria	6,1	18,4	7,4	22,0	6,8	20,2
Sicilia	7,5	17,8	8,4	21,9	8,0	19,9
Sardegna	7,5	19,2	8,1	23,8	7,8	21,6
Italia	7,6	19,2	8,8	24,1	8,2	21,7

Fonte: Istat

Degni di nota sono i singoli indici regionali di invecchiamento della popolazione femminile che appaiono per l'intero periodo osservato, sempre più alti di quelli della popolazione maschile.

I dati riportati nella Tab. 3, relativi alla velocità d'invecchiamento, mostrano che, per la popolazione complessiva (MF) del nostro paese questa è stata del 15,6%. In altre parole, l'indice di invecchiamento riscontrato nel 1951 è aumentato in media, seguendo il modello di accrescimento geometrico, di oltre un punto e mezzo percentuale ogni anno.

Tabella 3. Velocità d'invecchiamento della popolazione delle regioni italiane

Regioni	M	F	MF
Piemonte	11,6	12,7	12,2
Valle d'Aosta	13,2	15,4	14,4
Liguria	14,3	16,2	15,5
Lombardia	14,9	16,4	15,9
Trentino A.A.	13,6	14,6	14,1
Veneto	15,9	17,1	16,5
Friuli Venezia Giulia	16,3	15,8	16,0
Emilia-Romagna	15,1	16,4	15,8
Toscana	14,0	15,4	14,8
Umbria	16,2	18,6	17,5
Marche	16,7	17,4	17,2
Lazio	17,0	17,6	17,3
Abruzzo	15,4	16,6	16,0
Molise	15,8	17,2	16,7
Campania	15,8	15,8	15,9
Puglia	15,9	17,7	16,9
Basilicata	18,1	19,6	19,0
Calabria	17,6	17,4	17,4
Sicilia	13,8	15,3	14,5
Sardegna	15,0	17,2	16,2
Italia	14,8	16,1	15,6

Fonte: Nostre elaborazioni su dati Istat

E' da osservare che la Basilicata nel 1951 aveva il più basso indice di invecchiamento (6,6%). Al 2015 essa risulta la regione con la velocità maggiore di invecchiamento (19,0%) a differenza del Piemonte che, al censimento del 1951, presentava il più alto indice di invecchiamento della popolazione (11,4%) subendo tuttavia, nell'intervallo considerato, un maggiore rallentamento. Risultati, questi, che confermano l'ipotesi in base alla quale la tradizionale divaricazione demografica tra Nord e Sud potrebbe derivare soprattutto dalla difformità nei modelli sociali ed economici. Invero, un sistema di welfare robusto consente uno sviluppo territoriale maggiore anche a livello sanitario.

Apprezzabile è il risultato dei dati relativi ai tassi di variazione media annua dell'indice di invecchiamento ottenuto per le F: difatti i valori sono quasi sempre più alti di quelli dei M. Un'ulteriore conferma che la compagine femminile ha contribuito all'aumento del livello d'invecchiamento di tutta la popolazione italiana.

1.2 "Stima" futura del grado di invecchiamento

Valutati i livelli critici dell'invecchiamento raggiunti dalla popolazione italiana, nel periodo compreso tra il censimento del 1951 ed il 1° gennaio 2015, sarebbe interessante stimare come si potrebbe evolvere il fenomeno nel futuro.

Avendo a disposizione la serie storica dei predetti indici negli anni compresi tra il censimento del 1951 e quello del 2011, elencati nella Tab. 4 e rappresentati graficamente nella Fig.1, abbiamo eseguito una estrapolazione per gli anni futuri dal 2021 al 2041.

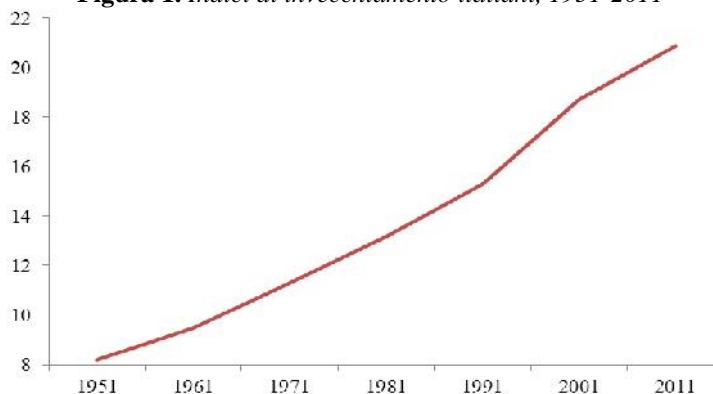
A tal proposito, non è superfluo sottolineare che i dati mostrano come la nostra popolazione, fino al censimento del 1961, con un indice di invecchiamento pari al 9,5% poteva considerarsi ancora "giovane". Nel tempo lo stesso indice viene aumentando sino a pervenire, al censimento del 2011, ad un valore di circa il 21%

Tabella 4. *Indici di invecchiamento in Italia. 1951-2011*

Anni di censimento	Valore dell'indice
1951	8,2
1961	9,5
1971	11,3
1981	13,2
1991	15,3
2001	18,7
2011	20,9

Fonte: Istat

Figura 1. *Indici di invecchiamento italiani, 1951-2011*



Per effettuare l'extrapolazione per gli anni futuri è necessario procedere alla scelta della funzione idonea a rappresentare l'evoluzione nel tempo dell'indice di invecchiamento.

Dalla rappresentazione grafica è intuitivo dedurre che la curva degli indici di invecchiamento può essere espressa, in forma analitica, da una funzione del tipo

$$y = \alpha \cdot \beta^x,$$

dove y è l'indice di invecchiamento relativo all'anno x mentre α e β sono i due parametri della funzione esponenziale¹. Per il calcolo di tali parametri occorre trasformare $y = \alpha \cdot \beta^x$ in una funzione lineare attraverso i logaritmi naturali.

Linearizzando si ha infatti

$$\ln y = \ln \alpha + x \cdot \ln \beta,$$

da cui, posto

$$\ln y = T \quad \ln \alpha = A \quad \ln \beta = B,$$

si può scrivere

$$T = A + Bx.$$

Per determinare i valori di A e B possiamo ricorrere al metodo dei minimi quadrati. Dopo aver posto $x=0, 1, \dots, 6$, in corrispondenza, rispettivamente, del 1951, 1961...2011, ponendo detti censimenti equidistanti, si ha la funzione interpolante cercata

$$y^* = 8,2 \cdot 1,17^x.$$

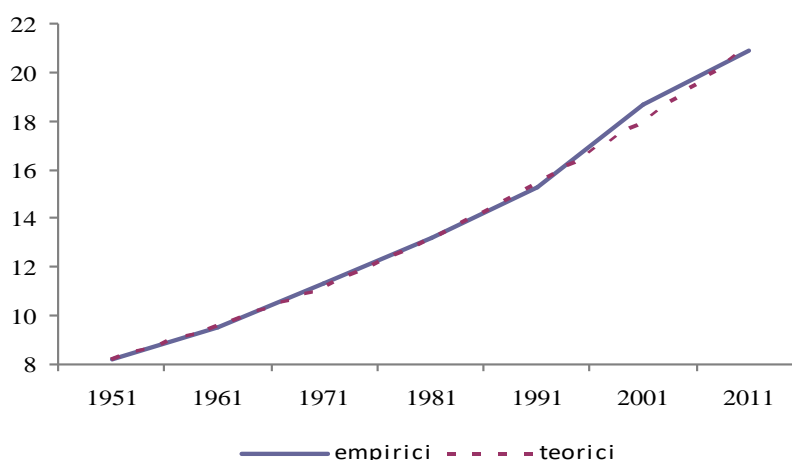
Sostituendo in essa $x=0, 1, \dots, 6$ si hanno, appunto, gli indici di invecchiamento teorici nell'intervallo di tempo 1951-2011, posti a confronto con i valori empirici in Tab.5. (cfr. anche Fig. 2)

Tabella 5. *Indici di invecchiamento empirici e teorici. 1951-2011*

Anni di censimento	Indici empirici	Indici teorici
1951	8,2	8,2
1961	9,5	9,6
1971	11,3	11,2
1981	13,2	13,1
1991	15,3	15,4
2001	18,7	18,0
2011	20,9	21,0

Fonte: Nostre elaborazioni su dati Istat

¹ A tal proposito è stato effettuato un confronto con l'interpolazione della funzione lineare, che ha presentato un indice di scostamento quadratico del 4,3%, sensibilmente maggiore di quello risultante dall'applicazione della funzione esponenziale (2,1%).

Figura 2: *Indici di invecchiamento teorici ed empirici, 1951-2011*

Va osservato che la funzione interpolante si adatta benissimo alla curva empirica degli indici di invecchiamento, in quanto l'indice di scostamento quadratico da noi calcolato risulta pari a 2,1%. Quindi possiamo asserire che l'adattamento, essendo molto inferiore al 10%, è ottimo. La funzione interpolante scelta conferma dunque che l'indice di invecchiamento nel nostro paese tende a crescere in progressione maggiore di quella lineare.

Inoltre, facendo riferimento sempre alla funzione interpolante, derivando rispetto a x e dividendo per y , si ha il *tasso di incremento istantaneo* relativo di y :

$$\frac{y'}{y} = \ln 1,17 = 0,15700 ,$$

ossia un incremento istantaneo relativo di circa il 16%.

A questo punto, con opportune sostituzioni (cfr. Tab. 6), otteniamo gli indici di invecchiamento teorici futuri, rispettivamente per il 2021 (24,6%), per il 2031 (28,8%) e per il 2041 (33,7%). Detti valori mostrano un incremento percentuale, al 2041, di oltre il 60% rispetto al 2011.

Tabella 6. *Estrapolazione futura degli indici di invecchiamento*

Date	Valore dell'indice
2001	18,0
2011	21,0
2021	24,6
2031	28,8
2041	33,7

Fonte: Nostre elaborazioni su dati Istat

2. Conseguenze del grado d'invecchiamento: modificazioni strutturali

L'andamento crescente del grado di invecchiamento giustifica le modificazioni che intervengono nella struttura per sesso e per età di una popolazione.

Infatti, la struttura per sesso ed età della popolazione italiana all'1.1.2015 e quindi il suo livello d'invecchiamento è dipesa dai seguenti fattori: dalla struttura per sesso ed età al censimento del 1951, dall'evoluzione della fecondità e della mortalità in rapporto all'età dal 1951 all'1.1.2015 e dalle migrazioni avvenute nello stesso periodo in osservazione.

Non considerando il fenomeno migratorio, fisiologicamente instabile, possiamo affermare che l'aumento del processo d'invecchiamento della popolazione italiana è determinato dall'andamento, nel periodo considerato, della fecondità e della mortalità per età.

In questa sede, però, tratteremo esclusivamente delle modificazioni intervenute nella struttura per età e sesso dovute all'incidenza della fecondità e della mortalità rimandando ad ulteriori approfondimenti il percorso evolutivo delle componenti naturali già citate.

Convenzionalmente la struttura per età di una popolazione è divisa in tre grandi classi: la classe 0-14 relativa al periodo della formazione, la classe 15-64 riferita al periodo dell'attività lavorativa e quella 65+ al pensionamento. Nella Tab.7, è riportata la distribuzione percentuale della popolazione, appartenente alle tre grandi classi, relativa ai censimenti intercorrenti tra il 1951 ed il 2011 e quella della stima della popolazione effettuata dall'Istat all'1.1.2015.

Dall'analisi tabellare possiamo osservare come, nel tempo, la proporzione dei giovani tende a decrescere a fronte della percentuale altalenante di popolazione alle età centrali che non presenta variazioni sensibili né nel tempo né nello spazio. Mentre la popolazione anziana presenta incrementi crescenti.

Tabella 7. Distribuzione percentuale della popolazione per grandi classi, 1951-1.1.2015

Anni	Classi di età			Totale
	0-14	15-64	65+	
1951	26,1	65,7	8,2	100,0
1961	24,5	66,0	9,5	100,0
1971	24,4	64,3	11,3	100,0
1981	21,4	65,4	13,2	100,0
1991	15,9	68,8	15,3	100,0
2001	14,2	67,1	18,7	100,0
2011	14,0	65,1	20,9	100,0
1.1.2015	13,8	64,5	21,7	100,0

Fonte: Nostre elaborazioni su dati Istat

Tale situazione è dovuta all'azione congiunta della caduta della fecondità e all'aumento della sopravvivenza. Questo è posto in luce, come già visto, anche dagli indici d'invecchiamento sempre crescenti dal 1951 all'1.1.2015. In dettaglio, all'ultima data osservata, il valore è pari al 21,7%. Vale a dire che ci sono circa 22 anziani di oltre 65 anni per 100 unità appartenenti alla popolazione totale. Ciò esprime che in 63 anni di storia demografica italiana l'indice di invecchiamento si è incrementato di circa il 165% con la conseguenza di un "forte invecchiamento" della popolazione.

Degno di interesse (cfr. Fig. 3) è il punto d'incontro tra le curve dei giovanissimi (classe 0-14 anni) e degli anziani (classe 65+). Ciò simboleggia l'uguaglianza raggiunta nel 1991 dalla proporzione dei giovanissimi e quella degli over 65.

Figura 3. Dinamiche temporali di giovani ed anziani in Italia

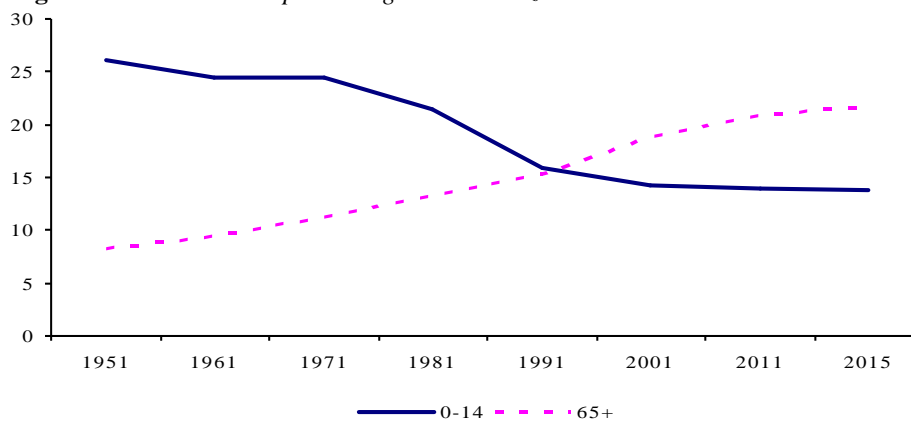


Tabella 8. Distribuzione percentuale della popolazione M e F per grandi classi; 1951-1.1.2015

Anni	Classi di età						Totale
	0-14		15-64		65+		
	M	F	M	F	M	F	
1951	13,3	12,8	31,9	33,8	3,7	4,5	100,0
1961	12,5	12,0	32,4	33,6	4,0	5,5	100,0
1971	12,5	11,9	31,7	32,6	4,7	6,6	100,0
1981	11,0	10,4	32,2	33,2	5,4	7,8	100,0
1991	8,1	7,8	34,2	34,6	6,2	9,1	100,0
2001	7,3	6,9	33,4	33,7	7,7	11,0	100,0
2011	7,2	6,8	32,3	32,8	8,9	12,0	100,0
1.1.2015	7,1	6,7	32,1	32,4	9,3	12,4	100,0

Fonte: Nostre elaborazioni su dati Istat

Analizzando le differenze tra i sessi, nelle stesse grandi classi e nello stesso periodo in esame (Tab.8), si rileva uno squilibrio. Effettivamente è ben evidente che nella

classe 0-14 c'è una prevalenza del sesso maschile rispetto a quello femminile, la classe 15-64 è sostanzialmente in equilibrio mentre la situazione cambia nella classe 65+ dove esiste una prevalenza del sesso femminile rispetto a quello maschile.

A cosa è imputabile questo squilibrio tra i sessi che si nota nelle due grandi classi estreme? Per comprenderlo è necessario porre a confronto le strutture per sesso e per età della popolazione dell'Italia al censimento del 1951 e all'1.1.2015 e, ovviamente, calcolare il rapporto di mascolinità $M/F \cdot 100$ alle varie età. (Tab.9, valori espressi in migliaia)

Il rapporto di mascolinità alla nascita equivale a 105-106 nascite maschili per 100 nascite femminili: rapporto costante nel tempo e nello spazio.

Tale eccedenza dei maschi decresce al crescere dell'età o per effetto della mortalità che è più elevata di quella delle femmine (supermortalità maschile) o per effetto delle migrazioni in quanto chi migra è generalmente di sesso maschile ed appartiene alle età centrali. Ciononostante, potrebbe derivare anche da un effetto combinato dei due fenomeni, ecco perchè ad una certa classe d'età si ha l'equilibrio dei sessi.

Tabella 9. *Popolazione residente, per classi di età e sesso, al censimento del 1951 e all'1.1.2015*

Classi di età	1951		Rapporto di mascolinità	1.1.2015		Rapporto di mascolinità
	M	F		M	F	
0-4	2.218,7	2.113,4	105,0	1.363,9	1.288,8	105,8
5-9	1.981,1	1.892,8	104,7	1.475,5	1.390,6	106,1
10-14	2.141,4	2.074,4	103,2	1.474,1	1.390,0	106,1
15-19	2.036,8	2.003,6	101,7	1.479,7	1.384,2	106,9
20-24	2.062,5	2.038,6	101,2	1.581,4	1.501,2	105,3
25-29	1.959,9	2.023,8	96,8	1.659,0	1.623,4	102,2
30-34	1.381,7	1.476,7	93,6	1.814,8	1.797,0	101,0
35-39	1.675,2	1.757,4	95,3	2.128,5	2.125,3	100,2
40-44	1.664,6	1.714,1	97,1	2.406,0	2.425,7	99,2
45-49	1.409,1	1.476,8	95,4	2.462,8	2.509,0	98,2
50-54	1.182,6	1.327,3	89,1	2.279,5	2.366,3	96,3
55-59	950,2	1.179,2	80,6	1.950,4	2.068,6	94,3
60-64	836,9	1.041,6	80,3	1.748,7	1.881,6	92,9
65-69	688,9	843,5	81,7	1.699,0	1.862,4	91,2
70-74	515,5	605,3	85,2	1.368,2	1.594,1	85,8
75-79	333,7	397,9	83,9	1.194,4	1.523,5	78,4
80-84	154,4	195,7	78,9	814,4	1.232,6	66,1
85-89	54,8	76,9	71,3	427,3	836,8	51,1
90+	10,7	17,7	60,5	173,7	492,6	35,3
Totale	23.258,8	24.256,7	95,9	29.501,6	31.294,0	94,3

Fonte: Nostre elaborazioni su dati Istat

Nelle nostre due popolazioni (Tab.9) l'equilibrio dei sessi viene raggiunto rispettivamente nella classe 20-24 per la popolazione riferita al censimento del 1951 mentre nella popolazione stimata all'1.1.2015 l'equilibrio dei sessi si ha nella classe 35-39.

Come interpretare il perché, seppure in momenti diversi, nella stessa popolazione l'equilibrio dei sessi si raggiunge in due classi differenti? Evidentemente nella popolazione italiana del 1951 la mortalità maschile era più alta di quella femminile. Inoltre, il genere maschile subiva forti flussi emigratori con la conseguenza immediata di una riduzione del numero dei maschi e, quindi, il rapporto di mascolinità determinava un equilibrio dei sessi verso le classi più basse della struttura per età.

È bene ricordare che con il passare del tempo il nostro paese da essere a forte pressione emigratoria è diventato a forte pressione immigratoria, sia di ritorno da paesi europei ed extraeuropei e sia da paesi in via di sviluppo. Ancora, si è ridotto il divario tra i sessi per quanto concerne la mortalità determinando, in tal modo, una presenza maggiore dei maschi nella popolazione che ha creato un equilibrio dei sessi verso le classi più alte della struttura per età con un diverso rapporto di mascolinità.

Quindi possiamo asserire che lo squilibrio tra i sessi nella grande classe 0-14, sin dal 1951, è dovuto essenzialmente ad una maggiore sopravvivenza maschile rispetto a quella femminile come si deduce dal confronto percentuale della popolazione maschile e femminile dal 1951 all'1/1/2015 (Tab. 8) ma, se calcoliamo i rapporti di mascolinità, nel periodo in osservazione (Tab. 10), ci accorgiamo che questo squilibrio diventa maggiore con il passare del tempo sino ad arrivare ad un valore di 106,0 (1/1/2015).

Tabella 10. *Rapporti di mascolinità per grandi classi. 1951-1.1.2015*

Anni	Classi di età		
	0-14	15-64	65+
1951	104,3	94,5	82,3
1961	104,6	96,5	73,2
1971	105,1	97,1	71,8
1981	105,1	97,3	69,5
1991	105,4	98,7	68,3
2001	105,5	99,1	70,0
2011	106,0	98,4	73,7
1.1.2015	106,0	99,1	75,3

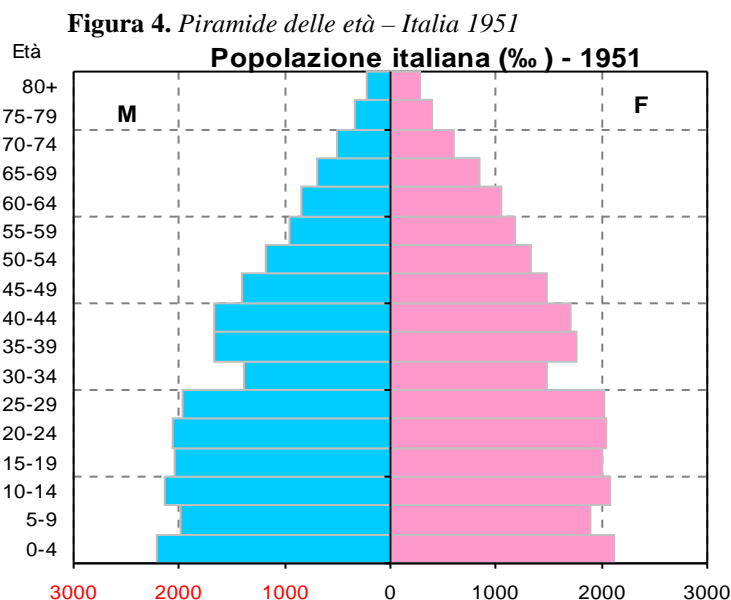
Fonte: Nostre elaborazioni su dati Istat

Questo significa che ai nostri giorni, la differenza tra i sessi nella classe 0-14 è a vantaggio dei maschi dovuta soprattutto ad una maggiore sopravvivenza ma forse anche ad una immigrazione minorile maschile il che giustifica lo spostamento del

rapporto di mascolinità verso le classi più alte della struttura per età nella popolazione all'1/1/2015 (cfr. Tab.9).

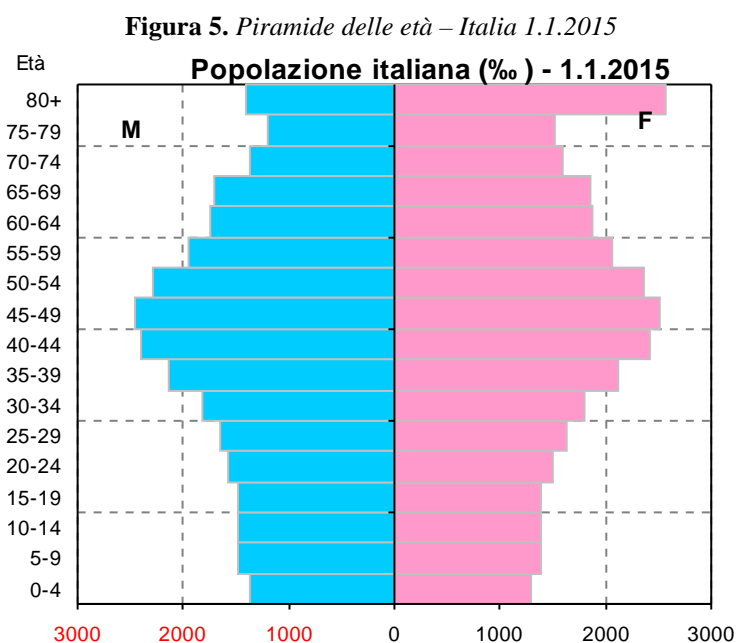
Malgrado ciò, nonostante l'equilibrio dei sessi riscontrato nella struttura per sesso e per età all'1/1/2015 si sia spostato verso le età più alte, sintomo di un aumento della sopravvivenza maschile ed anche, eventualmente, di un ingresso immigratorio maschile minorile, nella nostra popolazione continua a persistere una maggiore mortalità maschile rispetto a quella femminile tale da giustificare lo squilibrio che si nota nella classe di età 65+ (cfr. Tab.8 e Tab. 10).

Abbiamo anche rappresentato graficamente, attraverso le piramidi delle età del 1951 e dell'1/1/2015, le variazioni intervenute in un arco di tempo di 63 anni, nelle rispettive strutture per età (Figg.4 e 5).



In letteratura, per opinione comune, la suddetta rappresentazione grafica ci mostra l'intreccio esistente tra la struttura di una popolazione e la storia delle generazioni che la compongono, determinandone anche la forma.

E' indubbio il cambio di forma delle due piramidi delle età. Si è passati, invero, da una piramide con un profilo classico (Fig.4) con base allargata dovuta ad una natalità sostenuta e progressivo restringimento fino alle età più anziane derivante dall'effetto combinato di una mortalità lineare tra le classi ed emigrazione, ad una piramide con un profilo riduttivo (Fig.5) caratterizzata da una base più ristretta rispetto alla fascia centrale che a sua volta si allarga per effetto di un forte flusso di immigrati, e la parte terminale che si restringe con lentezza.



Oltre alla ben visibile progressiva “femminilizzazione” della struttura della popolazione per effetto di una maggiore sopravvivenza femminile, possiamo, allo stesso modo, notare un processo di invecchiamento dal basso della piramide dovuto alla bassa fecondità della popolazione italiana.

Nel presente scritto si è condiviso un concetto ben noto in Demografia riguardante il raggruppamento di tutta la popolazione considerata anziana nella classe aperta 65+. In questa maniera, però, non facciamo altro che rendere “invisibili” i vecchi e i grandi vecchi che riteniamo non debbano più “rimanere nascosti” sia per il loro straordinario numero e sia per il peso che hanno dal punto di vista economico, politico e sociale.

In passato il numero delle persone con più di 80 e 85 anni era ridotto e quindi il loro peso proporzionale sul totale della popolazione era trascurabile. Oggi, invece, è cresciuta l’importanza numerica delle persone molto vecchie e quindi si rende improcrastinabile, all’interno della classe 65+, una distinzione che potrebbe articolarsi in tre classi:

- 1) quella delle persone anziane con età compresa tra 65 e 74 anni;
- 2) quella delle persone vecchie con età compresa tra 75 e 84 anni;
- 3) quella dei grandi vecchi con età superiore a 85 anni.

La distribuzione percentuale delle suddette classi sul totale della popolazione ai vari anni in osservazione riportata nella Tab.11 mostra che il segmento della popo-

lazione appartenente alla classe 85+, si presenta in crescita rispetto agli altri con un incremento percentuale, rispetto al 1951, del 966,7%.

Partendo da questo eccezionale aumento è possibile ipotizzare una “quarta età” dovuta alla graduale crescita d’importanza delle fasce di età anziana che comporta un processo di invecchiamento dall’alto. Il predetto processo deve farci riflettere, attentamente, sulle linee di sviluppo da adottare intercettando i bisogni reali per far fronte ai cambiamenti attesi.

Tabella 11. *Distribuzione percentuale delle classi anziane, 1951-1.1.2015*

Classi di età	Anni								Incremento percentuale
	1951	1961	1971	1981	1991	2001	2011	1.1.2015	
65-74	5,6	6,2	7,4	8,5	8,6	10,3	10,5	10,7	91,1
75-84	2,3	2,8	3,2	3,9	5,4	6,2	7,5	7,8	239,1
85+	0,3	0,5	0,7	0,8	1,3	2,2	2,8	3,2	966,7

Fonte: Nostre elaborazioni su dati Istat

3. Conclusioni

Le considerazioni formulate ci inducono a ritenere l’invecchiamento della popolazione un fenomeno multidimensionale che pone sicuramente diversi problemi da affrontare, ma non deve essere sottovalutato l’aspetto sostanzialmente positivo di una popolazione che rispecchia una migliore salute.

Il progressivo innalzamento della vita media ha prodotto notoriamente il passaggio a un sistema demografico dominato da malattie degenerative legate alla senescenza. Già i punti fondamentali esposti con la Seconda Assemblea Mondiale sull’Invecchiamento², riguardanti le sfide poste dal rapido invecchiamento della popolazione, ritennero indispensabile l’integrazione delle persone anziane nella società, la salute e il benessere per la terza età e la garanzia di un ambiente favorevole per il loro sostegno e capacità di iniziativa.

Invecchiare nel ventunesimo secolo: un traguardo o una sfida? Dovremmo auspicare innanzitutto una svolta epocale nelle politiche sociali che sia in grado di mutare i comportamenti degli individui e delle famiglie. Difatti recentemente si è molto dibattuto sulle ripercussioni che l’invecchiamento avrà sul mercato del lavoro, sull’integrazione sociale della popolazione anziana, sulla sostenibilità dei sistemi di welfare e sanitari. Proprio per questo l’Europa ha iniziato a diffondere la cultura “dell’invecchiamento attivo” ritenuto un insieme di misure per migliorare la

² Madrid, Spagna, 2002

vita e la partecipazione delle generazioni più anziane attraverso strumenti come la formazione continua e il coinvolgimento in iniziative di socializzazione e volontariato.

Tuttavia, per far sì che gli anziani non siano considerati semplicemente come beneficiari dei servizi sociali ma come partecipanti attivi del processo di sviluppo con pieni diritti si richiede, innanzitutto, un cambiamento di comportamenti, atteggiamenti e abitudini non appartenenti alla nostra realtà.

Nella realtà italiana è soprattutto la famiglia, quando presente, a sostenere il carico assistenziale. Diverse ricerche hanno dimostrato come sia la famiglia³ ad essere l'elemento di sussidiarietà rispetto allo Stato e, quindi, ai servizi socio-sanitari. Si tratta di un carico assistenziale che le famiglie, purtroppo, faranno sempre più fatica a sopportare, soprattutto per l'evoluzione dei nuclei familiari (meno figli, meno matrimoni, aumento delle famiglie unipersonali) che sta modificando in modo progressivo il tessuto sociale italiano.

Nel nostro Paese la famiglia ha svolto tradizionalmente un ruolo di ammortizzatore sociale che si è affiancato ai principali strumenti di *politica pubblica*. Però i cambiamenti strutturali delle reti di solidarietà familiare e le crescenti difficoltà del mercato del lavoro ne hanno indebolito la sua funzione sussidiaria. Le varie letture prospettiche in chiave demografica del fenomeno invecchiamento hanno messo in luce che, lo stesso, potrebbe divenire una vera e propria emergenza sociale con conseguenze sulle prospettive di sviluppo e sostenibilità del sistema Paese. Tutto questo andrà ad incidere non solo sugli aspetti economici, ma anche sugli ambiti familiari, sfortunatamente, sorretti da un sistema di welfare tra i meno efficaci a livello europeo.

Difatti il modello della cosiddetta *famiglia forte di Reher*, nel Centro-Sud europeo, contrassegnato da un particolare incastro di obblighi reciproci tra generazioni, è ormai in declino proprio per la nota fragilità del supporto tra le generazioni.

Pertanto in un prossimo futuro sarà necessario monitorare la crescita in termini assoluti e relativi della popolazione anziana, ma sarà altrettanto doveroso valutare con riguardo le future condizioni di salute di tale popolazione e le esigenze del vivere quotidiano determinando, in tal modo, gli effetti concreti dello sviluppo di tale fenomeno.

³ Ad es., secondo un'indagine del Censis condotta nel 2004, il 75,1% degli anziani italiani, in caso di malattia o invalidità, riceve aiuto dai figli, il 41,6% dal coniuge/convivente, il 20,6% da altri parenti e il 4,1% dai vicini. Solo l'1% degli intervistati in caso di necessità riceve aiuto dai servizi sanitari territoriali e lo 0,8% dai servizi sociali (CENSIS, 2005).

Riferimenti bibliografici

- Cagianò de Azevedo R., Capacci G. (2004). *Invecchiamento e svecchiamento della popolazione europea*, Aracne editrice, Roma.
- Caselli G., Egidi V. (2011). Una vita più lunga e più sana, in: Golini A., Rosina A. (a cura di), *Il secolo degli anziani. Come cambierà l'Italia*, Il Mulino, Bologna.
- CENSIS (2005). *Analisi comparativa dei principali servizi per gli anziani non autosufficienti*, Fondazione Censis, Roma.
- EUROSTAT (2012). *Demographic outlook. National reports on the demographic developments in 2010*, Population and social conditions, Methodologies & Working papers.
- EUROSTAT (2014). EUROPOP-2013. *Population projections at national level*, in http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database.
- ISTAT (2011). *Il futuro demografico del Paese. Previsioni regionali della popolazione residente al 2065*, Statistiche Report, Roma.
- ISTAT (2015). *Annuario statistico*. Roma.
- Livi Bacci M. (1999). *Introduzione alla Demografia*, Terza edizione. Loescher, Torino.
- Reher D.S. (1998). Family ties, in: *Western Europe: persistent contrast*, *Population and Development Review*, Vol. 24, n.2, pp. 203-234.
- Sacco G., Sacco P. (2016).- The valency of the natural and social components as determinants of a population's future development, *ESJ - European Scientific Journal*, Vol. 12, n. 8 , March, (ISSN 1857-7431), pagg. 284-297.
- Saraceno C. (2003). *Mutamenti della famiglia e politiche sociali in Italia*, Il Mulino, Bologna.

Il ricorso ai servizi sanitari in Italia e in Puglia

Lucia Mongelli*

Istat, sede territoriale per la Puglia

Riassunto: In questa nota vengono analizzati i dati sul ricorso ai servizi sanitari in Italia e in Puglia per comprendere se, in un periodo di congiuntura economica sfavorevole, sono cambiati i comportamenti di tutela alla salute.

Keywords: servizi sanitari; condizioni di salute; visite mediche generali; visite mediche specialistiche.

1. Introduzione

La disponibilità di dati aggiornati rappresenta una condizione indispensabile per l'analisi dello stato di salute di una popolazione. L'Istat, con il sostegno del Ministero della Salute e delle Regioni, con la Regione Piemonte come capofila, ha realizzato l'indagine multiscopo "Condizioni di salute e ricorso ai servizi sanitari", condotta tra il 2012 e il 2013.

L'indagine permette di comprendere le condizioni di salute, le modalità di accesso ai servizi sanitari, i comportamenti di tutela della salute e di produrre stime a livello regionale e sub-regionale. L'indagine è campionaria e le informazioni sono state raccolte direttamente presso gli individui,

In questa nota ci limiteremo a considerare gli aspetti riguardanti il ricorso alle visite mediche generali e specialistiche sia con riferimento all'intero Paese, sia con riferimento alla Puglia.

*Corresponding author: mongelli@istat.it

2. La distribuzione degli intervistati

Le famiglie intervistate in Italia sono circa 60 mila per un totale di circa 119 mila individui, ossia lo 0,2 % della popolazione nazionale. Il 51,8% sono femmine ed il 48,2 % sono maschi. L'età media delle donne è di 43,0 anni, quella degli uomini è 45,9 anni (Tab. 1).

In Puglia sono stati intervistati 8.283 individui, il 51,3% sono femmine ed il 48,7% sono maschi. L'età media delle donne è di 44,4 anni, quella degli uomini è 41,5 anni.

Tabella 1. *Intervistati ed età medie in Italia e in Puglia nel 2013, secondo il sesso*

Sesso	Intervistati		%		Età medie (anni)	
	Italia	Puglia	Italia	Puglia	Italia	Puglia
Maschi	57.437	4.030	48,2	48,7	43,0	41,5
Femmine	61.636	4.253	51,8	51,3	45,9	44,4
Totale	119.073	8.283	100,0	100,0	44,5	43,0

3. Il ricorso alle visite mediche

L'indagine ha riguardato il ricorso alle visite mediche generiche e specialistiche con riferimento alle ultime 4 settimane e a quelle specialistiche, escluso quelle odontoiatriche, con riferimento all'ultimo anno precedenti l'intervista.

Nella seguente Tab. 2 sono riportati i risultati relativi al ricorso alle visite mediche nel confronto tra Puglia e intero Paese.

Tabella 2. *Percentuali di intervistati che hanno effettuato visite mediche nelle ultime 4 settimane e negli ultimi 12 mesi e numeri medi delle stesse in Italia in Puglia nel 2013*

Visite mediche	Italia	Puglia	Italia	Puglia
	Ultime 4 settimane		Ultimi 12 mesi	
	% intervistati			
Sì	32,6	32,0	50,0	43,9
No	67,4	68,0	50,0	56,1
Totale	100,0	100,0	100,0	100,0
	Numero medio di visite			
	1,9	2,1	1,2	1,3

Con riferimento alle visite mediche nelle ultime 4 settimane, poco meno di un terzo degli intervistati vi si è sottoposto, con una lieve prevalenza in Italia (32,6%) rispetto alla Puglia (32,0%). Con riferimento all'ultimo anno, in Italia metà degli

intervistati hanno subito una visita medica, abbastanza inferiore l'analogha incidenza di quanti hanno subito visita medica in Puglia (43,9%).

Il numero medio di visite nelle ultime 4 settimane è stato lievemente superiore in Puglia (2,1 visite) che in Italia (1,9 visite), anche nell'ultimo anno il numero medio di visite mediche in Puglia è lievemente superiore (1,3) rispetto all'Italia (1,2). Queste ultime come già detto sono solo visite specialistiche con esclusione di quelle odontoiatriche.

4. Le motivazioni delle visite

L'indagine mette in luce che in Italia il 43,5% delle visite mediche è dovuto a malattie e disturbi, poco meno di un quarto, ossia il 23,6%, è dovuto a prescrizione di ricette e poco più di un quinto, ossia il 20,7%, a controllo dello stato di salute in assenza di malattie e disturbi. Con un numero medio di visite pari quasi a 2 (Tab. 3).

I dati della Puglia confermano in parte la situazione nazionale: il 40,5% delle visite, ossia poco meno che nell'intero Paese, viene effettuato in presenza di malattie e disturbi, il 31,7% delle visite, ossia molto di più che in Italia, per la prescrizione di ricette e il 18,8% per il controllo dello stato di salute anche in questo caso una quota inferiore a quella nazionale.

Tabella 3. *Incidenze percentuali delle visite mediche effettuate dagli intervistati nelle ultime 4 settimane e numeri medi delle stesse in Italia e in Puglia nel 2013, secondo il motivo prevalente*

Motivo prevalente	% di visite mediche		Numero medio di visite	
	Italia	Puglia	Italia	Puglia
Malattie e disturbi	43,5	40,5	1,7	1,8
Controllo dello stato di salute	20,7	18,8	1,4	1,4
Prescrizione di ricette	23,6	31,7	1,5	1,6
Altri certificati o atti amministrativi	2,0	1,1	1,2	1,3
Altro motivo	10,2	7,9	5,0	5,0
Totale	100,0	100,0	1,7	1,7

5. Le visite mediche nelle ultime 4 settimane

Nella seguente Tab. 4 sono riportate per tipologia le incidenze percentuali degli intervistati che nelle ultime 4 settimane hanno effettuato visite mediche, le percentuali di visite e le percentuali di visite a pagamento.

Tabella 4. Percentuali di intervistati che hanno effettuato visite mediche nelle ultime 4 settimane e percentuali di visite a pagamento in Italia e in Puglia nel 2013, secondo il tipo di visita

Tipo di visita	% intervistati con visite mediche sul totale delle visite mediche		% di visite mediche		% di visite mediche a pagamento	
	Italia	Puglia	Italia	Puglia	Italia	Puglia
Generica (medico di famiglia o pediatra di base)	48,4	53,0	51,9	59,6	8,4	10,0
Pediatria	3,7	3,4	3,5	2,7	2,2	2,6
Geriatria	0,6	0,8	0,6	0,7	0,6	1,4
Cardiologica	5,4	6,4	4,8	5,8	6,6	10,8
Ostetrico-ginecologica	4,1	3,4	3,8	2,7	9,5	7,8
Oculistica	6,1	5,7	5,1	4,2	10,8	11,4
Odontoiatrica	6,6	5,6	7,0	5,9	24,7	22,0
Otorino-laringoiatrica	2,1	2,0	1,8	1,6	3,3	3,4
Ortopedica	5,4	4,8	5,3	4,3	10,0	10,2
Neurologica	2,1	2,1	1,9	1,6	2,5	3,3
Psichiatrica e psicologica	1,0	1,0	1,1	0,8	1,6	0,4
Urologica	1,8	1,6	1,6	1,2	2,4	2,7
Gastro-enterologica	1,6	1,6	1,4	1,2	2,5	2,9
Dietologica	0,7	0,6	0,6	0,4	0,9	0,4
Dermatologica	2,0	1,4	1,7	1,2	3,4	2,9
Altre visite specialistiche	8,3	6,6	8,1	5,9	10,6	7,7
Totale	100,0	100,0	100,0	100,0	100,0	100,0

Come può agevolmente vedersi le visite generiche primeggiano sia in Italia che in Puglia sia in termini di intervistati (rispettivamente 48,4% e 53,0%) che in termini di visite (51,9% e 59,6%). Il ricorso alle visite generiche è quindi in Puglia di alcuni punti superiore a quello dell'intero Paese. Le maggiori incidenze delle visite rispetto agli intervistati è la conseguenza del fatto che, sia pure nel breve periodo di 4 settimane, talvolta vengono effettuate più visite generiche.

Molto inferiori sono le incidenze delle visite specialistiche. Primeggiano per incidenza di intervistati le visite odontoiatriche (6,6% in Italia e 5,6% in Puglia), quelle oculistiche (6,1 e 5,7%), quelle cardiologiche (5,4 e 6,4%) e quelle ortopediche (5,8 e 4,8%).

Il ricorso alle visite è particolarmente contenuto per le specializzazioni dietologiche (0,7 e 0,6%), geriatiche (0,6 e 0,8%), psichiatriche e psicologiche (1,0 e 1,0%) e gastro-enterologiche (1,6 e 1,6%). Alle stesse conclusioni si perviene, sia

pure con incidenze in genere lievemente inferiori, considerando le percentuali di visite.

Molto diversa è la graduatoria delle incidenze percentuali delle visite a pagamento. Perdono il primato le visite generiche che si riducono alla sesta parte di quelle della totalità delle visite. Primeggiano invece le visite odontoiatriche che si attestano ad una su 4-5 della totalità delle visite a pagamento (24,7 e 22,0%), seguono le visite oculistiche poco più di una su 9-10 (10,8 e 11,4%), quelle ortopediche (10,0 e 10,2%) e quelle cardiologiche (6,6 e 10,8%).

Tra le visite a pagamento scarso invece risulta il peso delle visite dietologiche (0,9 e 0,4%), geriatriche (0,6 e 1,4%), psichiatriche e psicologiche (1,6 e 0,4%) e urologiche (2,4 e 2,7%). Le incidenze delle visite specialistiche a pagamento sono inferiori in Puglia rispetto all'Italia per le sole specializzazioni ostetrico-ginecologiche, odontoiatriche, psichiatriche e psicologiche, dietologiche e dermatologiche.

Tabella 5. *Percentuali di intervistati che hanno effettuato visite mediche nelle ultime 4 settimane, incidenze di quelle a pagamento e numeri medi di vissute in Italia e in Puglia nel 2013, secondo la tipologia di visita medica*

Tipologia di visita medica	% intervistati con visite		% di visite a pagamento sul totale delle visite		Numero medio di visite per intervistato che ha fatto visite	
	Italia	Puglia	Italia	Puglia	Italia	Puglia
Generica (medico di famiglia o pediatra di base)	64,2	71,1	3,8	3,5	1,5	1,8
Pediatrica	4,9	4,6	15,0	19,9	1,3	1,2
Geriatrica	0,9	1,1	24,6	41,0	1,3	1,3
Cardiologica	7,1	8,1	32,7	39,4	1,3	1,4
Ostetrico-ginecologica	5,4	4,5	59,4	59,9	1,3	1,3
Oculistica	8,1	7,7	49,0	58,0	1,2	1,1
Odontoiatrica	8,7	7,5	82,5	78,4	1,5	1,7
Otorino-laringoiatrica	2,8	2,7	42,6	44,4	1,2	1,3
Ortopedica	7,1	6,5	44,3	50,2	1,4	1,4
Neurologica	2,8	2,9	31,3	43,3	1,3	1,2
Psichiatrica, psicologica	1,3	1,3	33,9	11,1	1,6	1,3
Urologica	2,4	2,1	34,5	47,1	1,3	1,2
Gastro-enterologica	2,1	2,1	41,0	51,5	1,3	1,2
Dietologica	0,9	0,8	37,2	22,7	1,2	1,1
Dermatologica	2,7	1,8	47,1	50,0	1,2	1,4
Altre visite specialistiche	11,0	8,9	31,0	27,4	1,4	1,4
Totale	32,6	32,0	44,8	44,1	1,9	2,1

Nella Tab. 5 sono riportate per tipo di visita le incidenze di intervistati con visita, quelle delle visite a totale pagamento e le medie del numero di visite per intervistato.

Cominciamo con le visite generiche dal medico di famiglia o dal pediatra di base.

Nel 2013 in Italia il 64,2% degli intervistati, nelle ultime 4 settimane, si è sottoposto ad un visita medica generica, in Puglia tale incidenza è ben più elevata 71,1%. Non dissimile nelle due aree l'incidenza delle visite generiche a pagamento: 3,8% in Italia contro il 3,5% in Puglia. Anche più elevato in Puglia che in Italia il numero medio di visite generiche per intervistato: 1,8 contro 1,5.

Passiamo ora alle visite specialistiche.

Le divergenze tra Puglia e Italia non sono molto rilevanti. In entrambe le aree primeggiano le visite odontoiatriche (8,7 e 7,5% rispettivamente in Italia e in Puglia), quelle oculistiche (8,1 e 7,7%), quelle cardiologiche (7,1 e 8,1%) e quelle ortopediche (7,1 e 6,5%). Scarso invece il ricorso in entrambe le aree alle visite dietologiche (0,9 e 0,8%) psichiatriche e psicologiche (1,3 e 1,3%), gastroenterologiche (2,1 e 2,1%) e urologiche (2,4 e 2,1%).

Con poche eccezioni (geriatriche, cardiologiche e neurologiche) le incidenze degli intervistati che hanno fatto ricorso a visite specialistiche in Puglia sono lievemente inferiori a quelle dell'Italia.

Due sole considerazioni sul numero medio di visite: osserviamo valori medi più elevati per le visite pediatriche e per quelle odontoiatriche e non si riscontrano differenze significativa tra Italia e Puglia.

Le incidenze delle visite specialistiche a pagamento su quelle totali sono tutte superiori a quelle delle visite generiche. Primeggiano per l'onere a carico dei pazienti le visite odontoiatriche (82,5 e 78,4%), quelle ostetrico-ginecologiche (59,4 e 59,9%), quelle oculistiche (49,0 e 58,0%) e quelle dermatologiche (47,1 e 50,0%). Più contenute, anche se sempre consistenti, le incidenze delle visite a pagamento pediatriche (15,0 e 19,9%), psichiatriche e psicologiche (33,9 e 11,1%), dietologiche (37,2 e 22,7%) e geriatriche (24,6 e 41,0%).

Con poche eccezioni (odontoiatriche, psichiatriche e psicologiche e dietologiche) le incidenze delle visite a pagamento in Puglia sono più elevate che nell'intero Paese.

Le incidenze delle visite a pagamento in Puglia sono notevolmente superiori a quelle dell'intero Paese in particolare per due specializzazioni: geriatriche e gastroenterologiche.

6. Le visite mediche negli ultimi 12 mesi

Nella Tab. 6 sono riportate le incidenze percentuali degli intervistati che negli ultimi 12 mesi hanno effettuato visite specialistiche (escluse quelle odontoiatriche) secondo la specializzazione.

Tabella 6. *Percentuali di intervistati che hanno effettuato visite specialistiche negli ultime 12 mesi in Italia e in Puglia nel 2013, secondo la tipologia dell'ultima visita*

Tipologia dell'ultima visita specialistica*	Italia	Puglia
Pediatria	5,8	7,1
Geriatrica	0,9	0,7
Cardiologica	13,8	17,8
Ostetrico-ginecologica	13,3	11,2
Oculistica	17,8	17,6
Otorino-laringoiatrica	4,5	4,6
Ortopedica	13,7	13,1
Neurologica	3,5	4,3
Psichiatrica e psicologica	1,3	1,4
Urologica	4,0	4,0
Gastro-enterologica	3,2	3,6
Dietologica	1,0	0,7
Dermatologica	4,6	3,8
Altre visite specialistiche	12,4	10,3
Totale	100,0	100,0

* sono escluse le visite odontoiatriche.

I dati differiscono da quelli della tabella precedente per tre ragioni: si riferiscono solo alle visite specialistiche, sono relativi agli ultimi 12 mesi, escludono le visite odontoiatriche e le modalità sono riferite all'ultima visita effettuata.

Primeggiano sia in Italia che in Puglia le visite oculistiche (rispettivamente 17,8 e 17,6%), cardiologiche (13,8 e 17,8%), ortopediche (13,7 e 13,1%) e ostetrico-ginecologiche (13,3 e 11,2%). Molto contenuto il ricorso alle visite per le specializzazioni geriatriche (0,9 e 0,7%), dietologiche (1,0 e 0,7%), psichiatriche e psicologiche (1,3 e 1,4%) e gastro-enterologiche (3,2 e 3,6%). Le incidenze delle visite

specialistiche sono in Puglia più elevate che nell'intero Paese per le specializzazioni pediatriche, cardiologiche, otorino-laringoiatriche, neurologiche, psichiatriche e psicologiche e gastro-enterologiche.

7. Le visite mediche specialistiche secondo la copertura della spesa

Nella seguente Tab. 7 sono riportate per specializzazioni le percentuali di intervistati che hanno effettuato visite specialistiche negli ultimi 12 mesi per tipologia, secondo la copertura della spesa dell'ultima visita.

Cominciamo dalla totalità delle visite.

Tabella 7. Percentuali di intervistati che hanno effettuato visite mediche specialistiche negli ultimi 12 mesi per tipologia in Italia e in Puglia nel 2013, secondo la copertura della spesa dell'ultima visita (sono escluse le visite odontoiatriche)

Tipologia dell'ultima visita specialistica*	Non ha pagato nulla		Ha pagato il ticket, con ricetta rossa		Ha pagato interamente, senza rimborsi dall'assicurazione		Ha pagato interamente con rimborso parziale o totale		Totale	
	Italia	Puglia	Italia	Puglia	Italia	Puglia	Italia	Puglia	Italia	Puglia
Pediatrica	74,8	68,8	8,8	7,3	15,5	22,7	1,0	1,2	100,0	100,0
Geriatrica	61,2	53,8	14,6	3,8	23,0	42,3	1,2	0,0	100,0	100,0
Cardiologica	46,5	46,0	19,1	14,6	32,7	39,2	1,7	0,3	100,0	100,0
Ostetrico-ginecologica	22,4	20,2	18,6	15,8	55,3	61,8	3,7	2,2	100,0	100,0
Oculistica	28,0	23,5	20,1	15,5	48,9	60,1	3,0	0,9	100,0	100,0
Otorino-laringoiatrica	32,5	41,9	26,4	15,6	38,7	42,5	2,5	0,0	100,0	100,0
Ortopedica	32,9	33,3	25,5	19,9	39,3	45,8	2,3	1,0	100,0	100,0
Neurologica	48,8	45,8	19,9	14,2	29,0	36,1	2,3	3,9	100,0	100,0
Psichiatrica e psicologica	56,3	78,4	11,3	7,8	31,0	11,8	1,4	2,0	100,0	100,0
Urologica	40,7	29,9	22,0	18,1	35,4	51,4	1,8	0,7	100,0	100,0
Gastro-enterologica	36,0	36,4	27,7	21,2	34,3	41,7	2,0	0,8	100,0	100,0
Dietologica	41,8	58,3	19,5	12,5	38,2	29,2	0,5	0,0	100,0	100,0
Dermatologica	21,9	29,7	31,7	21,0	42,5	47,8	3,8	1,4	100,0	100,0
Altre visite specialistiche	48,9	48,5	19,8	15,8	29,7	34,9	1,6	0,8	100,0	100,0
<i>Totale visite</i>	<i>38,0</i>	<i>38,2</i>	<i>20,8</i>	<i>15,6</i>	<i>38,8</i>	<i>45,1</i>	<i>2,4</i>	<i>1,1</i>	<i>100,0</i>	<i>100,0</i>

* sono escluse le visite odontoiatriche.

Circa 4 visite su 10 sono gratuite, la differenza tra Italia e Puglia è irrisoria (rispettivamente 38,0 e 38,2%). All'altro estremo gli intervistati che hanno pagato la visita interamente senza rimborso la incidenza è ancora pari in Italia (38,8%), ma ben superiore in Puglia (45,1%). Coloro che hanno pagato il ticket con ricetta rossa sono uno su 5 (20,8%) in Italia e uno su 6 (15,6%) in Puglia. Marginale l'incidenza di coloro che hanno pagato interamente con rimborso parziale o totale: solo il 2,4% in Italia e l'1,1% in Puglia. In sintesi, rispetto all'intero Paese, la quota di quanti pagano interamente le visite specialistiche è notevolmente superiore in Puglia, mentre quella di chi paga solo il ticket è inferiore.

Ponendo l'attenzione sulle specializzazioni si evince che la quota di visite mediche specialistiche gratuite è particolarmente rilevante per talune specializzazioni: pediatriche (rispettivamente 74,8 in Italia e 68,8% in Puglia), psichiatriche e psicologiche (56,3 e 78,4%), geriatriche (61,2 e 53,8%) e dietologiche (41,8 e 58,3%). Più frequentemente a totale carico dei pazienti sono le visite mediche specialistiche ostetrico-ginecologiche (55,3 e 61,8%), oculistiche (48,9 e 60,1%), ortopediche (39,3 e 45,8%) e otorino-laringoiatriche (38,7 e 42,5%). Le specialità per le quali si paga più frequentemente solo il ticket è in Italia per le visite dermatologiche (31,7%), gastroenterologiche (27,7%), otorino-laringoiatriche (26,4%) e ortopediche (25,5%); in Puglia il ticket viene pagato più frequentemente per le visite gastro-enterologiche (21,2%), dermatologiche (21,0%), ortopediche (19,9%) e urologiche (18,1%).

Limitando l'attenzione alle incidenze delle visite a integrale pagamento senza rimborso i dati della Tab. 7 rivelano la situazione frequentemente più onerosa in Puglia che in Italia delle visite specialistiche: geriatriche, oculistiche, urologiche e neurologiche.

8. Le visite mediche specialistiche secondo la scelta della struttura utilizzata e le relative motivazioni

Nella seguente Tab. 8 sono riportate le percentuali di intervistati che hanno effettuato visite mediche specialistiche negli ultimi 12 mesi per tipologia, secondo la struttura utilizzata per l'ultima visita. Le modalità sono solo due (struttura pubblica o privata convenzionata, struttura privata non convenzionata) per cui le percentuali sono complementari, il che consente di limitare il commento ad una sola di esse.

Cominciamo con la totalità delle visite.

Come può vedersi tanto in Italia che in Puglia la struttura prevalentemente utilizzata è quella pubblica o privata convenzionata. Tale prevalenza inoltre è più

marcata nell'intero Paese che in Puglia, 70,3% contro 62,6%. Verosimilmente questo divario dipende dalla disponibilità delle strutture sanitarie nelle due aree a confronto.

Tabella 8. Percentuali di intervistati che hanno effettuato visite mediche specialistiche negli ultimi 12 mesi, per tipologia specialistica, in Italia e in Puglia nel 2013, secondo la struttura presso la quale è stata effettuata l'ultima visita

Tipologia dell'ultima visita specialistica*	Pubblica o privata convenzionata		Privata non convenzionata		Totale	
	Italia	Puglia	Italia	Puglia	Italia	Puglia
Pediatria	87,3	81,2	12,7	18,8	100,0	100,0
Geriatrica	85,2	65,4	14,8	34,6	100,0	100,0
Cardiologica	76,7	68,4	23,3	31,6	100,0	100,0
Ostetrico-ginecologica	52,8	45,3	47,2	54,7	100,0	100,0
Oculistiche	58,0	46,2	42,0	53,8	100,0	100,0
Otorino-laringoiatrica	71,4	64,1	28,6	35,9	100,0	100,0
Ortopedica	72,9	65,1	27,1	34,9	100,0	100,0
Neurologica	78,1	69,0	21,9	31,0	100,0	100,0
Psichiatrica e psicologica	77,8	86,3	22,2	13,7	100,0	100,0
Urologica	76,6	63,2	23,4	36,8	100,0	100,0
Gastro-enterologia	79,1	69,7	20,9	30,3	100,0	100,0
Dietologica	69,4	79,2	30,6	20,8	100,0	100,0
Dermatologica	66,2	58,7	33,8	41,3	100,0	100,0
Altre visite specialistiche	81,4	74,5	18,6	25,5	100,0	100,0
Totale complessivo	70,3	62,6	29,7	37,4	100,0	100,0

* sono escluse le visite odontoiatriche

Consideriamo ora l'utilizzo dei due tipi di struttura per tipologia della visita.

Nell'intero Paese le incidenze più elevate della scelta delle strutture pubbliche o convenzionate sono quelle delle visite pediatriche (87,3%), geriatriche (85,2%), gastro-enterologiche (79,1%) e neurologiche (77,8%). In Puglia le incidenze più elevate di scelta delle strutture pubbliche o convenzionate sono quelle delle visite psichiatriche e psicologiche (86,3%), pediatriche (81,2%), dietologiche (79,2%) e gastro-enterologiche (69,7%).

Il confronto tra Italia e Puglia mette in evidenza che solo per le visite psichiatriche e psicologiche e per quelle dietologiche il ricorso alle strutture pubbliche o convenzionate è superiore in Puglia.

Nella seguente Tab. 9 sono riportate le percentuali di intervistati che hanno effettuato visite specialistiche negli ultimi dodici mesi, secondo la motivazione della scelta della struttura utilizzata per l'ultima visita.

Emerge come motivazione prevalente la fiducia nella struttura utilizzata o il fatto che è stata consigliata (33,2% in Italia e poco meno, il 32,6%, in Puglia) in entrambe le aree, quindi, un intervistato ogni 3 ha indicato tale motivazione. Segue al secondo posto la motivazione logistica della vicinanza della struttura scelta, motivazione indicata dal 19,5% in Italia e, poco meno, dal 17,7%, in Puglia. Solo al terzo posto la motivazione economica (non pagare o pagare di meno), segno che per la salute questo tipo di motivazione è in terzo ordine. Va fatto notare che in Puglia la motivazione economica raccoglie un consenso appena maggiore che nel resto del Paese (15,6 in Puglia contro 13,5% in Italia). Al quarto posto la scelta di una struttura è in funzione dei medici che vi operano (13,1% in Italia e 12,0% in Puglia).

Tabella 9. *Percentuali di intervistati che hanno effettuato visite mediche specialistiche negli ultimi 12 mesi in Italia e in Puglia nel 2013, secondo le motivazioni della scelta della struttura utilizzata per l'ultima visita*

Motivazioni di scelta della struttura *	Italia	Puglia
La più vicina	19,5	17,7
Per non pagare o pagare di meno	13,5	15,6
Ha più fiducia in questa struttura o le è stata consigliata	33,2	32,6
Avrebbe dovuto aspettare troppo tempo per prenotare in un'altra struttura del SSN	10,8	13,2
Ha scelto uno specialista che lavora in questa struttura	13,1	12,0
Non esistono altre strutture nella zona di residenza	2,2	1,1
Ha scelto questa struttura perché è più accogliente o è possibile fissare l'ora dell'appuntamento	3,4	4,4
Altro	4,3	3,4
Totale	100,0	100,0

* potevano esserci più risposte

A poca distanza l'indicazione di ripiego (per non aspettare di prenotare in altra struttura migliore), 10,8% in Italia e un po' di più, ossia il 13,2%, in Puglia, quasi a testimoniare una maggiore difficoltà di scelta della struttura migliore in Puglia. Marginali, sia in Italia che in Puglia, le percentuali delle altre indicazioni (non esistono altre strutture in zona, struttura più accogliente o più facilmente fruibile o altra ancora).

9. Tempi di attesa

Nella seguente Tab. 10 sono indicati i tempi medi di attesa in giorni per l'ultima visita specialistica. Due sole considerazioni: la prima i tempi più lunghi in Italia ri-

spetto alla Puglia per fruire dell'ultima visita specialistica, poco meno di due mesi per una struttura pubblica o convenzionata e tre settimane in una struttura privata non convenzionata.

Stessa graduatoria in Puglia, ma con tempi di attesa più contenuti: poco più di 7 settimane in una struttura pubblica o convenzionata e solo due settimane in una struttura privata non convenzionata. Quindi un giudizio un po' più lusinghiero per la Puglia rispetto all'intero Paese.

Tabella 10. *Tempi medi di attesa in giorni dell'ultima visita effettuata negli ultimi 12 mesi in Italia e in Puglia nel 2013, secondo la struttura utilizzata*

Struttura	Italia	Puglia
Pubblica o privata convenzionata	54,8	52,6
Privata non convenzionata	21,2	14,8
Totale	50,5	47,2

Nella seguente Tab. 11 sono riportate per tipologia specialistica le percentuali di intervistati secondo le modalità di attesa ed i relativi tempi medi in giorni dell'ultima visita negli ultimi 12 mesi.

Cominciamo dai dati della totalità delle visite specialistiche.

Molto lusinghiero è il fatto che sia in Italia che in Puglia oltre la metà degli intervistati non ha dovuto aspettare per effettuare la visita specialistica; in Italia il 54,9% e in Puglia poco più il 55,1%. Ancora positivo è il fatto che al secondo posto ci sono le incidenze di coloro che non hanno dovuto aspettare perché la visita era stata prefissata: poco più di uno su 5, più precisamente il 21,5% in Italia ed il 20,7% in Puglia.

In sintesi, per tre quarti degli intervistati non vi è stata attesa. Un intervistato ogni 16-18 ha scelto di aspettare per scegliere la struttura di gradimento (6,0% in Italia e 5,5% in Puglia). In pratica hanno dovuto subire attesa per la visita specialistica uno ogni 5-6 intervistati, più precisamente il 17,6 ed il 18,7% rispettivamente in Italia e in Puglia.

Del tempo medio di attesa della totalità delle visite si è già detto a proposito della Tab. 10. In sostanza non vi sono differenze significative tra Puglia ed intero Paese.

Quanto all'analisi per tipologia specialistica limitiamo l'attenzione alle incidenze di quanti hanno subito attesa per la visita specialistica e alla durata media dell'attesa.

Innanzitutto segnaliamo che le attese di visite pediatriche riguardano solo quote modeste di intervistati 6,4 e 6,5% rispettivamente in Italia ed in Puglia. Le inciden-

ze più consistenti di attesa in Italia sono quelle delle visite gastro-enterologiche (23,0%), otorino-laringoiatriche (21,6%), urologiche (21,0%) e dermatologiche (20,8%).

Tabella 11. Percentuali di intervistati che hanno effettuato visite mediche specialistiche negli ultimi 12 mesi in Italia e in Puglia nel 2013 per tipologia, secondo le modalità di attesa e i relativi tempi medi dell'ultima visita

Tipologia dell'ultima visita specialistica*	No, non ho dovuto attendere, la data era stata comunicata nella precedente visita		No, non ho dovuto aspettare		Sì, ho dovuto attendere perché volevo una struttura di mio gradimento		Sì, ho dovuto attendere, perché non avevo alternativa		Attesa media (in giorni)	
	Italia	Puglia	Italia	Puglia	Italia	Puglia	Italia	Puglia	Italia	Puglia
Pediatria	19,1	14,6	71,5	74,6	3,0	4,2	6,4	6,5	34,1	19,8
Geriatrica	29,6	23,1	45,2	34,6	4,9	7,7	20,3	34,6	43,7	72,5
Cardiologica	24,0	21,8	50,5	52,0	5,7	5,9	19,7	20,3	48,5	49,4
Ostetrico-ginecologica	19,2	18,7	62,6	66,5	5,7	4,7	12,5	10,1	39,1	46,0
Oculistiche	17,7	17,2	56,5	58,4	7,7	7,4	18,1	17,1	61,9	51,5
Otorino-laringoiatrica	16,6	14,4	56,2	59,3	5,5	3,0	21,6	23,4	45,0	52,8
Ortopedica	20,6	19,0	54,1	53,1	6,1	4,8	19,1	23,0	42,3	37,6
Neurologica	26,6	31,0	47,3	46,5	5,6	3,9	20,4	18,7	50,9	41,2
Psichiatrica e psicologica	37,3	31,4	47,6	47,1	2,6	2,0	12,5	19,6	50,8	23,3
Urologica	23,7	19,4	48,7	47,2	6,6	6,9	21,0	26,4	47,7	33,7
Gastro-enterologica	21,7	28,8	48,9	50,0	6,4	4,5	23,0	16,7	45,8	29,4
Dietologica	30,2	37,5	50,4	33,3	4,1	12,5	15,2	16,7	60,9	85,3
Dermatologica	14,5	19,6	56,8	50,7	7,9	4,3	20,8	25,4	51,0	54,2
Altre visite specialistiche	28,1	27,3	48,3	43,4	5,7	6,4	17,8	22,8	61,3	61,9
Totale complessivo	21,5	20,7	54,9	55,1	6,0	5,5	17,6	18,7	50,5	47,2

* sono escluse le visite odontoiatriche

In Puglia le quote più consistenti di visite che hanno avuto attese sono quelle delle visite geriatriche (34,6%), urologiche (26,4%), dermatologiche (25,4%) e otorino-laringoiatriche (23,4%).

Quanto alla durata media di attesa particolarmente pesanti in Italia sono quelle che superano i due mesi: 61,9 giorni per le visite oculistiche, 60,9 giorni per le visite

dietologiche. In Puglia per le visite dietologiche occorre aspettare quasi 3 mesi 85,3 giorni, per quelle geriatriche quasi due mesi e mezzo, precisamente 72,5 giorni.

10. I motivi delle rinunce di visite specialistiche

Nella seguente Tab. 12 sono riportate le incidenze degli intervistati che negli ultimi 12 mesi hanno dovuto rinunciare o meno a qualche visita specialistica e, per quanti hanno rinunciato, le incidenze delle motivazioni della rinuncia.

Dalla tabella risulta che negli ultimi 12 mesi in Italia un intervistato ogni 9 (11,9%) ed in Puglia uno ogni 6 (17,4%) ha rinunciato ad una visita specialistica pur avendone bisogno. Emerge quindi una situazione meno soddisfacente per la Puglia.

Sia in Italia che in Puglia la motivazione prevalente della rinuncia è quella economica. Ancora una volta la situazione appare in Puglia più pesante: oltre la metà degli intervistati ha rinunciato alla visita specialistica perché non poteva pagarla; nell'intero Paese la situazione, pur grave, è meno pesante (45,3%).

Tabella 12. *Percentuali di intervistati che hanno rinunciato a visite mediche specialistiche* negli ultimi 12 mesi in Italia e in Puglia nel 2013 e percentuali di quanti hanno rinunciato, secondo la motivazione*

Modalità	Italia	Puglia
Ha rinunciato	11,9	17,4
Non ha rinunciato	88,1	82,6
<i>Totale intervistati</i>	<i>100,0</i>	<i>100,0</i>
Non poteva pagarla perché costava troppo	45,3	50,7
Lista d'attesa lunga	28,7	28,6
Scomodità (lontananza, mancanza di trasporti, orari scomodi)	7,7	7,5
Non poteva assentarsi dal lavoro	7,7	4,9
Doveva accudire i figli	4,5	4,5
Altro	6,0	3,8
<i>Totale intervistati che hanno rinunciato</i>	<i>100,0</i>	<i>100,0</i>

* sono escluse le visite odontoiatriche

La seconda motivazione della rinuncia, anch'essa consistente, è la lunghezza delle liste di attesa, poco più di un intervistato ogni 4, pressoché identica nelle due aree. Non trascurabile anche il peso della motivazione della scomodità sia in Italia (7,7%) che in Puglia (7,5%).

La modalità relativa al non potersi assentarsi dal lavoro ha in Puglia un peso inferiore che in Italia, in connessione alla struttura professionale, specialmente femminile nella Regione rispetto al Paese.

11. La soddisfazione per l'ultima visita specialistica

Nella seguente Tab. 13 sono riportati i punteggi medi in una scala crescente da 1 a 10 del grado di soddisfazione dell'ultima visita specialistica effettuata negli ultimi 12 mesi.

Come prima considerazione segnaliamo un'elevata e diffusa soddisfazione sia nell'intero Paese che in Puglia: non si scende al di sotto di una media di 7,4 e non si va al di là di 8,5. Quindi, si osserva anche una limitata variabilità.

Unica considerazione significativa è costituita dalla concordanza tra Italia e Puglia per la minore soddisfazione per le visite specialistiche geriatriche (7,8 e 7,4 rispettivamente per l'Italia e per la Puglia) e, per la maggiore soddisfazione per le visite specialistiche ostetrico-ginecologiche (8,5 per entrambe le aree).

Tabella 13. *Grado medio di soddisfazione (punteggi da 1 a 10) espresso dagli intervistati sull'ultima visita degli ultimi 12 mesi in Italia e in Puglia nel 2013, per tipologia della visita*

Tipologia dell'ultima visita specialistica*	Italia	Puglia
Pediatria	8,4	8,3
Geriatrica	7,8	7,4
Cardiologica	8,2	8,1
Ostetrico-ginecologica	8,5	8,5
Oculistica	8,3	8,4
Otorino-laringoiatrica	8,0	7,9
Ortopedica	7,9	7,6
Neurologica	7,9	8,1
Psichiatrica e psicologica	7,9	8,1
Urologica	8,1	7,8
Gastro-enterologica	8,0	7,9
Dietologica	8,2	7,6
Dermatologica	8,0	7,9
Altre visite specialistiche	8,2	8,2
Totale	8,2	8,1

* sono escluse le visite odontoiatriche

Nella seguente Tab. 14 sono riportati i punteggi medi sempre nella scala crescente da 1 a 10 del grado di soddisfazione dell'ultima visita degli ultimi 12 mesi secondo il tipo di struttura utilizzata.

Il livello di soddisfazione per le visite specialistiche fruite risulta decisamente alto; gli intervistati attribuiscono buoni voti (8,2 in Italia e 8,1 in Puglia), con sostanziali differenze tra quelle erogate dal Servizio Sanitario Nazionale e da strutture private convenzionate (8,0) rispetto al sistema privato (8,6). Ciò si riscontra soprattutto in Puglia dove il grado di soddisfazione per le visite erogate è notevolmente inferiore nelle strutture pubbliche o private convenzionate (7,8) rispetto alle strutture private (8,6), differenziandosi di quasi un punto.

Tabella 14. *Grado medio di soddisfazione (punteggi da 1 a 10) espresso dagli intervistati sull'ultima visita degli ultimi 12 mesi in Italia e in Puglia nel 2013, per struttura utilizzata*

Struttura	Italia	Puglia
Pubblica o Privata convenzionata	8,0	7,8
Privata non convenzionata	8,6	8,6
Totale	8,2	8,1

12. Considerazioni finali

In Italia il ricorso alle visite mediche nelle ultime quattro settimane riguarda un terzo della popolazione, con un numero medio di circa 2 visite; il ricorso alle visite mediche specialistiche nell'ultimo anno riguarda metà della popolazione, con un numero medio di 1,3 visite.

Il ricorso alle visite mediche è motivato prevalentemente da malattie o disturbi, consistente è anche la quota di visite mediche dovute al rilascio di ricette, solo al terzo posto le visite dovute al controllo dello stato di salute.

In Italia nelle ultime 4 settimane la gran parte delle visite sono generiche presso il medico di famiglia o il pediatra di base, ma per talune specialità (odontoiatria, oculistica, cardiologia e ortopedia) il ricorso appare consistente 7-8%. Le visite a pagamento sono poco meno della metà, ma per alcune specialità (odontoiatria, ostetricia e ginecologia) tale quota va ben al di là.

Le visite specialistiche di 4 specialità (oculistica, cardiologia, ortopedia, ostetricia e ginecologia) assommano poco meno del 60% di tutte le visite specialistiche.

Le visite specialistiche sono prevalentemente e con la stessa quota di circa il 40% a totale carico dei pazienti o del tutto gratuite. Solo un paziente ogni 5-6 paga solo il ticket. Contenuta è la variabilità di tali quote per specialità.

Sette intervistati su dieci per le visite specialistiche privilegiano le strutture pubbliche o convenzionate: le motivazioni prevalenti della scelta sono innanzitutto la fiducia e la vicinanza, ma non mancano scelte obbligate da motivi economici, da urgenza o da carenze di strutture adeguate.

I tempi di attesa per le visite specialistiche sono molto lunghi, specie nelle strutture pubbliche o convenzionate. Fortunatamente l'attesa riguarda solo un intervistato ogni 6-7. Infatti la quota più consistente, oltre la metà, è costituita da coloro che non hanno dovuto aspettare e per un quinto da coloro che avevano prefissato la data della visita.

Non trascurabile la quota di intervistati che ha dovuto rinunciare ad una visita specialistica necessaria per vari motivi, nell'ordine, onerosità, lunga attesa, impegni lavorativi, scomodità.

Gli intervistati sono molto soddisfatti delle visite mediche specialistiche, specie se fruite in una struttura privata non convenzionata.

Riferimenti bibliografici

- Istat (1999-2000). *Le condizioni di salute della popolazione*. <http://www.istat.it>
- Istat (2005). *Condizioni di salute, fattori di rischio e ricorso ai servizi sanitari*. <http://www.istat.it>
- Istat (2013). *La salute e il ricorso ai servizi sanitari attraverso la crisi, Anno 2012 (media settembre-dicembre)*, Statistiche Report.
- Istat (Anno 2014). *Tutela della salute e accesso alle cure. Anno 2013*.
- Istat (2015). *Le dimensioni della salute in Italia. Determinanti sociali, politiche sanitarie e differenze territoriali, Temi Letture Statistiche*.
- Istat (2016). *Sistema informativo e territoriale su sanità e salute*. Health for all – Italia. <http://www.istat.it>
- Ministero del Lavoro, della Salute e delle Politiche Sociali (2007-2008). *Relazione sullo stato sanitario del Paese*, Roma - Direzione Generale del Sistema Informativo.
- Ministero della Salute (2011). *Relazione sullo Stato Sanitario del Paese*. Direzione Generale del Sistema Informativo e Statistico Sanitario.
- Osservatorio nazionale sulla salute nelle regioni italiane (2013). *Rapporto OsservaSalute*. Università Cattolica del Sacro Cuore.

Fonte

Istat – Condizioni di salute e ricorso ai servizi sanitari. Anno 2013. File standard -
elaborazione dei dati: 15/01/2016

Il benessere equo e sostenibile dei territori pugliesi tra multidimensionalità e sintesi

Stefania Taralli¹, Monica Carbonara²

¹ISTAT, Ufficio territoriale per la Toscana, l'Umbria e le Marche;

²ISTAT, Ufficio territoriale per la Puglia

Riassunto: Lo studio del Benessere equo e sostenibile a livello sub-regionale consente di valutare non soltanto il livello di benessere e la posizione dell'area in esame ma anche il grado di coesione territoriale. Per rispondere a questi obiettivi si è seguita la via di una analisi che bilanciassero la necessaria ricerca di una sintesi con la capacità di leggere la multidimensionalità del benessere, allo scopo di mettere in luce i tratti distintivi dei *profili di benessere* dei territori in esame. Ciò ha consentito, tra l'altro, di scomporre il quadro regionale evidenziando i divari territoriali più rilevanti e significativi.

Keywords: well-being, analisi territoriale, sistemi di indicatori territoriali.

1. Obiettivi

Lo studio del benessere sociale, economico e ambientale dei territori è un fertile terreno di incontro tra domanda di conoscenza e offerta di informazione statistica, anche a fini valutativi. Infatti la *governance* territoriale richiede informazioni imparziali e condivise a partire dalle quali elaborare decisioni e valutazioni (Dente, 2011). Tra le sfide con cui ci si confronta in questo ambito di ricerca, una delle più rilevanti consiste nel dover gestire il *trade-off* tra la multidimensionalità e granularità delle informazioni e la necessità di ricondurre a sintesi le specificità locali entro un quadro più ampio e confrontabile a vari livelli.

In questo contesto, il presente lavoro si prefigge di tracciare il profilo del Benessere equo e sostenibile (Bes) della Puglia per valutarne la posizione nel contesto nazionale ed apprezzarne il grado di coesione territoriale, mettendo in luce disparità e vantaggi che ne connotano i territori di appartenenza.

Funzionale all'obiettivo conoscitivo è la scelta di applicare l'analisi dei gruppi o *cluster analysis*, una tecnica che può produrre confronti multidimensionali senza richiedere necessariamente una sintesi preliminare. Ciò è parso di particolare utilità vista la finalità valutativa dello studio, che richiede necessariamente la ricerca di una sintesi ma impone allo stesso tempo di valorizzare al massimo la complessità dell'oggetto di analisi, indirizzando verso strategie alternative ai più comuni metodi basati sulla costruzione di indicatori compositi.

Mentre, infatti, questi ultimi sono utili ad evidenziare soprattutto il *quanto*, il livello complessivo di benessere di un territorio, producendo essenzialmente punteggi su cui basare confronti aggregati, costruire *ranking*, o elaborare *rating*, valutare il benessere equo e sostenibile dei territori per orientare il *policy making* vuol dire essenzialmente analizzare *come* si determina e si caratterizza il benessere di ogni singolo territorio in confronto agli altri.

2. Il quadro di riferimento

Il Bes - Benessere Equo e Sostenibile - nasce nel 2010 per iniziativa congiunta di CNEL e Istat con l'obiettivo di misurare e valutare il progresso della società italiana. Nel 2011 parte il progetto "Bes delle province", un *framework* che unisce in un quadro coerente e consistente le indicazioni che emergono dall'iniziativa Cnel-Istat e le esigenze specifiche del livello territoriale/istituzionale di interesse. Il progetto assume tra i propri riferimenti il progetto "How's life in your region?" dell'Ocse (OECD, 2014), che promuove l'utilizzo degli indicatori di benessere a sostegno delle politiche locali, e le linee guida "Europa 2020 per le città e i territori" (EU COR 2012).

"Il tema della misurazione del progresso ha due componenti: la prima, prettamente politica, la seconda di carattere tecnico-statistico" (Cnel-Istat, 2010, p. 1).

Il Benessere territoriale non sfugge a questa duplice vocazione. Il *well-being* è la chiave di volta del nuovo ciclo dei fondi strutturali e il cardine su cui impennare la programmazione Europe 2020 a livello locale.

Con la recente riforma della Legge di Bilancio si va verso l'integrazione del quadro degli indicatori economici tradizionalmente allegati ai documenti di bilancio con nuovi indicatori di benessere da considerare per la fissazione degli obiettivi macroeconomici. Anche l'armonizzazione contabile della spesa della Pubblica Amministrazione introduce una articolazione per missioni e programmi che ha un'ampia convergenza con il Bes (Fiorillo et al, 2016).

Sulla scia di queste tendenze emergono quindi nuove opportunità di valorizzazione anche per le misure del Benessere a livello territoriale.

L’iniziativa Cnel-Istat assume come dettaglio territoriale massimo quello regionale, anche perché la disponibilità di indicatori provinciali risulta drasticamente limitata (Cnel-Istat, 2012). Tuttavia, date le caratteristiche del nostro Paese, la scalabilità sub-regionale degli indicatori è molto importante per riuscire a cogliere le diversità e specificità locali.

Con il progetto Bes delle province l’Istat, in collaborazione con gli Uffici di statistica degli Enti di area vasta, ha implementato un sistema di oltre 80 indicatori provinciali, coerenti con il *framework* nazionale, con l’obiettivo di comparare il livello e la struttura del benessere dei territori per individuare i punti di forza e di debolezza da porre all’attenzione dei *policy makers* (Taralli, 2013).

Il progetto ha evidenziato connessioni numerose e consistenti tra la *mission* degli Enti territoriali e i domini del benessere equo e sostenibile, rilevando in proposito una *vision* ampiamente condivisa dai *decision makers* locali (Cuspi-Istat, 2016; Bartoloni et al., 2016; Costa et al., 2016).

3. I dati e il metodo

3.1 Gli indicatori del Bes delle province

Il Benessere equo e sostenibile delle province è misurato per 11 domini, non essendo disponibili misure del Benessere soggettivo a questo dettaglio. Gli indicatori elementari, nella versione 2015 del dataset, sono 82, tra “Misure del benessere” - indicatori coincidenti o *proxy* delle misure del Bes nazionale - e “Altri indicatori generali”¹ (Tavola 1).

Tutti gli indicatori descritti nella Tavola 1 sono disponibili a livello provinciale, regionale e nazionale. I profili di benessere delle province pugliesi potrebbero perciò essere definiti a partire da una classica comparazione territoriale gerarchica, producendo per ciascuna provincia, dominio per dominio, il profilo delle differenze assolute e relative rispetto ai parametri regionale e nazionale. Questa impostazione, coerente con la gerarchia dei livelli amministrativi, non tiene conto appieno del peso

¹ Gli “Altri indicatori generali” sono stati introdotti per approfondire l’analisi del contesto di attuazione delle funzioni degli Enti locali, ferma restando la piena coerenza concettuale con il *framework* del Bes. La legenda degli indicatori è in allegato al presente lavoro. Per una descrizione analitica degli indicatori stessi si rinvia al Rapporto 2015 sul Bes delle province (Cuspi-Istat, 2016). Il glossario degli indicatori del Bes nazionale è consultabile in appendice ai rapporti Bes (cfr. ad esempio Istat 2015).

dei diversi territori in esame e della loro posizione nello spazio: le aree metropolitane e le province più grandi contribuiscono alla media regionale più delle altre province; talvolta la vicinanza tra territori produce reciproche influenze nella struttura e nei livelli di alcuni fenomeni che hanno un'importanza centrale nell'analisi del Bes (per esempio il mercato del lavoro).

Tavola 1. *Misure del Bes e Altri indicatori generali per dimensione*

Dimensioni del Bes	Misure del Bes	Altri indicatori generali
Salute	7	1
Istruzione e formazione	5	3
Lavoro e conciliazione dei tempi di vita	6	4
Benessere economico	4	5
Relazioni sociali	3	5
Politica e istituzioni	4	4
Sicurezza	3	3
Benessere soggettivo	-	-
Paesaggio e patrimonio culturale	3	1
Ambiente	5	2
Ricerca e Innovazione	2	5
Qualità dei servizi	5	2

Fonte: Il Benessere equo e sostenibile delle province. Rapporto 2015.

Per questo motivo si ritiene più efficace un approccio analitico che consideri anche le relazioni spaziali e faccia emergere i tratti specifici e distintivi delle strutture di benessere dei territori: le deprivazioni o i vantaggi sulle singole componenti del benessere, le convergenze o le divergenze rispetto ai profili degli altri territori presi a confronto.

3.2 L'analisi dei gruppi

Considerando la numerosità della popolazione in esame e la multidimensionalità del Bes, la strategia scelta in funzione degli obiettivi appena richiamati è l'analisi dei gruppi a partire da indicatori-chiave selezionati in base alla teoria. Oltre che agli obiettivi conoscitivi descritti in apertura (§ 1), la scelta di ricorrere agli indicatori chiave in alternativa alle più diffuse tecniche di riduzione della multidimensionalità *data driven*, è stata guidata dalle seguenti valutazioni:

- il costrutto del Bes è basato su un approccio “formativo”: gli indicatori elementari non sono tra loro interscambiabili, ma ciascuno di essi concorre direttamente a misurare una precisa componente del benessere equo e sostenibile (Istat, 2015, Mazziotta e Pareto, 2013);
- le caratteristiche del dataset del Bes delle province, che, come visto, contiene indicatori di diversa rilevanza e funzione informativa e presenta un diverso grado di copertura semantica dei domini e delle dimensioni (Taralli et al. 2015) richiedendo quindi una selezione preliminare degli indicatori per evitare distorsioni nell’analisi.

Sono quindi stati selezionati soltanto indicatori che esprimono misure del livello di benessere (escludendo gli indicatori di vulnerabilità o di equità) e *non sostituibili*, perché relativi a componenti non rappresentate da altri indicatori già inclusi. Si è inoltre avuto cura di bilanciare il set di indicatori chiave con riferimento alla copertura delle 11 dimensioni misurate nel Bes delle province.

Il risultato di queste operazioni è un insieme di 26 indicatori chiave, distribuiti nelle 11 dimensioni del Bes delle province (Tavola 2). La tavola evidenzia a colpo d’occhio la minore copertura di alcune dimensioni, dovuta anche alla minore consistenza della base informativa di partenza, e alcune apparenti ridondanze, date dalla selezione, per la stessa dimensione, di indicatori tra loro complementari (e perciò molto correlati) che sono stati comunque inclusi senza ulteriori trasformazioni preliminari, in considerazione della loro importanza nel costrutto del Bes.

Per quanto riguarda la standardizzazione, poiché gli indicatori utilizzati differiscono significativamente anche per la variabilità che li connota, e dato che quasi mai la forma della loro distribuzione si accosta alla distribuzione normale (Taralli et al., 2015) si è optato per il metodo max-min, anche e soprattutto allo scopo di preservare al massimo l’informazione sulla variabilità territoriale, che rappresenta il focus dell’analisi.

Ulteriori notazioni possono riguardare il metodo di raggruppamento, la robustezza e la significatività del *clustering*, nonché il contributo degli indicatori chiave all’analisi dei gruppi. Si è utilizzata l’analisi gerarchica aggregativa per individuare il numero di gruppi ottimale (a prescindere dal miglior partizionamento), e l’analisi “k means” per il *clustering* finale².

² La prima analisi ha indicato in 4 il numero ottimale dei gruppi, segnalando la struttura Ward come la più efficiente rispetto all’obiettivo (di massimizzare la devianza tra gruppi – SSB e minimizzare la devianza nei gruppi - SSW). Il confronto tra gli algoritmi di assegnazione ha dato esito positivo restituendo gruppi di composizione identica. L’analisi “k means” è stata perciò impostata per ottenere 4 gruppi, realizzati con la distanza euclidea ed il metodo Ward.

Tavola 2. Indicatori chiave utilizzati nella cluster analysis e valori della statistica $F^{(*)}$ per singolo indicatore

Dimensione	Indicatore	Valore di F
I - Salute	I.1 Speranza di vita -maschi	25,36
	I.2 Speranza di vita -femmine	30,75
	I.8 Tasso standardizzato di mortalità evitabile (0-74 anni)	51,66
II - Istruzione	II.1 Giovani che hanno abbandonato precocemente gli studi	7,83
	II.2 Persone in età lavorativa con livello di istruzione non elevato	11,39
	II.6 Livello di competenza alfabetica degli studenti	49,57
	II.7 Livello di competenza numerica degli studenti	33,42
III - Lavoro	III.1 Tasso di mancata partecipazione al lavoro 15-74	270,70
	III.3 Tasso occupazione totale 20 – 64	266,30
IV - Benessere economico	IV.1 Stima del reddito disponibile lordo per famiglia	78,38
	IV.5 Ammontare medio del patrimonio familiare	164,40
V - Relazioni sociali	V.7 Diffusione delle istituzioni non profit	22,16
VI - Politica e istituzioni	VI.2 Tasso di partecipazione alle elezioni regionali	23,55
	VI.3 Grado di inclusività nelle istituzioni locali- presenza donne a livello comunale	56,50
	VI.4 Grado di inclusività nelle istituzioni locali- presenza dei giovani a livello comunale	17,61
VII - Sicurezza	VII.3 Delitti violenti denunciati	4,32
VIII - Paesaggio e patrimonio culturale	VIII.1 Consistenza del tessuto urbano storico	60,83
	VIII.3 Strutture museali fruibili	5,40
IX - Ambiente	IX.2 Superamento limiti inquinamento aria - PM10	6,46
	IX.4 Acqua potabile erogata giornalmente	6,73
	IX.6 Energia prodotta da fonti rinnovabili	0,73
X - Ricerca e innovazione	X.1 Propensione alla brevettazione	35,58
	X.7 Specializzazione produttiva in settori ad alta intensità di conoscenza	1,47
XI - Qualità dei servizi	XI.1 Bambini 0-2 anni che usufruiscono di servizi per l'infanzia	62,50
	XI.4 Raccolta differenziata di rifiuti urbani	28,47
	XI.5 Densità di reti urbane di TPL nei capoluoghi di provincia	2,78

(*) $F = \text{SSB}_i / \text{SSW}_i$ per ogni i da 1 a $k=4$

Fonte: Il Benessere equo e sostenibile delle province. Rapporto 2015.

Confermata la significatività del raggruppamento³, la sensitività e la robustezza del modello di aggregazione scelto sono state valutate testando tutte le soluzioni da 3

³ L'analisi della varianza ha confermato la significatività della cluster analysis svolta dal momento che il valore della F di Snedecor, pari a 33,76, è risultato superiore al valore critico della distribuzione F di Fischer con $\alpha=5\%$ per $n-k=106$ e $k-1=3$ gradi di libertà.

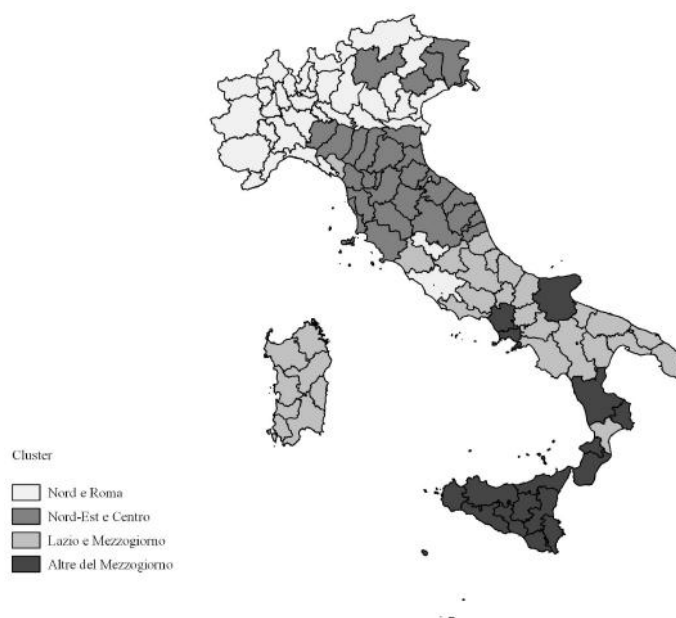
a 5 gruppi e applicando diversi algoritmi di assegnazione. Il confronto tra gli *output* ottenuti nei vari scenari di raggruppamento ha dato esito positivo in quanto i gruppi definiti nel modello adottato si sono confermati sufficientemente stabili anche di fronte a cambiamenti di scenario.

Infine, per apprezzare la capacità discriminante degli indicatori chiave e il loro contributo nel contesto del modello, si è svolta un'applicazione del "Test F" a mero scopo descrittivo, calcolando il rapporto tra la devianza tra i gruppi e la devianza nei gruppi per ciascuno dei 26 indicatori chiave singolarmente considerati, con l'obiettivo di individuare gli indicatori che più degli altri caratterizzano la composizione e i profili dei gruppi, e che quindi spiegano meglio l'assegnazione di una provincia ad un gruppo piuttosto che a un altro. I risultati di questo esercizio, che indicano per valori crescenti di F un maggiore contributo dell'indicatore al partizionamento ottenuto, sono riportati nella Tavola 2.

3.3 I profili di benessere dei gruppi

La *cluster analysis*, applicata su 26 indicatori chiave, ha restituito un quadro robusto e significativo individuando quattro gruppi di diversa numerosità e distribuzione territoriale. Il cartogramma (Figura 1) ne illustra la distribuzione spaziale e la composizione.

Figura 1. *Distribuzione dei territori nei Gruppi*



Fonte: Il Benessere equo e sostenibile delle province. Rapporto 2015.

La collocazione dei gruppi nello spazio geografico traccia un quadro piuttosto ordinato, a suggerire come in generale province limitrofe tendono ad assomigliarsi in termini di struttura del Bes e, quindi, a collocarsi nello stesso *cluster*. D'altra parte, però, la rappresentazione cartografica mette in luce l'esistenza di *enclaves*, di intersezioni e di fratture che indicano come la collocazione geografica e la contiguità territoriale non siano le uniche chiavi di lettura utili.

L'articolazione dei gruppi produce una mappa in cui alla tradizionale segmentazione del territorio italiano tra Nord, Centro e Mezzogiorno si sovrappongono dualismi e specificità territoriali che tagliano trasversalmente quella partizione, evidenziando dinamiche contrastanti anche all'interno del territorio pugliese.

Come visto, i gruppi sono stati analizzati in base a un sottoinsieme di indicatori chiave del Bes, individuati per la loro rilevanza e affidabilità. Province simili tra loro in base ai livelli di Bes misurati su questi indicatori sono classificate in uno stesso gruppo, province tra loro diverse per i livelli di Bes misurati dagli indicatori chiave confluiscono in gruppi diversi; il raggruppamento infatti è costruito sotto il vincolo della massima differenza tra i gruppi e della massima omogeneità nei gruppi.

Una volta individuati i gruppi attraverso la *cluster analysis*, il profilo di benessere di ciascun gruppo è stato definito analizzando dimensione per dimensione l'intero set degli 82 indicatori del Bes delle province (Allegato 1).

L'analisi è stata condotta sempre in termini comparativi, confrontando i valori medi del *cluster* con quelli nazionali sia in valore assoluto che attraverso numeri indice territoriali.

L'esposizione dei profili di benessere al livello territoriale prescelto può svilupparsi lungo due direttrici: la descrizione dei tratti comuni ai territori che ricadono nell'area considerata, funzionale a connotare meglio i *cluster* individuati, e una analisi degli indicatori di Bes delle province che, nel gruppo in esame, esprimono la maggiore dispersione, evidenziando divergenze territoriali significative, come tali da approfondire.

Questa analisi è stata svolta in maniera sistematica per tutti i Gruppi e per tutte le Province e Città metropolitane italiane ed ha messo in evidenza importanti differenze tra i profili di benessere delle 4 macro-aree in esame.

La Tavola 3 offre una sintesi descrittiva dei risultati dell'esercizio svolto con riferimento all'intero territorio nazionale, i cui risultati analitici sono riportati nel Rapporto 2015 sul Bes delle province (Cuspi-Istat, 2016).

I profili di benessere dei quattro gruppi differiscono significativamente per struttura, verso ed entità degli scostamenti dalla media nazionale.

L'area in assoluto più avvantaggiata, per numero ed entità delle differenze rispetto alla media-Italia, su tutti i domini del Bes e per la gran parte delle misure, è il Gruppo "Nord-est e Centro" (G.2), seguito dal Gruppo "Nord e Roma" (G. 1) che tuttavia lo supera per i migliori livelli riscontrati nei domini Lavoro e conciliazione dei tempi di vita e Benessere economico.

Il Gruppo "Altre del mezzogiorno" (G. 4) è in assoluto il più svantaggiato in tutti i domini del Bes, rispetto ai quali mostra i livelli più bassi in assoluto su pressoché tutte le misure considerate. Anche il Gruppo "Lazio e Mezzogiorno" risulta tendenzialmente svantaggiato rispetto alla media-Italia, ma con differenze negative più contenute e un profilo più articolato rispetto al Gruppo 4. Nel Gruppo 3 si notano anche performance tendenzialmente positive nei domini Paesaggio e patrimonio culturale e Ambiente e livelli di benessere vicini alla media-Italia nei domini Salute e Relazioni sociali.

Tavola 3. Sintesi dei profili di Bes dei Gruppi

Dimensione	G1	G2	G3	G4
I. Salute	<	>	=	><
II. Istruzione e formazione	><	=	><	><
III. Lavoro e conciliazione dei tempi di vita	>	><	<	<
IV. Benessere economico	>	>	<	<
V. Relazioni sociali	>	>	=	<
VI. Politica e istituzioni	><	><	<	<
VII. Sicurezza	>	=	<	<
VIII. Paesaggio e patrimonio culturale	>	>	>	<
IX. Ambiente	><	><	>	><
X. Ricerca e innovazione	>	>	<	<
XI. Qualità dei servizi	=	>	<	<

Legenda: > posizione migliore rispetto all'Italia; = simile; < peggiore

Fonte: Il Benessere equo e sostenibile delle province. Rapporto 2015.

4. Il benessere della Puglia nel contesto italiano

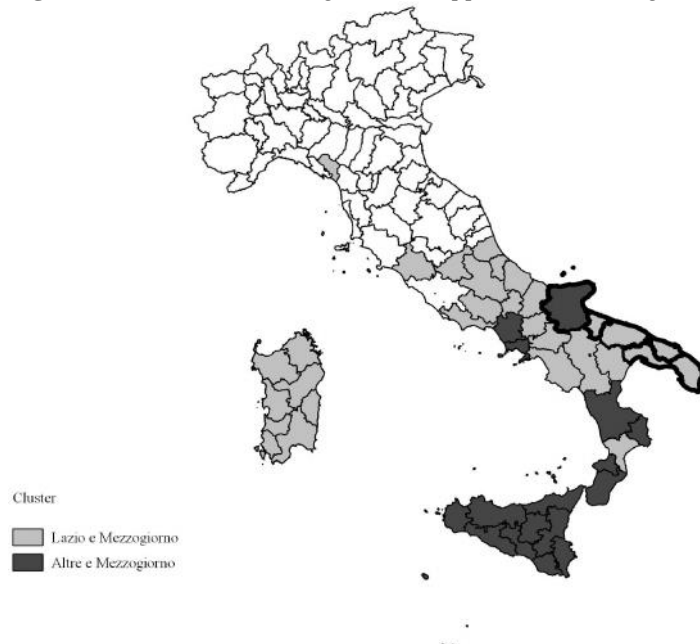
4.1 Il confronto con il Mezzogiorno e le convergenze tra territori

Lo stesso esercizio comparativo proposto al paragrafo precedente può essere svolto focalizzando il territorio regionale e valutando l'articolazione del Bes a livello sub-regionale - cioè di Province e Città metropolitane - non soltanto in relazione alle corrispettive misure nazionali ma anche in relazione al profilo del cluster (o dei cluster) di appartenenza.

In questo modo, gli indicatori di Bes a livello provinciale, con il supporto dell'analisi dei gruppi, consentono di valutare l'equità del benessere di una regione spingendosi al di là e/o al di sotto dei confini regionali.

In particolare per la Puglia l'analisi dei gruppi ha evidenziato un dualismo territoriale da approfondire. La gran parte del territorio, infatti, si colloca nel Gruppo Lazio e Mezzogiorno, che comprende tutte le province pugliesi tranne quella di Foggia. Quest'ultima, invece, posizionandosi nel gruppo che comprende Napoli, quasi tutte le province calabresi e tutte le siciliane (Gruppo Altre del Mezzogiorno), mostra un profilo di benessere sostanzialmente diverso rispetto sia alle altre province della regione di appartenenza che ai territori limitrofi (Figura 2).

Figura 2. I territori della Puglia nei Gruppi Lazio e Mezzogiorno e Altre del Mezzogiorno



Fonte: Il Benessere equo e sostenibile delle province. Rapporto 2015.

La figura 3 riporta, per gli 11 domini considerati, gli indici di confronto territoriale delle province pugliesi rispetto alla media-Italia e al cluster di appartenenza. In particolare, poiché i territori pugliesi si dividono tra due gruppi, ma quasi tutte le province tranne quella di Foggia ricadono nel Gruppo “Lazio e Mezzogiorno”, si è scelto di utilizzare la media di quest’ultimo gruppo come termine di confronto, oltre alla media-Italia, e di analizzare separatamente la provincia di Foggia dall’insieme delle “Altre province pugliesi”, che sono considerate in termini di “profilo medio” (Figura 3). Una lettura analitica dei dati relativi a queste ultime province sarà comunque svolta per illustrare le differenze più significative in tutti i casi in cui la media non risulta sufficientemente significativa⁴.

Nell’ambito del dominio *Salute* nelle “Altre province pugliesi” la speranza di vita (I.1 e I.2) e la mortalità evitabile (I.8) sono in linea con la media dell’Italia e denotano una posizione tendenzialmente migliore rispetto alla media del Gruppo Lazio e Mezzogiorno. La speranza di vita dei maschi (I.1: 80,2 anni), anzi è maggiore di entrambi i valori di confronto. Anche la mortalità per tumori (I.5: 8,2 per 10mila abitanti) è decisamente meno incidente sia in confronto al Gruppo Lazio e Mezzogiorno (8,7 per 10mila abitanti), sia, ancor più, rispetto all’Italia (8,9 per 10mila abitanti). Ciò che caratterizza in negativo il profilo di benessere delle “Altre province pugliesi” nel dominio salute è la mortalità stradale dei giovani (I.4: 1,0 per 10mila abitanti), più elevata sia della media del gruppo di confronto (0,9 per 10mila abitanti) che della media-Italia (0,8 per 10mila abitanti). Il profilo della provincia di Foggia evidenzia svantaggi significativi su diverse componenti del dominio salute: la speranza di vita, che è inferiore ai valori di confronto sia per i maschi (79,7 anni) che per le femmine (84,2 anni) e la mortalità evitabile (6,0 decessi per 10mila abitanti), che invece è decisamente superiore a tutti gli altri valori considerati. Inoltre, nell’anno⁵ considerato, Foggia risulta particolarmente penalizzata dalla mortalità infantile (I.3: 4,3 morti per mille nati vivi) e dalla mortalità dei giovani per incidenti di trasporto (I.4: 1,2 morti per 10mila abitanti). Entrambi gli indicatori superano tutti i parametri di confronto. Il quadro è mitigato in parte dalla migliore posizione relativa della provincia rispetto alla mortalità per demenza degli anziani (I.6: 24,6 decessi per 10mila abitanti) e a quella per tumore della popolazione adulta (I.5: 8,2 per 10mila abitanti), che in entrambi i casi sono apprezzabilmente inferiori ai valori di riferimento.

⁴ La significatività della media è valutata con il coefficiente di variazione. Il criterio di significatività è fissato nel valore soglia di 0,2.

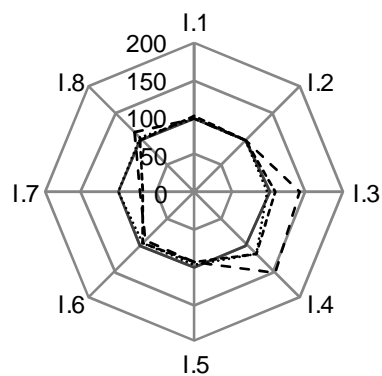
⁵ I dati si riferiscono all’aggiornamento 2015 della base di indicatori. Per gli anni di riferimento, diversi per ciascun indicatore, si rinvia al Rapporto Bes delle province 2015.

Figura 3. Indici di confronto territoriale per dimensione del Bes - Provincia di Foggia, Altre province pugliesi, Gruppo Lazio e Mezzogiorno, Italia (Italia=100)

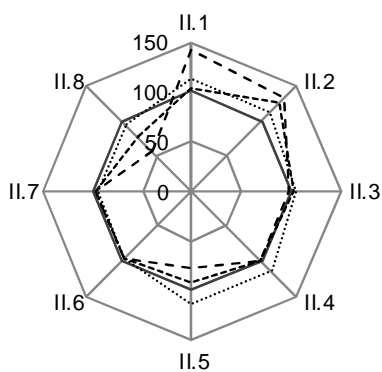
Legenda grafici

- Italia (valore)
- Gruppo Lazio e Mezzogiorno (media)
- Altre province pugliesi (media)
- - - Provincia di Foggia

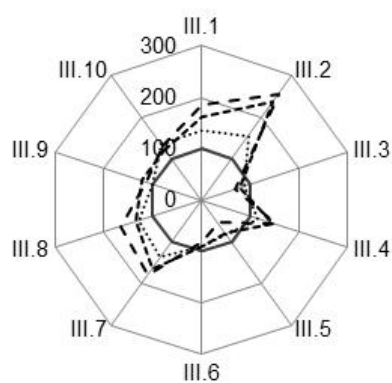
I. Salute



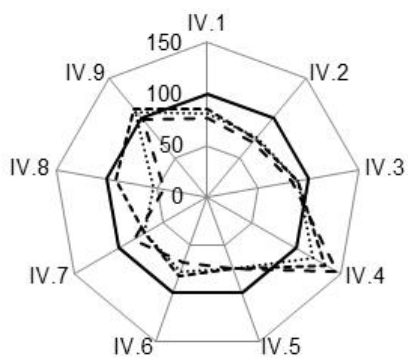
II. Istruzione e formazione



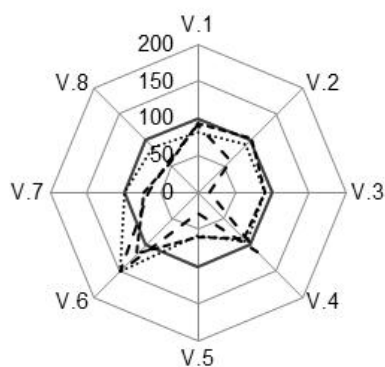
III. Lavoro e conciliazione tempi di vita



IV. Benessere economico

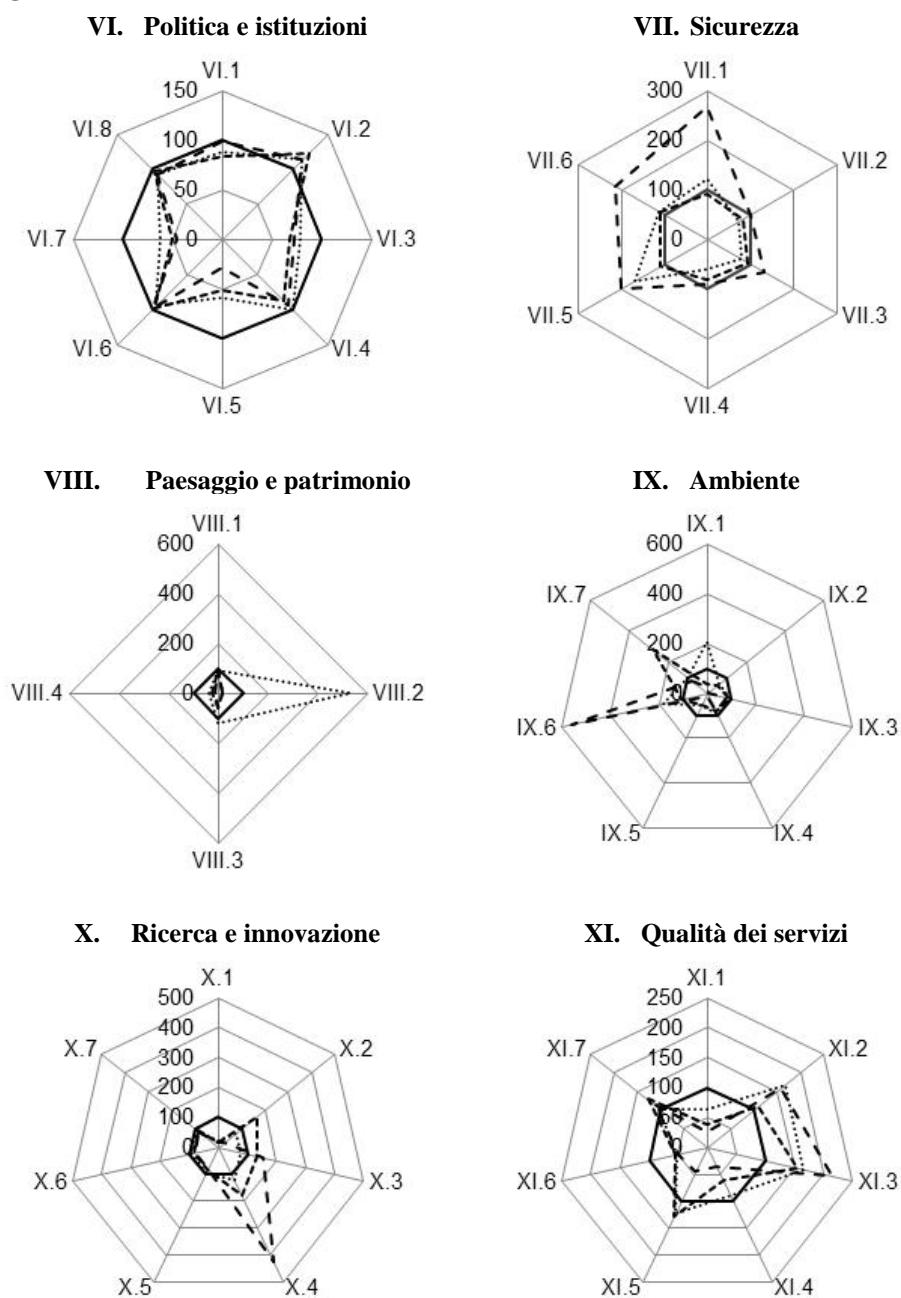


V. Relazioni sociali



continua

Figura 3. *Continuazione.*



Fonte: Il Benessere equo e sostenibile delle province. Rapporto 2015.

Nell'ambito del dominio *Istruzione* i principali svantaggi che caratterizzano le "altre province" della Puglia rispetto alla media nazionale sono la notevole quota di persone in età lavorativa con livello di istruzione non elevato (II.2: 49,8%), che supera apprezzabilmente anche la media del gruppo di riferimento (45%), la maggiore incidenza di *early school leavers*, cioè di giovani adulti che non hanno conseguito il diploma superiore (II.1: 16,3%), e la partecipazione all'istruzione terziaria (II.4: 38,5%) e all'apprendimento permanente (II.8: 5,8%). Questi due ultimi indicatori sono entrambi più bassi anche rispetto alla media del Gruppo Lazio e Mezzogiorno (II.4: 44,6%; II.8: 6,9%). I livelli di competenza alfabetica (II.6: 196,5) e numerica (II.7: 199) degli studenti della secondaria di II grado, invece, sono più elevati rispetto alla media del gruppo di confronto (rispettivamente 192,4 e 192,7). Nel profilo della provincia di Foggia l'incidenza della popolazione adulta con bassa istruzione (II.2: 52,6%) e degli *early school leavers* (II.1: 22,3%) è ancor più accentuata e i livelli di competenza alfabetica (II.6: 192,1) e numerica (II.7: 196,3) degli studenti ancor più bassi delle altre province pugliesi. Si confermano in questo caso i minori livelli di partecipazione all'istruzione terziaria (II.4: 38,4%) e alla formazione continua (II.8: 4,2%) già rilevati per il resto della Puglia, ai quali si aggiunge un tasso di partecipazione all'istruzione secondaria più basso di tutti i valori presi a confronto (II.3: 93,5%).

Nell'ambito del dominio *Lavoro e conciliazione dei tempi di vita*, le "Altre province" della Puglia si trovano in posizione di svantaggio rispetto alla media-Italia per pressoché tutti gli indicatori considerati. Le differenze più marcate sono nei tassi di mancata partecipazione (III.1: 37,1%) e di occupazione (III.2: 46,3%) sia generali che femminili (III.2: pari a 19 punti percentuali; III.4: pari a -28,6 punti percentuali) e nei maggiori livelli di disoccupazione giovanile (III.8: 42,4%), che segnano distanze evidenti sia dalla media del Gruppo Lazio e Mezzogiorno (40,6%) che dalla media-Italia (31,6%). Il profilo della provincia di Foggia ricalca quello delle altre province pugliesi, evidenziando svantaggi più accentuati su tutte le misure già citate: la partecipazione al mercato del lavoro, sia generale (III.1: 42,7%), che femminile (la differenza rispetto ai maschi è di 20,4 punti percentuali, III.2) e la disoccupazione generale (III.7: 22,8%) e giovanile (III.8: 52,4%). Si aggiunge anche un livello particolarmente basso dell'occupazione giovanile (III.5: 15,5%) che denota una situazione di rilevante penalizzazione dei giovani residenti nella provincia anche rispetto ai propri corregionali.

Considerando il dominio *Benessere economico*, il profilo delle "altre province pugliesi", pur in un quadro di generale svantaggio in confronto alla media-Italia, rivela livelli generali di reddito (IV.1: 34.090,49 €) e ricchezza (IV.5: 271,62 €)

relativamente migliori del Gruppo Lazio e Mezzogiorno. I redditi da lavoro dipendente (IV.2: 15.539,75 €) e da pensione (IV. 3: 15.191,42 €) sono, invece, più bassi della media del gruppo ed è apprezzabilmente maggiore l'incidenza delle pensioni di basso importo (IV.4: 15%). Per le famiglie in cui queste componenti di reddito sono preponderanti, quindi, lo svantaggio rispetto alla media nazionale è ancora più marcato che nel complesso della macro-area di confronto. In un generale contesto di retribuzioni relativamente più basse, va notata anche la maggiore differenza di genere nella retribuzione media dei lavoratori dipendenti (IV.6: -6.311,9 €). Il profilo della provincia di Foggia evidenzia ulteriori svantaggi rispetto alle altre province pugliesi: sono ancor più accentuate le differenze sui livelli di reddito da lavoro dipendente (IV.2: 14.667,6 €) e da pensione (IV.3: 14.192,3 €) ed è maggiore l'incidenza di pensionati con pensioni di basso importo (IV.4: 16,4%), mentre i livelli di ricchezza (IV.5), e soprattutto di reddito complessivo delle famiglie (IV.1) sono ben al di sotto di tutti gli altri valori di confronto.

Per quanto riguarda le *Relazioni sociali*, gli indicatori relativi all'integrazione scolastica delle persone con disabilità (V.1-V.4) sono migliori nelle "Altre province pugliesi" rispetto al Gruppo Lazio e Mezzogiorno e non distanti dai valori nazionali, eccezion fatta per la minore presenza di alunni disabili nelle scuole (V.4: 2,5%). Le acquisizioni della cittadinanza italiana da parte dei cittadini stranieri residenti (V.5, 1,6%) sono state nell'anno considerato molto più basse che nel complesso del Paese, ma in linea con il Gruppo Lazio e Mezzogiorno. Il dato rispecchia peraltro anche la diversa natura e tradizione storica dei flussi immigratori che interessano il Sud del Paese rispetto a quelli del Centro-Nord.

Gli indicatori della diffusione delle cooperative sociali (V.6: 2,6 per 10mila abitanti), della diffusione delle istituzioni non profit (V.7: 37,3 per 10mila abitanti), e la quota di volontari per 100 abitanti di 14 anni e più (V.8: 5,7%) tratteggiano, invece, un profilo piuttosto distante sia da quello nazionale che dalla media del Gruppo Lazio e Mezzogiorno: tutti i valori degli indicatori considerati sono infatti nettamente inferiori ai dati di confronto. In questo quadro il profilo della provincia di Foggia appare piuttosto in linea con quello delle altre province Pugliesi, al netto di alcune differenze di lieve entità.

Nell'ambito del dominio *Politica e istituzioni* le "Altre province" della Puglia spiccano per l'elevato tasso di partecipazione alle ultime elezioni regionali (VII.2: 64%), dato ampiamente superiore a quello relativo al Gruppo Lazio e Mezzogiorno, che a sua volta supera apprezzabilmente il corrispettivo valore nazionale (52,0%). La partecipazione alle elezioni europee (VI.1) mostra un gradiente territoriale opposto, con le "Altre province pugliesi" (49%) ben al di sotto del valore nazionale (58,7%) e

di quello del gruppo (51,4%). Gli indicatori relativi al grado di inclusione delle donne (VI.3) e dei giovani (VI.4) nelle istituzioni locali risultano inferiori ad entrambi i parametri di confronto territoriale, denotando un particolare svantaggio delle “Altre province pugliesi” sotto questo profilo: in questi territori le donne non raggiungono in media il 18% degli amministratori comunali e i giovani amministratori (con meno di 40 anni) sono poco più di uno su 4 (28%). Gli indicatori di finanza locale, quali il grado di finanziamento interno (VI.5: 0,1 euro per un euro di entrata) e la capacità di riscossione delle amministrazioni provinciali (VI. 6: 0,7 euro per un euro di entrata) e comunali (VI.8: 0,7 euro per un euro di entrata) sono, invece, sostanzialmente in linea con la media del gruppo Lazio e Mezzogiorno, che a sua volta si discosta apprezzabilmente dal valore nazionale per la minore capacità di riscossione delle amministrazioni comunali, pari nel gruppo a 0,1 euro per un euro di entrata a fronte dei 0,2 relativi al complesso dei Comuni italiani. Il profilo della provincia di Foggia si distingue da quello delle altre province pugliesi per il minore squilibrio tra la partecipazione alle elezioni per il parlamento europeo e alle regionali (VI.1: 57,5%; VI.2: 59,1%) e per una presenza di donne e giovani nelle amministrazioni comunali lievemente maggiore. Nell’anno considerato spicca invece, in negativo, la limitata capacità di riscossione dell’amministrazione provinciale.

Per quanto riguarda il dominio *Sicurezza*, nelle altre province pugliesi il tasso di omicidi è in linea con il tasso nazionale (VII.1: 0,8 per 100mila abitanti) e quello di criminalità violenta (VII.3) inferiore, con 20,9 delitti violenti ogni 10mila abitanti a fronte dei 22,3 denunciati in Italia. Anche il livello dei delitti diffusi (VII.4: 207,3 per 10mila abitanti) è più basso che in Italia (258,7 per 10mila abitanti). Entrambi gli indicatori superano, comunque, la media del Gruppo Lazio e Mezzogiorno (VII.3: 17,9 per 10mila abitanti; VII.4: 155,1 per 10mila abitanti), rispecchiando anche differenze territoriali nella propensione a denunciare i reati. Gli indicatori di sicurezza stradale sono in linea con la media-Italia per quanto riguarda la mortalità stradale generale (VII.5: 2%), e quindi descrivono una situazione relativamente migliore rispetto al complesso del Gruppo Lazio e Mezzogiorno (3,2%), mentre invece evidenziano una situazione piuttosto critica per la mortalità stradale in ambito extraurbano che nelle “Altre province pugliesi”, come nel complesso del Gruppo Lazio e Mezzogiorno, si attesta a 5,1 morti per 100 incidenti contro il 4,6 nazionale. Tutti gli indicatori del dominio sicurezza evidenziano per Foggia un profilo peggiore del complesso delle province Pugliesi, con valori più elevati di tutti gli indicatori di criminalità (VII.1-VII.4) e di sicurezza stradale. Spiccano in particolare, nell’anno considerato, gli elevati livelli del tasso di mortalità sulle strade extraurbane (VII.6: 9,9) e del tasso di omicidi (VII.1: 2,2), doppi rispetto alle altre province pugliesi.

Le “Altre province della puglia” risultano regolarmente penalizzate su tutte le misure di benessere del dominio *Paesaggio e patrimonio culturale*. La consistenza del tessuto urbano storico in buone condizioni di conservazione (VIII.1: 66,8%), in linea con la media del Gruppo Lazio e Mezzogiorno, è più bassa del dato nazionale (71,8%). Particolarmente bassa è la densità territoriale di parchi urbani e verde di interesse storico (VIII.2: 0,7%) sia in confronto alla media-Italia (5,1%) che al valore medio del Gruppo Lazio e Mezzogiorno (27,2%). Quest’ultimo valore è tuttavia trainato dal dato della provincia di Matera, i cui “Sassi”, patrimonio UNESCO, contribuiscono ad innalzarne il livello con la loro notevole estensione territoriale. Senza il contributo di Matera, anche quest’ultimo indicatore posizionerebbe le province del gruppo su un livello medio del 3,3%, inferiore al dato nazionale di confronto, ma comunque quasi 5 volte maggiore di quello delle “Altre province pugliesi”. Molto bassa risulta anche la dotazione di strutture museali fruibili, pari a 0,4 per 10mila abitanti contro lo 0,8 nazionale e l’1,0 del gruppo di confronto. A fronte di questa minore dotazione strutturale la quota di visitatori (VII.4, 4.169,9 per 10mila abitanti), che è circa un quarto di quella nazionale, mette in luce anche una scarsa valorizzazione e fruizione del patrimonio disponibile, svantaggio che è ancor più marcato nel complesso dei territori del Gruppo Lazio e Mezzogiorno, dove, nonostante una dotazione di strutture maggiore della media-Italia, la media dei visitatori è circa un terzo di quella nazionale. A Foggia il patrimonio di strutture fruibili è lievemente superiore a quello delle altre province pugliesi (VIII.3: 0,5 per 10mila abitanti) ma ciononostante la fruizione si riduce ulteriormente, scendendo a quota 0,5 visitatori annui per 10mila abitanti. Per contro la consistenza del tessuto urbano storico in buone condizioni e la disponibilità di parchi urbani di interesse pubblico sono qui lievemente superiori alla media delle altre province pugliesi.

Il profilo delle province pugliesi nel dominio *Ambiente* mostra tratti ambivalenti. Tra i vantaggi evidenti sia rispetto al Gruppo Lazio e Mezzogiorno che all’Italia vanno segnalati i minori superamenti giornalieri dei limiti di inquinamento da PM10 (IX.2: 21,4 contro i 44,0 italiani), i minori consumi domestici di elettricità (IX.3: 1.032,3 Kwh per abitante) e di acqua potabile (IX.4: 196,3 litri per abitante). La quota di elettricità prodotta da fonti rinnovabili è maggiore qui che in Italia (IX.6: 41,6%) ma non raggiunge i livelli, molto più elevati, che caratterizzano il complesso dei territori del Gruppo Lazio e Mezzogiorno, più che doppi della media nazionale (69,4%). Per contro nelle “Altre province pugliesi” risulta più bassa la disponibilità di piste ciclabili e di verde urbano e, soprattutto, si rileva un elevatissimo afflusso in discarica di rifiuti urbani, che raggiunge le 102,3 tonnellate per kmq di superficie, un valore triplo rispetto alla media-Italia e alla stessa media del gruppo di confronto. Ri-

petto alle altre province pugliesi, il profilo di Foggia, al netto della ancor più ridotta disponibilità di verde urbano (IX.1: 8,4 mq per abitante) e di piste ciclabili (I.5: 1,7 Km per 100 Kmq), appare ancor più “verde” caratterizzandosi per livelli di consumi domestici idrici ed elettrici ancora più contenuti (191,8 litri per abitante, 896,7 Kwh per abitante), per un afflusso in discarica di rifiuti urbani di gran lunga più ridotto (29,4 tonnellate per kmq, un terzo rispetto alle altre province), e per una notevolissima produzione di energia elettrica da fonti rinnovabili in rapporto ai consumi (215%), cinque volte maggiore di quella delle altre province.

Nel dominio *Ricerca e innovazione*, le “Altre province pugliesi” rivelano un profilo di svantaggio generalizzato su tutte le misure considerate. Il tasso di brevettazione europea (X.1), pari a 12,8 domande per milione di abitanti, è inferiore a quello del gruppo di confronto (15,6), già a sua volta notevolmente più basso della media-Italia (75,2). L’incidenza dei brevetti nei settori high-tech (X.2: 14,4%), ICT (X.3: 20,3%) e biotech (X.4: 5,7%) supera invece sia la media del gruppo che quella nazionale, denotando una particolare diffusione di questi settori di ricerca e sviluppo, pur nel quadro di ridotta capacità brevettuale descritto. Nell’anno considerato i nuovi flussi di laureati in discipline scientifico-tecnologiche sia triennali, che totali, sono risultati inferiori sia a quelli del Gruppo Lazio e Mezzogiorno che ai nazionali. Anche la specializzazione produttiva in settori ad alta intensità di conoscenza (X.7) si attesta sui 24,8 addetti per 100 totali, un valore più basso del nazionale (28,2%) e del gruppo di confronto (25,3%). La provincia di Foggia appare in ulteriore svantaggio rispetto a tutti gli indicatori del dominio e in particolare in riferimento al tasso di brevettazione (X.1: 11,8 domande per milione di abitanti), ai flussi di nuovi laureati in S&T (X.5: 5,6 per mille abitanti) e alla specializzazione produttiva nei settori produttivi *knowledge intensive* (X.7: 23,5%).

La *Qualità dei servizi* nelle “Altre province della puglia” si attesta su livelli bassi anche in confronto alla media del Gruppo Lazio e mezzogiorno su quasi tutte le misure considerate, ad eccezione degli indicatori relativi all’assistenza ospedaliera e al trasporto pubblico locale. Il tasso di emigrazione ospedaliera in altra regione (XI.2: 6,8%) è di poco superiore al nazionale (6,3%) ma apprezzabilmente inferiore a quello del gruppo di confronto (10,5%). Gli indicatori relativi al TPL dei capoluoghi di provincia denotano invece una densità di reti urbane (XI.5: 157,0 km per 100 kmq) e un’offerta di posti-km (XI.6: 2,5 posti per Km) superiori alla media del gruppo anche se ben al di sotto dei valori nazionali di confronto (4,6 posti per km). Gli indicatori della qualità dei servizi di *public utilities* e di servizi per l’infanzia marcano tutti posizioni di svantaggio, più evidenti rispetto all’Italia, ma piuttosto significative anche in confronto alla media del Gruppo Lazio e Mezzogiorno. Molto bassa è la

quota di bambini di 0-2 anni che usufruiscono di servizi per l'infanzia, pari a 5,0% contro l'8,6% del gruppo e il 13,5% nazionale; la raccolta differenziata di rifiuti urbani (XI.4: 26,4%) è molto più bassa di quella del gruppo (37,9%), già a sua volta distante dal dato nazionale (42,3%); le interruzioni lunghe di servizio elettrico senza preavviso (i cosiddetti *blackout*) sono quasi tre all'anno contro le 1,9 medie nel Paese. L'indice di sovraffollamento degli istituti di pena (XI.7: 138,8%), infine, denota una situazione particolarmente critica negli istituti penitenziari delle "Altre province pugliesi" sia in relazione al complesso del Gruppo Lazio e Mezzogiorno (112,5%) che all'Italia (108,3%). In confronto, il profilo della provincia di Foggia evidenzia livelli di qualità dei servizi peggiori su tutte le misure considerate e in particolare nella minore presa in carico dei servizi per l'infanzia (XI.1: 3,3), nella frequenza dei blackout, che sale a 4 nell'anno, nella raccolta differenziata, che non raggiunge il 15% dei rifiuti totali, nella densità delle reti di trasporto pubblico locale che a Foggia (55,2 Km per 100 Km²) è meno di un terzo di quella degli altri capoluoghi pugliesi.

4.2 Divergenze e disparità territoriali

L'analisi ha messo in evidenza numerose e rilevanti divergenze tra il livello e la struttura del benessere della provincia di Foggia e quello complessivo delle altre province pugliesi, che costituisce il principale e più rilevante elemento di criticità per la coesione territoriale della regione.

La significatività delle medie utilizzate per l'analisi dei profili di benessere delle "Altre province pugliesi", e quindi la convergenza tra questi territori, è alta per la larga maggioranza delle misure considerate. Tuttavia un esame sistematico delle differenze tra i singoli profili territoriali consentirebbe comunque di specificare ulteriormente le componenti del Bes e le aree su cui intervenire per migliorare la coesione e il livello complessivo di benessere del territorio pugliese.

A completamento dell'analisi svolta, è quindi utile approfondire l'illustrazione delle divergenze più rilevanti riscontrabili tra le "Altre province pugliesi", riscontrabili in particolare su alcune componenti del benessere.

Nel dominio *Istruzione e competenze* una differenza apprezzabile tra le province pugliesi riguarda la quota di giovani-adulti che non hanno conseguito il diploma superiore e non sono inseriti in un percorso formativo (II.1), la cui incidenza varia tra il 21,0% di Lecce e il 9,7% di Taranto, raggiungendo il picco massimo a Foggia, come già evidenziato (22,3%).

Nell'ambito del *Lavoro e conciliazione dei tempi di vita* il tasso di mancata partecipazione al lavoro discrimina apprezzabilmente, passando dal 33,6% di Bari al

39,1% della provincia Barletta-Andria-Trani, così come il tasso di occupazione giovanile che passa dal 32,9% di Brindisi al 50,9% di Lecce e il tasso di rischiosità per infortuni sul lavoro che dal 25,5% di Lecce passa al 36,2% di Taranto.

Nelle *Relazioni sociali* si denotano differenze nell'incidenza di scuole con percorsi interni ed esterni privi di barriere che varia tra il 17,2% di Brindisi al 26,7% di Bari, nella percentuale di acquisizioni di cittadinanza italiana sul totale stranieri che passa dall'1% di Taranto al 2,3% di Bari, raggiungendo il valore più basso a Foggia (0,7%), nella diffusione delle istituzioni non profit che varia da 2,2 per 10mila abitanti a Taranto a 3,2 a Brindisi, nella percentuale di volontari che passa da 4,2% a Barletta-Andria-Trani a 7,6% a Lecce.

Nel dominio *Sicurezza* le province pugliesi si discostano tra loro per il tasso di omicidi che varia da 0,2 per 100mila abitanti di Brindisi a 1,4 di Bari e raggiunge il valore massimo a Foggia (2,2 per 100mila abitanti), per i delitti violenti denunciati e diffusi più bassi a Taranto (rispettivamente 346,1 per 10mila abitanti e 171,6 per 10mila abitanti) e più frequenti a Bari (rispettivamente 489,3 per 10mila abitanti e 274,4), per la mortalità stradale in ambito extraurbano che varia dal 3,9% di Barletta-Andria-Trani al 7,1% di Taranto, raggiungendo il picco a Foggia (9,9%).

Nel dominio *Paesaggio* si osservano differenze nella percentuale di densità di verde storico e parchi urbani di notevole interesse pubblico, più bassa a Lecce e Taranto (0,1%) e più alta a Brindisi (2,7%), nelle strutture museali fruibili che variano da 0,2 per 10mila abitanti a Taranto a 0,6 a Brindisi e nel numero di visitatori che passa da 1,530,2 per 10mila abitanti a Lecce a 10.063,6 a Barletta-Andria-Trani.

Nel dominio *Ambiente* le province pugliesi sono discriminate dalla disponibilità di verde urbano che varia da 3,1 mq per abitante a Taranto a 32,2 mq a Barletta-Andria-Trani, dai giorni di superamento dei limiti del PM10 che passa da 8 giorni a Brindisi a 44 giorni a Barletta-Andria-Trani e raggiunge il valore minimo a Foggia (2,0), dalla disponibilità delle piste ciclabili più bassa a Brindisi (1,6 km per 100 Km²) e più elevata a Barletta-Andria-Trani (18,9 Km per 100 Km²), dalla percentuale di energia prodotta da fonti rinnovabili che varia da 13,4% a Taranto a 48,5% a Bari e raggiunge il valore massimo a Foggia (215%), dall'afflusso di rifiuti urbani in discarica che passa da 21,8 tonnellate per Km² a Bari a 193,2 tonnellate a Barletta-Andria-Trani.

Nella *Ricerca e innovazione* la propensione alla brevettazione discrimina le province pugliesi e varia da 8 per milione di abitanti a Taranto a 21,5 per milione di abitanti a Bari.

Nella *Qualità dei servizi* si denotano distanze marcate per la percentuale di bambini che usufruiscono di servizi per l'infanzia che passa da 2,8% a Barletta-

Andria-Trani a 8,6% a Brindisi, per la raccolta differenziata di rifiuti urbani più bassa a Taranto (15,7%) e più alta a Brindisi (42,4%), per la densità di reti urbane di TPL che varia da 67,8 km per 100 kmq a Brindisi a 247 km a Bari, per i posti-km offerti dal TPL più bassi a Lecce (1,6 posti per km) e più elevati a Taranto (3,6), per l'indice di sovraffollamento degli istituti di pena che varia da 101,5% di Barletta-Andria-Trani a 171,9% di Taranto.

5. Conclusioni

Le analisi svolte sugli indicatori del Bes delle province ne hanno evidenziato le potenzialità informative non soltanto per valutare livello e struttura del benessere nei singoli territori, ma anche per individuare relazioni spaziali e strutture di benessere comuni tra territori, e per esplorare le differenze territoriali nelle determinanti dei livelli di benessere.

La lettura degli indicatori di Bes delle province pugliesi indica che la Puglia si caratterizza innanzitutto per un dualismo territoriale tra Foggia e le altre province che condiziona negativamente la coesione territoriale. Infatti, la provincia di Foggia è risultata frequentemente svantaggiata rispetto a tutte le altre province su molte componenti e in numerosi domini del benessere. Per contro le altre province pugliesi hanno mostrato un buon grado di convergenza sulla gran parte delle misure di benessere, pur con alcune eccezioni importanti come nei domini *Lavoro*, *Benessere economico*, *Paesaggio e patrimonio culturale*, *Ricerca e innovazione*, *Qualità dei servizi*. In generale la posizione della regione evidenzia vantaggi e svantaggi relativi sia rispetto all'Italia che al Gruppo Lazio e Mezzogiorno, area omogenea individuata sulla base della valutazione multidimensionale del benessere, e che include tutte le province pugliesi tranne Foggia.

I principali vantaggi si riscontrano nei domini *Salute*, *Sicurezza e Ambiente*; gli svantaggi più evidenti nei domini *Lavoro*, *Benessere economico*, *Paesaggio e patrimonio culturale*, *Ricerca e innovazione* e *Qualità dei servizi*, che, come visto, sono anche connotati dalle divergenze territoriali più rilevanti e diffuse, rappresentando i principali ambiti in cui è possibile migliorare la coesione e l'equità territoriale del Benessere della Puglia.

Bibliografia

- Bartoloni, M., Carbonara, C., Pallotti, A. (2016). *Misurare il benessere equo e sostenibile a livello locale*. Comunicazione alla XXXVII Conferenza Italiana Di Scienze Regionali, Ancona (in corso di pubblicazione).
- Cnel-Istat (2012). *La misurazione del Benessere Equo e Sostenibile*. Documenti del Comitato Cnel-Istat sugli indicatori di progresso e benessere. <http://www.misuredelbenessere.it/fileadmin/upload/docPdf/Bes.pdf>
- Costa, R., Marchesich, E., Pallotti, A. (2016). *Una comparazione tra progettualità e incisività dell'azione degli enti di area vasta. L'esperienza del progetto Bes delle province*. Comunicazione alla XXXVII Conferenza Italiana di Scienze Regionali, Ancona, <http://www.aisre.it/images/aisre/578e3a75ebc6d3.66848590/Costa.pdf>
- Cuspi-Istat (2016). *Il Benessere equo e sostenibile delle province. Rapporto 2015*. <http://www.besdelleprovince.it/bes-volume-nazionale/>.
- Dente, B. (2011). *Le decisioni di policy*. Bologna, Il Mulino.
- EU COR - European Union Committee of The Regions (2012). *Delivering on the Europe 2020 Strategy, Handbook for Local and Regional Authorities*.
- Fiorillo, F., Muscillo, C., Taralli, S. (2016). *Misure di benessere dei territori e programmazione strategica: il livello comunale*. Comunicazione alla XXXVII Conferenza Italiana Di Scienze Regionali, Ancona (in corso di pubblicazione).
- Istat, (2013). *Bes 2013. Il benessere equo e sostenibile in Italia*. Roma, Istat.
- Istat (2015). *Bes 2015. Il Benessere Equo e sostenibile in Italia*. Roma, Istat. <http://www.misuredelbenessere.it/>
- Mazziotta, M., Pareto, A. (2013). *Methods for constructing Composite Indices: One for All or All for One?* Rivista Italiana di Economia, Demografia e Statistica, Vol. LXII, n.2, pp. 67-80.
- OECD (2014). *How's Life in Your Region? Measuring Regional and Local Well-being for Policy Making*.
- Taralli S. (2013). *Indicatori del Benessere Equo e Sostenibile delle province: informazioni statistiche a supporto del policy-cycle e della valutazione a livello locale*. Rassegna Italiana di Valutazione, Franco Angeli, Milano, XVI, 55, 171:187.
- Taralli, S., Capogrossi, C., Perri, G. (2015). *Mesasuring equitable and sustainable well-being (Bes) for policy-making at local level (NUTS3)*. Rivista Italiana di Economia Demografia e Statistica Volume LXVIV n. 3/4 Luglio-Dicembre.

Allegato 1. *Gli indicatori del Bes delle province per dimensione*

Dimensione	Indicatori
Salute	I.1 Speranza di vita alla nascita - Maschi
	I.2 Speranza di vita alla nascita - Femmine
	I.3 Tasso di mortalità infantile
	I.4 Tasso standardizzato di mortalità per incidenti di trasporto (15-34 anni)
	I.5 Tasso standardizzato di mortalità per tumore (20-64 anni)
	I.6 Tasso di mortalità per demenza (65 anni e+)
	I.7 Tasso di mortalità per suicidio
	I.8 Tasso standardizzato di mortalità evitabile (0-74 anni)
Istruzione e formazione	II.1 Giovani che hanno abbandonato precocemente gli studi
	II.2 Persone in età lavorativa con istruzione non elevata*
	II.3 Partecipazione all'istruzione secondaria superiore
	II.4 Partecipazione all'istruzione terziaria (19-25 anni)
	II.5 Partecipazione all'istruzione terziaria S&T (19-25 anni)
	II.6 Livello di competenza alfabetica degli studenti
	II.7 Livello di competenza numerica degli studenti
	II.8 Persone in età lavorativa in formazione permanente
Lavoro e conciliazione dei tempi di vita	III.1 Tasso di mancata partecipazione al lavoro (15-74 anni)
	III.2 Differenza di genere nel tasso di mancata partecipazione (F-M)
	III.3 Tasso di occupazione (20-64 anni)
	III.4 Differenza di genere nel tasso di occupazione (F-M)
	III.5 Tasso di occupazione giovanile (15-29 anni)
	III.6 Giornate retribuite nell'anno - lavoratori dipendenti
	III.7 Tasso di disoccupazione (15-74 anni)
	III.8 Tasso di disoccupazione giovanile (15-29 anni)
	III.9 Tasso di rischiosità per infortuni sul lavoro
	III.10 Tasso di rischiosità grave per infortuni sul lavoro
Benessere economico	IV.1 Reddito lordo disponibile per famiglia
	IV.2 Retribuzione media annua lavoratori dipendenti
	IV.3 Importo medio annuo delle pensioni
	IV.4 Pensionati con pensione di basso importo
	IV.5 Ammontare medio del patrimonio familiare
	IV.6 Differenza di genere nella retribuzione media dei lavoratori dipendenti (F-M)
	IV.7 Differenza di generazione nella retribuzione media dei lavor. dipend.
	IV.8 Provvedimenti di sfratto da abitazioni emessi
	IV.9 Tasso di ingresso in sofferenza dei prestiti bancari alle famiglie
Relazioni sociali	V.1 Scuole con percorsi privi di barriere (sia interni che esterni)
	V.2 Scuole con soli percorsi interni privi di barriere
	V.3 Scuole con soli percorsi esterni privi di barriere
	V.4 Presenza di alunni disabili nelle scuole
	V.5 Acquisizioni della cittadinanza italiana nell'anno
	V.6 Diffusione delle cooperative sociali
	V.7 Diffusione delle istituzioni non profit
	V.8 Volontari per 100 abitanti di 14 anni e più

Politica e istituzioni	VI.1 Tasso di partecipazione alle elezioni europee
	VI.2 Tasso di partecipazione alle elezioni regionali
	VI.3 Percentuale di donne nelle amministrazioni comunali
	VI.4 Percentuale di giovani (<40 anni) nelle amministrazioni comunali
	VI.5 Amministrazioni provinciali: grado di finanziamento interno
	VI.6 Amministrazioni provinciali: capacità di riscossione
	VI.7 Comuni: grado di finanziamento interno
	VI.8 Comuni: capacità di riscossione
Sicurezza	VII.1 Tasso di omicidi
	VII.2 Delitti denunciati
	VII.3 Delitti violenti denunciati
	VII.4 Delitti diffusi denunciati
	VII.5 Morti per 100 incidenti stradali
	VII.6 Morti per 100 incidenti stradali sulle strade extraurbane (escluse le autostrade)
Paesaggio e patrimonio culturale	VIII.1 Consistenza del tessuto urbano storico in buone condizioni
	VIII.2 Densità di verde storico e parchi urbani di notevole interesse pubblico
	VIII.3 Strutture museali fruibili
	VIII.4 Visitatori delle strutture museali fruibili
Ambiente	IX.1 Disponibilità di verde urbano
	IX.2 Superamento dei limiti di inquinamento dell'aria - PM10 (n° massimo)
	IX.3 Consumo di elettricità per uso domestico
	IX.4 Acqua potabile erogata giornalmente
	IX.5 Densità di piste ciclabili
	IX.6 Energia prodotta da fonti rinnovabili
	IX.7 Afflusso in discarica di rifiuti urbani (anche da fuori provincia)
Ricerca e innovazione	X.1 Propensione alla brevettazione (domande presentate)
	X.2 Incidenza dei brevetti nel settore High-tech
	X.3 Incidenza dei brevetti nel settore ICT
	X.4 Incidenza dei brevetti nel settore delle biotecnologie
	X.5 Flussi di nuovi laureati in S&T residenti
	X.6 Flussi di nuovi laureati in S&T residenti (totale)
	X.7 Specializzazione produttiva in settori ad alta intensità di conoscenza
Qualità dei servizi	XI.1 Bambini di 0-2 anni che usufruiscono di servizi per l'infanzia
	XI.2 Emigrazione ospedaliera in altra regione
	XI.3 Interruzioni di servizio elettrico senza preavviso
	XI.4 Raccolta differenziata di rifiuti urbani
	XI.5 Densità di linee urbane di trasporto pubblico locale nei capoluoghi di provincia
	XI.6 Posti per chilometro offerti dal trasporto pubblico locale nei capoluoghi di provincia
	XI.7 Indice di sovraffollamento degli istituti di pena

Fonte: Il Benessere equo e sostenibile delle province. Rapporto 2015.

Antroponimia a Ruvo di Puglia tra XII e XIV secolo: alcune analisi statistiche

Vito Ricci*

Università degli studi di Bari Aldo Moro
Direzione Affari Istituzionali, Unità Operativa Statistiche di Ateneo

Riassunto: In questo contributo si prende in esame il *corpus* onomastico desunto da un obituario trecentesco proveniente dalla Cattedrale di Ruvo di Puglia che registra le morti dei benefattori del Capitolo in un lasso temporale che va dalla fine del XII secolo alla metà del XIV. Questi dati di tipo qualitativo possono essere oggetto di analisi statistica che permette di fornire una descrizione dello stock onomastico anche in termini quantitativi relativamente ad alcuni aspetti come la concentrazione, la variabilità, la densità, etc. Sono utilizzate come variabili di segmentazione il genere e l'origine del nome che hanno permesso di costruire dei sottogruppi per effettuare delle comparazioni. Infine si è utilizzato un modello teorico per la descrizione dei dati attraverso la stima dei parametri con il metodo dei minimi quadrati e il calcolo di una misura di *goodness of fit*.

Keywords: antroponimia; statistica linguistica; modello di Pareto; Ruvo di Puglia.

1. Introduzione

Una fonte interessante e utile ai fini dello studio delle vicende socio-economiche di Ruvo nel periodo dal XII al XIV secolo è costituita dalla *Matricula maioris ecclesie Rubensis* (Ficco, 2005). Si tratta di un obituario o necrologio, un libro o registro nel quale venivano annotati i nomi dei defunti con il giorno della settimana e la data di morte (*obitus*) e le donazioni pro anima, in natura o in denaro, che venivano effettuate a favore del Capitolo della Cattedrale al fine di ricevere suffragi da parte di quest'ultimo. Per tale scopo nell'obituario si teneva conto solo dell'anniversario

* Corresponding author: vito.ricci@uniba.it.

della morte e raramente veniva riportato l'anno della morte del defunto. I vari *obiti* sono in genere segnati in un calendario, da mani diverse, a seconda dell'epoca dell'inserzione della nota necrologica. È questo anche il caso dell'obituario di Ruvo, nel quale mancano completamente riferimenti cronologici dei decessi, anche se dal registro si possono ricavare delle informazioni che permettono di definire un lasso temporale degli *obiti* tra la fine del XII secolo e la prima metà del XIV.

L'obituario di Ruvo, contenendo un elenco di persone defunte che avevano lasciato dei beni al Capitolo della Cattedrale, è una fonte utilizzabile per un'indagine sull'antroponimia; si tratta di un ramo dell'onomastica che ha per oggetto lo studio degli antroponimi, ossia dei nomi propri attribuiti alle persone. Questo tipo di analisi in ambito medievale ha avuto una certa diffusione negli ultimi decenni del XX secolo, soprattutto per merito di ricercatori francesi¹ e in seguito si è affermata anche tra gli studiosi italiani. Non mancano dei contributi aventi per oggetto delle realtà del Mezzogiorno italiano o, più in particolare della nostra regione. Nella maggior parte dei casi si tratta di analisi di tipo dinamico che trattano dell'evoluzione degli antroponimi in un lasso temporale più o meno ampio, mentre nel nostro caso, non avendo indicazioni di natura temporale in riferimento agli antroponimi presenti nel necrologio, l'analisi sarà essenzialmente di tipo statico. In particolare ci occuperemo dell'esame dei nomi di persona, distinti secondo il genere e l'origine.

2. Lo stock onomastico

Gli individui presenti nella *Matricula* sono in numero di 670 (372 maschi pari al 55,5% e 298 femmine pari al 44,5%) che può essere considerato un campione abbastanza discreto se si considera comunque limitato ad un ambito geografico piuttosto ristretto (una città di piccole dimensioni che nel 1321 contava 431 fuochi pari a circa 2.000 anime). Occorre tenere presente che per essere inseriti nell'obituario occorre pagare una quota che non tutti potevano permettersi, soprattutto i ceti popolari. Il numero degli individui prende il nome di dimensione del *corpus* onomastico (n). Sin dagli anni Trenta del XX secolo lo storico francese Bloch metteva l'accento sull'importanza dello studio dei nomi osservando che «la scelta dei nomi di battesimo, la loro natura, la loro frequenza relativa, sono altrettanti tratti che, convenientemente interpretati, rivelano delle correnti di pensiero o di sentimenti ai quali lo storico non dovrebbe restare indifferente» (Bloch, 1932). In tali analisi so-

¹ Basta fare riferimento ai diversi convegni su *Genèse médiévale de l'anthroponymie moderne* organizzati dal 1986 al 2008.

vente trova spazio l'impiego di tecniche statistiche (Chareille, 2008, Chareille 2012). Nelle Tabelle 1 e 2 si riportano i nomi maschili e quelli femminili più diffusi a Ruvo come risultano dal necrologio, dopo aver raggruppato i nomi con forme o grafie simili (lemmatizzazione). I nomi maschili distinti sono in numero di 129 (47,8% del totale), mentre quelli femminili in numero di 141 (52,2%). Si osserva quindi nel *corpus* una leggera prevalenza di nomi muliebri.

Tabella 1. *Nomi propri maschili più diffusi a Ruvo tra XII e XIV secolo.*

Nome proprio	Occorrenze	
	N.	%
<i>Iohannis</i>	38	10,2
<i>Nicolaus</i>	25	6,7
<i>Petrus</i>	25	6,7
<i>Angelus</i>	16	4,3
<i>Guillielmus</i>	12	3,2
<i>Robertus</i>	10	2,7
Altri	246	66,1
Totale	372	100,0

Tabella 2. *Nomi propri femminili più diffusi a Ruvo tra XII e XIV secolo.*

Nome proprio	Occorrenze	
	N.	%
<i>Margarita</i>	17	5,7
<i>Iohanna</i>	10	3,4
<i>Gilia</i>	8	2,7
<i>Maria</i>	7	2,3
<i>Churamaria</i>	7	2,3
<i>Perrona</i>	7	2,3
<i>Beatrix</i>	7	2,3
<i>Gemma</i>	7	2,3
<i>Pasca</i>	6	2,0
Altri	222	74,5
Totale	298	100,0

Nello stock esaminato, il nome proprio maschile con la maggiore diffusione è Giovanni (10% del totale degli individui di genere maschile presenti nella *Matricula*), seguito al secondo posto da Pietro e Nicola, più distanziati Angelo, Guglielmo e Roberto; il quadro onomastico di Ruvo, non risulta dissimile da quello del resto del

Mezzogiorno nel Trecento (Martin, 1996). Tra le donne il nome proprio più comune era Margherita, seguito da Giovanna e *Gilia*, ipocoristico di Egidia, e poi Maria, il nome composto *Churamaria*, *Perrona*, Beatrice, Gemma e *Pasca* (Pasqua).

3. La descrizione dello stock onomastico attraverso gli indicatori statistici

Il *corpus* dei nomi può essere analizzato da un punto di vista quantitativo attraverso una serie di indicatori (Chareille, 2008, Chareille, 2012), mettendo in evidenza diverse caratteristiche. Tra gli indicatori più semplici che servono da base per costruire quelli più complessi troviamo il numero di individui che sono presenti nel *corpus* (n) e il numero di nomi distinti ad esso associato (K_n). In seguito si calcola un indicatore che esprime il numero medio di individui per ciascun nome calcolato rapportando il numero di individui n al numero dei nomi propri K_n :

$$\frac{n}{K_n}$$

Esprime una misura della densità dei nomi, potendo variare tra 1 (minima omogeneità, si verifica quanto tutti gli individui hanno un nome diverso l'uno dall'altro) e il numero degli individui (massima omogeneità, si ha quando tutti gli individui portano lo stesso identico nome proprio). La densità è direttamente proporzionale alla dimensione del collettivo osservato. Nel nostro campione assume un valore pari a 2,88 per i maschi e 2,11 per le femmine, mostrando una maggiore densità dei nomi all'interno del primo gruppo. Il reciproco moltiplicato per 100 costituisce un altro indicatore che misura il numero di nomi ogni 100 individui:

$$\frac{K_n}{n} * 100$$

Nel nostro stock abbiamo che esso vale 34,7% per il gruppo dei maschi e 47,3% per quello delle femmine. La ricchezza (eterogeneità) del *corpus* onomastico si può misurare in diversi modi. Il primo è l'indice di eterogeneità di Gini (Girone, 2009):

$$IE = 1 - \sum_{i=1}^{K_n} f_i^2$$

dove f_i è la frequenza relativa del nome i -esimo, e K_n è il numero totale di nomi distinti. Esso misura il grado di disuguaglianza, variando tra 0, nel caso di massima omogeneità (tutti portano il medesimo nome), a $(K_n-1)/K_n$, nel caso di massima

concentrazione (tutti portano un nome diverso); rapportando IE al suo valore massimo si ottiene una misura relativa che varia nell'intervallo $[0,1]$. Tale indicatore è pari a 0,98 per il gruppo maschile e 0,99 per quello femminile; i nomi maschili presentano una eterogeneità leggermente inferiore rispetto a quelli femminili. Tale aspetto indica una maggiore ricchezza dello stock femminile. Un indicatore alternativo per misurare l'eterogeneità del *corpus* dei nomi è l'indice di entropia di Shannon (Girone, 2009):

$$S = - \sum_{i=1}^{K_n} f_i \log_e f_i$$

rapportando S al suo valore massimo pari a $\log_e K_n$ si ottiene una misura relativa che varia tra 0 e 1. L'indice di entropia normalizzato risulta pari a 0,87 per il gruppo dei nomi maschili e a 0,93 per quello femminile, mettendo in evidenza come il secondo collettivo sia più eterogeneo e come esso presenta una maggiore ricchezza onomastica. L'indice di entropia consente di percepire meglio il divario in termini di variabilità tra i due collettivi rispetto all'indice di Gini che presenta un lieve scarto tra i due gruppi.

Collegato all'eterogeneità dello stock onomastico è la presenza di nomi *hapax*, ovvero di quei nomi che presentano una frequenza assoluta unitaria; si tratta dei nomi rari. Lo stock da un lato si compone di un numero di nomi particolarmente diffusi, mentre dall'altro vi sono quelli piuttosto singolari e infrequenti. Nello stock relativo a Ruvo di Puglia il numero di nomi *hapax*, indicato in genere in letteratura con K_1 , è pari a 84 per il genere femminile (corrispondente al 60% del totale dei nomi femminili), mentre per il genere maschili ammonta a 67 (corrispondente al 51,9% del totale dei nomi maschili). Si riscontra quindi una maggiore incidenza dei nomi rari tra le donne rispetto agli uomini, circostanza questa che conferma la maggiore ricchezza di questo gruppo.

L'aspetto che si contrappone ai nomi rari è la misura della concentrazione che può computarsi attraverso l'indice di Gini (Girone, 2009):

$$G = \frac{\sum_{i=1}^{K_n-1} (P_i - Q_i)}{\sum_{i=1}^{K_n-1} P_i}$$

dove P_i è la frequenza cumulata relativa sino al nome i -esimo, mentre Q_i è la frequenza cumulata relativa del rango dell' i -esimo nome. Il suo campo di variazione è compreso tra 0, in caso di equi-distribuzione ovvero quando ogni individuo porta un nome diverso, e 1, in caso di massima concentrazione ovvero quando tutti gli individui portano un medesimo nome. È un aspetto che risulta opposto all'eteroge-

neità: alla massima variabilità corrisponde l'equi-distribuzione, mentre alla minima variabilità corrisponde la massima concentrazione.

Nel nostro caso G è pari a 0,42 per il collettivo femminile e a 0,52 per quello maschile, mostrando una maggiore concentrazione dei nomi per quest'ultimo raggruppamento. Tale indicatore è preferibile ad altri utilizzati in genere in letteratura (Chareille, 2012), come ad esempio la quota percentuale rappresentata dai primi 2, 5 e 10 nomi più diffusi, in quanto tiene conto di tutta la distribuzione e non solo di una piccola parte della stessa.

Un'altra metrica è il tasso di omonimia che misura la probabilità che due individui scelti casualmente in un dato insieme portino il medesimo nome. Tale probabilità è nulla se e solo se tutti i nomi del *corpus* sono distinti, mentre è pari a 1 se e solo se tutti gli individui dell'insieme portano lo stesso nome (Chareille, 2012):

$$TO = \frac{\sum_{i=1}^{K_n} n_i(n_i - 1)}{n(n - 1)}$$

dove n_i è la frequenza assoluta del nome i -esimo. Il tasso di omonimia è pari al 2,6% per il collettivo maschile e al 1,1% per quello femminile. Esso mostra come a Ruvo nel periodo in esame era molto più probabile trovare due uomini con il medesimo nome che non due donne.

Nella Tabella 3 si riporta il riepilogo di tutti gli indicatori menzionati in precedenza distinti in base al genere.

Gli indicatori proposti permettono di descrivere lo stock onomastico sotto gli aspetti della numerosità, estensione, densità, ricchezza/eterogeneità, omonimia e concentrazione.

Tabella 3. Indicatori statistici del corpus onomastico secondo il Genere.

Indicatore		Genere		Totale
		M	F	
n	N. individui	372	298	670
K_n	N. nomi	129	141	270
K_n/n	N. individui/N. nomi	2,88	2,11	2,48
n/K_n*100	N. nomi/N. individui*100	34,7	47,3	40,3
IE	Indice di eterogeneità di Gini	0,98	0,99	0,99
S	Indice di entropia di Shannon	0,87	0,93	0,91
K_1	N. nomi <i>hapax</i>	67	84	151
K_1/K_n*100	Incidenza percentuale nomi <i>hapax</i> su totale nomi	51,9	59,6	55,9
G	Indice di concentrazione di Gini	0,52	0,42	0,32
TO	Tasso percentuale di omonimia	2,6	1,1	1,0

4. La classificazione dei nomi

È possibile raggruppare i nomi in base all'origine come riportato nella Tabella 4, in cui sono presenti le frequenze assolute.

Tabella 4. *Distribuzione osservata dei nomi dello stock secondo l'Origine e il Genere.*

Origine	Genere		Totale
	M	F	
Araba	2	1	3
Biblica	6	1	7
Cristiana	37	19	56
Germanica	9	7	16
Greca	17	27	44
Latina	32	57	89
Longobarda	4	2	6
Normanna	8	2	10
Slava	0	2	2
Incerta	14	23	37
Totale	129	141	270

Se esaminiamo i profili riga (Tabella 5) si osserva una prevalenza in termini relativi dei nomi maschili in tutte le origini, tranne che per quelli di origine Greca, Latina, Slava e Incerta ove prevalgono i nomi femminili.

Tabella 5. *Distribuzione percentuale dei nomi dello stock secondo il Genere, per Origine.*

Origine	Genere		Totale
	M	F	
Araba	66,7	33,3	100,0
Biblica	85,7	14,3	100,0
Cristiana	66,1	33,9	100,0
Germanica	56,3	43,8	100,0
Greca	38,6	61,4	100,0
Latina	36,0	64,0	100,0
Longobarda	66,7	33,3	100,0
Normanna	80,0	20,0	100,0
Slava	0,0	100,0	100,0
Incerta	37,8	62,2	100,0
Totale	47,8	52,2	100,0

Tabella 6. Distribuzione percentuale dei nomi dello stock secondo l'Origine, per Genere.

Origine	Genere		Totale
	M	F	
Araba	1,6	0,7	1,1
Biblica	4,7	0,7	2,6
Cristiana	28,7	13,5	20,7
Germanica	7,0	5,0	5,9
Greca	13,2	19,1	16,3
Latina	24,8	40,4	33,0
Longobarda	3,1	1,4	2,2
Normanna	6,2	1,4	3,7
Slava	0,0	1,4	0,7
Incerta	10,9	16,3	13,7
Totale	100,0	100,0	100,0

Analizzando i profili colonna (Tabella 6) in termini percentuali, i nomi di origine latina sono quelli più numerosi, 33%, seguiti da quelli di origine cristiana e poi da quelli di origine greca/bizantina. Il 13,7% dei nomi è di origine incerta, molto spesso si tratta di nomi formati nel Medioevo e poi scomparsi. L'incidenza dei nomi latini risulta maggiore per il genere femminile (40,4%), mentre tra gli uomini prevalgono i nomi cristiani (28,7%). Se calcoliamo l'indice di quadratico di contingenza (Cramer) per misurare la connessione (Girone, 2009) tra le variabili qualitative Origine e Genere si ottiene un valore pari a 0,2 che esprime un legame estremamente labile tra le due mutabili essendo il campo di variazione dell'indice tra 0 (indipendenza) e 1 (massima connessione).

Nella Tabella 7 è riportato il riepilogo degli indicatori statistici calcolati in riferimento a ciascuna modalità della mutabile Origine.

I diversi sottogruppi presentano differenze abbastanza sensibili per quanto concerne i parametri dimensionali (n e K_n) ciò influisce di conseguenza sul numero medio di individui per nome che varia da un minimo di 1,43 (nomi di origine Biblica) ad un massimo di 4,5 (nomi di origine Cristiana). Si osserva una correlazione positiva tra n e K_n con coefficiente di Pearson pari a 0,859 e mediamente un incremento di 10 individui nel campione comporta una crescita di 3 nomi nel *corpus*. Per quanto riguarda l'eterogeneità e la ricchezza del *corpus* onomastico, i gruppi più omogenei e meno ricchi sono quelli dei nomi di origine Normanna, Araba e Cristiana; se si esamina l'incidenza percentuale dei nomi *hapax* il valore massimo si riscontra per i nomi di origine Incerta, seguono poi quelli di origine Latina.

Tabella 7. *Indicatori statistici del corpus onomastico secondo l'Origine.*

Indicatore	Origine										Totale
	Araba	Biblica	Cristiana	Germanica	Greca	Latina	Longobarda	Normanna	Slava	Incerta	
n	9	10	252	27	83	180	9	38	3	59	670
K _n	3	7	56	16	44	89	6	10	2	37	270
K _n /n	3,00	1,43	4,50	1,69	1,89	2,02	1,50	3,80	1,50	1,59	2,48
n/K _n *100	33,3	70,0	22,2	59,3	53,0	49,4	66,7	26,3	66,7	62,7	40,3
IE	0,85	0,98	0,96	0,97	0,98	0,99	0,98	0,88	0,89	0,98	0,99
S	0,85	0,97	0,84	0,94	0,93	0,92	0,97	0,81	0,92	0,94	0,91
K ₁	1	4	19	10	27	56	3	4	1	26	151
K ₁ /K _n *100	33,3	57,1	33,9	62,5	61,4	62,9	50,0	40,0	50,0	70,3	55,9
G	0,31	0,17	0,36	0,24	0,27	0,28	0,17	0,31	0,25	0,23	0,32
TO	36,1	6,7	5,3	5,4	2,7	1,9	8,3	18,3	33,3	2,7	1,0

I gruppi che presentano la maggiore concentrazione sono quelli dei nomi di origine Cristiana, seguiti da quelli di origine Araba e Normanna. I raggruppamenti nei quali è più elevato il tasso di omonimia sono quelli di dimensione ridotta (nome di origine Araba e di origine Slava). Gli indicatori statistici consentono di affermare come all'interno delle diverse tipologie di nomi individuate in base all'origine degli stessi vi siano delle differenze non solo quantitative, ma anche di tipo qualitativo.

Nell'analisi sull'onomastica a Ruvo di Puglia nel periodo in esame si è cercato di seguire le linee d'indagine suggerite da Bloch; si è trattato della frequenza e la diffusione dei nomi, anche attraverso dei dati di natura quantitativa; si è indagato sulla natura e l'origine dei nomi, mettendo in evidenza come lo stock onomastico su una base latino-cristiana (che copre più del 50% dei nomi) sia caratterizzato dalla stratificazione delle diverse dominazioni straniere succedutesi nel Mezzogiorno (arabi, bizantini, longobardi, normanni, svevi) che hanno lasciato anche una traccia ben visibile nell'onomastica. Più difficile risulta trarre delle notizie in merito alla scelta dei nomi di battesimo, i dati permettono poco di esprimersi su questo punto evidenziato da Bloch; in alcuni casi sembra esservi come criterio di scelta quello di attribuire ai figli i nomi dei nonni; questa modalità era ancora molto utilizzata sino ad alcuni decenni fa.

5. Il modello di Pareto

Nel campo dell'antroponimia è possibile mutuare dei modelli teorici che vengono utilizzati nel settore della statistica linguistica, come la legge empirica di Zipf sulla relazione tra il rango di un nome e la sua frequenza (Chareille, 2008). In un ordinamento decrescente di nomi in base alla frequenza, la frequenza è inversamente proporzionale al rango occupato dal nome nella graduatoria. Sia n_i la frequenza assoluta associata al nome i -esimo in una graduatoria decrescente e r_i il relativo rango, secondo la legge di Zipf abbiamo:

$$n_i = \frac{C}{r_i}$$

dove C è una costante che varia in base alla dimensione del *corpus*. In altri termini C è uguale alla frequenza del nome più diffuso n_1 . La legge di Zipf prevede un decremento progressivo della frequenza di una nome proporzionale all'aumentare del suo rango. Sebbene questa legge sia stata oggetto di numerose critiche da parte di alcuni Autori che ne hanno messo in evidenza i limiti, essa continua ad avere grande applicazione nell'ambito della linguistica e in altri contesti delle scienze sociali attraverso una sua generalizzazione detta modello di Pareto (Chareille, 2008):

$$n_i = A \cdot r_i^{-b}$$

nel caso in cui $b=1$ esso coincide con la formula proposta da Zipf. In scala doppio-logaritmica il legame assume una forma lineare:

$$\log(n_i) = \log(A) - b \cdot \log(r_i)$$

I parametri A e b possono essere facilmente stimati con il metodo dei minimi quadrati (Girone, 2009). Il primo parametro è influenzato dall'ampiezza del *corpus* onomastico, mentre il secondo è un indice inverso della ricchezza del *corpus*, all'aumentare di b si riduce la ricchezza dello stock onomastico, collegato alla maggiore o minore concentrazione dei nomi. Nelle Figure 1 e 2 si riporta il diagramma di Pareto in scala naturale in scala doppio-logaritmica per lo stock completo dei nomi desunti dall'obituario di Ruvo. Dai due grafici si osserva subito come il legame tra rango e frequenza assoluta venga descritto molto bene dal modello di Pareto; in secondo luogo emerge un altro particolare che si riscontra spesso nell'applicazione di tale modello nel campo dell'antroponimia (Chareille, 2008): i nomi che occupano i primi posti nella graduatoria decrescente risultano avere delle frequenze osservate inferiori rispetto a quelle previste dal modello teorico, in lette-

ratura si definisce tale fenomeno come appiattimento iniziale o convessità sommitale.

Figura 1. *Diagramma di Pareto (stock onomastico totale, scala naturale).*

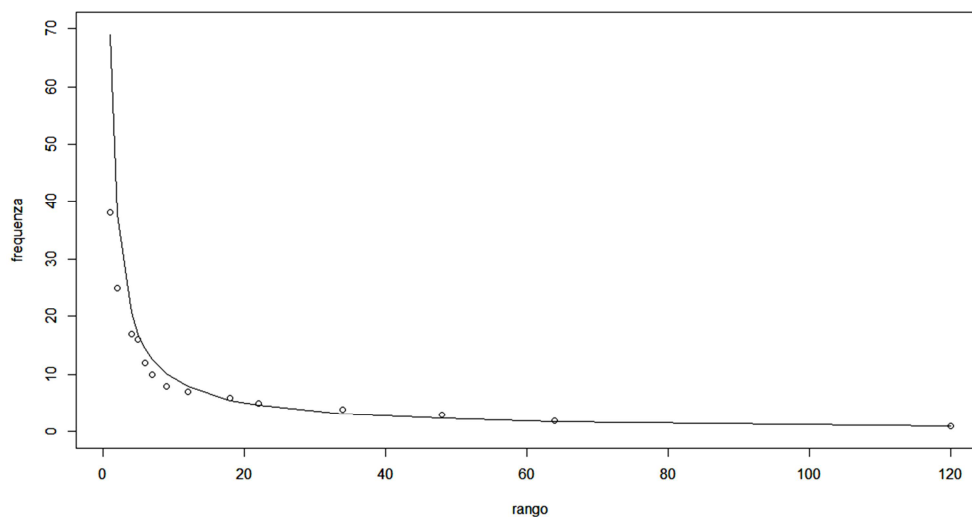
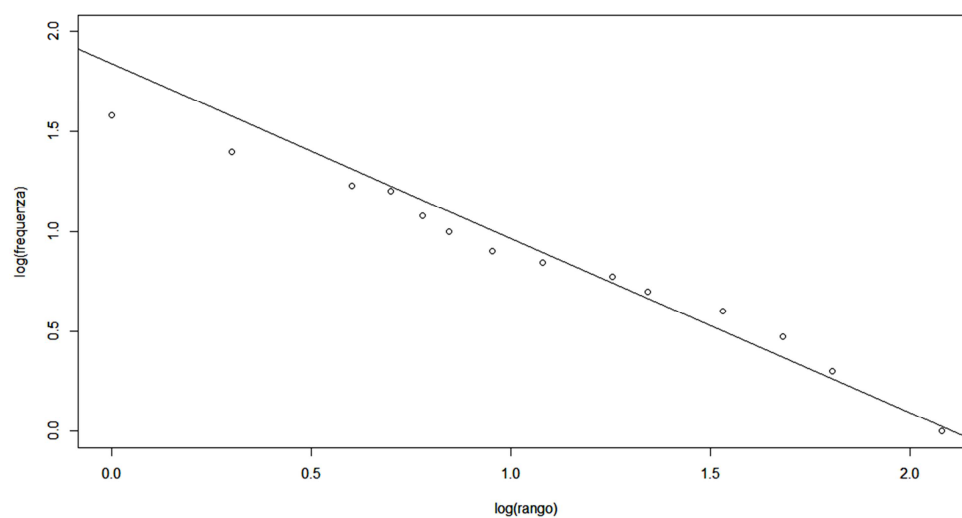


Figura 2. *Diagramma di Pareto (stock onomastico totale, scala doppio-logaritmica).*



Come si è scritto, il parametro A è correlato alla dimensione dello stock e lo si può considerare come un parametro dimensionale; il parametro b , che nel diagramma in scala doppio-logaritmica è dato dal coefficiente angolare della retta, invece riflette la ricchezza dello stock, essendo correlato negativamente con l'etero-

geneità dei nomi presenti nello stock: tale relazione è pienamente verificata se si confronta b con l'indice di Shannon, al crescere del primo il secondo si riduce.

Tabella 8. *Stima dei parametri del modello di Pareto in alcuni stock onomastici.*

Stock onomastico	A	b	R ²
Tutto	69,0084	0,8718	0,9734
Nomi maschili	57,8522	0,9708	0,9910
Nomi femminili	28,4649	0,8063	0,9435
Nomi di origine cristiana	64,0040	1,0970	0,9626
Nomi di origine latina	22,1080	0,8620	0,9597
Nomi di origine greca	10,2920	0,7950	0,9725

Nel presente studio, il parametro b varia tra 0,795 per il gruppo dei nomi di origine Greca, che è quello che presenta la maggiore ricchezza (maggiore eterogeneità), e 1,097 per il gruppo dei nomi di origine Cristiana, che è quello meno ricco e più omogeneo, che è anche l'unico a presentare $b > 1$, mentre in tutti gli altri casi tale parametro è sempre inferiore all'unità. L'indice di determinazione risulta molto elevato, con valori superiori a 0,94 mettendo in evidenza come il modello di Pareto ben si adatta a descrivere il legame tra rango e frequenza assoluta dei nomi presenti nell'obituario di Ruvo. Il risultato migliore ($R^2=0,991$) si verifica in corrispondenza del gruppo dei nomi di genere maschile. I valori del parametro b non risultano molto dissimili da quelli stimati per altri stock onomastici medievali sebbene di dimensioni maggiori (Chareille, 2008).

6. Conclusioni

Partendo da un obituario del XIV secolo si è costruito il *corpus* dei nomi che da esso si è derivato; i nomi sono stati distinti in base al genere e in relazione all'origine. Sono stati utilizzati alcuni indicatori statistici per descrivere le caratteristiche principali dell'intero stock onomastico e dei suoi sottogruppi in termini di numerosità, estensione, densità, ricchezza/eterogeneità, omonimia e concentrazione.

Tra gli aspetti emersi vi è la differenza tra i nomi secondo in genere: il gruppo dei nomi femminili è risultato essere più eterogeneo, e quindi più ricco, rispetto a quello maschile. I nomi femminili a Ruvo nel Medioevo presentavano una varietà maggiore di quelli maschili che invece si concentravano intorno ad alcune modalità ricorrenti come ad esempio i nomi di alcuni Santi (Giovanni, Pietro, Nicola). Que-

sta circostanza comportava una maggiore probabilità di trovare due omonimi tra gli uomini rispetto alle donne.

Se invece si considera l'origine dei nomi, essa riflette la storia della Puglia medievale con la presenza di nomi che rimontano le radici nei principali popoli stranieri che hanno dominato il Mezzogiorno italiano; tuttavia i nomi di origine latina e quelli di origine cristiana nel complesso rappresentano il 53,7% del totale dei nomi, mostrando come lo zoccolo duro sia costituito dall'eredità romana e dalla successiva cristianizzazione. I diversi gruppi distinti in base all'origine dei nomi presentano caratteristiche differenti come messo in evidenza dalle metriche calcolate.

In ultimo si è applicato il modello di Pareto per descrivere il legame tra la frequenza dei nomi e il posto occupato nella graduatoria decrescente. Si sono stimati i parametri del modello per il complesso dello stock onomastico e per alcuni sotto-raggruppamenti e in tutte le circostanze esso ha presentato un grado di adattamento molto buono.

Riferimenti bibliografici

- Bloch, M. (1932). Noms de personne et histoire sociale. *Annales d'histoire économique et sociale*, 4: 67-69.
- Chareille, P. (2008). *Genèse médiévale de l'antroponymie moderne. 6. Le nom: histoire et statistiques: quelles méthodes quantitatives pour une étude de l'antroponymie médiévale*. Université de Tours, Tours.
- Chareille, P. (2012), Anthroponymie et statistique: quelques outils d'analyse. In: A. Addobbati, R. Bizzocchi, G. Salinero (eds.), *L'Italia dei cognomi: l'antroponimia italiana nel quadro mediterraneo*. PLUS-Pisa University Press, Pisa: 169-201.
- Ficco, F. (2009). La "Matricula" della Cattedrale di Ruvo. Un obituario inedito del '300. *Odegitria*, 12: 315-379.
- Girone, G. (2009). *Statistica*. Cacucci, Bari.
- Martin, J. M. (1996), L'Italie méridionale. In: M. Bourin, J. M. Martin, F. Menant (eds.), *L'antroponymie document de l'histoire sociale des mondes méditerranéens médiévaux*. Actes du colloque international organisé par l'École française de Rome avec le concours du GDR 955 du C.N.R.S. «Genèse médiévale de l'antroponymie moderne» (Rome, 6-8 octobre 1994). Rome: 29-39.

Degenze ospedaliere per malattie cardiovascolari in Puglia

Najada Firza¹, Rossana Mancarella^{2*}

¹Università "Nostra Signora del Buon Consiglio" di Tirana (AL),

²Agenzia Regionale per la Tecnologia e l'Innovazione - Regione Puglia

Riassunto: Nel presente lavoro è stata effettuata, usando tecniche classiche di esplorazione statistica uni/multivariata, una analisi di durata dei ricoveri per patologie cardiovascolari, aggregate per gruppo clinico di diagnosi (CCS, Clinical Classification Software), valutando l'influenza delle varie patologie e di altri fattori prognostici caratterizzanti ogni paziente ricoverato (età, sesso, provincia di residenza e altro ancora) sulla durata della degenza. La tecnica utilizzata è ben nota in ambito epidemiologico, ma è qui gestita in modo alternativo.

Keywords: degenza ospedaliera; analisi di durata; metodo di Kaplan-Meier; modello di Cox.

1. Introduzione

La presente ricerca ha lo scopo di quantificare i principali fattori che influenzano la durata di ricovero negli ospedali, anche allo scopo di utilizzare efficacemente le loro risorse. L'analisi è stata operata su un campione rappresentativo di oltre 11.500 ricoveri ordinari, estratto casualmente da una popolazione originaria di pazienti affetti da una o più malattie cardiovascolari e ricoverati negli ospedali pugliesi dal 2009 al 2014 in modalità ordinaria: dato l'oggetto dell'indagine, sono ovviamente esclusi i ricoveri in Day Hospital e in Day Surgery, nonché le cure ambulatoriali e in Pronto Soccorso non seguite da ricovero. Il database originale è stato ottenuto

* Autore corrispondente: r.mancarella@arti.puglia.it.

Il lavoro qui descritto è frutto di un progetto comune, ma R. Mancarella ha provveduto alla redazione dei paragrafi 1 e 2, mentre N. Firza ha redatto i paragrafi 3 e 4.

Gli autori ringraziano il A.R.E.S.-Puglia, che ha fornito i dati utilizzati per le analisi statistiche.

dalla registrazione completa delle Schede di Dimissione Ospedaliera (SDO) di tutti i pazienti ricoverati negli ospedali pugliesi durante il periodo su definito.

La durata dell'ospedalizzazione dei pazienti è stata assunta come variabile oggetto al fine di sviluppare un'analisi di sopravvivenza, qui meglio definita "analisi di durata", dato che l'interesse non è incentrato sul tempo di sopravvivenza degli ammalati, ma sul tempo trascorso tra il ricovero e la dimissione (pur se, ovviamente, l'eventuale decesso del ricoverato va preso in considerazione nell'analisi in quanto "interrompe" l'osservazione: si tratta dunque di un "dato censurato").

Pur se cambia l'ottica di analisi, la tecnica utilizzata nel presente lavoro è comunque parallela a quella delle classiche indagini epidemiologiche che tentano di mettere in evidenza l'esistenza di relazioni tra variabili esplicative (fattori prognostici) ed eventi obiettivo, che sono sia variabili a risposta binaria (sano-ammalato, vivo-morto, e così via) che variabili temporali (tempi di sopravvivenza, durata della malattia, o altro).

2. Materiali e metodi

Un'indagine statistica ha bisogno, in generale, di tre fasi:

1. una fase di progettazione preliminare per l'indagine, che deve definire;
 - la popolazione da indagare e la variabile risposta che rappresenta l'evento da studiare;
 - le unità di rilevazione e la tecnica di raccolta di dati relativi a queste;
 - l'analisi dei costi eventuali dell'indagine, in termini di denaro e di tempo;
 - le ipotesi da verificare attraverso l'individuazione di variabili esplicative che possono influenzare la variabile di risposta;
2. la fase dell'indagine statistica vera e propria;
3. l'analisi e l'interpretazione dei risultati.

Come accennato introduttivamente, la popolazione di indagare è costituita dai pazienti ricoverati presso le strutture ospedaliere pugliesi (pubbliche e private), negli anni dal 2009 al 2014, con almeno una diagnosi relativa a patologie cardiovascolari. La variabile risposta è la durata del ricovero, tenendo conto di eventuali decessi che determinano logicamente l'interruzione del ricovero per motivi opposti a quelli auspicabili (dimissione dei pazienti per guarigione almeno parziale).

Le unità di rilevazione sono le Schede di Dimissione Ospedaliera registrate dal personale sanitario al momento della dimissione del paziente tenendo conto delle informazioni presenti in cartella clinica. Le variabili osservate sono in primo luogo le variabili di identificazione dei ricoverati (sesso, età, residenza) e poi la durata di

ricovero, gli elementi caratteristici del medesimo (struttura sanitaria in cui viene registrato il ricovero, reparto di ammissione e di dimissione, trasferimenti e così via) e infine le variabili cliniche, ossia diagnosi, procedure diagnostiche e interventi effettuati, e infine l'esito del ricovero.

La raccolta dei dati viene dunque effettuata a livello amministrativo, e la qualità dei dati medesimi viene costantemente monitorata dall'Agenzia Regionale Sanitaria pugliese (AReS-Puglia). Di tale amplissimo database (oltre 4 milioni di schede registrate in 6 anni) sono state prese in considerazione solo le schede relative a pazienti che avessero almeno una diagnosi relativa a patologie cardiovascolari (categoria IX della classificazione internazionale delle malattie, 10^{ma} revisione, codici da I00 a I99 di ICM/10)¹. Tale sub-popolazione, benché meno ampia di quella relativa a tutti i ricoveri del periodo, ammonta comunque a oltre un milione e trecentomila SDO, di cui poco meno di 1.150.000 sono i ricoveri ordinari.

In un primo momento è stata effettuata un'analisi descrittiva della popolazione di ricoveri ordinari, replicata poi, per verifica, su un campione rappresentativo pari a circa il 10% della popolazione medesima; detta analisi mostra innanzitutto (Tab. 1) che in media sopravvive oltre il 96% dei pazienti ricoverati per diagnosi cardiovascolare in Puglia nel periodo 2009-2014, pur se con un trend di mortalità in leggero aumento.

Tabella 1. *Distribuzione degli esiti dei ricoveri per diagnosi cardiovascolare in Puglia, 2009-2014.*

anno	Ricoveri ordinari registrati			%	%
	Sopravvivenenti	Deceduti	Totale		
2009	201.511	6.493	208.004	18,1	3,1
2010	199.040	6.553	205.593	17,9	3,2
2011	186.230	6.806	193.036	16,8	3,5
2012	176.458	7.299	183.757	16,0	4,0
2013	174.833	7.242	182.075	15,9	4,0
2014	168.476	7.580	176.056	15,3	4,3
Totale	1.106.548	41.973	1.148.521	100,0	3,7

La dinamica dei decessi su riportata non deve però allarmare, scontando innanzitutto il sempre più frequente ricorso prima al day hospital e poi ai servizi ambulatoriali per il trattamento delle forme patologiche più lievi (evitando dunque il ricovero in struttura ospedaliera e conseguentemente la rilevazione tramite SDO), ma

¹ Le malattie cardiovascolari sono la principale causa di morte in Italia e in Europa, con un notevole impatto in termini di morbilità e mortalità, e un numero di ricoveri ospedalieri sempre molto alto. Il Sistema Sanitario Nazionale dedica moltissime risorse alla prevenzione, diagnosi e cura dei tali patologie.

anche il crescente invecchiamento della popolazione pugliese, e sempre dunque la maggiore esposizione a potenziali eventi cardiocircolatori fatali.

Ai fini della presente ricerca, ciò che interessa è la quota non eccessiva di decessi ospedalieri nella popolazione osservata, e il mantenimento di tale condizione nel campione estratto, al fine di garantirne una buona rappresentatività unitamente al mantenimento della dinamica temporale. La Tab. 2 conferma queste due condizioni, pur se con alcune variazioni casuali. Il campione così determinato può essere dunque assunto come base di riferimento per l'analisi dei ricoveri.

Tabella 2. *Distribuzione degli esiti dei ricoveri per diagnosi cardiovascolare nel campione estratto, 2009-2014.*

anno	Ricoveri ordinari campionati			%	%
	Sopravvivenuti	Deceduti	Totale		
2009	1.969	66	2.035	17,9	3,2
2010	2.040	78	2.118	18,7	3,7
2011	1.944	63	2.007	17,7	3,1
2012	1.741	70	1.811	16,0	3,9
2013	1.658	71	1.729	15,2	4,1
2014	1.569	73	1.642	14,5	4,4
Totale	10.921	421	11.342	100,0	3,7

L'ipotesi da verificare è che alcune delle variabili relative al ricovero influenzino la durata del ricovero medesimo, e, se ciò appare logico in riferimento alle patologie individuali e agli interventi che si rendono necessari per farvi fronte, meno logico e più interessante è l'eventualità che alcune delle altre variabili relative al ricovero influenzino in modo rilevante la durata del ricovero.

Si vuole qui studiare quali sono le caratteristiche più importanti per la permanenza in struttura ospedaliera, tenendo conto della limitata ma persistente influenza dei decessi fra i ricoverati.

La metodologia di indagine è dunque una analisi di durata, utilizzando le tecniche della più nota analisi di sopravvivenza. In particolare, viene qui utilizzato il modello semiparametrico di Cox per l'identificazione delle variabili che influenzano la durata del ricovero, riducendo eventuali effetti confondenti.

Il modello di Cox è, in campo biomedico, uno dei più noti e importanti modelli per l'analisi di sopravvivenza, sotto l'ipotesi di rischi proporzionali nel tempo.

Esso viene detto "semiparametrico" in quanto non viene fatta nessuna ipotesi sulla forma della funzione di sopravvivenza, e quindi sulla funzione di rischio, mentre viene parametrizzata l'influenza delle variabili esplicative (o fattori prognostici) sul rischio medesimo.

I suoi fondamenti teorici partono appunto dal concetto di funzione di rischio, la cui stima generale (con eventi di interruzione, default o decesso, indicati genericamente con d_i , non necessariamente equispaziati nel tempo) è data, in un gruppo di n unità statistiche con r tempi distinti di interruzione, da:

$$\hat{H}_k(t) \approx \sum_{i=1}^k \frac{d_i}{n_i} = \sum_{i=1}^k \hat{q}_i, \quad \text{con } t_k \leq t < t_{k+1} \text{ e } k=1, 2, \dots, r,$$

ove $r < n$, e n_i pari al numero di unità sopravvissute fino all'evento d_i (Kaplan & Meier, 1958).

Il modello di Cox nasce per verificare, tramite un campione, l'influenza che sui tempi di sopravvivenza possono avere alcune caratteristiche, sia quantitative che qualitative (variabili esplicative o "fattori prognostici"), possedute dalle unità campionate. Esso definisce $h_i(t)$ la funzione di rischio al tempo t (per $t \geq 0$) per una qualsiasi unità con determinati valori dei fattori prognostici e $h_0(t)$ la funzione di rischio al tempo t per una qualsiasi unità che abbia valore zero dei fattori prognostici (tale condizione implica che i fattori quantitativi siano riscaldati in modo che sia nullo il loro valore minimo, e che i fattori qualitativi siano definiti con variabili indicatrici nulle in corrispondenza delle modalità di base, spesso assunte come relative al minimo rischio). La funzione $h_0(t)$ è detta *funzione di rischio baseline*.

Ciò premesso, ipotizzando che in ogni istante la funzione di rischio $h_i(t)$ sia proporzionale alla funzione $h_0(t)$ in ragione di un parametro ψ , il modello base di Cox è espresso dalla seguente formulazione²:

$$h(t) = \psi h_0(t) = \exp(\psi) h_0(t), \quad \psi \in \mathbb{R}.$$

Considerando ora in modo esplicito una variabile X relativa a un fattore prognostico, con determinazioni x_i che assumano valore 0 se l'unità appartiene al gruppo base, per l'*i*.mo individuo del campione la formulazione precedente diventa

$$h_i(t) = \exp(x_i) h_0(t) = \exp(x_i) h_0(t), \quad \psi \in \mathbb{R}.$$

Logicamente, se sulla sopravvivenza delle unità campionate influiscono più fattori, detto modello è generalizzato da:

$$h_i(t) = \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_m x_{mi}) h_0(t), \quad \beta_j \in \mathbb{R}.$$

Si tenga conto che il modello di Cox presuppone che:

² Si noti che $\psi = h(t)/h_0(t)$ rappresenta il rischio di default al tempo t di una qualsiasi unità statistica rispetto ad un'altra unità appartenente al gruppo base, per cui viene detto *rischio relativo* oppure *rapporto di rischio*. Naturalmente, se $\psi < 1$ (e, quindi, $\beta = \ln \psi < 0$) il rischio di default per detta unità è più piccolo rispetto a quello dell'unità appartenente al gruppo base. Il contrario avviene se $\psi > 1$ e $\beta > 0$. Cfr., ad es., Delvecchio, 2001.

- le variabili esplicative abbiano un effetto moltiplicativo sul rischio, ovvero sia che la funzione *link* sia di tipo esponenziale (ipotesi, questa, generalmente verosimile quando il rapporto di rischio sia positivo);
- il rapporto fra le funzioni di rischio di individui con diversi valori dei fattori prognostici sia costante fra i vari strati individuati dai fattori;
- l'invarianza nel tempo dei valori delle variabili esplicative: se ciò non avviene, sta a significare che non è più legittimo registrare i valori dei fattori prognostici (o di un set di essi) solo all'inizio dello studio; esiste, tuttavia, anche un modello per covariate non indipendenti dal tempo.

Per quanto detto, risulta sempre opportuno verificare preliminarmente almeno l'assenza dell'effetto stratificazione.

Si classificano i dati di sopravvivenza secondo i livelli dei fattori e si considerino tutti i sottogruppi (strati) definiti sulla base di tutte le J combinazioni dei livelli dei vari fattori, in modo che gli individui dello stesso strato siano caratterizzati dai valori delle stesse esplicative (nel caso di covariate quantitative, occorre effettuare una preliminare riduzione in classi).

In ogni sottogruppo si stima la funzione di sopravvivenza col metodo di Kaplan-Meier e si riportano in un riferimento cartesiano i punti $(t, \ln \hat{H}_j(t))$, dove $j=1,2,\dots,J$ individua il sottogruppo dove avviene l'analisi. Se i vari diagrammi (uno per ciascun valore di j) che si ottengono congiungendo i punti $(t, \ln \hat{H}_j(t))$ appaiono paralleli, allora i rischi dei vari gruppi possono considerarsi perfettamente proporzionali.

Il modello di Cox, comunque, è abbastanza robusto da fornire risultati affidabili anche quando il parallelismo delle linee non è perfetto; al limite, può essere sufficiente verificare che esse non si intersechino, se non, eventualmente, agli estremi (Delvecchio, 2001).

3. Analisi di durata dei ricoveri con il modello semi-parametrico di Cox

Per l'identificazione delle possibili variabili esplicative da includere nel modello si è dunque condotta una analisi preliminare con il metodo di Kaplan-Meier. I risultati grafici di tale analisi vengono riportati conclusivamente solo per i fattori prognostici significativamente influenzanti la durata del ricovero (a meno dei decessi).

I fattori prognostici che hanno prodotto curve di rischio e sopravvivenza grossomodo parallele (implicando dunque rischi all'incirca proporzionali), e che sono quindi stati inclusi nel modello di Cox sono descritti nel prospetto seguente.

a. CCS diagnostici cardiovascolari:

96. *Malattie delle valvole cardiache.*

98. *Ipertensione essenziale.*
 100. *Infarto miocardico acuto.*
 107. *Arresto cardiaco e fibrillazione ventricolare.*
 108. *Insufficienza cardiaca congestizia, non da ipertensione.*
- b. CCS di interventi e procedure cardiovascolari**
44. *Bypass dell'arteria coronarica.*
 47. *Cateterismo cardiaco diagnostico, arteriografia coronarica.*
 48. *Inserzione, revisione, sostituzione, rimozione di pacemaker cardiaco o defibrillatore automatico.*
 54. *Altro cateterismo vascolare, escluso cuore.*
- c. Classificazioni nosologiche (aggiuntive a patologie cardiovascolari)**
1. *Malattie infettive e parassitarie.*
 2. *Neoplasie.*
 13. *Malattie del sistema muscoloscheletrico e del tessuto connettivo.*
 16. *Lesioni e avvelenamenti.*
 17. *Sintomi, segni, condizioni indefinite, altri fattori influenzanti lo stato di salute.*
- d. Classificazioni operative (diverse da interventi cardiovascolari)**
8. *Operazioni sul sistema sanguigno e linfatico.*
 14. *Operazioni sull'apparato muscolo-scheletrico.*

Non proporzionali fra loro (o persino tempo-dipendenti, in qualche caso) sono invece tutte le caratteristiche “non cliniche” dei ricoverati e dei ricoveri.

Come si evince dalla Tab. 3, il modello ottenuto applicando alla variabile obiettivo “durata della degenza” (con censure derivate dagli eventuali decessi ospedalieri) la tecnica *forward stepwise* rispetto ai suddetti fattori prognostici, con probabilità di esclusione $\alpha=0,05$, risulta molto significativo. Il test di massima logverosimiglianza assicura, infatti, una probabilità quasi infinitesimale ($p<0,0000001$) che le relazioni identificate nel modello siano dovute al caso.

Anche gli effetti dei singoli fattori prognostici sono tutti statisticamente significativi, alcuni in misura eccezionale. Le variabili che incrementano la probabilità di “rischio” (di degenze brevi, comprendendo purtroppo anche i decessi) riducendo la “sopravvivenza”, ossia la durata della degenza, sono:

- CCS di *Arresto cardiaco* (“rischio” oltre 25 volte superiore alla mancanza di diagnosi appartenenti a tale classificazione);
- diagnosi appartenente alla Categoria nosologica “*Sintomi, segni, condizioni indefinite, altri fattori influenzanti lo stato di salute*” (probabilità 1,8 volte maggiore rispetto all’assenza di una di dette diagnosi);

Tabella 3. Fattori prognostici del modello di Cox per l'analisi della durata dei ricoveri.

Prognostic factors		SE	p-values	Exp()
Malattie delle valvole cardiache	-1,143	0,286	<0,001	0,319
Iperensione essenziale	-0,730	0,168	<0,001	0,482
Arresto cardiaco e fibrillazione ventricolare	3,234	0,108	<0,001	25,385
Insufficienza cardiaca congestizia, non da ipertensione	0,463	0,125	<0,001	1,589
Cateterismo cardiaco diagnostico, arteriografia coronarica	-0,837	0,284	<0,005	0,433
Inserzione, revisione, sostituzione, rimozione di pacemaker cardiaco o defibrillatore automatico	-1,417	0,582	<0,020	0,242
Altro cateterismo vascolare, escluso cuore	0,560	0,188	<0,005	1,750
Malattie del sistema muscoloscheletrico e del tessuto connettivo	-0,826	0,266	<0,005	0,438
Sintomi, segni, condizioni indefinite, altri fattori influenzanti lo stato di salute	0,594	0,125	<0,001	1,811

Test max logverosimiglianza: $\chi^2=3491,8$ (g.d.l.=10); $p<0,0000001$

- intervento di *Cateterismo vascolare non cardiaco* (probabilità 1,75 volte maggiore dei casi senza interventi);
- CCS di *Insufficienza cardiaca congestizia, non da ipertensione* (probabilità quasi 1,6 volte maggiore della sua mancanza).

Dette conclusioni sono tratte, come è evidente, dall'ultima colonna della tabella, che riporta il parametro di proporzionalità $\text{Exp}() =$ (detto anche *odds ratio*).

Alcuni fattori prognostici, presentando coefficienti di segno negativo, risultano invece "protettivi" dal rischio di degenze brevi (o decesso), e dunque sono statisticamente connessi a lunghe degenze; andando da quello più connesso a quello che invece influenza meno la durata della degenza, detti fattori sono:

- *Inserzione, revisione, sostituzione, rimozione di pace-maker cardiaco o defibrillatore automatico* (rispetto a chi non subisce interventi di questo tipo, la probabilità di degenze brevi si riduce a meno di un quarto);
- CCS di *Malattie delle valvole cardiache* (probabilità meno di un terzo rispetto a chi non soffre di tali patologie);
- interventi di *Cateterismo cardiaco diagnostico o arteriografia coronarica* e diagnosi della Categoria nosologica *Malattie del sistema muscoloscheletrico e del tessuto connettivo* riducono in misura quasi uguale la probabilità di degenze brevi (43-44% rispetto ai casi privi di tali fattori);
- anche il CCS di *Iperensione essenziale* riduce in maniera significativa (a meno della metà) la probabilità di brevi degenze.

Sono state ricercate anche possibili interazioni tra i fattori prognostici, ma alcune incongruenze (esclusione degli effetti dei singoli fattori in seguito all'inclusione

delle interazioni fra i medesimi) hanno sconsigliato questa ulteriore analisi. La Tab. 4 riporta, invece, i risultati di una semplice analisi *logit*³ di regressione della variabile di status (sopravvive/deceduto) dalle medesime variabili considerate nel precedente modello di Cox, allo scopo di comprendere il ruolo delle censure (appunto i decessi) all'interno di tale modello. Dal confronto con la Tab. 3, appare evidente che in riferimento al semplice evento morte perdono qualsiasi significatività statistica gli interventi di *Inserzione, revisione, sostituzione, rimozione di pacemaker cardiaco o defibrillatore automatico* e le *Malattie del sistema muscolo-scheletrico e del tessuto connettivo*, mentre risultano molto più rilevanti gli altri fattori prognostici atti a incrementare il rischio di decesso.

I fattori “protettivi” hanno invece, in termini di odds ratio, effetti leggermente inferiori ma nel complesso non molto dissimili da quelli del modello di Cox. Si può dunque presumere che la struttura di quest'ultimo modello sia influenzata solo parzialmente dall'andamento dei decessi ospedalieri, e che i suoi risultati siano sufficientemente affidabili ai fini della determinazione dei fattori prognostici.

Tabella 4. *Modello di regressione logit per l'analisi dei decessi ospedalieri*

Prognostic factors		SE	p-values	Exp()
Malattie delle valvole cardiache	-1,290	0,261	<0,001	0,275
Ipertensione essenziale	-0,881	0,169	<0,001	0,414
Arresto cardiaco e fibrillazione ventricolare	5,388	0,201	<0,001	218,671
Insufficienza cardiaca congestizia, non da ipertensione	1,184	0,147	<0,001	3,266
<i>Cateterismo cardiaco diagnostico, arteriografia coronarica</i>	-0,981	0,266	<0,001	0,375
<i>Altro cateterismo vascolare, escluso cuore</i>	3,015	0,205	<0,001	20,388
Sintomi, segni, condizioni indefinite, altri fattori influenzanti lo stato di salute	1,368	0,148	<0,001	3,926

Test max logverosimiglianza: $\chi^2=13314,3$ (g.d.l.=7); $p<0,000000001$

Partendo dai risultati qui descritti, osserviamo ora gli effetti dei fattori prognostici significativi sui tempi di degenza, utilizzando i grafici ricavati dagli stimatori di Kaplan-Meier precedentemente utilizzati per identificare le variabili che avevano titolo di essere incluse nel modello di Cox.

Innanzitutto, i ricoveri per patologie appartenenti al CCS “Malattie delle valvole cardiache” presentano durata che arriva a 100 giorni, con alta sopravvivenza e distribuzione dei decessi molto concentrata nel primo mese.

³ Per quanto riguarda i fondamenti dell'analisi logit, peraltro ormai usuale in molti ambiti di ricerca e a cui si riallaccia anche il modello semiparametrico di Cox, si veda, ad es., Delvecchio (2015).

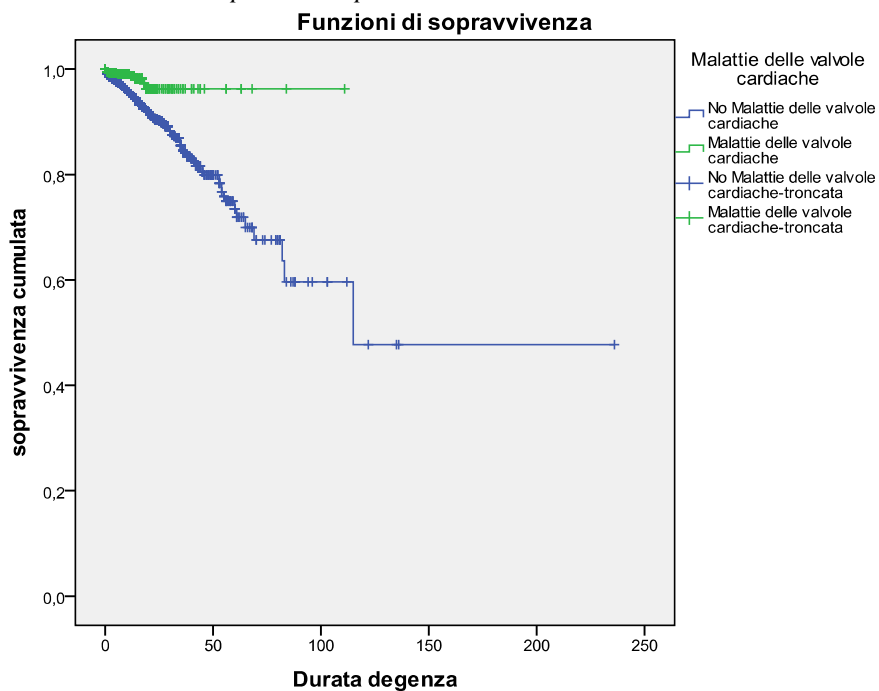
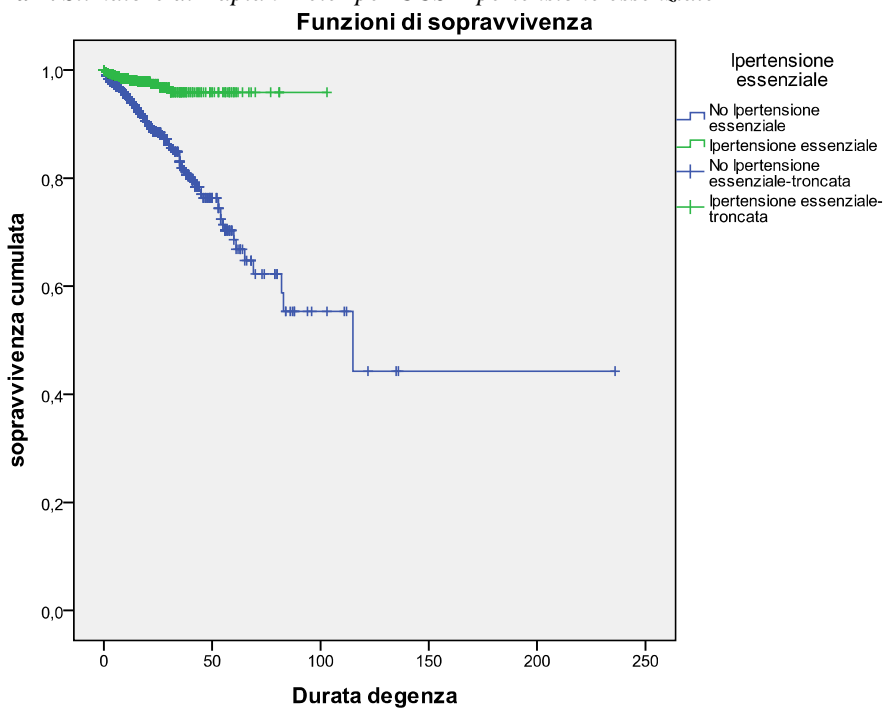
Figura 1. Stimatore di Kaplan Meier per CSS “Malattie delle valvole cardiache”**Figura 2.** Stimatore di Kaplan Meier per CCS “Ipertensione essenziale”

Figura 3. Stimatore di Kaplan Meier per CCS “Arresto cardiaco e fibrillazione ventricolare”
Funzioni di sopravvivenza

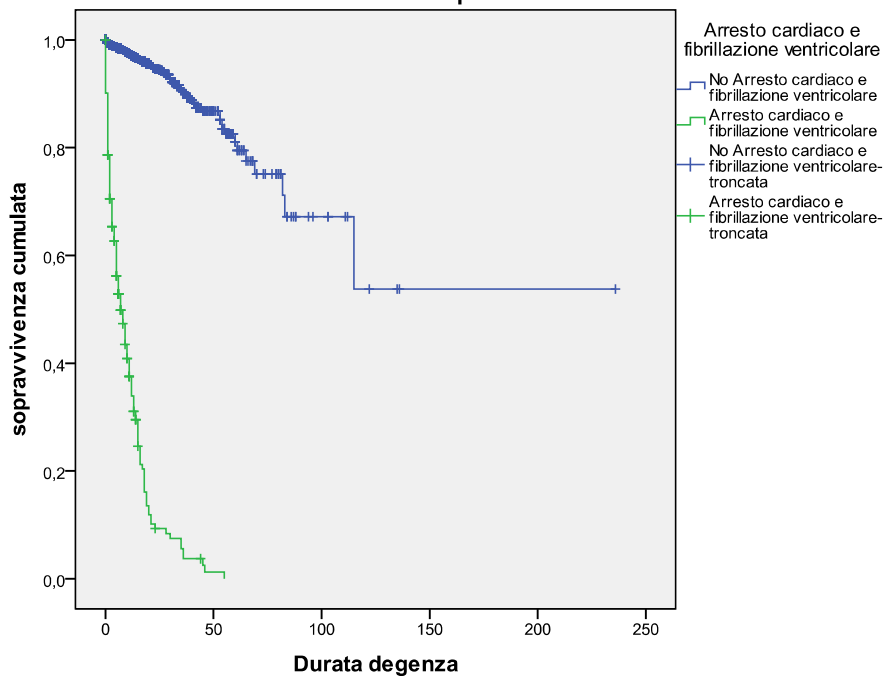


Figura 4. Stimatore K-M per CCS “Insufficienza cardiaca congestizia, non da ipertensione”
Funzioni di sopravvivenza

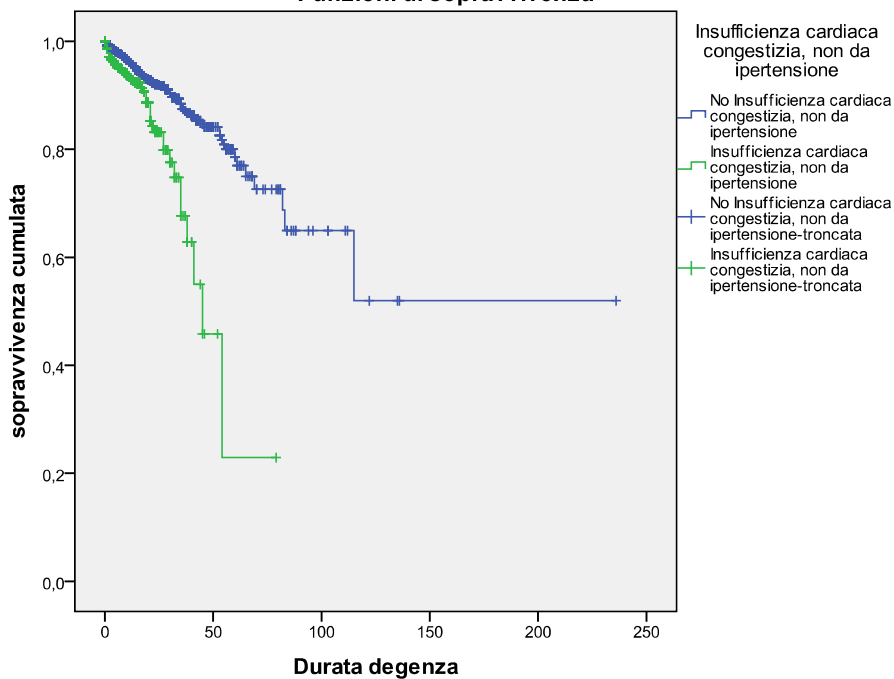


Figura 5. Stimatore K-M per “Cateterismo cardiaco diagnostico, arteriografia coronarica”
Funzioni di sopravvivenza

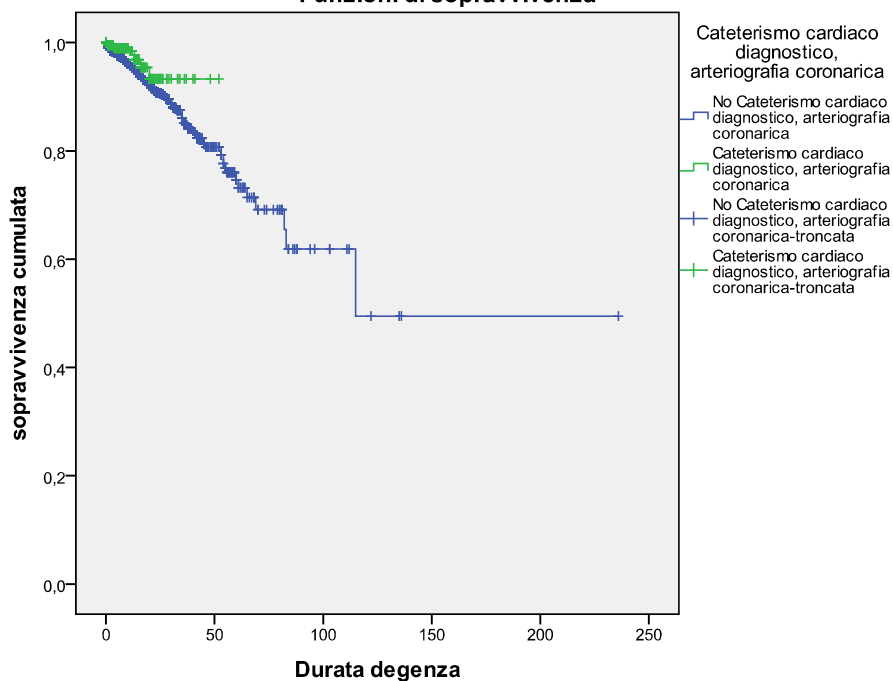


Figura 6. Stimatore di Kaplan Meier per “Altro cateterismo vascolare, escluso cuore”
Funzioni di sopravvivenza

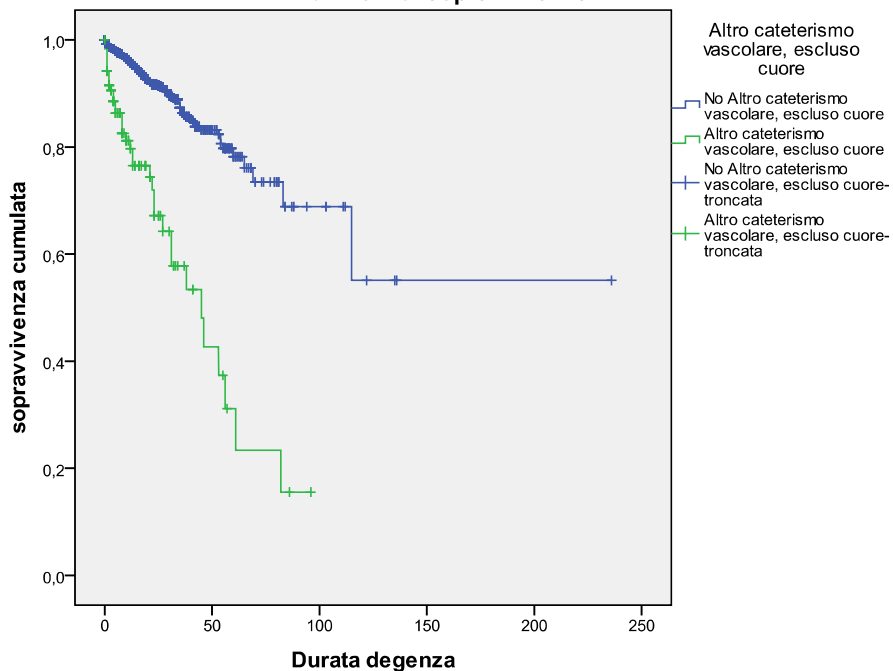
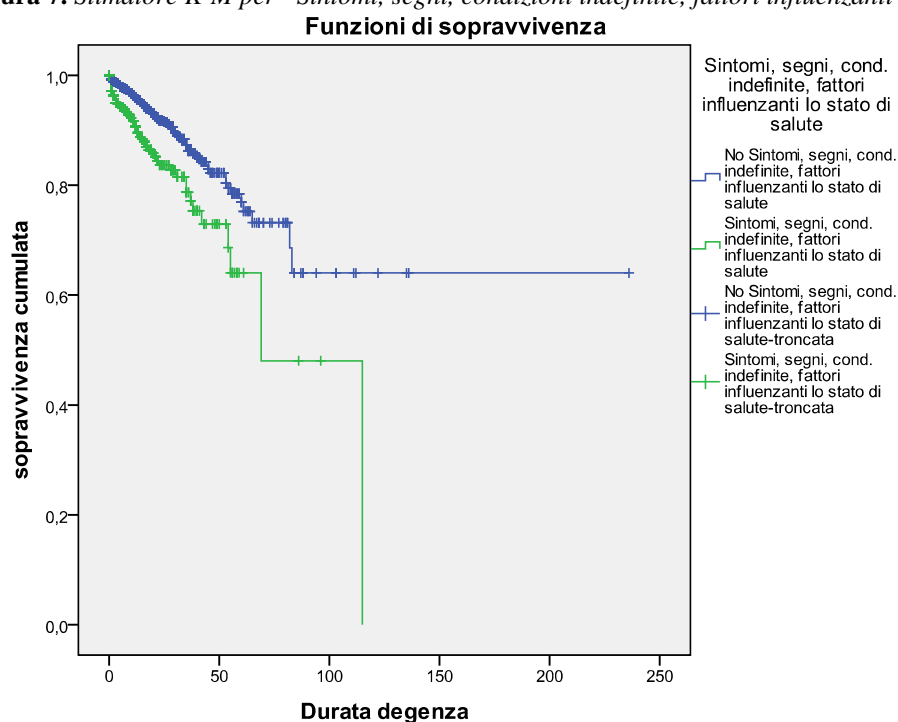


Figura 7. Stimatore K-M per “Sintomi, segni, condizioni indefinite, fattori influenzanti”

All'incirca le stesse osservazioni possono essere fatte a riguardo delle patologie appartenenti al CCS “Ipertensione essenziale” e per gli interventi di “Cateterismo cardiaco diagnostico, arteriografia coronarica” (le cui degenze non superano, tuttavia, i 50-60 giorni di durata). Del tutto opposte le dinamiche relative agli altri fattori prognostici, in particolare per le degenze relative al CCS di “Arresto cardiaco e fibrillazione ventricolare” (la cui curva di durata è ripidissima, con molti decessi concentrati nei primi 10 giorni di ricovero e che comunque non supera i 55 giorni).

4. Conclusioni

Questo lavoro ha voluto verificare statisticamente se la durata dei ricoveri ospedalieri per patologie appartenenti alla categoria nosologica delle *Patologie cardiovascolari*, in Puglia, nell'ultimo periodo con dati disponibili (2009-2014) risultasse influenzata da fattori non legati esclusivamente a motivi clinici.

La risposta a tale quesito, almeno per questa importantissima categoria nosologica, in base alle analisi qui effettuate risulta negativa, in quanto, di

tutti i possibili fattori prognostici, quelli risultati significativi sono tutti relativi alle categorie diagnostiche o ad alcuni interventi effettuati.

Certamente, è possibile che nei ricoveri ospedalieri per categorie diagnostiche meno esiziali vi sia più spazio per scelte meno dipendenti da motivazioni cliniche, come, ad esempio, l'età dei pazienti, o anche legate a motivi di convenienza economica. Potrebbe essere utile, dunque, verificare anche la durata delle degenze successive a dette categorie diagnostiche tramite una ulteriore analisi dei dati parallela a quella qui portata a termine.

Riferimenti bibliografici (esempi)

- Cox, DR (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society B*, 34: 187-220.
- Cox, DR (1975). Partial likelihood. *Biometrika*, 62: 269-276.
- Cox, DR; Snell, EJ (1981). *Applied statistics: Principles and examples*. Chapman & Hall, Londra.
- Delvecchio, F (2001). *Analisi dei tempi di sopravvivenza. Appunti a uso degli studenti*, Dipartimento di Scienze Statistiche dell'Università di Bari.
- Kaplan, EL; Meier, P (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* 53: 457-481.
- Tripepi, G; Catalano, F (2004). Kaplan-Meier analysis. *Giornale Italiano di Nefrologia*; 21 (6): 540-6.
- Van Dijk PC; Jager KJ; Zwinderman AH; Zoccali C; Dekker FW. (2008). The analysis of survival data in nephrology: basic concepts and methods of Cox regression. *Kidney Int*; 74 (6): 705-9.
- Provenzano, F; Tripepi, G; Zoccali, C (2010). How to measure effects in clinical research (corrected). *Giornale Italiano di Nefrologia*; 27 (3): 296-300.
- Provenzano, F; Versace, MC; Tripepi, R; Zoccali, C; Tripepi, G (2010). Confounding in epidemiology. *Giornale Italiano di Nefrologia*; 27 (6): 664-7.

Sull'integrazione tra fonti amministrative e statistiche per le imprese

Agata Maria Madia Carucci¹, Giovanni Vannella^{2*}

¹Istat; ²Università degli studi di Bari Aldo Moro

Riassunto: Oggetto del presente lavoro è l'analisi delle modalità attraverso cui le fonti amministrative riescono a migliorare l'offerta informativa statistica aziendale, verificando quindi lo stato di attuazione del regolamento sulle statistiche strutturali delle imprese basato su una costante integrazione tra fonti amministrative e statistiche al fine di ridurre il carico statistico per le imprese e migliorare la qualità dell'informazione prodotta.

Keywords: FRAME SBS; Rilevazioni PMI e SCI; Integrazione; Fonti amministrative; Censimenti permanenti

1. Introduzione

L'offerta di informazione statistica, come è noto, si differenzia a seconda che sia "ufficiale" o meno, ovvero a seconda che sia prodotta dai soggetti che fanno parte del Sistema Statistico Nazionale (SISTAN) o Europeo (Biggeri, 2012) o da altri soggetti ed enti terzi.

In un complesso caratterizzato da crescente dinamismo e concorrenza globalizzata, la produzione e diffusione di statistiche aziendali desta una sempre maggiore attenzione da parte dei numerosi operatori e per una pluralità di fini (fiscali, di programmazione economica, etc.) che richiedono informazioni sempre più tempestive, dettagliate e precise.

* Corresponding author: giovanni.vannella@uniba.it

Sebbene il lavoro sia il frutto del lavoro congiunto dei due autori, a A. M. M. Carucci sono attribuiti i paragrafi 3 e 4, a G. Vannella i paragrafi 1, 2, ed 5

Nell'ambito dell'offerta di tali statistiche, particolare rilevanza hanno, nel contesto delle pubblicazioni prodotte nell'ambito del SISTAN, le rilevazioni PMI e SCI dell'Istat.

Si ricorda che l'indagine PMI è una indagine di tipo campionario sulle piccole e medie imprese (1-99 addetti) e sull'esercizio di arti e professioni, mentre l'indagine SCI è una rilevazione di tipo censuario sul totale delle imprese con 100 ed oltre addetti.

L'offerta statistica aziendale, da tempo esistente, ha avuto un notevole impulso negli ultimi anni in quanto a livello comunitario l'Eurostat ha profuso un impegno particolarmente sentito finalizzato all'accrescimento di statistiche strutturali armonizzate a livello europeo. Tale impegno è stato "formalizzato" con l'introduzione del "Regolamento sulle Statistiche Strutturali sulle imprese" (di seguito SBS) n. 58/97, implementato con i successivi emendamenti fino ad arrivare all'attuale vigente Regolamento n. 295/2008.

I suddetti regolamenti presentano una doppia finalità: da un lato mirano ad ottenere un maggiore dettaglio informativo sugli aspetti strutturali del mondo imprenditoriale, maggiormente coerente con le mutate condizioni del mercato globale, dall'altro puntano a ridurre le incombenze comunicative delle imprese cercando di integrare le informazioni provenienti dalle indagini di cui sopra con le informazioni di tipo amministrativo già esistenti.

Il secondo aspetto è particolarmente rilevante nel caso del territorio italiano, essendo lo stesso caratterizzato da una forte presenza di imprese di piccole e piccolissime dimensioni con una media di addetti per impresa molto bassa ed una significativa incidenza rispetto al valore aggiunto totale.

Aspetto, quest'ultimo, che determina un onere piuttosto rilevante per le imprese che sopportano oneri burocratici particolarmente marcati al fine di ottemperare agli obblighi di comunicazione ai fini statistici, che tuttavia mal si conciliano con la necessità di disporre di statistiche strutturali contestualmente tempestive e sufficientemente dettagliate e precise.

Già in un precedente lavoro degli autori del presente contributo (Carucci, 2015), si è avuto modo di affrontare l'aspetto dell'utilizzo delle fonti amministrative nella statistica ufficiale, limitatamente all'ambito delle informazioni dichiarate dalle imprese negli "studi di settore".

Scopo del presente lavoro è illustrare le modalità attraverso cui si sta procedendo con l'implementazione della produzione di tali statistiche strutturali al fine di poter evidenziare i punti di forza delle stesse oggi utilizzabili dai principali *stakeholders* operanti sul territorio.

Per ultimo, si analizzerà il ruolo delle fonti amministrative nell'implementazione dei "nuovi" censimenti permanenti che, con d.lgs 179/2012, sostituiscono i tradizionali censimenti a cadenza decennale.

Il censimento diventa dunque, l'aggiornamento di un registro statistico attraverso l'integrazione di fonti amministrative con il supporto di indagini campionarie ad hoc.

2. Il FRAME SBS: Aspetti generali e diffusione dei dati

L'universo di riferimento per le statistiche strutturali sulle imprese è fornito annualmente dall'Archivio Statistico Asia (Archivio Statistico delle Imprese Attive) che l'Istat realizza partendo da una ampia base informativa di dati derivanti sia da informazioni statistiche, sia da fonti di tipo amministrativo.

L'archivio Asia è suddivisibile in due macro archivi:

- Imprese con meno di 100 addetti;
- Imprese con più di 100 addetti.

Fino al 2011, l'Istat ha risposto agli adempimenti del regolamento SBS con i dati rilevati dalle rilevazioni PMI e SCI.

A partire dal 2012, il regolamento SBS viene soddisfatto attraverso l'elaborazione dei dati del nuovo sistema FRAME SBS.

Tale archivio contiene variabili strutturali (attività economica, localizzazione territoriale, numero di addetti e dipendenti) ed economiche (fatturato, costi di acquisto di beni e servizio, costo del personale, valore aggiunto ed altre variabili più di dettaglio). Occorre in tal senso distinguere le informazioni presenti nel FRAME a seconda che riguardino le piccole o medie imprese con meno di 100 addetti, o le grandi imprese.

Per quanto concerne le prime, le informazioni del FRAME provengono dall'utilizzo delle fonti amministrative tenendo conto delle risultanze delle rilevazioni campionarie PMI; per quanto concerne le seconde, le informazioni provengono dalla rilevazione (di carattere censuario) SCI.

Tali dati, come si diceva in premessa, vengono ottenuti senza creare ulteriori appesantimenti comunicativi a carico delle imprese, bensì derivano da un processo di integrazione di stime applicate alle varie fonti amministrative esistenti (ad es. Inps, Camere di commercio, Agenzia delle entrate).

L'integrazione è il risultato di una lunga disamina relativa alle fonti amministrative esistenti ed utilizzabili unite alla risoluzione delle problematiche esistenti in

materia di protezione dei dati personali, aspetto questo del tutto evidente laddove si pensi, a solo titolo di esempio, all'utilizzo delle informazioni che le imprese trasmettono ai fini fiscali (Boselli, 2016).

Diversi enti provvedono alla produzione ed alla trasmissione dei dati. Oltre all'Istat, agiscono in tale ambito:

- Ivass (per quanto concerne le assicurazioni)
- La Banca d'Italia (per quanto concerne gli Enti creditizi)
- La Covip (per quanto concerne i Fondi pensione)

Le variabili più significative rilevate sono:

- numero di imprese;
- fatturato;
- valore della produzione;
- margine lordo sui beni destinati alla rivendita;
- valore aggiunto al costo dei fattori;
- margine operativo lordo;
- acquisto complessivo di beni e servizi;
- acquisto di merci da rivendere senza trasformazione lordi in beni materiali;
- variazione delle scorte di beni e servizi;
- acquisto di merci da rivendere senza trasformazione;
- variazione delle scorte di beni e servizi destinati alla rivendita;
- costo del personale;
- retribuzione lorda;
- ore lavorate;
- investimenti lordi in beni materiali;
- numero di persone occupate;
- numero di dipendenti;

La disponibilità dei dati è ottemperata attraverso la pubblicazione degli stessi su I.Stat ed attraverso i file di microdati.

In dettaglio, dal data warehouse I.Stat partendo dal tema "imprese" è possibile accedere al sistema informativo già aggiornato con i dati SBS attraverso il sotto tema "competitività".

Per quanto concerne i microdati, l'Istat rende disponibili due differenti tipologie di file

- FRAME – SBS – Sistema integrato di dati amministrativi e dati di indagine per la stima degli aggregati economici e delle imprese;
- TEC – FRAME SBS – Struttura e performance economica delle imprese esportatrici.

3. Struttura e costruzione del FRAME SBS

A partire dai dati 2012, i regolamenti statistici comunitari sulle imprese sono soddisfatti attraverso l'elaborazione dei dati:

- del nuovo sistema *FRAME* – imprese con meno di 100 addetti (file di microdati di fonte amministrativa trattati statisticamente, costruito per le imprese con meno di 100 addetti) combinati con quelli della rilevazione diretta PMI per le variabili non disponibili attualmente dalle fonti amministrative;
- della rilevazione SCI.

Data la struttura produttiva italiana, caratterizzata da imprese piccole e piccolissime, il *FRAME* – imprese con meno di 100 addetti rappresenta oltre il 99% delle imprese (Tab. 1).

Tabella 1. *Composizione universo di riferimento (FRAME-Sbs)*

Fonte	Imprese
Frame	4.253.338
Sci	10.540
TOTALE	4.263.878

Fonte: Istat

Il nuovo sistema *FRAME* per le imprese con meno di 100 addetti (più di 4 milioni di imprese) è basato sul trattamento statistico delle informazioni provenienti dalle fonti amministrative riportate in Tab. 2.

Tabella 2. *FRAME - Imprese con meno di 100 addetti per fonte*

Fonte	Imprese	Valore Aggiunto	Addetti
Bilanci civilistici	16,20%	53,80%	38,20%
Studi di settore	67,20%	36,60%	49,90%
Modello Unico	12,30%	3,10%	5,90%
Modello Irap	1,80%	4,40%	3,30%
Integrazione	2,50%	2,10%	2,70%

Fonte: Istat

La scelta della fonte amministrativa per l'integrazione segue un approccio gerarchico che definisce come prima fonte da utilizzare i Bilanci civilistici che ven-

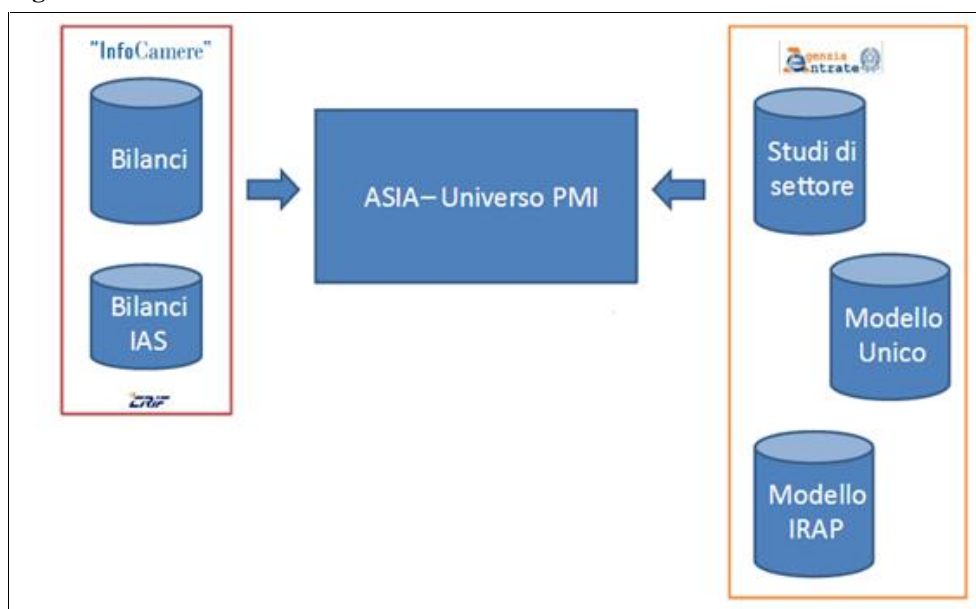
gono utilizzati per il 16% delle imprese sebbene rappresentino più del 50% in termini di Valore Aggiunto e poco meno del 40% in termini di addetti.

Gli Studi di settore sono la fonte utilizzata per quasi il 70% delle imprese che rappresentano meno che il 40% del Valore Aggiunto e il 50% degli addetti.

Il Modello Unico copre il 12% delle imprese, che rappresenta però il 3% in termini di Valore aggiunto e poco più del 5% in termini di addetti. Residuale risulta il contributo del modello Irap e dell'integrazione.

L'alimentazione del FRAME segue lo schema riproposto nella Fig. 1 (Luzi, 2014

Figura 1. *Struttura del FRAME*



Come si vede, da un lato le Camere di Commercio e dall'altro l'Agenzia delle entrate contribuiscono alla compilazione di tutte le variabili economiche riferite all'Universo di riferimento ASIA-PMI. Sebbene la copertura in termini di imprese e di Valore aggiunto e addetti sia particolarmente alta, più complessa risulta l'analisi di affidabilità di ciascuna fonte in termini di singole voci economiche.

Per questa ragione, risulta ancora indispensabile e centrale l'ausilio dell'indagine campionaria PMI. Da essa vengono ricostruite tutte le variabili economiche utili per la stima dei principali indicatori economici delle imprese.

In particolare, di seguito (Tab. 3) si mostra la copertura, o mancata copertura, delle fonti per le principali variabili economiche. Se per le variabili legate ai ricavi

e al costo del lavoro la copertura è particolarmente alta, la stessa cosa non può dirsi per le variabili di costo.

Tabella 3. *FRAME - Imprese con meno di 100 addetti: copertura delle principali variabili economiche*

	Fonte	No fonte
Ricavi	97,3	2,7
Variazioni rimanenze per prodotti finiti	97,4	2,6
Variazioni dei lavori in corso	97,4	2,6
Incrementi di immobilizzazioni	97,4	2,6
Altri ricavi e proventi	97,4	2,6
Costo per acquisto di beni	69,9	30,1
Spese per servizi	69,2	30,1
Godimento di beni di terzi	69,1	30,9
Variazione delle rimanenze materie prima e merci	42,4	57,6
Oneri diversi di gestione	70,8	29,2
Costo del lavoro	100,0	0,0

Fonte: Istat

L'utilizzo del FRAME-SBS e, dunque, delle fonti amministrative nelle statistiche strutturali sulle imprese, rappresenta un importante valore aggiunto in termini di ricchezza e accuratezza informativa sebbene meriti, come tutte le fonti amministrative, la dovuta attenzione in fase di costruzione dell'archivio e analisi delle variabili.

In particolare, la disponibilità di dati a livello micro, oltre a ridurre l'onere statistico sulle imprese, permette di:

- produrre stime molto dettagliate (per classi dimensionali, forma giuridica, collocazione territoriale ecc.);
- migliorare la coerenza complessiva delle statistiche economiche, in quanto il FRAME è alla base sia dei dati SBS che della Contabilità Nazionale;
- produrre analisi su scala longitudinale, potendo disporre di un panel di notevoli dimensioni per n-anni (il FRAME è stato predisposto dall'anno 2010)
- mettere in coerenza dati provenienti da fonti diverse;
- rappresentare meglio alcuni settori economici in cui la stima campionaria risentiva della forte variabilità che caratterizzava le imprese del settore.

Di contro, la dipendenza dalle fonti amministrative potrebbe comportare

- difficoltà nella tempistica richiesta dai regolamenti EUROSTAT per problemi legati alla fornitura dei dati amministrativi;
- rischi nella continuità informativa delle fonti amministrative (ad esempio, le novità legate agli Studi di settore potrebbero comportare l'eliminazione degli stessi nei prossimi anni o la semplificazione dei modelli di rilevazione);
- problemi legati alle modifiche normative sulla riservatezza dei dati statistici.

Di seguito (Tabb. 4 e 5) si riporta la distribuzione per numero imprese, addetti, fatturato e valore aggiunto delle imprese presenti nel FRAME-SBS per attività economica e classe di addetti.

Tabella 4. *FRAME – SBS: Distribuzione per settore di attività economica. Anno 2014*

Settori di attività economica	Imprese	Addetti	Fatturato	Valore Aggiunto
Industria in senso stretto	418.284	3.957.596	1.166.264	245.962
Costruzioni	529.103	1.356.571	170.612	46.551
Servizi	3.316.491	10.301.175	1.607.359	395.703
Totale	4.263.878	15.615.342	2.944.235	688.217

Fonte: Istat

Tabella 5. *FRAME – SBS: Distribuzione per classe di addetti. Anno 2014*

Classe di addetti	Imprese	Addetti	Fatturato	Valore Aggiunto
0-9	4.065.829	7.358.830	756.977.455	209.212.408
10-19	124.461	1.630.165	312.525.137	69.517.945
20-49	49.571	1.474.802	350.355.799	76.052.590
50-249	20.639	2.007.811	634.499.470	120.920.397
250e oltre	3.378	3.143.734	889.876.672	212.513.180
Totale	4.263.878	15.615.342	2.944.234.533	688.216.520

Fonte: Istat

Questi, ovviamente, sono solo due degli innumerevoli aspetti che possono essere analizzati con il nuovo archivio integrato. Le principali statistiche descrittive sono annualmente sintetizzate nella statistica report “Risultati economici delle impre-

se” (Istat, 2016) che dall’ottobre 2016 comprende anche analisi sui gruppi di impresa.

La disponibilità del FRAME-SBS ha introdotto cambiamenti molto rilevanti anche nelle metodologie di stima dei principali aggregati di Contabilità Nazionale.

Sino al 2011, il livello del valore aggiunto per settore di attività economica era ottenuto applicando un valore medio unitario per classe dimensionale d’impresa al totale addetti. Grazie alla disponibilità di dati a livello micro, è stato possibile adottare un approccio di livello di tipo additivo e che distingue la parte regolare dell’economia da quella non osservata per quasi tutti i settori di attività economica (tranne agricoltura, assicurazioni, attività finanziarie e locazione per i quali si adotta un approccio differente).

Al valore aggiunto regolare si aggiunge poi la componente di valore aggiunto sottodichiarata, quella generata dal lavoro irregolare e dalle attività illegali. Anche per la stima del valore aggiunto sottodichiarato il ruolo della nuova base dati micro è centrale, in quanto ha permesso una articolazione molto fine delle procedure di individuazione e correzione della sottodichiarazione, a livello settoriale, dimensionale, territoriale e di tipologia societaria.

Non da ultimo, particolarmente importante risulta l’apporto del FRAME-SBS ai conti economici territoriali. Si è passati da una stima di tipo top-down in cui si definiva l’aggregato economico a livello di intera economia e lo si ripartiva attraverso adeguati indicatori su base regionale, ad un approccio di tipo bottom-up dove la stima nasce a livello micro e per aggregazioni successive si giunge alla grandezza economica su base regionale.

4. Le fonti amministrative e i censimenti

Dal 2012, con l’art. 3 del d.lgs 179/2012, convertito con modificazioni in legge 221/2012, è stato introdotto nell’ordinamento italiano il Censimento permanente (Cfr. <http://www.istat.it/it/censimento-permanente>).

Le fonti amministrative sono entrate dunque “a gamba tesa” nella statistica ufficiale. Tale innovazione ha riguardato non solo le rilevazioni sulle imprese ma l’intera articolazione censuaria (popolazione, imprese, istituzioni pubbliche, istituzioni private non profit, imprese agricole).

I censimenti del 2011 hanno dunque chiuso l’era dei tradizionali censimenti a cadenza decennale. In particolare, il Censimento permanente prevede la realizzazione dei registri statistici basati sull’utilizzo congiunto di fonti amministrative e sullo

svolgimento di rilevazioni statistiche a supporto. L'obiettivo è di aumentare quantità e qualità dell'offerta informativa, valorizzandone la multidimensionalità e riducendo l'onere statistico a carico dei cittadini.

Come per i tradizionali censimenti anche quelli permanenti rileveranno informazioni sulle seguenti categorie di istituzioni:

- Popolazione e abitazioni;
- Imprese;
- Istituzioni pubbliche;
- Istituzioni non profit;
- Aziende agricole

Il Censimento permanente della popolazione e delle abitazioni, come i tradizionali censimenti a cadenza decennale, produrrà informazioni sulle caratteristiche strutturali e condizioni sociali e economiche della popolazione a livello nazionale, regionale e locale; permetterà l'aggiornamento e la revisione delle anagrafi comunali della popolazione residente; fornirà informazioni sulla consistenza numerica e sulle caratteristiche delle abitazioni.

Nel 2015 l'Istat ha svolto le prime due rilevazioni finalizzate alla definizione della metodologia statistica da utilizzare a regime per lo svolgimento del censimento permanente in Italia:

- Rilevazione sperimentale sul conteggio della popolazione (C-Sample);
- Rilevazione sperimentale per l'acquisizione di dati socio-economici sulla popolazione (D-Sample).

Obiettivo della rilevazione C-Sample è contare la popolazione residente e misurare l'errore delle Liste anagrafiche comunali (LAC). Come suggerito dal nome dell'indagine, essa è campionaria e coinvolge tutte le regioni italiane, per un totale di 134 comuni e circa 80.000 famiglie.

La rilevazione D-sample rappresenta una delle attività preparatorie necessarie a valutare e a mettere a punto l'organizzazione e le tecniche più idonee per la realizzazione del nuovo Censimento permanente della popolazione e delle abitazioni. Coinvolge 148 comuni e circa 86.000 famiglie.

Come il 9° Censimento generale dell'industria e dei servizi e Censimento delle istituzioni non profit del 2011, anche il Censimento permanente si articola su tre differenti rilevazioni: imprese, istituzioni non profit e istituzioni pubbliche.

Esso prevede lo svolgimento periodico di rilevazioni a supporto dei registri statistici delle unità economiche con l'obiettivo di integrare le informazioni non presenti nelle fonti, analizzare la qualità delle informazioni amministrative e correggere eventuali errori nella fonte.

La rilevazione di supporto al Censimento sulle imprese è la Rilevazione campionaria di controllo a copertura dell'archivio statistico sulle imprese attive (ASIA) e aggiornamento delle unità locali.

L'indagine, con cadenza annuale, rileva alcune informazioni di stratificazione (esistenza in vita, localizzazione, attività economica, natura giuridica, occupazione) e l'elenco aggiornato delle unità locali delle imprese di grandi dimensioni.

Anche per le istituzioni pubbliche è stato realizzato il Registro delle istituzioni pubbliche, attraverso l'integrazione delle più svariate fonti amministrative. Esso offre una fotografia strutturale del settore pubblico.

La rivelazione annuale di controllo del Registro è la Rilevazione di controllo della copertura del registro delle istituzioni pubbliche e aggiornamento delle unità locali il cui obiettivo è l'aggiornamento dell'elenco delle unità locali delle istituzioni pubbliche; l'acquisizione di nuove informazioni non presenti nelle fonti e la verifica di copertura di particolari istituzioni pubbliche appartenenti al campo di osservazione, per le quali si dispone di scarse informazioni da archivi amministrativi.

Lo stesso approccio è stato seguito per le istituzioni private non profit la cui rilevazione di supporto è la rilevazione campionaria sulle istituzioni non profit 2016. Per questo settore istituzionale le fonti amministrative sono spesso deboli sebbene sia stato costruito il registro statistico delle istituzioni non profit.

La rilevazione coinvolge un campione di circa 40 mila unità estratto dal registro statistico delle istituzioni non profit e i dati raccolti sono relativi a: forma giuridica e organo direttivo; numerosità e tipologia di soci; attività svolte; destinatari dei servizi erogati; risorse umane (volontari, addetti, collaboratori) e loro caratteristiche (categorie professionali e formazione); risorse economiche; reti di relazione; attività di comunicazione e di raccolta fondi.

Anche per le imprese agricole la rilevazione campionaria di supporto al registro statistico delle aziende agricole viene effettuata per verificare la copertura del registro statistico delle aziende agricole e stimarne alcune variabili di interesse. La numerosità è di circa 22 mila unità agricole e prevede l'intervista diretta del rilevatore all'unità agricola.

5. Aspetti conclusivi

Alla luce di quanto su esposto, si può ritenere che il processo di produzione di informazione statistica aziendale stia vivendo un momento di forte cambiamento sia dal punto di vista contenutistico, sia dal punto di vista di strette modalità di raccolta dei dati.

Lo scenario che si sta manifestando con la produzione delle statistiche strutturali sulle imprese, nonché dei censimenti permanenti si sta muovendo proprio nel verso della maggiore ed ulteriore soddisfazione di quei requisiti tipici dell'informazione statistica ufficiale di tempestività, accuratezza e dettaglio informativo sempre più sentiti come impellenti da parte del mondo economico.

Certamente occorre ancora lavorare parecchio per ridurre le distorsioni informative provenienti da fonti di tipo amministrativo spesso tra di loro non coordinate (si veda, a solo titolo di esempio la recente querelle relativa all'efficacia del job act sul mercato del lavoro italiano), ma l'integrazione delle stesse nell'ambito dei processi rappresentati nell'ambito del presente lavoro fanno ben sperare in una informazione sempre più rigorosa e sempre meno costosa sia in termini di costi che in termini di impegno, aspetto questo particolarmente sentito specialmente allorché l'ottemperanza all'obbligo informativo statistico è posta a capo del sistema imprenditoriale di per sé già notoriamente oberato.

Riferimenti bibliografici

- Biggeri, L.; Bini, M.; Coli, A.; Grassini, L.; Maltagliati, M. (2012). *Statistica per le decisioni aziendali*, Pearson ed, Milano
- Carucci, A.M.M.; Costanzo, A.; Vannella, G. (2015). *Studi di settore tra lotta all'evasione e statistica*, in: Studi in ricordo di Mario Cecchi, Università degli Studi di Bari Aldo Moro
- Boselli, C.; Brunetti, S.; Cammarota, M.; De Giorgi, V.; D'Urzo, A.; Ricci, M.; Pazzini, R.; Seri, G.; Siesto, G.; Virgili, L.(2016). *Il processo di diffusione dei dati sulle statistiche strutturali delle imprese (FRAME – SBS): aspetti normativi e metodologici connessi all'ampliamento del dettaglio informativo*, Istat working papers nr. 14, 2016
- Luzi, O.; Sanzo, R.; Righi, P.; Guarnera, U.; Oropallo, F.; Puggioni, A.(2014). *Le innovazioni metodologiche e le fonti per la costruzione del nuovo sistema*, in: Workshop "Nuove informazioni statistiche per misurare la struttura e la performance delle imprese italiane", Istat Roma.
<http://www.istat.it/it/files/2014/11/Luzi-Le-innovazioni-metodologiche.pdf>
- Istat (2016). *Risultati economici delle imprese*, Statistiche report, Roma.

Misurare la Salute: un'applicazione dell'indice di Mazziotta-Pareto

Angela Maria D'Uggento*

*Dipartimento di Scienze economiche e metodi matematici,
Università degli Studi di Bari Aldo Moro*

Riassunto: Da alcuni anni, il dibattito su come misurare il Benessere coinvolge i responsabili politici, i ricercatori e l'opinione pubblica in quanto i cambiamenti negli stili di vita e quelli in ambito economico, sociale, sanitario ed ambientale hanno evidenziato la necessità di sviluppare nuove misure statistiche in grado di supportare i decisori, le imprese e i cittadini nelle loro scelte. Un contributo in tal senso è venuto dalla comunità dei ricercatori partecipanti al progetto ISTAT BES-Benessere Equo e Solidale nel quale si declina il concetto di Benessere attraverso 12 dimensioni, una delle quali è, per l'appunto, la salute.

Dopo alcune osservazioni circa gli indicatori semplici che compongono la dimensione della Salute, viene ripreso l'indice sintetico proposto da Mazziotta e Pareto per la misura della salute su base regionale e discusso il ranking derivante dalla sua applicazione.

Keywords: Salute; indicatore sintetico; Mazziotta Pareto Index, Benessere equo e sostenibile; analisi statistiche multivariate.

1. Introduzione

La salute è una condizione indispensabile per il benessere individuale e, conseguentemente, della collettività; può, infatti, influenzare tutte fasi della vita di un individuo, modificandone le condizioni di vita, i comportamenti, le relazioni sociali, le opportunità e le prospettive, così influenzando, di riflesso, anche le famiglie e, in generale, coinvolgendo le reti sociali di cui l'individuo fa parte.

* Corresponding author: angelamaria.duggento@uniba.it.

Si ringrazia la dott.ssa Alessia Bellini per aver collaborato alla ricerca effettuando alcune elaborazioni statistiche nell'ambito della sua tesi di laurea triennale.

Secondo quanto emerge dai recenti dati diffusi dall'ISTAT, in Italia le condizioni di salute hanno registrato negli ultimi anni progressivi miglioramenti, anche se con ritmi contenuti a causa della crisi economica che ha inevitabilmente influenzato anche questa dimensione della vita umana.

I fattori che maggiormente hanno inciso sul miglioramento delle condizioni di salute degli Italiani dal 2009 al 2013 sono risultati l'incremento degli indici della speranza di vita e dello stato di salute fisica, a fronte di un peggioramento, però, di quello dello stato di salute psicologica.

A fronte di questa situazione a livello di Paese, permangono tuttavia ampie disuguaglianze territoriali tra Nord e Sud del Paese e, addirittura, sempre secondo i dati ISTAT, esse risultano in lieve aumento. In particolare, le forti disuguaglianze territoriali aumentano poiché la situazione del Mezzogiorno, anche per effetto della crisi economica, è più svantaggiata relativamente alla speranza di vita, alla qualità della vita media nelle età anziane e alla mortalità infantile.

Sono altrettanto marcate le differenze geografiche, a vantaggio del Centro-Nord, rispetto alla salute fisica e psicologica e ai fattori di rischio legati agli stili di vita. Nel Mezzogiorno le criticità riguardano una maggiore diffusione dei comportamenti poco salutari quali sedentarietà, eccesso di peso e scorrette abitudini alimentari, con l'unico elemento positivo che è il ridotto consumo di alcol a rischio, maggiormente diffuso al Nord.

Da un punto di vista metodologico, la salute è un concetto multidimensionale e numerosi sono i fattori che contribuiscono a definire lo stato di salute degli individui; l'Organizzazione Mondiale della Sanità (OMS) definisce la salute "stato di completo benessere fisico, psichico e sociale e non semplice assenza di malattia" e la colloca tra i più importanti diritti fondamentali dell'uomo, ritenendola cruciale nel contrasto alla povertà nonché presupposto necessario per lo sviluppo economico di un Paese.

Il contributo fornito dal progetto ISTAT-BES Benessere Equo e Sostenibile pone l'Italia all'avanguardia nel panorama internazionale nel dibattito sullo sviluppo di indicatori atti a misurare anche lo stato di salute degli abitanti di un Paese. L'elemento innovativo di tale filone di ricerca è un concetto di stato di salute fortemente improntato sugli aspetti della qualità della vita, seguendo l'approccio proposto negli studi sulla *Health related quality of life* promosso dall'OMS.

Inoltre, gli indici selezionati per comporre la dimensione Salute del BES mirano a valutare lo stato di salute in un'ottica di benessere che sia al contempo equa e sostenibile e, quindi, non sono concentrati soltanto sull'approccio tradizionale, basato

sulla presenza/assenza della malattia, sulla descrizione delle sue caratteristiche o sull'incidenza dei nuovi casi, ma lo integrano con valutazioni sulla qualità della "vita in assenza della malattia".

Dal 2016, investendo in questa direzione, l'Istat ha affiancato alle analisi sul Benessere equo e sostenibile (Bes) un primo set di indicatori finalizzati a monitorare i target fissati nell'Agenda 2030 delle Nazioni Unite sullo sviluppo sostenibile, corrispondenti ai *Sustainable Development Goals*¹ (SDGs). Inoltre, gli indicatori sul Benessere equo e sostenibile sono stati ampiamente consultati in fase di redazione dei documenti di programmazione, sia per tenere conto dei trend degli ultimi anni che per la valutazione dell'impatto delle politiche sulle 12 dimensioni del benessere della popolazione.

Ad attestare l'attenzione dei *policy makers* verso il benessere è stato costituito un Comitato di alto livello, a cui partecipa anche l'Istat, con il "compito di selezionare gli indicatori di benessere equo e sostenibile sui quali stimare gli effetti derivanti dai provvedimenti inseriti nella Legge di bilancio, con la presentazione annuale al Parlamento degli esiti sia del monitoraggio sull'andamento degli indicatori sia della valutazione di impatto delle policy".

Riconoscendo alla salute un ruolo prioritario nel determinare il benessere, la misurazione della condizione di salute di un individuo e, per aggregazione, di una popolazione, può avvenire secondo un approccio *oggettivo*, in cui il personale sanitario rileva il corteo sintomatologico ravvisando la presenza o assenza di uno stato morboso, oppure secondo un approccio *soggettivo*, basato sulle percezioni del proprio stato di salute da parte di un individuo.

Ben lungi dalla pretesa di affrontare una trattazione esaustiva di un tema così complesso, il presente lavoro intende illustrare una rivisitazione di indicatore sintetico per la misura della salute con un'applicazione al contesto italiano.

2. Materiali e metodi

La fonte dei dati utilizzati per l'analisi è il dataset del progetto ISTAT BES - Il Benessere equo e sostenibile in Italia, anno 2015. Il Rapporto BES offre un quadro

¹ I 17 obiettivi (*Sustainable Development Goals*) per lo sviluppo sostenibile fissati dalle Nazioni Unite nell'Agenda 2030 costituiscono il set di indicatori che la comunità politica nazionale e sovranazionale ha selezionato per rappresentare i propri valori, priorità e obiettivi. Molti di questi indicatori coincidono con quelli del Bes. "Gli indicatori SDGs possono essere, quindi, interpretati come l'ultimo atto di un dibattito iniziato molti anni fa, di cui il Bes è il capitolo italiano".

abbastanza completo sull'evoluzione dei principali fenomeni sociali, economici e ambientali avvenuti nel nostro Paese in tale anno². Il punto di partenza è la multidimensionalità del benessere, declinato in 12 dimensioni o "domini" e analizzato attraverso 130 indicatori individuati da una commissione di esperti di settore ed alimentati da fonti diverse, ma tutte ufficiali.

Le 12 dimensioni del benessere sono: salute, istruzione e formazione, lavoro e conciliazione dei tempi di vita, benessere economico, relazioni sociali, politica e istituzioni, sicurezza, benessere soggettivo, paesaggio e patrimonio culturale, ambiente, ricerca e innovazione, qualità dei servizi.

La dimensione salute è, a sua volta, declinata dall'ISTAT in 14 indici che, prendendo in parte spunto dalla classificazione adottata nel rapporto con riferimento a tutti gli indici, si potrebbero suddividere in:

- 1) indicatori globali di outcome, che forniscono informazioni complessive sul fenomeno;
- 2) indicatori specifici per ciascuna fase del ciclo di vita, che integrano l'informazione complessiva con quella relativa ai rischi che si possono manifestare specificatamente in ogni fase;
- 3) indicatori relativi a fattori di rischio o di protezione della salute, direttamente riconducibili agli stili di vita.

L'analisi è stata condotta seguendo le tradizionali fasi: costruzione della base informativa, così da sviluppare il quadro di riferimento concettuale in cui collocare le misurazioni statistiche rilevanti per la salute nel BES2015; costruzione di un indicatore sintetico, avendo selezionato il metodo tra le numerose tecniche per la combinazione di indicatori elementari, e contemporaneamente avendo effettuato il trattamento dei dati; individuazione di gruppi omogenei, anche attraverso la Cluster Analysis (con metodo *k-means*). Il software utilizzato per le analisi è Statistical Package for Social Sciences (SPSS).

La Tab. 1, che riporta i 14 indici selezionati dall'ISTAT con indicazione della fonte, della periodicità di rilevazione e dell'anno di riferimento del dato utilizzato, pone in luce un lieve disallineamento nell'aggiornamento dei dati che vanno dal 2012 al 2014. L'ISTAT, consapevole di tale criticità, sta già lavorando alla predisposizione di un dataset completo e tempestivamente aggiornato.

² Il rapporto BES è giunto alla sua quarta edizione, diffusa dall'ISTAT il 16 dicembre 2016, quindi non è stato possibile utilizzare i dati più aggiornati per effettuare l'applicazione.

Tabella 1. *Gli indicatori del Bes Salute 2015 per fonte, periodicità e anno di pubblicazione*

Indici	Fonte	Periodicità	Anno
Speranza di vita alla nascita	Istat-Indagine sui decessi e sulle cause di morte	Annuale	2013
Speranza di vita in buona salute alla nascita	Istat-Indagine sui decessi e sulle cause di morte	Annuale	2013
Indice di stato fisico (PCS)	Istat-"Condizioni di salute e ricorso ai servizi"	Quinquennale	2013
Indice di stato psicologico (MCS)	Istat-"Condizioni di salute e ricorso ai servizi"	Quinquennale	2013
Tasso di mortalità infantile	Istat-Indagine sui decessi e sulle cause di morte; Rilevazione degli iscritti in anagrafe per nascita	Annuale	2012
Tasso std di mortalità per accidenti di trasporto (15-34)	Istat-Indagine sui decessi e sulle cause di morte; Rilevazione sulla "Popolazione residente comunale per sesso, anno di nascita e stato civile"	Annuale	2012
Tasso std di mortalità per tumore (19-64)	Istat-Indagine sui decessi e sulle cause di morte; Rilevazione sulla "Popolazione residente comunale per sesso, anno di nascita e stato civile"	Annuale	2012
Tasso std di mortalità per demenze e malattie del sistema nervoso (65 anni e più)	Istat-Indagine sui decessi e sulle cause di morte; Rilevazione sulla "Popolazione residente comunale per sesso, anno di nascita e stato civile"	Annuale	2012
Speranza di vita senza limitazioni nelle attività quotidiane a 65 anni	Istat-Indagine sui decessi e sulle cause di morte	Annuale	2013
Eccesso di peso. Tasso std per età	Istat-Aspetti della vita quotidiana Istat-Condizioni di salute e ricorso ai servizi sanitari	Annuale	2014
Fumo. Tasso std per età	Istat-Aspetti della vita quotidiana Istat-Condizioni di salute e ricorso ai servizi sanitari	Annuale	2014
Alcol. Tasso std per età	Istat-Aspetti della vita quotidiana	Annuale	2014
Sedentarietà. Tasso std per età	Istat-Aspetti della vita quotidiana	Annuale	2014
Alimentazione. Tasso std per età	Istat-Aspetti della vita quotidiana	Annuale	2014

3. Risultati

Per le analisi statistiche sono stati utilizzati i dati regionali, preliminarmente accorpando le due province autonome di Trento e Bolzano e prendendo in considerazione l'ultimo dato disponibile della serie storica per i 14 indici che compongono la dimensione della salute.

Tabella 2. Il dataset del Bes Salute 2015

Regioni	Speranza di vita alla nascita		Speranza di vita in buona salute alla nascita		Indice di stato fisico (PCS)		Indice di stato psicologico (MCS)		Tasso di mortalità infantile (*)		Tasso std di mortalità per accidenti di trasporto (*)		Tasso std di mortalità per tumore (*)		Tasso std di mortalità per demenze e malattie sistemiche (*)		Speranza di vita senza limitazioni nelle attività quotidiane 65 anni		Eccesso di peso. Tasso std per età (*)		Fumo. Tasso std per età (*)		Alcol. Tasso std per età (*)		Sedentarietà. Tasso std per età (*)		Alimentazione. Tasso std per età	
	2013	2014	2013	2014	2013	2014	2013	2014	2012	2013	2012	2013	2012	2013	2012	2013	2012	2013	2012	2013	2012	2013	2012	2013	2012	2013	2012	2013
Piemonte	82,1	82,1	57,9	51,7	49,0	9976,8	9999,1	9990,9	9969,6	9,9	61,1	79,0	81,1	67,7	23,6													
Valle d'Aosta	82,1	81,9	59,1	51,4	49,4	9974,5	9998,4	9991,0	9961,0	10,4	61,4	81,2	76,9	16,1														
Liguria	81,9	82,6	60,8	51,7	50,1	9968,9	9999,4	9990,8	9968,3	10,2	59,4	80,7	84,6	19,9														
Lombardia	82,6	83,2	60,7	51,7	49,4	9975,9	9999,2	9991,0	9971,3	10,4	58,7	79,7	82,9	17,8														
Trentino Alto Adige	83,2	82,7	65,5	51,8	50,3	9966,9	9999,3	9991,6	9973,0	10,4	61,3	81,3	75,7	16,4														
Veneto	82,7	82,0	59,0	51,4	49,3	9973,7	9998,9	9991,4	9969,3	9,5	56,6	83,1	80,2	20,6														
Friuli-Venezia Giulia	82,0	82,8	60,4	51,6	49,6	9977,6	9999,0	9990,6	9975,1	10,6	57,4	78,9	78,8	19,2														
Emilia-Romagna	82,8	82,7	60,2	51,3	49,2	9973,8	9999,0	9991,6	9973,0	11,2	54,3	79,7	83,3	21,4														
Toscana	82,7	82,7	59,6	51,9	48,9	9978,5	9999,2	9991,4	9973,9	10,5	58,0	82,0	83,1	20,5														
Umbria	82,7	83,0	60,0	51,2	49,3	9972,4	9999,2	9991,3	9973,1	9,0	55,2	77,9	84,5	28,3														
Marche	83,0	81,8	59,0	51,4	48,0	9983,3	9999,2	9991,7	9971,1	9,7	54,9	79,3	85,2	21,2														
Lazio	81,8	82,2	57,5	51,2	49,4	9969,6	9998,9	9990,5	9975,8	8,8	58,8	78,4	87,1	20,6														
Abruzzo	82,2	82,1	58,8	51,3	49,0	9959,8	9999,4	9992,2	9969,1	8,6	51,6	78,8	82,9	17,3														
Molise	82,1	80,5	57,5	51,0	49,7	9974,3	9999,7	9992,0	9976,8	8,8	49,3	81,7	81,2	14,6														
Campania	80,5	82,3	55,7	50,9	48,3	9957,9	9999,7	9990,0	9979,1	6,9	47,4	77,7	88,1	16,5														
Puglia	82,3	82,0	55,0	50,6	48,7	9966,4	9999,0	9991,8	9973,3	7,6	50,0	81,1	86,4	8,5														
Basilicata	82,0	81,8	52,8	50,5	48,9	9975,4	9998,8	9992,2	9980,0	7,9	50,4	80,5	84,3	6,2														
Calabria	81,8	81,2	52,3	49,9	48,7	9954,8	9999,3	9991,8	9978,0	6,7	49,7	83,5	87,2	10,0														
Sicilia	81,2	82,0	56,2	50,8	48,9	9957,7	9999,3	9991,1	9973,1	7,7	51,3	79,8	89,3	13,3														
Sardegna	82,0	82,0	53,9	50,2	49,4	9977,5	9999,0	9989,9	9967,8	8,0	59,4	79,8	80,6	20,4														

Fonte: nostre elaborazioni su dati Istat, BES 2015. Gli indici con (*) sono stati trasformati con polarità positiva

Sono state effettuate le tradizionali attività di pre-processing sui dati, in particolare gli indici sono stati resi omogenei rispetto alla loro direzione, rendendoli tutti a polarità positiva, quindi è stata calcolata la matrice di correlazione fra tutti gli indici per esplorarne le relazioni; infine, per eliminare l'influenza delle differenti unità di misura e della variabilità, i dati sono stati standardizzati.

La Tabella 2 illustra il dataset con i dati per regione e i relativi indicatori elementare positivi, o resi tali con la trasformazione.

Osservando la matrice di correlazione (Tabella 3) si evince che gli indici mostrano differenti livelli di correlazione; le relazioni più forti si registrano tra "Speranza di vita senza limitazioni nelle attività a 65 anni" e "Indice di stato fisico" ($\rho=0,838$) e tra "Speranza di vita senza limitazioni nelle attività a 65 anni" e "Speranza di vita in buona salute" ($\rho=0,868$), ma anche tra la Sedentarietà e un certo numero di indici.

Per la sintesi degli indici semplici si è ritenuto di adottare la metodologia proposta nel recente Rapporto BES, l'Indice MPI di Mazziotta e Pareto (Mazziotta, Pareto, 2013), dotata di rigore statistico ma al contempo semplice nelle modalità di calcolo e con la caratteristica di soddisfare determinati requisiti, quali la comparabilità spaziale, consentendo di confrontare valori di sintesi tra unità territoriali; la non sostituibilità degli indici elementari, in quanto è possibile assegnare un peso identico a ciascun indice; l'immediata fruizione e interpretazione dei risultati di *output*; la robustezza dei risultati ottenuti.

Come anticipato, l'indice proposto da Mazziotta e Pareto è stato utilizzato per calcolare l'Indice composito, a livello regionale, per quasi tutte le dimensioni, compresa la salute, nelle ultime due edizioni del rapporto BES, anni 2015 e 2016.

Dal momento che la presente analisi e quella del BES si pongono il comune obiettivo di individuare una misura dello stato di salute di un collettivo, l'adozione della stessa metodologia consente di svolgere qualche riflessione in parallelo tra i risultati a cui esse giungono, pur adottando percorsi leggermente diversi.

La costruzione dell'MPI è basata sulla trasformazione degli indici elementari in variabili con media 100 e scostamento quadratico medio 10 e aggregando con la media aritmetica, "penalizzata" dalla variabilità «orizzontale» degli stessi indici, secondo la seguente formula:

$$MPI_i^{+/-} = M_{zi} \pm S_{zi}cv_i$$

Tabella 3. Matrice di correlazione tra i 14 indici della dimensione Salute

Indici	Speranza di vita alla nascita	Speranza di vita in buona salute alla nascita	Indice stato fisico (PCS)	Indice di stato psicologico (MCS)	Tasso di mortalità infantile (●)	Tasso di mortalità per incidenti di trasporto (●)	Tasso std di mortalità per demenze e malattie del sistema nervoso (●)	Speranza di vita senza limitazioni nelle attività quotidiane a 65 anni	Eccesso di peso. Tasso std per età (●)	Fumo. Tasso std per età (●)	Alcol. Tasso std per età (●)	Sedentarietà. Tasso std per età (●)	Alimentazione. Tasso std per età
Speranza di vita alla nascita	1,000	0,551*	0,509*	0,081	0,379	-0,135	0,415	0,564**	0,240	0,172	-0,453*	0,573**	0,399
Speranza di vita in buona salute alla nascita	0,551*	1,000	0,855**	0,574**	0,265	0,074	-0,170	0,868**	0,562**	-0,067	-0,458*	0,774**	0,446*
Indice stato fisico (PCS)	0,509*	0,855**	1,000	0,399	0,415	0,045	-0,180	0,838**	0,646**	0,028	-0,502*	0,779**	0,451*
Indice di stato psicologico (MCS)	0,081	0,574**	0,399	1,000	0,142	-0,039	-0,302	0,508*	0,593**	0,109	-0,633**	0,569**	0,115
Tasso di mortalità infantile (●)	0,379	0,265	0,415	0,142	1,000	-0,427	-0,163	0,568**	0,412	-0,032	-0,493*	0,558*	0,436
Tasso std di mortalità per incidenti di trasporto(●)	-0,135	0,074	0,045	-0,039	-0,427	1,000	0,127	-0,229	-0,365	-0,051	0,299	-0,318	-0,150
Tasso std di mortalità per tumore (●)	0,415	-0,170	-0,180	-0,302	-0,163	0,127	1,000	-0,177	-0,502*	0,405	0,046	-0,234	-0,386
Tasso std di mortalità per demenze e malattie del sistema nervoso (●)	-0,331	-0,401	-0,395	-0,317	-0,209	0,145	0,219	1,000	-0,655**	-0,037	0,473*	-0,454*	-0,378

Tabella 3. *(continuazione)*

Indici	Speranza di vita alla nascita	Speranza di vita in buona salute alla nascita	Indice stato fisico (PCS)	Indice stato psicologico (MCS)	Tasso di mortalità infantile (●)	Tasso di mortalità per incidenti di trasporto (●)	Tasso std di mortalità per demenze e malattie del sistema nervoso (●)	Speranza di vita senza limitazioni nelle attività quotidiane a 65 anni	Eccesso di peso. Tasso std per età (●)	Fumo. Tasso std per età (●)	Alcol. Tasso std per età (●)	Sedentarietà. Tasso std per età (●)	Alimentazione. Tasso std per età
Speranza di vita senza limitazioni nelle attività quotidiane a 65 anni	0,564**	0,868**	0,838**	0,508*	0,568**	-0,229	-0,372	1,000	0,629**	0,024	-0,601**	0,875**	0,505*
Eccesso di peso. Tasso std per età (●)	0,240	0,562**	0,646**	0,593**	0,412	-0,365	-0,655**	0,629**	1,000	0,022	-0,597**	0,714**	0,427
Fumo. Tasso std per età (●)	0,172	-0,067	0,028	0,109	-0,032	-0,051	-0,037	0,024	0,022	1,000	-0,240	0,121	-0,422
Alcol. Tasso std per età (●)	-0,453*	-0,458*	-0,502*	-0,633**	-0,493*	0,299	0,473*	-0,601**	-0,597**	-0,240	1,000	-0,786**	-0,152
Sedentarietà. Tasso std per età (●)	0,573**	0,774**	0,779**	0,569**	0,558*	-0,318	-0,454*	0,875**	0,714**	0,121	-0,786**	1,000	0,474*
Alimentazione. Tasso std per età	0,399	0,446*	0,451*	0,115	0,436	-0,150	-0,378	0,505*	0,427	-0,422	-0,152	0,474*	1,000

* la correlazione è significativa a livello di 0,05 (a due code); ** la correlazione è significativa a livello di 0,01 (a due code)

Fonte: nostre elaborazioni su dati Istat, BES 2015. Gli indici con (●) sono stati trasformati con polarità positiva.

La funzione di aggregazione *media aritmetica* dei valori standardizzati viene “corretta” mediante un coefficiente di penalità che dipende, per ciascuna unità, dalla variabilità degli indicatori rispetto al valor medio (definita appunto “variabilità orizzontale”).

Tale variabilità, misurata attraverso il coefficiente di variazione (*cv*), consente di “penalizzare” il punteggio delle unità che, a parità di media aritmetica, presentano un maggiore squilibrio tra i valori degli indicatori. La penalità può essere sommata o sottratta, a seconda del tipo di fenomeno studiato, ad esempio povertà o sviluppo. I valori così ottenuti risultano compresi nell’intervallo 70-130.

Nel rapporto BES2015, l’Indice composito della salute ottenuto applicando l’MPI è basato sull’aggregazione di soli 5 indici elementari, scelti tra i 14 disponibili in quanto considerati indicatori globali di outcome; tali indici sono:

1. Speranza di vita alla nascita;
2. Speranza di vita in buona salute alla nascita;
3. Indice di stato fisico (PCS);
4. Indice di stato psicologico (MCS);
9. Speranza di vita senza limitazione nelle attività quotidiane a 65 anni.

Nella presente analisi, invece, si ritiene opportuno utilizzare tutti i 14 indici del dominio Salute per il calcolo dell’indice sintetico, sia perché tale scelta è scevra dall’introduzione di una componente di soggettività insita nella scelta fatta nel Rapporto BES, ancorché frutto di una scelta ragionata, e sia perché attribuisce uguale peso a tutti i 14 indici elementari selezionati dall’ISTAT i quali, come dimostrato dalla tabella di correlazione, non sono risultati nemmeno troppo “ridondanti”.

Peraltro, in base allo scopo dell’analisi, si ritiene che gli stili di vita costituiscano una componente determinante, ancorché predittiva, dello stato di salute di un individuo e, quindi, che non si possa prescindere dal prendere in considerazione i relativi indici nella sua definizione.

Una volta calcolato l’MPI a 14 indici (MPI_{14}) per ciascuna regione, è possibile individuare le unità territoriali che presentano una condizione di salute al di sopra di quella media (fissata pari a 100) e quelle che si collocano al di sotto della media, con valori dell’indice valori minori di 100.

Dalla Tabella 4 appaiono subito evidenti notevoli differenze, con un *range* dell’indice che oscilla tra il valore più basso (89,78), registrato per la Campania e quello più alto, pari a 106,32, del Trentino Alto Adige.

Tabella 4. *Indice di Mazziotta-Pareto sui 14 indici della Salute (MPI_{14}) per le regioni italiane*

<i>Regione</i>	<i>MPI_{14}</i>	<i>Regione</i>	<i>MPI_{14}</i>
Piemonte	101,05	Marche	101,73
Valle d'Aosta	96,78	Lazio	98,79
Liguria	103,27	Abruzzo	97,75
Lombardia	103,59	Molise	101,17
Trentino Alto Adige	106,32	Campania	89,78
Veneto	102,35	Puglia	94,20
Friuli-Venezia Giulia	101,45	Basilicata	94,30
Emilia-Romagna	103,36	Calabria	92,18
Toscana	105,33	Sicilia	93,42
Umbria	101,76	Sardegna	94,74

Fonte: elaborazioni a cura della dott.ssa A. Bellini su dati ISTAT

Al fine di esplorare in maniera più approfondita i raggruppamenti di regioni “simili” dal punto di vista delle condizioni di salute è stata effettuata una *cluster analysis* tramite metodo *k-means*³ con l’obiettivo di minimizzare la varianza totale intra-cluster. Come si evince dalla Tabella 5, sono emersi tre *cluster*.

Tabella 5. *Attribuzione delle regioni ai cluster (metodo k-means)*

Cluster	Regione	Distanza	Cluster	Regione	Distanza
1	Piemonte	20,788	2	Umbria	21,168
1	Valle d'Aosta	35,632	2	Marche	31,240
1	Liguria	24,958	2	Lazio	25,945
1	Lombardia	11,516	2	Abruzzo	27,235
1	Trentino Alto Adige	36,363	2	Molise	34,498
1	Veneto	20,821	2	Sardegna	36,779
1	Friuli-Venezia Giulia	20,869	3	Campania	41,326
1	Emilia-Romagna	20,527	3	Puglia	19,589
1	Toscana	20,955	3	Basilicata	31,297
			3	Calabria	27,401
			3	Sicilia	18,900

Fonte: elaborazioni a cura della dott.ssa A. Bellini su dati ISTAT

³ K-means (Forgy, 1965; MacQueen, 1967) è un algoritmo *non supervisionato* atto a classificare un insieme di dati attraverso un determinato numero di cluster k . Esso funziona in modo ricorsivo, unendo in uno spazio m -dimensionale (individuato attraverso le m variabili) tutte le osservazioni prossime ai baricentri di k cluster K , ricalcolando poi passo per passo tali insiemi (detti centroidi) al fine di trovare la migliore allocazione delle unità in termini di distanza tra loro. È un algoritmo non gerarchico perché permette di unire un'unità a un cluster estraendolo da un cluster precedente, se la sua distanza dal nuovo centroide è minore..

Il primo cluster comprende le regioni “in buona salute”, risultano essere Trentino Alto Adige (106,32), Toscana (105,33), Lombardia (103,59), Emilia Romagna (103,36), Liguria (103,27), Veneto (102,35). Le regioni con uno stato di salute nella media (Cluster 2) sono risultate Umbria (101,76), Marche (101,73), Friuli Venezia Giulia (101,45), Molise (101,17), Piemonte (101,05), Lazio (98,79), Abruzzo (97,75), Valle d’Aosta (96,78), Sardegna (94,74). Infine, nell’ultimo gruppo si collocano le regioni in cui si registra uno stato di salute al di sotto della media: Basilicata (94,30), Puglia (94,20), Sicilia (93,42), Calabria (92,18) e Campania (89,78).

4. Conclusioni

La salute è uno dei diritti fondamentali dell’individuo e la sua tutela uno degli obiettivi più importanti per i governanti di un Paese. Secondo quanto emerge dai dati ISTAT, in Italia le condizioni di salute hanno registrato nell’ultimo quinquennio progressivi miglioramenti, anche se con ritmi contenuti a causa della crisi economica. A segnare il passo sono stati gli incrementi negli indici della speranza di vita e dello stato di salute fisica, mentre è peggiorato lo stato di salute psicologica. Si confermano, tuttavia, notevoli disuguaglianze territoriali tra Nord e Sud del Paese, in particolare dovute ad una peggiore situazione relativamente alla speranza di vita, alla qualità della vita media nelle età anziane, alla mortalità infantile, alla salute fisica e psicologica e ai fattori di rischio legati all’adozione di non corretti stili di vita. Purtroppo, nel Mezzogiorno sono risultati molto più diffusi comportamenti poco salutari quali sedentarietà, eccesso di peso e scorrette abitudini alimentari a cui fa da contrappeso solo il minor consumo di alcol a rischio rispetto al Nord.

Alla luce di queste chiavi di lettura si possono meglio interpretare i risultati dell’analisi statistica condotta allo scopo di individuare, attraverso un indice sintetico, le differenti condizioni di salute registrate nelle regioni italiane. L’indice sintetico utilizzato, denominato MPI_{14} , è una rivisitazione di quello proposto da Mazziotta e Pareto che viene utilizzato per definire l’Indice composito nelle varie dimensioni del Benessere nei rapporti BES 2015 e 2016.

A differenza dell’ MPI del BES, basato su 5 indici elementari, l’ MPI_{14} è ottenuto per aggregazione di tutti gli indici che definiscono il dominio della Salute con l’assunzione di uguale peso per ciascuno. Applicando l’ MPI_{14} le regioni si sono raggruppate in modo da determinare tre cluster caratterizzati rispettivamente da buone-medie-scarse condizioni di salute.

Una più dettagliata analisi a livello regionale evidenzia delle dinamiche abbastanza differenziate non solo a livello di macro ripartizioni geografiche ma anche tra le regioni appartenenti alla stessa ripartizione. Si consideri, ad esempio, il differente livello delle condizioni di salute registrato nelle due Isole, rispettivamente medio per la Sardegna e basso per la Sicilia.

Si ritiene interessante approfondire ulteriormente l'analisi "scomponendo" le determinanti delle condizioni di salute in ciascuna regione ma ciò costituirà oggetto di lavori futuri.

Come auspicato dal Comitato per gli indicatori di benessere e dall'ISTAT, la comunità scientifica è invitata a contribuire alla definizione di metodologie per l'individuazione di strumenti utili a supportare i decisori politici nel valutare l'efficacia delle politiche sia a livello nazionale sia a livello locale. In particolare, come previsto dalla recente riforma della Legge di Bilancio, tali strumenti saranno impiegati per l'individuazione delle priorità e delle principali criticità del Paese e, "in prospettiva, per valutare ex-ante gli effetti degli interventi di politica economica sul benessere della popolazione. In quest'ambito, l'esperienza del Bes costituirà ovviamente un punto di riferimento essenziale".

Riferimenti bibliografici

- Carbonara M., D'Uggento A.M., Toma E. (2015) Subjective component is essential for the well-being definition. *Book of Abstracts of the International Conference IES 2015 "Statistical methods for service assessment"*, University of Bari Aldo Moro, pag. 59
- Delvecchio, F. (2015). *Statistica per l'analisi dei fenomeni sociali*. Cleup, Padova.
- Delvecchio, F. (1995). *Scale di misura e indicatori sociali*. Cacucci Editore, Bari.
- Fabbris, L. (1997). *Statistica multivariata - Analisi esplorativa dei dati*, McGraw Hill, Milano.
- Forgy E.W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics* 21, pag. 768-769.
- ISTAT (2016). *Rapporto BES 2015*. ISBN 978-88-458-1876-9 (print); ISBN 978-88-458-1875-2 (online). <http://www.istat.it/it/benessere-e-sostenibilità/misure-del-benessere/il-rapporto-istat-sul-bes>
- MacQueen J. B. (1967). Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1, pag. 281-297.

-
- Massoli, P., Mazziotta, M., Pareto, A., Rinaldelli, C. (2014). *Indici compositi per il BES*. Giornate della ricerca, 10-11 novembre 2014. Istat. Roma
- Pareto, A., Mazziotta M. (2013). *A non-compensatory composite index for measuring well-being over time*, Editura Pro Universitaria, Bucarest (RO)
- World Health Organization (2016). *Investing in Health: a summary of the findings of the Commission on Macroeconomics and Health*. <http://www.who.int/macrohealth>

L'aggiornamento di strutture di Contabilità Nazionale disaggregate

Vittorio Nicolardi*, Caterina Marini

*Dipartimento di Scienze Economiche e Metodi Matematici
Università degli Studi di Bari "Aldo Moro"*

Riassunto: La necessità da parte dei policy maker di disporre di basi di dati sempre più ampie e disaggregate, nonché la necessità di disporre di tali dati in tempi ragionevolmente brevi ha prodotto come effetto, da parte degli Enti di produzione e diffusione di dati statistici, in particolare gli Istituti Nazionali di Statistica, l'incremento delle rilevazioni su base campionaria e l'implementazione di nuovi metodi statistici per soddisfare tali richieste. Nel campo delle metodologie statistiche, in particolare, si adottano tecniche econometriche molto sofisticate tese ad aggiornare le basi di dati disaggregate anche quando i risultati delle rilevazioni campionarie o i dati contabili di alcuni Settori istituzionali non siano disponibili. Nel presente lavoro, si propone una tecnica econometrica finalizzata ad aggiornare strutture contabili disaggregate anche in presenza di informazioni incomplete. Tale tecnica, basandosi su una variante del metodo di bilanciamento proposto da Nicolardi (1998) per strutture contabili di grandi dimensioni, attualmente ancora in uso presso l'ISTAT, viene applicata al Conto dei beni e servizi della NAM 2012 con dati 2013. I risultati ottenuti sono risultati molto buoni e incoraggianti per un'estensione della metodologia applicata all'intera struttura della NAM e ad altre strutture della Contabilità Nazionale, sia matriciali, sia non matriciali.

Keywords: National Accounting; National Accounts Matrix; Econometric modeling; National economic and social accounts updating and balancing.

1. Introduzione

La sempre maggiore richiesta di dati statistici necessari per una corretta programmazione degli interventi di politica economica e sociale che proviene dai policy maker sia pubblici, sia privati, ha reso necessario l'ampliamento delle basi di dati

* Corresponding Author: vittorio.nicolardi@uniba.it.

da rendere disponibili al pubblico da parte degli Istituti Nazionali di Statistica. Già a partire dalla pubblicazione del SNA 1993 / SEC 1995 e maggiormente con il recente SNA 2008 / SEC 2010 (EUROSTAT, 1995; EU – EUROSTAT, 2013; CEC - EUROSTAT, IMF, OECD, UN, WB 1993; EC, IMF, OECD, UN, WB 2009), si è focalizzata l'attenzione degli enti produttori e fornitori di dati statistici, in particolare statistico-economici, verso la compilazione di strutture di dati sempre più complesse e con un maggiore livello di disaggregazione, nonché verso la pubblicazione di dati a tempi sempre più ravvicinati, onde venire incontro anche alla necessità di disporre di dati sempre aggiornati necessari ad una verifica in tempo reale dei modelli di interpretazione e di previsione degli scenari economici e ad una loro eventuale ritaratura.

In particolare, il problema dell'aggiornamento in tempi ravvicinati delle strutture di dati complesse rende necessaria l'adozione di opportune tecniche econometriche che permettano di ovviare alla mancata disponibilità di tutti i dati statistico-economici utili. Ciò è dovuto principalmente al fatto che molti dei dati disaggregati che sarebbero necessari per tale aggiornamento provengono da ricerche campionarie alquanto complesse, che per loro natura non possono fornire dati in tempi ristretti. Ovviamente, e come è d'altronde prassi per tutti gli Istituti Nazionali di Statistica, man mano che i risultati rivenienti dalle rilevazioni campionarie si rendono disponibili, le diverse strutture di dati vengono aggiornate e rese disponibili agli utilizzatori finali.

In questo paper si vuole affrontare il problema dell'aggiornamento delle strutture di dati disaggregate utilizzando una variante del metodo di bilanciamento di strutture contabili comunque definite (Nicolardi 1998, 2000, 2010, 2011, 2012, 2013; Marini, Nicolardi 2010) ed attualmente in larga parte ancora utilizzato dall'ISTAT (ISTAT, 2011).

In particolare, in questo contesto, la metodologia di bilanciamento viene applicata ad una National Accounts Matrix (NAM), la quale costituisce una struttura contabile matriciale disaggregata particolarmente adatta al bilanciamento dei flussi di contabilità nazionale e alla stima di valori contabili mancanti o alla correzione di valori contabili con insufficiente livello di significatività.

Il presente lavoro è così strutturato: nella Sezione 2 viene descritta la struttura della National Accounts Matrix; nella Sezione 3 si illustra il metodo di bilanciamento che viene utilizzato come base per l'aggiornamento e il bilanciamento della NAM oggetto di studio; nella Sezione 4 si descrive il processo di aggiornamento e bilanciamento della NAM 2013; nella Sezione 5 vengono presentate alcune considerazioni conclusive.

2. La National Accounts Matrix (NAM)

Una delle rappresentazioni matriciali più semplici previste dai sistemi di contabilità nazionale è costituita dalla Matrice dei Conti Nazionali (National Accounts Matrix – NAM, d'ora innanzi), la quale costituisce, sostanzialmente, una ricompilazione della sequenza dei conti nazionali in forma classica (T-accounts) secondo una struttura contabile matriciale. La NAM può essere compilata a diversi livelli di disaggregazione a seconda dell'utilizzo cui è destinata. Essa offre una visione sintetica delle variabili macroeconomiche di un paese e mostra le interrelazioni esistenti tra le varie fasi del processo economico, evidenziando l'origine delle entrate e la destinazione delle uscite di ogni singolo conto.

La NAM si presenta strutturalmente come una matrice quadrata in cui ogni coppia riga-colonna descrive la struttura di ogni singolo conto della sequenza dei conti del SNA, in base alla convenzione secondo cui in riga vengono rappresentate le voci di entrata (risorse o acquisizioni di attività) e in colonna le voci di uscita (usi o acquisizioni di passività) e le eventuali voci di saldo del conto stesso.

La principale caratteristica della NAM, così come, in generale, di tutti i conti in forma matriciale, è che le voci che vengono di norma riportate due volte nei conti tradizionali sono, invece, riportate una sola volta nella matrice, più precisamente nella cella di intersezione della riga del conto nel quale la voce economica è registrata tra le entrate e della colonna del conto in cui è, invece, registrata tra le uscite.

In particolare, una NAM aggregata è di norma limitata alla sola sequenza dei Conti non finanziari per il totale dell'economia, mentre nelle versioni maggiormente disaggregate la NAM viene compilata adottando per ogni conto la classificazione maggiormente appropriata, in modo tale da evidenziare non solo le connessioni tra le varie fasi del circuito economico, ma anche le interdipendenze tra gli operatori coinvolti nelle diverse operazioni economiche. Oltre ciò, la struttura della NAM può essere ulteriormente ampliata estendendola anche ai Conti finanziari.

In una NAM disaggregata, inoltre, alcuni flussi possono essere riportati in forma matriciale evidenziando i settori di origine e quelli di destinazione dei flussi stessi, determinando, così, un significativo aumento della capacità di analisi della NAM stessa. In tale direzione, ad esempio, una NAM opportunamente disaggregata può essere utilizzata per fornire una più ampia informazione sui conti nazionali per Settore istituzionale, presentandoli con un maggior livello di dettaglio rispetto a quello presente nelle pubblicazioni standard.

A volte, però, non sono disponibili dati disaggregati relativi ad alcune tipologie di transazione, o, addirittura, possono essere disponibili solo i totali delle voci di

entrata o di uscita di un conto senza alcun dettaglio relativo ai flussi posti in essere dagli operatori economici coinvolti nelle diverse operazioni economiche. In questi casi si utilizza un cosiddetto *conto dummy*, nel quale i dati disponibili sono disposti in modo da salvaguardare la sequenza dei saldi contabili a livello aggregato, anche se l'informazione a livello disaggregato non risulterà egualmente completa.

Nella sua struttura tipica più aggregata la NAM presenta, per ogni coppia sequenziale riga-colonna, rispettivamente, il Conto dei beni e servizi (detto anche Conto di equilibrio delle risorse e degli impieghi o Conto di equilibrio dei beni e servizi), il Conto della produzione, il Conto della generazione del reddito, compreso il Conto delle Imposte indirette nette sui prodotti e l'IVA, il Conto della distribuzione del reddito, eventualmente distinto in Conto dell'allocazione dei redditi primari e Conto della distribuzione secondaria del reddito, il Conto dell'utilizzo del reddito disponibile, il Conto capitale, eventualmente scomposto nelle componenti della Formazione lorda del capitale, della Variazione scorte e degli oggetti di valore e del Saldo del conto economico, e il Conto del Resto del Mondo, eventualmente distinto tra Parte corrente e Conto capitale.

3 La metodologia di bilanciamento

Diverse sono le tecniche che sono state proposte in ambito scientifico per attuare il bilanciamento di sistemi contabili di medie e grandi dimensioni, ma un ruolo di rilievo viene ancora assunto non solo in letteratura, ma anche nelle pratiche applicazioni dal metodo di bilanciamento di Byron R.P. (1978).

In particolare, Byron recuperando il metodo di bilanciamento proposto nel 1960 da R. Stone, che grazie alle sue caratteristiche di flessibilità e linearità ben si adattava a risolvere problemi di gestione di mole di dati anche molto rilevanti per schemi contabili complessi, ne propone una variante molto più snella e più elastica, per quanto ugualmente solida, per il bilanciamento di sistemi econometrici di grandi dimensioni.

Sia l'approccio di Stone, sia l'approccio alternativo di Byron affrontano il problema del bilanciamento dei sistemi contabili attraverso il calcolo di stime consistenti per gli aggregati interessati, come se si trattasse di un problema di stime vincolate, attraverso una combinazione lineare ponderata delle stime iniziali degli aggregati stessi. Se da un lato Stone suggeriva di stimare il vettore di s dati contabili che soddisfacesse i vincoli contabili come previsti dalla teoria economica attraverso la definizione di uno stimatore vincolato generalizzato che tenesse conto anche

dei livelli di significatività delle stime stesse, Byron affronta il problema attraverso una funzione di perdita quadratica vincolata nella sua classica formulazione di funzione lagrangiana ed il metodo del gradiente coniugato usato per stimare il sistema.

Pertanto, indicando con \mathbf{x} il vettore ($s \times 1$) dei dati contabili, con \mathbf{G} la matrice ($k \times s$) di vincoli contabili, con \mathbf{h} un vettore ($k \times 1$) di valori noti, con $\hat{\mathbf{x}}$ un vettore di partenza di stime non bilanciate di \mathbf{x} , con \mathbf{V} una prima stima della matrice di covarianze ($s \times s$) di $\hat{\mathbf{x}}$ (che contiene anche l'informazione circa il livello di significatività delle stime $\hat{\mathbf{x}}$), con $\tilde{\mathbf{x}}$ il vettore stimato che soddisfa i vincoli contabili e con il vettore dei moltiplicatori di Lagrange, la funzione di perdita quadratica vincolata definita da Byron per affrontare il problema del bilanciamento è definita nel seguente modo:

$$Z = \frac{1}{2}(\tilde{\mathbf{x}} - \hat{\mathbf{x}})' \mathbf{V}^{-1}(\tilde{\mathbf{x}} - \hat{\mathbf{x}}) - (\mathbf{G}\tilde{\mathbf{x}} - \mathbf{h}) \quad (1)$$

nella quale la stima $\tilde{\mathbf{x}}$ deve risultare quanto più vicina in termini di perdita quadratica alla stima $\hat{\mathbf{x}}$ e contemporaneamente soddisfare anche i vincoli contabili $(\mathbf{G}\tilde{\mathbf{x}} - \mathbf{h})$. È importante precisare che la matrice \mathbf{V} è normalmente usata in forma diagonale assumendo nulle le covarianze delle stime $\hat{\mathbf{x}}$ in quanto l'inefficienza delle stime, che risulta essere una diretta conseguenza di tale semplificazione, è trascurabile nella maggior parte dei casi anche perché le metodologie applicate per la compilazione dei sistemi di contabilità nazionale solitamente non consentono la stima delle correlazioni tra stime di flussi diversi.

Le due condizioni del primo ordine da verificare nella (1) al fine di procedere con la stima del sistema riguardano i due stimatori direttamente coinvolti nella funzione:

$$(a) \quad \tilde{\lambda} = (\mathbf{G}\mathbf{V}\mathbf{G}')^{-1}(\mathbf{G}\hat{\mathbf{x}} - \mathbf{h}); \quad (b) \quad \tilde{\mathbf{x}} = \hat{\mathbf{x}} - \mathbf{V}\mathbf{G}'\tilde{\lambda}; \quad (2)$$

Per quanto il sistema (2) sia equivalente allo stimatore generalizzato vincolato definito da Stone¹, la formulazione proposta da Byron consente di considerare condizioni meno restrittive per risolvere il problema della stima e, allo stesso tempo, mostra una semplificazione del processo di stima stesso attraverso una riduzione del sistema di equazioni lineari coinvolto. La funzione di perdita quadratica vincolata (1) e le due condizioni del primo ordine (2) definiscono il punto di partenza per la definizione dell'algoritmo per l'applicazione del metodo del gradiente coniugato che Byron utilizza per stimare i sistemi contabili di grandi dimensioni. Difatti, par-

¹ Lo stimatore vincolato generalizzato proposto da Stone che consente di ottenere una stima di \mathbf{x} che soddisfi i vincoli contabili presenti in \mathbf{G} è definito da: $\tilde{\mathbf{x}} = \hat{\mathbf{x}} - \mathbf{V}\mathbf{G}'(\mathbf{G}\mathbf{V}\mathbf{G}')^{-1}(\mathbf{G}\hat{\mathbf{x}} - \mathbf{h})$.

tendo dalla linearità del sistema di equazioni (2a) e considerato che la matrice $(\mathbf{G}\mathbf{V}\mathbf{G}')$ è simmetrica definita positiva² quando il sistema contabile è esattamente definito, la stessa relazione (2a) può essere riscritta nel seguente modo:

$$(\mathbf{G}\mathbf{V}\mathbf{G}') \mathbf{x} = \mathbf{A} \mathbf{y} = \mathbf{q} \quad (3)$$

la cui soluzione è possibile attraverso l'algoritmo del gradiente coniugato di seguito definito:

$$\begin{aligned} \mathbf{x}_0 &= \mathbf{x}_0 = -\mathbf{A}^{-1} \mathbf{q} \\ \mathbf{r}_i &= \mathbf{y}'_i / \mathbf{A}'_i \mathbf{A}_i \\ \mathbf{x}_{i+1} &= \mathbf{x}_i + \mathbf{r}_i \mathbf{A}_i \\ \mathbf{x}_{i+1} &= \mathbf{x}_i - \mathbf{r}_i \mathbf{A}_i \\ \mathbf{S}_i &= \mathbf{y}'_{i+1} / \mathbf{A}'_{i+1} \mathbf{A}_{i+1} \\ \mathbf{x}_{i+1} &= \mathbf{x}_{i+1} + \mathbf{S}_i \mathbf{A}_{i+1} \end{aligned} \quad (4)$$

dove \mathbf{x}_0 è il vettore dei valori iniziali di \mathbf{x} , e \mathbf{A}_i e \mathbf{A}_{i+1} sono i vettori ($k \times 1$) direzionali definiti in base al gradiente, ed i pedici i e $i+1$ si riferiscono al passo dell'iterazione. Fissato un ν sufficientemente piccolo, il processo iterativo verrà ripetuto sino a quando la condizione $(\mathbf{y}'_i - \mathbf{A}'_i \mathbf{x}_i) < \nu$ non è verificata. Successivamente, la stima $\tilde{\mathbf{x}}$ si ottiene sostituendo $\tilde{\mathbf{x}}$ con \mathbf{x}_i nel sistema di equazioni (2b). Tutto il processo può essere velocizzato nel raggiungimento di una soluzione di convergenza operando un'opportuna normalizzazione della matrice $(\mathbf{G}\mathbf{V}\mathbf{G}')$ (Byron, 1978).

Per quanto robusto e consistente come metodo, anche Byron non risolse uno dei problemi più importanti che già si palesava in modo evidente nel metodo di Stone e che riguardava il calcolo e la gestione in termini statistici della matrice $(\mathbf{G}\mathbf{V}\mathbf{G}')$ dovuto alla gran mole di dati che la caratterizza.

Uno dei metodi ancora in uso presso l'ISTAT per la sua validità di calcolo e che risolve il problema generato dalla dimensione della matrice $(\mathbf{G}\mathbf{V}\mathbf{G}')$ è il metodo proposto da Nicolardi (1998, 2000). Tale metodo si basa essenzialmente sulla scomposizione in blocchi della matrice contabile e tale operazione consente non solo di ridurre drasticamente la quantità di dati da utilizzare per il bilanciamento del sistema stesso, ma anche di calcolare il valore della matrice $(\mathbf{G}\mathbf{V}\mathbf{G}')$ senza dover definire la matrice \mathbf{G} elemento per elemento.

² Qualora la matrice $(\mathbf{G}\mathbf{V}\mathbf{G}')$ fosse non definita positiva, un'opportuna riformulazione del sistema di equazioni può risolvere il problema. Il problema può presentarsi qualora la matrice \mathbf{G} non sia a rango colonna pieno oppure quando le varianze e covarianze degli elementi di una generica equazione del sistema sia tutte pari a zero.

Il punto di partenza del metodo è proprio l'algoritmo del gradiente coniugato. Difatti, si nota nella (4) che la matrice $(\mathbf{G}\mathbf{V}\mathbf{G}')$ è implicitamente presente nella seconda e nella quarta relazione dell'algoritmo attraverso il prodotto \mathbf{A} nel seguente modo:

$$\mathbf{A} = \mathbf{W}(\mathbf{G}\mathbf{V}\mathbf{G}') \tag{5}$$

dove \mathbf{W} è proprio la matrice di normalizzazione $(k \times k)$ proposta da Byron. Si può dimostrare che calcolare il prodotto nella (5) partendo dall'ultima coppia di elementi e procedendo verso la prima produce risultati successivi che sono sempre costituiti da vettori. In termini matematici, questa deduzione può essere espressa nel seguente modo:

$$\mathbf{G}'_{(s \times k)} \mathbf{f}_{(k)} = \mathbf{v}_{(s)}; \mathbf{V}_{(s \times s)} \mathbf{v}_{(s)} = \mathbf{g}'_{(s)}; \mathbf{G}_{(k \times s)} \mathbf{g}'_{(s)} = \mathbf{w}_{(k)}; \mathbf{W}_{(k \times k)} \mathbf{w}_{(k)} = \mathbf{a}_{(k)} \tag{6}$$

Considerato che \mathbf{V} e \mathbf{W} sono matrici diagonali e possono essere memorizzate in forma vettoriale, la riduzione nella quantità di dati da memorizzare per calcolare \mathbf{A} nella (5) è molto evidente, per quanto resta ancora irrisolto nella (6) il problema della grande dimensione che la matrice \mathbf{G} può raggiungere per il calcolo del prodotto \mathbf{G}' . Al fine di risolvere questo problema, è interessante notare come spesso le matrici contabili abbiano una configurazione a blocchi così da poter riformulare il prodotto \mathbf{G}' mediante una combinazione lineare di una opportuna scomposizione di \mathbf{G} in blocchi senza dover necessariamente determinare \mathbf{G} (Nicolardi 1998, 2000). Un semplice schema di bilanciamento viene definito di seguito sia in termini scalari nella (7a), sia in termini matriciali nella (7b) dove ognuna delle tre generiche matrici di dati contabili \mathbf{A} , \mathbf{B} e \mathbf{C} ha dimensione $(n \times m)$, ed i vettori \mathbf{i}_n e \mathbf{i}_m sono vettori composti da soli elementi unitari rispettivamente di dimensione n e m :

$$(a) \begin{cases} a_{ij} = b_{ij} + c_{ij} \quad \forall i, j \\ \sum_j a_{ij} = \sum_j b_{ij} + \sum_j c_{ij} \quad \forall i \\ \sum_i a_{ij} = \sum_i b_{ij} + \sum_i c_{ij} \quad \forall j \\ \sum_{i,j} a_{ij} = \sum_{i,j} b_{ij} + \sum_{i,j} c_{ij} \end{cases}; \quad (b) \begin{cases} \mathbf{A} - \mathbf{B} - \mathbf{C} = \mathbf{0} \\ \mathbf{A}\mathbf{i}'_m - \mathbf{B}\mathbf{i}'_m - \mathbf{C}\mathbf{i}'_m = \mathbf{0} \\ \mathbf{i}_n\mathbf{A} - \mathbf{i}_n\mathbf{B} - \mathbf{i}_n\mathbf{C} = \mathbf{0} \\ \mathbf{i}_n\mathbf{A}\mathbf{i}'_m - \mathbf{i}_n\mathbf{B}\mathbf{i}'_m - \mathbf{i}_n\mathbf{C}\mathbf{i}'_m = \mathbf{0} \end{cases} \tag{7}$$

È importante precisare che seppur implicite nella prima equazione di (7a) e (7b), le equazioni 2-4 sono normalmente usate esplicitamente per evidenziare i vincoli sulle distribuzioni marginali delle matrici contabili. Inoltre, considerato che nella (7b) solo q matrici, vettori e scalari sono formati da dati iniziali, mentre i restanti non sono altro che una loro trasformazione, come si evince dal prodotto $\mathbf{A}\mathbf{i}'_m$,

si desume come sia possibile calcolare i vettori \mathbf{g} e \mathbf{h} nella (6) senza usare esplicitamente la matrice \mathbf{G} per mezzo delle formulazioni (7b).

In particolare, al fine di calcolare il prodotto \mathbf{G}' nella (6), è necessario scomporre \mathbf{G} in k blocchi \mathfrak{S}_i (matrici, vettori e scalari) in numero pari a quello delle macro-equazioni della (7b), tale che ognuno dei blocchi abbia una dimensione pari alle matrici dei dati, o corrispondenti trasformazioni, presenti in ognuna delle macro-equazioni stesse. Nell'esempio riportato nella (7b), \mathbf{G} è scomposto in $k = 4$ blocchi con dimensione rispettivamente pari a $(n \times m)$, $(n \times 1)$, $(1 \times m)$ e (1×1) . Conseguentemente, la soluzione per \mathbf{G}' è data dai q blocchi \mathfrak{R}_i ottenuti come combinazione lineare dei blocchi \mathfrak{S}_j , corrispondenti, questi ultimi, alle macro-equazioni nelle quali le q matrici di partenza dello schema (7b), opportunamente ordinate in relazione alla posizione sequenziale nello schema di bilanciamento e senza ripetizione, sono presenti. Nel caso analizzato, $q = 3$ poiché le sole matrici di dati iniziali sono \mathbf{A} , \mathbf{B} e \mathbf{C} , e il vettore ordinato di matrici iniziali è $\mathbf{g} = (\mathbf{A}, \mathbf{B}, \mathbf{C})$. Definita una matrice $(q \times k)$ nella quale ogni i -esima riga contiene il numero di identificazione delle macro-equazioni dove la i -esima matrice in \mathbf{G} è presente, l' i -esimo blocco \mathfrak{R}_i può essere ottenuto come segue:

$$\mathfrak{R}_i = \left(\sum_{j \in \Lambda_i} \mathfrak{S}_j \right)^{el} \quad i = 1, \dots, q; \quad = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}; \quad \begin{cases} \mathfrak{R}_1 = +\mathfrak{S}_1 + \mathfrak{S}_2 \mathbf{i}'_m + \mathbf{i}_n \mathfrak{S}_3 + \mathfrak{S}_4 \mathbf{i}_n \mathbf{i}'_m \\ \mathfrak{R}_2 = -\mathfrak{S}_1 + \mathfrak{S}_2 \mathbf{i}'_m - \mathbf{i}_n \mathfrak{S}_3 - \mathfrak{S}_4 \mathbf{i}_n \mathbf{i}'_m \\ \mathfrak{R}_3 = -\mathfrak{S}_1 + \mathfrak{S}_2 \mathbf{i}'_m - \mathbf{i}_n \mathfrak{S}_3 - \mathfrak{S}_4 \mathbf{i}_n \mathbf{i}'_m \end{cases} \quad (8)$$

dove i indica la i -esima riga della matrice \mathbf{G} , l'apice *el* indica che tutte le somme devono essere calcolate elemento per elemento³. Inoltre, ogni blocco \mathfrak{S}_j deve avere lo stesso segno algebrico dell' i -esima matrice in \mathbf{G} per ognuna delle k equazioni.

L'operazione di vettorizzazione sequenziale dei q blocchi permette di calcolare

Per calcolare, infine, \mathbf{g} nella (6), è necessario che lo stesso \mathbf{G} venga scomposto in q blocchi \mathfrak{S}_i con dimensioni corrispondenti alle i -esime matrici di dati in \mathbf{G} . Di conseguenza, la soluzione per \mathbf{G}' è data dai k blocchi \mathfrak{R}_i ottenuti come combinazione lineare dei blocchi \mathfrak{S}_j la cui posizione è corrispondente a quella delle

³ In particolare: 1) la somma elemento per elemento di una matrice \mathbf{A} e un vettore colonna \mathbf{b} è la matrice $\mathbf{C} = \mathbf{A} + \mathbf{b} \mathbf{i}'$; 2) la somma elemento per elemento di una matrice \mathbf{A} e un vettore riga \mathbf{b} è la matrice $\mathbf{C} = \mathbf{A} + \mathbf{i} \mathbf{b}$; 3) la somma elemento per elemento di una matrice \mathbf{A} e uno scalare b è la matrice $\mathbf{C} = \mathbf{A} + b (\mathbf{i} \mathbf{i}')$; 4) la somma elemento per elemento di un vettore \mathbf{a} e uno scalare b è il vettore $\mathbf{c} = \mathbf{a} + b \mathbf{i}$. La definizione delle somme elemento per elemento di due matrici o due vettori non è data perché ovvia.

matrici \mathfrak{I}_j in ognuna delle macro-equazioni di (7b). Definita una matrice $(k \times q)$ nella quale ogni i -esima riga contiene il numero di identificazione relativo alla posizione di ogni j -esima matrice di dati in (7b) per ognuna delle j -esime equazioni in (7b), l' i -esimo blocco \mathfrak{R}_i può essere ottenuto come segue:

$$\mathfrak{R}_i = \left(\sum_{j \in \Lambda_i} \mathfrak{I}_j^{tr} \right) \quad i = 1, \dots, k; \quad = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}; \quad \begin{cases} \mathfrak{R}_1 = +\mathfrak{I}_1 - \mathfrak{I}_2 - \mathfrak{I}_3 \\ \mathfrak{R}_2 = +\mathfrak{I}_1 \mathbf{i}_m - \mathfrak{I}_2 \mathbf{i}_m - \mathfrak{I}_3 \mathbf{i}_m \\ \mathfrak{R}_3 = +\mathbf{i}'_n \mathfrak{I}_1 - \mathbf{i}'_n \mathfrak{I}_2 - \mathbf{i}'_n \mathfrak{I}_3 \\ \mathfrak{R}_4 = +\mathbf{i}'_n \mathfrak{I}_1 \mathbf{i}_m - \mathbf{i}'_n \mathfrak{I}_2 \mathbf{i}_m - \mathbf{i}'_n \mathfrak{I}_3 \mathbf{i}_m \end{cases} \quad (9)$$

dove i indica la i -esima riga della matrice \mathfrak{I}_j , l'apice tr indica che la stessa trasformazione dell'elemento ij -esimo dello schema (7b) deve essere applicata al blocco \mathfrak{I}_j . Inoltre, ogni blocco \mathfrak{I}_j deve avere lo stesso segno algebrico della j -esima matrice in (7b) per ognuna delle i -esime equazioni.

L'operazione di vettorizzazione sequenziale dei k blocchi permette di ottenere " " .

4. L'aggiornamento della NAM

La metodologia di bilanciamento esposta viene utilizzata al fine di aggiornare la NAM 2012 pubblicata dall'ISTAT (ISTAT, 2016) utilizzando i dati di contabilità nazionale relativi all'anno 2013 pubblicati sul database online dell'ISTAT I.Stat (<http://dati.ISTAT.it/>).

La struttura della NAM pubblicata dall'ISTAT é visibile in forma aggregata in Tab. 1, dove ogni singola cella, ad eccezione delle celle [3,9], [8c,10], [8d,10] e [9,4], rappresenta il valore totale o la voce di saldo di un conto sottostante.

Per lo sviluppo della metodologia di bilanciamento ci si è limitati a considerare il Conto dei beni e servizi della NAM. A livello disaggregato, in questo conto sono incluse in riga i Margini distributivi di commercio e di trasporto (cella [1,1]), il cui valore si annulla per il totale dell'economia, gli Impieghi intermedi (cella [1,2]), i Consumi finali (cella [1,7]), gli Investimenti fissi lordi (cella [1,8a]), la Variazione delle scorte e Oggetti di valore (cella [1,8b]) e le Esportazioni fob di beni e servizi (cella [1,9]); in colonna sono inclusi i Margini distributivi di commercio e di trasporto (cella [1,1]), la Matrice di produzione interna valutata a prezzi base (cella [2,1]), le Imposte indirette nette sui prodotti compresa l'IVA (cella [4,1]) e le Importazioni fob di beni e servizi (cella [9,1]).

Tabella 1. Struttura della National Accounts Matrix pubblicata dall'ISTAT per l'economia italiana.

	Parte corrente										Capitale			Resto del Mondo	
	1	2	3	4	5	6	7	8a	8b	8c	8d	9	10	Partite correnti	Partite in conto capitale
Beni e servizi	Produzione	Generazione del reddito	Imposte indirette nette sui prodotti e IVA	Allocazione dei redditi primari	Distribuzione secondaria del reddito	Utilizzo del reddito	Investimenti fissi lordi	Variazione delle scorte e attività non finanziarie	Acquisizioni meno cessioni di attività finanziarie	Saldo del conto economico	Resto del Mondo				
1	MTC														
2	CI						IFL					ESP			
3	VAPB											YLMA			
4	INP											YDMA			
5		YGLN	INPN	YCSR	YNL	TCSR						TCMA			
6															
7															
8a															
8b															
8c															
8d															
9	IMP		YLINMP	INPMP	YCMP	TCMP	TCCMP								
10															

MTC: Margini trasporto e commercio; CI: Consumi intermedi; CF: Consumi finali; IFL: Investimenti fissi lordi; VSOV: Variazione delle scorte compresi gli oggetti di valore; ESP: Esportazioni beni e servizi; OUT: Output; VAPB: Valore aggiunto ai Prezzi base; YLMA: Redditi da lavoro dipendente dal RdM; INP: Imposte nette sui prodotti; YGLN: Reddito generato nazionale; INPN: Imposte nette sui prodotti nazionali; YCSR: Redditi da Capitale tra settori residenti; YDMA: Redditi da Capitale dal RdM; YNL: Reddito nazionale lordo; TCSR: Trasferimenti correnti tra settori residenti; TCMA: Trasferimenti correnti dal RdM; YDNL: Reddito disponibile nazionale lordo; RFFP: Rettifica per variazioni dei diritti netti delle famiglie sulle riserve dei fondi pensione; SNL: Risparmio nazionale lordo; TCCSR: Trasferimenti in conto capitale tra settori residenti; TCCMA: Trasferimenti in conto capitale dal RdM; VSOV: Variazione delle scorte compresi gli oggetti di valore; ANFNPN: Acquisizioni nette di attività non finanziarie non prodotte nazionali; ANFNPM: Acquisizioni di attività non finanziarie non prodotte con il RdM; IAN: Indebitamento / Accredito nazionale; IAM: Indebitamento / Accredito del RdM; IMP: Importazioni beni e servizi; YLINMP: Redditi da lavoro dipendente al RdM con imposte nette su produzione al RdM; INPMP: Imposte nette sui prodotti al RdM; YCMP: Redditi da Capitale al RdM; TCMP: Trasferimenti correnti al RdM; TCCMP: Trasferimenti in conto capitale al RdM; SOCM: Saldo operazioni correnti con il RdM.

Dal punto di vista strutturale, le matrici di dati che descrivono gli impieghi dell'economia hanno come unità di classificazione di riga le prime 20 sezioni dei prodotti CPA⁴, mentre in colonna le diverse matrici di dati hanno come unità di classificazione le prime 20 branche di attività economica della classificazione NACE A*21⁵ per la matrice dei Consumi intermedi, i Settori istituzionali Famiglie, Amministrazioni pubbliche e Istituzioni sociali private per la matrice dei Consumi finali, i Settori istituzionali Società non-finanziarie, Società finanziarie, Amministrazioni pubbliche, Famiglie produttrici, Famiglie consumatrici e Istituzioni sociali private per la matrice degli Investimenti fissi lordi. La Variazione delle scorte e Oggetti di valore e le Esportazioni di beni e servizi ai prezzi fob sono, invece, costituiti da due vettori che riportano i rispettivi valori per le 20 sezioni dei prodotti CPA.

Le matrici dei dati che descrivono le risorse hanno come unità di classificazione di colonna ancora le 20 sezioni dei prodotti CPA, mentre in riga la matrice di Produzione presenta le 20 branche NACE. Per quanto riguarda, invece, le Imposte indirette nette sui prodotti compresa l'IVA e le Importazioni di beni e servizi ai prezzi fob, questi sono costituiti da 2 vettori.

Infine, la matrice dei Margini distributivi ha sia in riga che in colonna come unità di classificazione le 20 sezioni dei prodotti CPA e presenta, però, solo due righe significative in corrispondenza dei prodotti CPA *Servizi di vendita all'ingrosso e al dettaglio; servizi di riparazione di autoveicoli e motocicli e Servizi di trasporto e magazzinaggio*.

La necessità di utilizzare un metodo econometrico per l'aggiornamento della NAM per l'anno 2013 è dovuta al fatto che i dati di contabilità nazionale presenti su I.Stat sono disaggregati soltanto per branca e per Settore istituzionale, mentre non sono pubblicati i dati disaggregati in base ai prodotti CPA. Di conseguenza, per ottenere una NAM correttamente bilanciata in base a quanto stabilito dal SNA2008 / SEC2010, è necessario stimare i valori dei diversi flussi contabili per i prodotti CPA, nonché quelli dei flussi incrociati prodotti CPA – branche NACE o Settori istituzionali delle matrici dianzi descritte.

In particolare, occorre rilevare come le matrici dei Consumi intermedi e della Produzione contengano al loro interno le informazioni relative ai coefficienti di spesa e ai coefficienti di mercato delle Supply-Use Matrices e delle Symmetric Input-Output Matrices. Quindi, all'interno del processo di bilanciamento occorre evi-

⁴ Non viene considerata la sezione 21 *Servizi forniti da organizzazioni ed organismi extraterritoriali* della classificazione CPA 2008.

⁵ Non viene considerata la branca 21 *Attività di organizzazioni e organismi extraterritoriali* della NACE Rev. 2.

tare che le strutture di tali coefficienti vengano stravolte, principalmente quando l'aggiornamento dei dati riguarda due anni contigui, come nel caso in esame.

4.1 I dati per l'aggiornamento della NAM

Al fine di aggiornare la NAM 2012 all'anno 2013, si sono innanzitutto individuati i dati già pubblicati dall'ISTAT per quest'ultimo anno. In particolare, l'ISTAT rende disponibili i flussi di contabilità nazionale, in parte anche disaggregati per branca NACE o per Settore istituzionale, attraverso la banca dati pubblica I.Stat.

Per le finalità del presente lavoro, dalla banca dati I.Stat sono stati ricavati i dati relativi alla Produzione a prezzi base e i Consumi intermedi a prezzi di acquisto per branca NACE e gli Investimenti fissi lordi per Settore istituzionale. Per ciò che attiene, invece, le Esportazioni fob, le Importazioni fob e le Imposte indirette nette sui prodotti compresa l'IVA, sono disponibili nel database dell'ISTAT soltanto i valori totali per l'intera economia. Per ciò che attiene, comunque, i valori delle Imposte indirette nette sui prodotti compresa l'IVA per branca NACE si è ottenuta una prima stima dei valori per branca applicando i coefficienti di incidenza sulla produzione a prezzi base ricavati dalla NAM 2012.

Tutti i valori dei flussi sono stati considerati ai prezzi correnti dell'anno 2013.

Prima di procedere al bilanciamento, inoltre, tutte le matrici dei dati 2012 sono state in prima istanza aggiornate al 2013 utilizzando i coefficienti del 2012 e i valori di branca del 2013. Ad esempio, la matrice dei Consumi intermedi 2013 è stata inizialmente stimata utilizzando i coefficienti di spesa 2012 e i dati della produzione per branca NACE 2013.

4.2 La procedura di aggiornamento e bilanciamento della NAM

Definite tutte le strutture di dati necessari per la compilazione del Conto dei beni e servizi della NAM, si è proceduto all'aggiornamento e al bilanciamento della stessa per l'anno 2013. A tale scopo, si è innanzitutto definito il sistema di equazioni di bilanciamento. Si deve evidenziare, in tal senso, che tale sistema non assume la forma semplice dell'equazione di equilibrio della NAM in quanto, a causa della diversità che caratterizza le fonti delle diverse strutture di dati, è necessario utilizzare uno schema che definisca anche i vincoli di somma per prodotto CPA, per branca NACE e vincoli di totale delle stesse strutture di dati. In tal modo è possibile, attraverso la definizione di una opportuna struttura di varianze, mantenere inalterati nel processo di aggiornamento tutti i dati contabili già pubblicati relativi all'anno 2013 e rivalutare i rimanenti dati rinvenienti dalla NAM 2012 rispettando, al contempo, tutti i vincoli di bilancio di ogni singola struttura di dati.

La struttura delle varianze è, inoltre, stata definita in maniera da evitare stravolgimenti nei coefficienti strutturali delle diverse matrici dati. Ad esempio, in riferimento ai coefficienti di spesa definiti attraverso la matrice dei Consumi intermedi, si è evitato che vi fossero differenze tra i coefficienti 2012 e i coefficienti 2013 troppo marcate e non attribuibili a cambiamenti di tecnologia produttiva, sostanzialmente improbabili per due anni contigui.

In linea generale, quindi, è stata attribuita varianza nulla a tutti i dati ricavati dal database I.Stat, varianza relativamente bassa alle strutture di dati contenenti informazioni strutturali rilevanti e varianze relativamente più elevate per tutte le strutture di dati per le quali non è stato possibile utilizzare informazioni strutturali per l'anno 2013. Il valore relativamente più elevato di varianza è stato attribuito al vettore della Variazione delle scorte, al quale anche nelle procedure usuali di stima delle poste di contabilità nazionale viene sempre attribuito un valore *residuale* rispetto agli altri flussi⁶. Naturalmente, il processo di attribuzione delle varianze alle diverse poste contabili ha seguito un iter alquanto complesso di affinamento per step successivi.

I risultati della procedura di aggiornamento e bilanciamento del Conto dei beni e servizi della NAM sono risultati molto buoni e incoraggianti per una estensione della metodologia all'intera struttura della NAM e ad altre strutture matriciali complesse. In Tab. 2 sono riportati i dati aggregati delle matrici e dei vettori contenuti nel suddetto conto⁷. È appena il caso di ricordare che il valore 0 delle celle Beni e servizi è dovuto al fatto che a livello aggregato per l'intera economia il valore dei margini di commercio e di trasporto si annulla.

Tutti i valori contabili desunti dal database I.Stat, come era da attendersi dalla corretta applicazione del metodo di bilanciamento, non sono stati rivalutati dalla procedura di bilanciamento, mentre l'attento studio delle strutture di varianze ha contribuito a non cambiare in maniera non significativa dal punto di vista economico e produttivo le matrici dei coefficienti strutturali rivenienti dalla NAM 2012.

Difatti, le maggiori rivalutazioni si sono verificate in corrispondenza delle celle delle diverse matrici contenenti valori contabili esigui, per le quali anche una rivalutazione di poche centinaia di migliaia di euro può determinare una variazione in termini relativi nella procedura di bilanciamento relativamente molto significativa. Anche in tal caso, comunque, l'incidenza di tali rivalutazioni in termini di coefficienti strutturali è risultata entro limiti ragionevoli. Ad esempio, per ciò che attiene

⁶ Comunque, il valore della varianza attribuito non è stato tale da stimare il vettore della Variazione delle scorte in maniera completamente residuale rispetto alle equazioni di equilibrio contabile, onde evitare che una parte rilevante dei residui contabili fosse "scaricata" in tale vettore.

⁷ Per ovvi motivi di spazio non sono riportati nel presente lavoro le strutture di dati disaggregate.

la matrice dei coefficienti di spesa, questi hanno mostrato un livello di rivalutazione media, in termini assoluti, di poco superiore all'1%, e nella maggior parte dei casi le rivalutazioni maggiori si sono concentrate nelle branche dei servizi dove, a tutti gli effetti, i coefficienti di spesa possono essere alquanto variabili anche tra periodi di tempo contigui.

Tabella 2. Valori contabili aggregati relativi al conto dei beni e servizi della NAM. Anni 2012 e 2013. Milioni di € a prezzi correnti.

Partite contabili 2012		Impieghi	Risorse
Parte corrente	Beni e servizi	0	0
	Produzione	1.683.002	3.132.431
	Generazione del reddito	-	-
	Imposte indirette nette sui prodotti e IVA	-	165.245
	Allocazione dei redditi primari	-	-
	Distribuzione secondaria del reddito	-	-
	Utilizzo del reddito	1.310.581	-
Capitale	Investimenti fissi lordi	296.166	-
	Variazione delle scorte. scorte e oggetti di valore	-8.010	-
	Acquisizione meno cessioni di attività non finanziarie non prodotte	-	-
	Saldo del conto economico		
RdM	Partite correnti	461.172	445.236
	Partite in conto capitale	-	-
Partite contabili 2013		Impieghi	Risorse
Parte corrente	Beni e servizi	0	0
	Produzione	1.640.245	3.084.351
	Generazione del reddito	-	-
	Imposte indirette nette sui prodotti e IVA	-	160.493
	Allocazione dei redditi primari	-	-
	Distribuzione secondaria del reddito	-	-
	Utilizzo del reddito	1.296.296	-
Capitale	Investimenti fissi lordi	273.364	-
	Variazione delle scorte. scorte e oggetti di valore	-1302	-
	Acquisizione meno cessioni di attività non finanziarie non prodotte	-	-
	Saldo del conto economico		
RdM	Partite correnti	463.129	426.888
	Partite in conto capitale	-	-

5. Conclusioni

La crescente richiesta di dati statistici fondamentali per la corretta definizione di politiche economico-sociali da parte degli attori istituzionali preposti alla programmazione di interventi economico-finanziari sui territori interessati determina, in modo diretto, la necessità da parte degli Istituti Nazionali di Statistica di rendere disponibile al pubblico una sempre più crescente mole di dati. Pertanto, negli ultimi 20 anni se da un lato l'attenzione si è focalizzata sulla compilazione di strutture di dati sempre più complesse, dall'altro l'aggiornamento di tali dati in tempi quanto più ravvicinati possibili ha costituito oggetto di studio al fine di disporre non solo di verifiche in tempo reale della validità del dato stesso, ma anche di prime stime consistenti di dati non ancora disponibili a causa dei tempi tecnici necessari per la loro stessa fruizione.

Nel presente lavoro, si è affrontato il problema dell'aggiornamento di dati disaggregati utilizzando una variante del metodo di bilanciamento di strutture contabili di grandi dimensioni proposto da Nicolardi V. nel 1998, attualmente ancora largamente in uso dall'Istituto Nazionale di Statistica Italiano, e si è proceduto all'aggiornamento e bilanciamento di uno dei conti di contabilità nazionale che compongono la National Accounts Matrix (NAM).

In particolare, la procedura è stata utilizzata per l'aggiornamento e il bilanciamento del Conto dei beni e servizi della NAM 2012 al 2013 in base ai dati ISTAT disponibili su I.Stat.

La necessità di utilizzare un metodo econometrico per l'aggiornamento della NAM per l'anno 2013 è dovuta principalmente alla mancanza di dati disaggregati per lo stesso anno in base alla classificazione dei prodotti CPA. Considerata l'importanza di poter fornire nel più breve tempo possibile una stima robusta dei dati non ancora disponibili, si comprende l'importanza che un solido metodo di aggiornamento e bilanciamento delle stime ha all'interno della produzione statistica del dato di un paese.

I risultati della procedura di aggiornamento e bilanciamento applicata al Conto dei beni e servizi della NAM 2012 con dati 2013 sono soddisfacenti. Si evidenzia come tutti i valori contabili per l'anno 2013 già pubblicati non siano stati rivalutati dalla procedura stessa e, ancor più importante, le matrici dei coefficienti strutturali rivenienti dalla NAM 2012 non abbiano subito un cambiamento economicamente non significativo. Pertanto, si conclude che tutti i presupposti per l'estensione del metodo all'intera struttura della NAM siano stati verificati e risultati molto validi.

Riferimenti bibliografici

- Commission of the European Communities - EUROSTAT, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, World Bank (1993). *System of National Accounts 1993*, Brussels/Luxembourg, New York, Paris, Washington, D.C.
- European Communities, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, World Bank (2009). *System of National Accounts 2008*, New York.
- European Union – EUROSTAT (2013). *European system of accounts. ESA 2010*, Luxembourg: Publications Office of the European Union
- EUROSTAT (1995). *European System Of Accounts. ESA 1995*, Luxembourg.
- ISTAT (2011). *Metodologie di stima degli aggregati dei conti nazionali a prezzi correnti Anno base 2000, Metodi e Norme n. 51*, Roma.
- ISTAT (2016). *La matrice dei conti nazionali. Anno 2012*, ISTAT, Roma.
- Marini C., Nicolardi V. (2010) Simultaneously Balancing Supply-Use Tables at current and constant prices, *Annali del Dipartimento di Scienze Statistiche "Carlo Cecchi"*. Università degli Studi di Bari Aldo Moro, vol. IX, pp. 111-131, dicembre 2010.
- Nicolardi V. (1998). Un sistema di bilanciamento per matrici contabili di grandi dimensioni. *Quaderni di ricerca, 4/1998*. Istituto Nazionale di Statistica, Roma.
- Nicolardi V. (2000). Balancing Large Accounting System: an Application to the 1992 Italian I-O Table. *13th International Conference on Input-Output Techniques*, University of Macerata (Italy), August 21-25th, 2000. (Available from [http://www.iioa.org/Conference/13thdownable paper.htm](http://www.iioa.org/Conference/13thdownable%20paper.htm)).
- Nicolardi V. (2010). Simultaneous I-O tables balancing at current and constant prices. *45th Scientific Meeting of the Italian Statistical Society*, University of Padua (Italy), June 29th, 2011 - July 1th, 2011. (Available from <http://homes.stat.unipd.it/mgri/SIS2010/Program/contributedpaper/537-1312-1-DR.pdf>).
- Nicolardi V. (2011) Supply-Use Tables: Simultaneously Balancing at Current and Constant Prices. A New Procedure, *19th International Input-Output Conference*, Alexandria VA (USA), 13th – 17th June, 2011. Paper available at the web page: <http://www.iioa.org/Conference/19th-downable%20paper.htm>.
- Nicolardi V. (2012). A New Method for the Simultaneous Balancing of the Economic Accounting Systems at Current and Constant Prices, *JSM Proceedings, Business and Economic Statistics Section*. San Diego, CA: American Statistical Association.
- Nicolardi V. (2013). Simultaneously Balancing Supply–Use Tables at Current and Constant Prices: a New Procedure, *Economic Systems Research*, Routledge Ltd, London, UK.

Un indicatore sintetico per il bilancio di genere dell'Università di Foggia

Laura Antonucci*, Corrado Crocetta

Università di Foggia

Riassunto: Sulla base della crescente sensibilità al tema della responsabilità sociale, molte istituzioni, pubbliche e private, sentono ormai l'esigenza di integrare le informazioni contabili pubblicando bilanci sociali e di genere. Il presente lavoro descrive un nuovo indicatore sintetico, proposto in occasione della redazione del primo bilancio di genere dell'Università di Foggia, che consente di monitorare l'equilibrio di genere di ciascuna componente della comunità accademica.

Keywords: Bilancio di genere, Indicatore sintetico, Università di Foggia.

1. Introduzione

La Quarta Conferenza delle Donne, tenutasi a Pechino nel 1995, ha approvato la "Beijing Platform for Action" in cui si afferma, tra le altre cose, l'importanza e l'efficacia del bilancio di genere come documento per l'analisi delle esigenze, delle opportunità e delle occasioni di partecipazione ai processi decisionali ed, in generale, delle politiche di genere delle grandi organizzazioni.

In questo lavoro si riportano alcuni risultati metodologici sviluppati in occasione della redazione del primo bilancio di genere dell'Università di Foggia. Tale documento sintetizza l'approfondita analisi, condotta da un gruppo di lavoro sul tema delle pari opportunità, relativa alle diverse componenti della comunità accademica: studenti, personale docente, personale tecnico amministrativo e organi di governo e di controllo.

* Autore corrispondente: laura.antonucci@unifg.it.

Il presente lavoro è opera congiunta dei due autori, tuttavia a L. Antonucci vanno attribuiti i paragrafi 3, 4 e 5 ed a C. Crocetta i paragrafi 1 e 2

2. Le fonti

I dati utilizzati per le analisi sono stati raccolti nei primi 8 mesi del 2016 e si riferiscono al triennio 2013-2015. La maggior parte dei dati è stata tratta dal sistema informativo dell'ateneo ma sono state utilizzate anche diverse banche dati quali: Funzione Pubblica, MIUR, Consorzio ALMALAUREA e Consorzio Interuniversitario Cineca. Solo i dati relativi agli studenti iscritti sono riferiti agli anni accademici, mentre per tutte le altre componenti è stato utilizzato l'anno solare.

3. Il metodo di ponderazione proposto

Il bilancio di genere dell'Università di Foggia, pubblicato nel 2016, fa riferimento al periodo 2013 -2015 e considera le diverse componenti in cui si articola la comunità accademica dell'Università di Foggia analizzandone il peso in relazione al genere. Per sintetizzare l'enorme mole di dati forniti dal documento, si è pensato di costruire un cruscotto di indicatori in grado di monitorare il raggiungimento degli obiettivi strategici inerenti il rispetto della parità di genere.

Dovendo sintetizzare informazioni eterogenee come popolazione studentesca, corpo docente, personale tecnico amministrativo ed organi statutari di governo è importante definire dei criteri in grado di aggregare tali variabili e di attribuire opportuni pesi alle diverse componenti.

Il gruppo di lavoro ha deciso di rinunciare a criteri di ponderazione di tipo soggettivo, adottando un sistema oggettivo che fa riferimento a parametri di costo definiti a livello ministeriale. Il sistema di ponderazione scelto si ispira, infatti, al criterio del costo standard delle stesse. Pertanto il peso di ciascuna componente nel calcolo dell'indicatore sintetico è proporzionale al costo effettivo o a quello standard della componente considerata. Per gli organi statutari i pesi utilizzati sono rappresentati dalle indennità previste per il Rettore, per i componenti del senato accademico, del consiglio di amministrazione, del nucleo di valutazione ecc.

Per gli studenti, il peso attribuito per il calcolo dell'indice sintetico è rappresentato dal costo standard, calcolato in base al Decreto Interministeriale 893 del 9 dicembre 2014, e facendo riferimento alla tabella relativa all'anno 2015, che per l'Università di Foggia indica un costo di 6.394 Euro pro-capite. Per i docenti ed il personale tecnico amministrativo sono stati applicati i coefficienti di peso dei costi standard, definiti nel D.M. 713 del 9 agosto 2013, assumendo come base di calcolo il valore del punto organico che, in base al D. M. 619 del 5 agosto 2016, nel 2015 è pari a 114.610 euro.

Tabella 1. *Coefficienti di ponderazione dei ruoli pubblicati nel D.M. 713 del 9 agosto 2013.*

Ruolo	Peso dei punti organico
Professori Ordinari	1,00
Professori Associati	0,70
Ricercatori Universitari	0,50
Categoria EP	0,40
Categoria D	0,30
Categoria C	0,25
Categoria B	0,20
Categoria CEL (collab. esperti linguistici)	0,20

In pratica, l'indice sintetico proposto attribuisce a ciascun organo un peso pari al costo (effettivo per gli organi di indirizzo politico amministrativo e standard per tutte le altre componenti) che l'Università sostiene per remunerare i componenti di quell'organo o i propri dipendenti. Per facilitare la confrontabilità degli indici nei vari anni si è deciso di utilizzare come base di riferimento le voci di costo relative all'anno 2015, modificando solo il numero di componenti, studenti e dipendenti, distinti in base al sesso, per il triennio considerato.

4. Il bilancio di genere

I dati riportati nella **Tabella 2** riportano i costi delle diverse componenti dell'Università di Foggia, ricavati analizzando la composizione per sesso dei vari organi di indirizzo politico amministrativo previsti dallo statuto, la composizione del corpo docente, del personale tecnico amministrativo e della popolazione studentesca.

Considerando gli organi di governo dell'Ateneo, si nota che il peso dei compensi erogati ai componenti di sesso femminile è notevolmente inferiore rispetto a quello della componente maschile e varia fra il 21% del 2013 ed il 27% del 2015. Tale differenza è dovuta al fatto che le posizioni di Rettore, presidente del collegio dei revisori e di presidente del nucleo di valutazione sono ricoperte da uomini e che il collegio dei direttori di dipartimento, il senato accademico ed il consiglio di amministrazione evidenziano un consistente equilibrio di genere.

Proseguendo l'analisi dei costi relativi al corpo accademico si nota facilmente che i professori ordinari, associati ed i ricercatori a tempo determinato sono prevalentemente di sesso maschile mentre fra i ricercatori a tempo indeterminato prevale la componente femminile.

Tabella 2. Costi in Euro delle diverse componenti della Università di Foggia in relazione al genere, anni 2013 -2015

Componenti	2013		2014		2015	
	F	M	F	M	F	M
Rettore	-	33.466	-	33.466	-	33.466
Pro-Rettore	13.387	10.877	13.387	10.877	13.387	10.877
Dir. dipartimento	8.367	41.833	8.367	41.833	8.367	41.833
Senato accadem.	10.709	53.546	13.387	50.869	16.064	48.191
Cons. amministr.	4.351	39.156	8.701	34.805	13.052	26.104
Pres. N.V.I.	-	7.179	-	7.179	-	7.179
Comp. N.V.I	4.859	14.578	4.859	14.578	4.859	14.578
Pre. coll. revisori	-	9.571	-	9.571	-	9.571
Comp. coll. revis.	14.578	-	14.578	-	14.578	-
Organi governo	56.250	210.206	63.278	203.178	70.306	191.799
Prof. Ordinari	1.740.000	8.700.000	1.856.000	8.236.000	1.856.000	8.004.000
Prof. Associati	2.923.200	4.222.400	3.248.000	4.141.200	3.248.000	4.790.800
Ric. Universitari	5.278.000	4.988.000	4.988.000	4.814.000	4.814.000	4.408.000
Ricercatori TD	92.800	139.200	139.200	232.000	603.200	696.000
Totale docenti	10.034.000	18.049.600	10.231.200	17.423.200	10.521.200	17.898.800
Direttore generale	-	125.886	-	125.886	-	125.886
Categoria EP	603.200	788.800	603.200	788.800	556.800	788.800
Categoria D	2.088.000	1.426.800	2.088.000	1.357.200	1.983.600	1.357.200
Categoria C	3.103.000	2.407.000	3.016.000	2.407.000	2.987.000	2.291.000
Categoria B	417.600	371.200	417.600	371.200	417.600	348.000
Categoria CEL	162.400	46.400	162.400	46.400	162.400	46.400
Totale T.A.	6.374.200	5.166.086	6.287.200	5.096.486	6.107.400	4.957.286
Studenti	37.799.282	21.262.096	38.388.553	21.593.561	37.291.982	22.856.376
Totale ateneo	54.263.732	44.687.988	54.970.231	44.316.425	53.990.888	45.904.261

Fonte: Miur, anagrafe docenti; sito web www.unifg.it

Anche per il personale tecnico amministrativo si nota che nei ruoli apicali di direttore generale e di personale di elevata professionalità (categoria EP) prevale la componente maschile mentre in tutti gli altri la situazione è ribaltata.

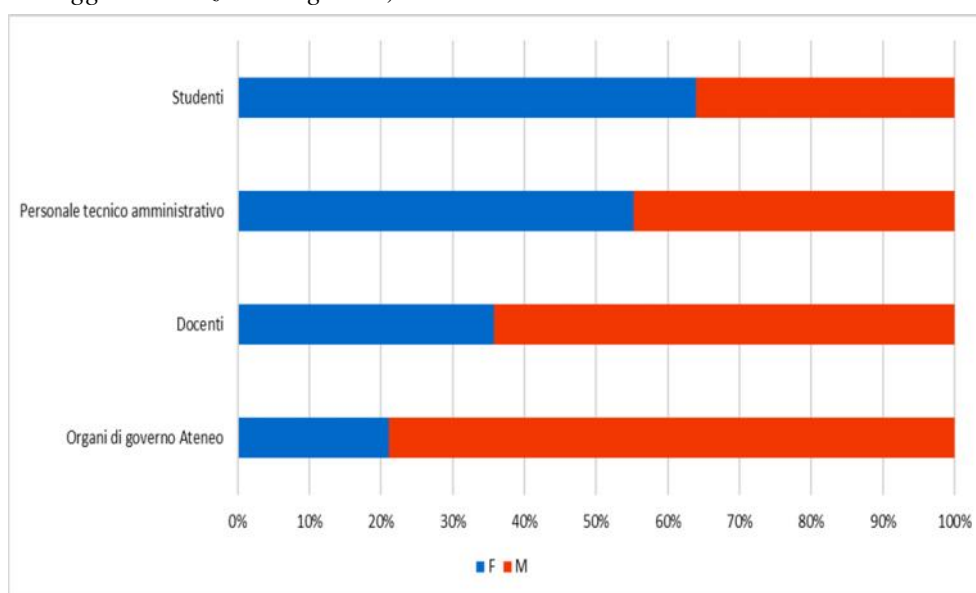
Analizzando la componente studentesca, come detto nei capitoli precedenti, si nota che le studentesse rappresentano la quota maggioritaria con tassi di femminilità pari al 64% nel 2013 e 2014 e al 62% nel 2015.

L'indicatore sintetico complessivo evidenzia che nell'Università di Foggia non si ravvisa un problema di scarsa presenza femminile, anzi al contrario, questa com-

ponente supera quella maschile in tutto il triennio considerato, ma, così come rilevato in molte altre università italiane, sembra esserci una sorta di *soffitto di cristallo*, ovvero una barriera invisibile, che impedisce alle donne di raggiungere le posizioni apicali. Tale barriera denota la presenza di ostacoli che impediscono alle donne di accedere a posizioni di responsabilità. Traducendo l'immagine del *soffitto di cristallo* possiamo affermare che le donne guardano in alto e non vedono ostacoli, perché l'atmosfera paritaria che sembra regnare nel loro ambiente le stimola ad una competizione aperta; ma nella loro carriera sembra esserci un ostacolo invisibile che impedisce loro di superare una determinata soglia. Nell'Università di Foggia componente femminile, pur essendo nel complesso più numerosa, ha un peso ridotto tra i professori ordinari, i dirigenti ed vertici decisionali. Per modificare tale situazione l'Ateneo ha previsto l'attivazione di alcune *azioni positive* in favore delle donne per permettere loro di avanzare e progredire sensibilmente.

Esiste quindi un problema di *segregazione verticale* che, inevitabilmente, comporta una *segregazione economica* che fa sì che le donne finiscano con il percepire un reddito inferiore rispetto a quello degli uomini.

Figura 1. Incidenza percentuale dei costi delle diverse componenti dell'Università di Foggia in relazione al genere, anno 2015



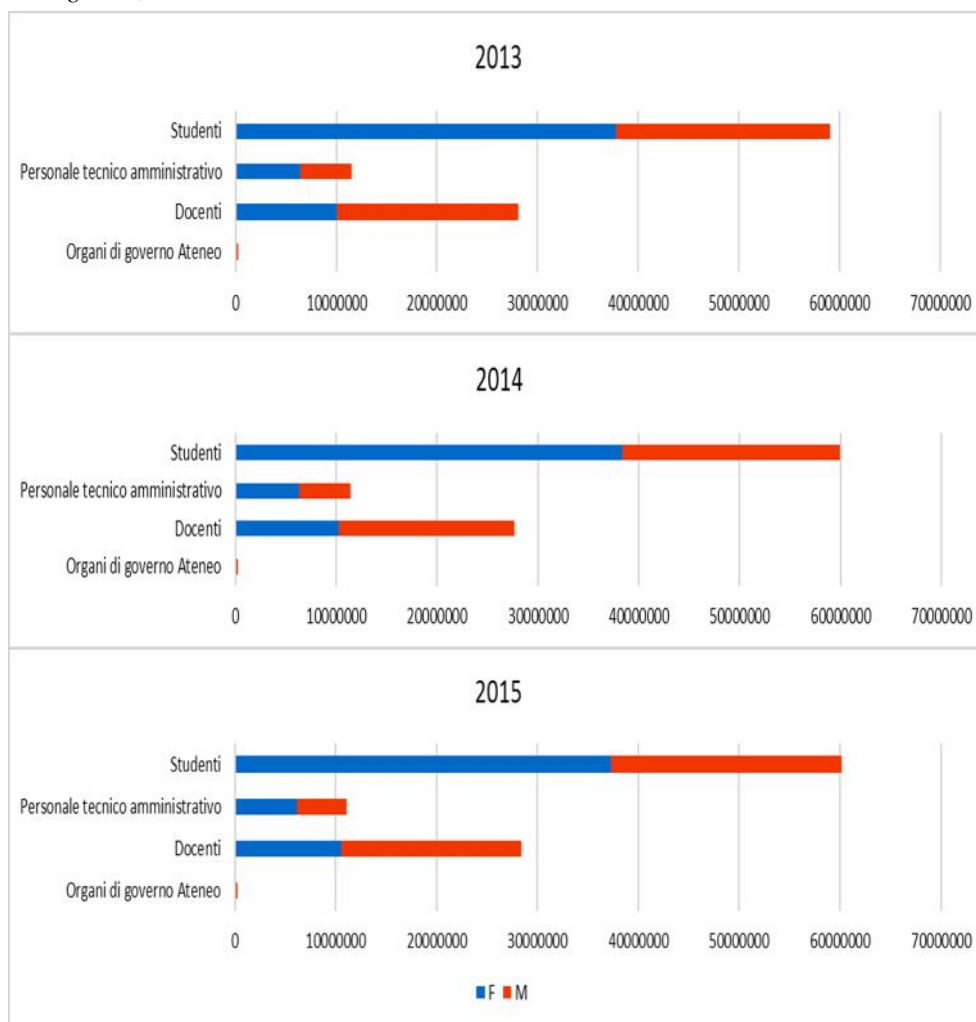
Fonte: Miur anagrafe docenti e sito web www.unifg.it

La **Figura 1** mostra con grande chiarezza l'esistenza di una piramide rovesciata per cui il peso della componente femminile decresce man mano che ci si avvicina

ai vertici. La letteratura sulla segregazione di genere spesso parte dalla collocazione occupazionale per spiegare le disparità salariali, infatti, come sostiene Bettio, la segregazione esiste a prescindere dal fatto che vi siano o meno differenze di produttività o che vi sia o meno la volontà di segregare.

Anche effettuando i confronti in termini di costi (in valore assoluto) sostenuti per le diverse componenti nei 3 anni considerati (**Figura 2**) si nota che il contributo della componente femminile diminuisce al crescere delle responsabilità e che tale tendenza non ha mostrato che lievi segni di miglioramento

Figura 2. Incidenza dei costi delle diverse componenti dell'Università di Foggia in relazione al genere, anni 2013- 15



Fonte: Miur anagrafe docenti e sito web www.unifg.it

5. Gli indicatori sintetici per l'equilibrio di genere

Volendo capire le motivazioni alla base delle situazione descritta nei paragrafi precedenti, si potrebbe far riferimento alla tesi di Polachek, per il quale le donne sono orientate verso lavori che richiedono minori investimenti in capitale umano e che quindi sono meno soggetti alla così detta *atrofia delle conoscenze*, causata da periodiche assenze dal mercato del lavoro, ma le vere motivazioni vanno ricercate nei processi decisionali delle famiglie che, stante l'assenza di un serie di servizi di supporto, scelgono di sacrificare le ambizioni lavorative di uno dei coniugi, solitamente quello a minore capacità reddituale.

Tabella 3. *Indicatori sintetici di incidenza della componente femminile sui costi complessivi delle diverse componenti della Università degli Studi di Foggia, anni 2013 -2015*

Indicatori sintetici	2013	2014	2015
Organi di governo Ateneo	0,21	0,24	0,27
Professori Ordinari	0,17	0,18	0,19
Professori Associati	0,41	0,44	0,40
Ricercatori Universitari	0,51	0,51	0,52
Ricercatori TD	0,40	0,38	0,46
Totale docenti	0,36	0,37	0,37
Direttore generale	-	-	-
Categoria EP	0,43	0,43	0,41
Categoria D	0,59	0,61	0,59
Categoria C	0,56	0,56	0,57
Categoria B	0,53	0,53	0,55
Categoria CEL	0,78	0,78	0,78
Totale tecnici amministrativi.	0,55	0,55	0,55
Studenti	0,64	0,64	0,62
Totale complessivo	0,55	0,55	0,54

Fonte: Miur anagrafe docenti e sito web www.unifg.it

La **Tabella 3** fornisce un quadro sintetico del fenomeno, mediante il tasso di femminilità, ottenuto aggregando le diverse componenti in base ai costi (effettivi o standard) e calcolando l'incidenza dei costi sostenuti in favore delle donne rispetto al totale.

Il confronto fra i vari anni consente di cogliere con immediatezza i progressi registrati sino ad ora, infatti se nel 2013 la spesa “al femminile” per la remunerazione degli organi di governo aveva una incidenza di 0,21 rispetto al totale, nel 2015 è passata a 0,27.

Piccoli progressi sono stati fatti anche sul fronte dei professori ordinari dove il tasso è passato dallo 0,17 allo 0,19. La componente docenti vede prevalere la componente femminile solo per i ricercatori universitari a tempo indeterminato ma, nel complesso, si caratterizza per una netta prevalenza del numero di maschi.

Poiché il direttore generale è un organo monocratico e nel triennio l’incarico è stato ricoperto da un uomo, l’indice è pari a zero.

Per quanto riguarda il personale di categoria EP, a seguito della diminuzione di due unità, avvenuta nel 2015, il tasso è leggermente peggiorato rispetto al 2013, anche se le disparità di genere sono meno pronunciate rispetto a quelle riscontrate per i professori ordinari. Man mano che si scorre verso il basso l’organigramma dell’Università di Foggia, si vede che l’indice supera il valore 0,5 testimoniando che la numerosità della componente femminile è maggiore rispetto a quella della componente maschile. Raggruppando tutto il personale tecnico amministrativo in un’unica classe si nota come, pur tendo conto delle diverse remunerazioni percepite, il saldo è leggermente favorevole alle donne.

Le studentesse prevalgono numericamente rispetto ai loro colleghi, anche se, nel triennio, il divario è lievemente diminuito.

In sintesi, l’indicatore complessivo indica l’assenza di disuguaglianze di genere all’interno dell’Ateneo, visto che tale valore oscilla fra 0,55 e 0,54 a testimonianza del fatto che, nella attribuzione delle risorse, la componente femminile riceve una quota leggermente superiore rispetto a quella maschile.

Ovviamente i valori indicati nella tabella risentono delle scelte fatte sia in merito al metodo di aggregazione (somma ponderata di costi) che al livello di granularità scelto, ma, pur con tutte le dovute cautele si può affermare che il bilancio di genere indica un sostanziale equilibrio e l’assenza di situazioni particolarmente gravi. Certo permane la situazione di segregazione verticale sopra descritta su cui non è facile intervenire nel breve periodo perché le cause di tali segregazione vanno ricercate nella:

- a) asimmetria dei ruoli e delle responsabilità domestiche (a causa del permanere di modelli familiari tradizionali e della assenza di serie politiche di conciliazione famiglia - lavoro).
- b) diverse motivazioni alla base dell’investimento in carriera (dovuta anche alla consapevolezza di avere minori opportunità).

- c) perdurare di una mentalità che tende a relegare le donne in ruoli di servizio e di cura, escludendole da posizioni di maggiore prestigio.
- d) presenza di modelli organizzativi che premiano la disponibilità di tempo a discapito dell'efficacia e del raggiungimento degli obiettivi.

Riferimenti bibliografici

Delvecchio, F. (2015). *Statistica per l'analisi dei fenomeni sociali*. Cleup, Padova.

Galizzi, G. (2013). *Il bilancio di genere negli enti pubblici territoriali. Origini, strumenti e implicazioni aziendali*, Franco Angeli, collana Economia - Ricerche, Milano

Ronda S., Broomhill R. (2002), *Budgeting for equality; the australian experience*, Feminist economics, p. 25

Walby S. (2006) *Gender budgeting: what is it? how to do it? why do it?*

Analisi dell'efficienza organizzativa nelle strutture sanitarie pugliesi tramite alberi di classificazione

Rossana Mancarella*

ARTI Agenzia Regionale per la Tecnologia e l'Innovazione - Regione Puglia

Riassunto: Nell'ambito del tema vastissimo *efficienza sanitaria* questo studio si propone di valutare l'efficienza dei servizi ospedalieri ponendo particolare attenzione ad uno specifico aspetto: *l'appropriatezza organizzativa dei ricoveri ospedalieri*. Una gestione adeguata dei ricoveri ospedalieri, infatti, potrebbe consentire una contrazione dei costi di assistenza sanitaria pur garantendo standard elevati ed una ottimizzazione finale del rapporto costi-benefici. Sarebbe auspicabile, dunque, riuscire a suggerire delle strategie opportune al fine di evitare il ricorso ai costosi ricoveri ordinari quando superflui, ossia adeguatamente sostituibili da altri tipi di assistenza quali per esempio il day hospital o l'assistenza ambulatoriale; naturalmente, quando il quadro clinico del paziente lo consenta. Questa indagine, avvalendosi del corposo database delle Schede di Dimissione Ospedaliera fornito dall'Agenzia Regionale Sanitaria della Regione Puglia, riguardante i tutti i pazienti ricoverati negli istituti di assistenza sanitaria pugliesi nel periodo 2009-2014, fornisce risultati interessanti sul meccanismo decisionale della gestione sanitaria in termini di appropriatezza organizzativa dei ricoveri ospedalieri.

Keywords: Assistenza sanitaria; Efficienza organizzativa; Day Hospital; Analisi di Segmentazione.

1. Introduzione

Questo lavoro si propone di valutare l'efficienza dei servizi ospedalieri mediante tecniche statistiche e computazionali multivariate prestando particolare attenzione all'adeguatezza organizzativa dei ricoveri ospedalieri. Nello specifico, questo stu-

* Corresponding Author: r.mancarella@arti.puglia.it

dio cerca di verificare tale adeguatezza negli istituti di ricovero e cura pubblici e privati pugliesi, mediante un'analisi di classificazione.

La gestione dei ricoveri ospedalieri può considerarsi *efficiente* quando riesce a ridurre i costi di assistenza sanitaria pur garantendo standard elevati, evitando, per esempio, inutili ricoveri ordinari, allorquando questi possono essere adeguatamente sostituiti da altri tipi di assistenza.

Questa indagine si avvale di un corposo database fornito dall'Agenzia Regionale Sanitaria della Regione Puglia (AReS), riportante gran parte delle informazioni salienti (ad eccezione di dati di identità) contenute nelle Schede di Dimissione Ospedaliera (SDO) di tutti i pazienti ricoverati negli istituti di assistenza sanitaria pugliesi nel periodo di tempo 2009-2014¹.

I dati inclusi nella SDO (tra cui più di 18.000 codici diagnostici e di procedura, a seguito della classificazione ICD-9-CM²) possono essere raggruppati in base a due diverse Classificazioni: il sistema DRG (Diagnosis Related Groups), e il sistema di raggruppamento ACC (Aggregated Clinical Codes, ora definiti come CCS³), basato su criteri clinici di affinità, e dunque adatto per analisi di tipo epidemiologico e di programmazione sanitaria, ma generalmente meno appropriato dei DRG ai fini della programmazione economica.

Tuttavia, i 289 CCS (o ACC) di ICD-9-CM sono troppi in questo contesto sperimentale, per cui si è deciso di condurre l'indagine soltanto sui 24 codici CCS legati alle malattie cardiovascolari, con un numero di ricoveri in ospedali pubblici e

¹ Sono disponibili, in realtà, anche i dati relativi al quinquennio precedente, utilizzati in precedenza per alcune ricerche-pilota anche dalla scrivente, ma si è ritenuto opportuno non farvi ricorso in questa ricerca sia per problemi di corretta standardizzazione delle diagnosi che per problemi legati alle cospicue riorganizzazioni del sistema ospedaliero pugliese avvenute in tale periodo.

² L'ICD-9-CM (Classificazione internazionale delle malattie, Clinical Modification) è un adattamento dell'OMS ICD-9, creato dal Centro nazionale degli Stati Uniti for Health Statistics (NCHS) e utilizzato per l'assegnazione di codici di diagnosi e procedura associata a pazienti ricoverati e ambulatoriali. L'ICD-9-CM (aggiornato annualmente) si basa sul ICD-9, ma prevede ulteriori dettagli di morbilità. Conta 14.087 codici di diagnosi e 3.900 codici di procedura. L'ICD-10-CM, sviluppato dal NCHS nel mese di ottobre del 2013, dopo un beta test di due anni, è in uso dall'ottobre 2015 in alcuni Stati negli U.S.A; esso conta quasi 70.000 codici di diagnosi e 72.000 codici di procedura, ed è un sistema di classificazione che comprende episodi di ospedalizzazione caratterizzati da un consumo di risorse omogeneo.

³ La classificazione ACC (in ICD-9-CM, 289 ACC diagnostici e 231 ACC di interventi/procedure) è stata sviluppata a partire dal 1993 dalla US Agency for Healthcare Quality e la Ricerca (AHRQ), al fine di raggruppare ICD-9-CM in un insieme limitato di classi relativamente omogenee, utile per i ricercatori e gli analisti. Questa classificazione è stata progressivamente ottenuta modificando la precedente CCHPR (Classificazione clinica per la ricerca sulle politiche sanitarie), seguendo criteri clinici di omogeneità e la frequenza delle dimissioni ospedaliere. È stato periodicamente aggiornato secondo la ICD. Dal 1999 il sistema è stato informatizzato e rinominato "Classification Clinical Software".

privati di tutta l'area pugliese, pari a 1.310.716 (va precisato che l'oggetto dell'indagine è rappresentato dai ricoveri e non dai pazienti, che possono essere ricoverati più di una volta nell'arco di tempo considerato).

La scelta è ricaduta sulla patologia cardiovascolare in quanto essa costituisce di gran lunga la prima causa di morte in Italia. Inoltre, la letteratura testimonia che il suo impatto in termini di mortalità, morbosità e anche sui ricoveri ospedalieri si mantiene elevato. Il quadro epidemiologico, dunque, suggerisce che un'ampia parte delle risorse del Sistema Sanitario debba essere dedicata alla prevenzione, diagnosi e cura delle malattie cardiovascolari e che l'organizzazione dell'attività diagnostica e terapeutica debba fortemente concentrarsi sulle principali patologie che affliggono la popolazione.

2. Regimi di ricovero previsti dal SSN

Prima di addentrarci nello studio, appare opportuno definire le varie tipologie di ricovero previste dal nostro Sistema Sanitario e nostro oggetto di interesse:

Il *day hospital* è un regime di ricovero che prevede il compimento dell'assistenza sanitaria unicamente nelle ore diurne, può svolgersi, eventualmente, in uno o più "accessi" programmati, purché tutti vengano registrati in un'unica cartella clinica aperta in concomitanza del primo accesso.

Durante tale ricovero: il quadro clinico del paziente è oggetto di valutazione combinata di più specialisti; il paziente può beneficiare di prestazioni specialistiche di notevole impegno professionale, sia mediche che chirurgiche, e viene sottoposto ad un periodo di osservazione a seguito di eventuali procedure diagnostiche o chirurgiche di tipo invasivo.

La peculiarità di tale regime di ricovero è l'assenza di pernottamento presso la struttura erogante l'assistenza sanitaria.

Il vantaggio del ricovero *day hospital* è la possibilità di trattare appropriatamente i pazienti, il cui quadro clinico lo consenta, con un minor impegno economico richiesto alla struttura erogante l'assistenza sanitaria, rispetto al ricovero ordinario.

In sostanza, a parità di prestazioni professionali e di appropriatezza di assistenza sanitaria, il regime *day hospital* sembrerebbe ottimizzare il rapporto costi-benefici, rispetto al regime ordinario.

Ciò che invece differenzia il regime di *day hospital*, rispetto al trattamento ambulatoriale, è il tempo necessario all'espletamento delle prestazioni e l'eventuale necessità del paziente di essere sottoposto durante il ricovero a più prestazioni mediche o infermieristiche e/o di necessitare di alcune ore di osservazione.

Infine, potremmo definire il *day surgery* come la versione chirurgica del *day hospital*. Il *day surgery* è, in effetti, un regime diurno destinato al ricovero di pazienti che necessitano di interventi chirurgici o di procedure diagnostiche invasive che richiedono l'utilizzo della sala operatoria, anche in questo caso, come per gli altri regimi diurni, non è previsto il pernottamento nella struttura erogante l'assistenza sanitaria.

Nel presente lavoro si è scelto di considerare il ricovero *day surgery* congiuntamente a quello *day hospital*, posto che entrambi mirano al medesimo scopo (riduzione del costo per la struttura e dei disagi per i pazienti) e soffrono dei medesimi limiti operativi (non essendo applicabili a casi di una certa gravità).

Dall'analisi della tabella 1 è possibile evincere la dinamica temporale della distribuzione dei ricoveri secondo il regime (Ordinario e Day Hospital).

Tabella 1. *Distribuzione dei ricoveri con diagnosi di patologie cardiovascolari in Puglia secondo il regime, per anno.*

Anno	Regime Ricovero				Totale	
	Ordinario		Day Hospital		%	Numero ricoveri
	%	Numero ricoveri	%	Numero ricoveri		
2009	88,3	208.004	11,7	27.653	100,0	235.657
2010	82,7	205.593	17,3	42.981	100,0	248.574
2011	83,9	193.036	16,1	36.921	100,0	229.957
2012	89,1	183.757	10,9	22.378	100,0	206.135
2013	91,4	182.075	8,6	17.051	100,0	199.126
2014	92,0	176.056	8,0	15.211	100,0	191.267
Totale	87,6	1.148.521	12,4	162.195	100,0	1.310.716

Fonte: Elaborazioni proprie su dati AReS.

Appare evidente che, nell'arco temporale considerato, viene preferito il regime di ricovero Ordinario rispetto a quello Day Hospital⁴ (87,6% VS 12,4%).

È interessante sottolineare come la distribuzione dei ricoveri secondo il regime conosca fasi oscillanti nel corso degli anni: i ricoveri in Day Hospital vengono preferiti a quelli Ordinari in poco meno del 12% dei casi nel 2009, si apprezza un incremento di tale valore di circa cinque punti percentuali nell'anno successivo, poi la distribuzione decresce prima lentamente (circa il 16% nel 2011) e in seguito molto più velocemente (circa 11% nel 2012, circa 8% nei restanti anni).

⁴ Il nostro campione appare evidentemente sbilanciato rispetto a quella che sarà la nostra variabile di interesse, "Regime di ricovero": severo limite, questo, per le successive analisi multivariate, ma vedremo nel seguito come esso sarà superato.

Va detto che l'incremento dei ricoveri in regime diurno era stato invece costante nel quinquennio precedente. Tale dinamica è presumibilmente da addebitare alla sempre maggiore diffusione dei servizi sanitari ambulatoriali e anche al ricorso (sostenuto anche da politiche sanitarie pubbliche) all'assistenza sanitaria domiciliare: politiche entrambe ancor meno costose del ricovero diurno, ma che sfuggono alle rilevazioni tramite SDO.

Per quel che concerne la distribuzione percentuale dei ricoveri secondo il regime (Ordinario o Day Hospital) e per diagnosi connesse al ricovero (Cardiocircolatorie e Altre diagnosi) appare evidente, nel periodo 2009-2014, la preponderante scelta del regime di ricovero Ordinario sia per le diagnosi di tipo cardiocircolatorio che per le altre, rispettivamente circa l'88% e il 73% dei ricoveri totali (Tabella 2).

Tabella 2. *Distribuzione percentuale dei ricoveri nelle strutture sanitarie pugliesi secondo il regime di ricovero, per diagnosi connesse. Anni 2009 – 2014.*

Anno	Regime di ricovero	Diagnosi connesse al ricovero		Totale ricoveri
		Cardiocircolatorie	Altre diagnosi	
2009	Ordinario	88,3	78,5	81,5
	Day Hospital	11,7	21,5	18,5
	Totale ricoveri	235.657	543.395	779.052
2010	Ordinario	82,7	68,8	72,8
	Day Hospital	17,3	31,2	27,2
	Totale ricoveri	248.574	612.475	861.049
2011	Ordinario	83,9	69,3	73,5
	Day Hospital	16,1	30,7	26,5
	Totale ricoveri	229.957	571.195	801.152
2012	Ordinario	89,1	72,8	77,6
	Day Hospital	10,9	27,2	22,4
	Totale ricoveri	206.135	504.775	710.910
2013	Ordinario	91,4	75,3	80,1
	Day Hospital	8,6	24,7	19,9
	Totale ricoveri	199.126	471.148	670.274
2014	Ordinario	92,0	76,7	81,3
	Day Hospital	8,0	23,3	18,7
	Totale ricoveri	191.267	444.959	636.226
Totale	Ordinario	87,6	73,3	77,5
	Day Hospital	12,4	26,7	22,5
	Totale ricoveri	1.310.716	3.147.947	4.458.663
Totale%		100,0	100,0	100,0

Fonte: Elaborazioni proprie su dati AReS.

Nello specifico, osservando la dinamica temporale del fenomeno, nel caso di diagnosi di tipo Cardiocircolatorio il regime di ricovero Ordinario viene scelto in percentuali che oscillano tra valori compresi tra circa l'88% e il 92% del totale

dei ricoveri ad eccezione degli anni 2010 e 2011, anni in cui tali percentuali calano fino a poco meno dell'84%.

Quando il ricovero è connesso ad Altre Diagnosi, viene preferito il regime Ordinario in circa il 78% dei casi nel 2009, nel 2010 si registra un calo di circa 10 punti percentuali, per lo stesso regime di ricovero, e negli anni successivi l'andamento della distribuzione percentuale dei ricoveri Ordinari appare crescente fino a sfiorare il valore del 77% dei ricoveri nel 2014.

Se osserviamo la distribuzione percentuale dei ricoveri secondo le diagnosi ad essi connesse (Cardiocircolatorie e Altre diagnosi) per regime (Ordinario o Day Hospital) è immediato constatare (Tabella 3) la presenza di un maggior numero di ricoveri sia ordinari che day hospital per le diagnosi classificate come non cardiocircolatorie, rispettivamente si registra circa il 67% e l'84% sul totale dei ricoveri.

In particolare, per quel che riguarda i ricoveri ordinari, osservandone la dinamica temporale, viene all'incirca confermata la tendenza generale; invece, per i rico-

Tabella 3. *Distribuzione percentuale dei ricoveri secondo le diagnosi ad essi connesse, per regime. Anni 2009 – 2014.*

Anno	Diagnosi connesse	Regime di ricovero		Totale ricoveri
		Ordinario	Day Hospital	
2009	Cardiocircolatorie	32,8	19,2	30,2
	Altre diagnosi	67,2	80,8	69,8
	Totale ricoveri	634.800	144.252	779.052
2010	Cardiocircolatorie	32,8	18,4	28,9
	Altre diagnosi	67,2	81,6	71,1
	Totale ricoveri	627.113	233.936	861.049
2011	Cardiocircolatorie	32,8	17,4	28,7
	Altre diagnosi	67,2	82,6	71,3
	Totale ricoveri	588.617	212.535	801.152
2012	Cardiocircolatorie	33,3	14,0	29,0
	Altre diagnosi	66,7	86,0	71,0
	Totale ricoveri	551.435	159.475	710.910
2013	Cardiocircolatorie	33,9	12,8	29,7
	Altre diagnosi	66,1	87,2	70,3
	Totale ricoveri	537.072	133.202	670.274
2014	Cardiocircolatorie	34,0	12,8	30,1
	Altre diagnosi	66,0	87,2	69,9
	Totale ricoveri	517.306	118.920	636.226
Totale	Cardiocircolatorie	33,2	16,2	29,4
	Altre diagnosi	66,8	83,8	70,6
	Totale ricoveri	3.456.343	1.002.320	4.458.663
Totale %		100,0	100,0	100,0

Fonte: Elaborazioni proprie su dati AReS.

veri in Day Hospital, nel corso dei sei anni considerati, le percentuali si scostano, seppur in maniera poco rilevante, dall'attitudine generale: negli anni 2009 e 2010 tale percentuale risulta pari a circa l'81% dei casi, nel 2011 cresce di circa un punto percentuale e la distribuzione continua ad esibire un andamento crescente anche negli anni successivi (l'86% nel 2012 e circa l'87% nel 2013 e 2014).

3. Analisi esplorativa: la distribuzione della popolazione dei ricoveri per patologie cardiovascolari secondo il regime di ricovero.

Appare opportuna ed utile una prima analisi esplorativa del nostro dataset, abbiamo ritenuto interessante conoscere la distribuzione percentuale dei ricoveri secondo il regime di ricovero (Ordinario e Day Hospital) distinto in base alle principali caratteristiche socio-demografiche, territoriali e cliniche del nostro campione nell'arco temporale considerato (2009/2014).

È interessante osservare come la percentuale dei ricoveri in Day Hospital, che ricordiamo sul totale del campione è pari a circa il 12%, cresca per le classi di età più basse della nostra distribuzione toccando una quota pari a circa il 27% dei casi per i pazienti di età inferiore a 44 anni e circa il 26% per la fascia di età 45-56 anni (Tabella 4). All'aumentare dell'età, pur aumentando il numero dei ricoveri, le percentuali di quelli diurni cala drasticamente a circa il 18% per i pazienti di età compresa tra i 57 e i 67 anni e poi tende a quasi dimezzarsi ad ogni fascia di età sempre più alta. In estrema sintesi, si osserva un corretto legame, inversamente proporzionale, tra le percentuali dei ricoveri in regime Day Hospital e l'età dei pazienti: quanto più il paziente invecchia tanto meno viene trattato in regime Day Hospital.

Per quel che riguarda la distribuzione del fenomeno per genere, si registra una sostanziale parità: le pazienti vengono ricoverate in regime diurno in circa il 13% dei casi mentre gli uomini in circa il 12% dei casi.

Riguardo la provincia di residenza, la percentuale dei ricoveri diurni dei pazienti baresi è di circa 5 punti al di sopra della media generale mentre per i foggiani tale valore appare grosso modo in linea con la tendenza generale.

Analoga tendenza viene rilevata, logicamente (data la quasi completa sovrapposibilità delle due variabili) per la ASL di provenienza dei ricoverati.

Quasi altrettanto sovrapponibile risulta la distribuzione dei ricoveri secondo la ASL della struttura di ricovero: in effetti, la mobilità dei ricoveri tra le ASL pugliesi risulta esistere, ma vi è una sorta di compensazione fra essi. Nella Tab. 4 si legge tuttavia una maggiore attrattività per le ASL di Bari e Foggia, e in parte Lecce, a discapito delle ASL di Brindisi, BAT e Taranto.

Tabella 4. Distribuzione percentuale dei ricoveri per patologia cardiovascolare in Puglia dal 2009 al 2014 secondo il regime di ricovero, per varie caratteristiche dei ricoverati e delle strutture di ricovero e cura.

Caratteristiche	Regime di ricovero		Numero di ricoveri
	Ordinario	Day Hospital	
	%	%	
Classi di età (espresse in anni)			
<44	72,5	27,5	86.950
45-56	74,2	25,8	143.089
57-67	82,4	17,6	271.980
68-77	90,7	9,3	375.294
78-83	94,5	5,5	238.802
>84	97,3	2,7	194.601
Genere			
Maschio	88,1	11,9	705.479
Femmina	87,1	12,9	605.237
Provincia di residenza			
Brindisi	90,0	10,0	124.985
Taranto	90,4	9,6	174.617
BAT	90,7	9,3	102.919
Bari	82,7	17,3	376.610
Foggia	87,3	12,7	251.847
Lecce	90,3	9,7	235.215
Fuori Puglia	91,9	8,1	40.513
Esteri	95,1	4,9	4.010
ASL di provenienza			
Brindisi	90,0	10,0	125.316
Taranto	90,4	9,6	174.785
BAT	90,7	9,3	103.004
Bari	82,7	17,3	376.645
Foggia	87,3	12,7	251.860
Lecce	90,3	9,7	235.285
Fuori regione	92,2	7,8	43.821
ASL di ricovero			
Brindisi	89,5	10,5	106.439
Taranto	91,3	8,7	156.443
BAT	92,3	7,7	82.308
Bari	83,5	16,5	441.844
Foggia	87,3	12,7	277.820
Lecce	90,7	9,3	245.862
Tipo di istituto			
Pubblico	86,5	13,5	890.203
Privato	89,3	10,7	238.565
Ecclesiastico	90,9	9,1	181.948
Totale ricoveri	87,6	12,4	1.310.716

Fonte: Elaborazioni proprie su dati AReS.

Tabella 5 *Distribuzione percentuale dei ricoveri negli ospedali pugliesi per tipologia di ricovero, per ogni CCS relativo a patologie cardiovascolari nelle diagnosi.*

CCS	Regime di ricovero		Numero ricoveri
	Ordinario	Day Hospital	
	%	%	
Malattie delle valvole cardiache	92,7	7,3	111.383
Peri-, endo- e miocarditi, cardiomiopatie (escluse quelle da tubercolosi o da MST)	94,2	5,8	53.481
Ipertensione essenziale	80,0	20,0	432.490
Ipertensione con complicaz. e ipertensione secondaria	92,5	7,5	311.830
Infarto miocardico acuto	99,4	0,6	46.533
Aterosclerosi coronarica e altre mal. ischemiche card.	95,7	4,3	264.614
Dolore toracico non specifico	98,7	1,3	34.178
Malattia cardiopolmonare	98,4	1,6	26.844
Altre e mal definite cardiopatie	90,8	9,2	32.245
Disturbi della conduzione	96,4	3,6	62.020
Aritmie cardiache	97,5	2,5	240.174
Arresto cardiaco e fibrillazione ventricolare	99,8	0,2	19.736
Insufficienza cardiaca congestizia, non ipertensiva	98,1	1,9	133.524
Vasculopatie cerebrali acute	98,6	1,4	57.856
Occlusione o stenosi delle arterie precerebrali	87,9	12,1	53.423
Altre e mal definite vasculopatie cerebrali	96,9	3,1	76.446
Ischemia cerebrale transitoria	99,1	0,9	20.477
Postumi delle malattie cerebrovascolari	90,9	9,1	24.025
Aterosclerosi delle arterie periferiche e viscerali	92,8	7,2	45.619
Aneurismi aortici, delle arterie periferiche e viscerali	93,9	6,1	25.868
Embolia o trombosi dell'aorta e delle arterie perifer.	98,3	1,7	5.403
Altre malattie circolatorie	84,7	15,3	47.817
Flebite, tromboflebite e tromboembolia	96,4	3,6	15.633
Varici degli arti inferiori	66,3	33,7	24.299
Altre malattie delle vene e dei vasi linfatici	79,5	20,5	23.296

Fonte: *Elaborazioni proprie su dati AReS.*

Gli Istituti di ricovero e cura sono istituti che coprono larga parte dei settori di attività del Servizio Sanitario Nazionale e si dividono in pubblici, privati ed ecclesiastici. I primi sono enti pubblici veri e propri, e si caratterizzano per la maggiore ingerenza dello Stato (nella veste del Ministero della Salute) sull'andamento della loro gestione (nomina in caso di necessità del commissario straordinario, controllo delle deliberazioni di spesa ed eventuale congruità con il finanziamento erogato, controllo del bilancio ecc.). Gli istituti privati, invece, hanno una maggiore libertà di azione, mentre quelli ecclesiastici sono controllati da ordini religiosi. Classifi-

cando gli Istituti per tipologia, il protocollo di Day Hospital risulta essere una scelta percentualmente più frequente nelle strutture pubbliche, seguite da quelle privati ed infine dalle strutture configurate come Enti Ecclesiastici.

Per un'analisi più puntuale dal punto di vista clinico possiamo osservare (Tabella 5) la distribuzione percentuale delle registrazioni secondo il regime di ricovero, per ciascun CCS appartenenti alla categoria nosologica "malattie del sistema circolatorio". Dalla tabella emerge che, in ordine, la patologia che più di tutte viene trattata in regime day hospital è quella delle *Varici degli arti inferiori* (quasi il 34% dei casi), seguita da *ipertensione essenziale e altre malattie delle vene e dei vasi linfatici* (circa il 20% dei ricoveri in entrambi i casi). Logicamente, le quote di ricovero in Day Hospital per le diagnosi più gravi sono irrisorie, e quasi nulle per *Arresto cardiaco e fibrillazione ventricolare* (ma anche *Infarto miocardico acuto e Ischemia cerebrale transitoria* fanno rilevare quote inferiori all'1% dei ricoveri).

4 Scelte metodologiche e risultati

In un ambito in cui occorre disporre e analizzare grandi e complesse masse di dati, oltretutto riferite a interi collettivi, allo scopo di ottenere informazioni utili e soprattutto interpretabili è essenziale utilizzare strumenti informatici e statistici appropriati. In questo campo, il confine tra Informatica e Statistica diventa labile, e spesso si è incerti se parlare di Informatica statistica (elaborazione degli algoritmi informatici adeguati ad implementare i metodi statistici) o di Statistica computazionale (utilizzo di strumenti informatici per l'analisi statistica).

Come recentemente illustrato da Paula Brito (Università di Porto, Portugal), relattrice invitata alla Conferenza Internazionale DSSR (Data Science & Social Research), tenutasi a Napoli il 17 febbraio 2016, per analizzare appropriatamente grandi basi di dati ("Big Data") occorre una disciplina al tempo stesso nuova e vecchia, la Scienza dei dati ("Data Science"), e che rappresenta "l'inevitabile riunione del pensiero informatico e statistico nell'era dei Big Data", utilizzando metodi che provengono da differenti campi come Statistica, Informatica, Teoria dell'informazione, Machine Learning, Pattern Recognition.

Lo scopo dichiarato di questo studio è identificare un algoritmo in grado di evincere informazioni e relazioni relative all'appropriatezza organizzativa dei ricoveri ospedalieri. Precedenti studi esplorativi (d'Ovidio, Mancarella e Toma, 2015) hanno sottolineato l'efficienza sia dell'analisi di segmentazione che, ancor più, delle reti neurali, ma resta il problema della completezza delle informazioni ricavate dai due metodi: il primo, per sua natura, cerca di identificare relazioni parziali, gerar-

chizzate, tra variabili antecedenti e variabile obiettivo, ottimizzando la capacità predittiva dell'analisi e la parsimonia delle combinazioni di relazioni (escludendo dunque tutte quelle variabili antecedenti che non apportano miglioramenti sufficienti alla capacità predittiva), ma in tale processo si perdono spesso le interazioni non gerarchiche tra le variabili. L'analisi delle reti neurali, per suo conto, presenta elevata capacità predittiva ma spesso anche un livello di complicazione altrettanto elevato, non disponendo di strumenti adeguati alla ricerca della parsimonia statistica.

In questo lavoro viene applicata l'analisi di segmentazione.

4.1 Problematiche legate alla struttura della popolazione e campionamento

La prima scelta metodologica scaturisce proprio dai risultati emersi dalla prima analisi esplorativa della popolazione osservata. Come già accennato, infatti, i due stati della variabile risposta non sono bilanciati nella popolazione. È nota la difficoltà intrinseca in ogni metodo di analisi relazionale quando il peso delle relazioni è più rappresentativo per uno status anziché l'altro⁵. Dunque, si è reso necessario un ricampionamento della popolazione al fine di rendere più semplice per gli algoritmi il riconoscimento dell'azione delle variabili predittive per ambo le modalità di ricovero.

È stato dunque effettuato un campionamento stratificato uniforme, dunque con probabilità diseguali per quel che riguarda le due modalità di ricovero, ma proporzionali al totale dei ricoveri per anno, ottenendo un campione sufficientemente bilanciato tra i due casi da consentire previsioni ottimali. Abbiamo scelto questo campionamento per la necessità di mantenere una proporzione della dinamica totale dei ricoveri nel tempo, pur costruendo in ogni anno campioni numericamente uguali per ciascun tipo di ricovero.

La procedura di campionamento è stata abbastanza complessa concettualmente, fin dalla determinazione dell'oggetto della procedura stessa: infatti, non solo era necessario estrarre dalla cospicua popolazione di ricoveri oggetto della ricerca (oltre un milione e 300mila) un campione che avesse un numero circa pari di ricoveri ordinari e di ricoveri diurni, ma la procedura di analisi scelta (analisi di segmentazione dei campioni) richiedeva di essere definita sulla base di subcampioni di *training* e di *test* estratti casualmente dal campione bilanciato, in ragione, rispettivamente, del 60% e del 40% (mantenendo circa uguali all'interno di detti subcampioni le quote di ricoveri ordinari e diurni).

Si è dunque stabilito che il "campione base" su cui andava calibrata l'intera procedura dovesse essere il subcampione di training, e precisamente la quota di esso

⁵ Cfr. ad es., Fabbris, 1997; King & Zeng, 2001

da estrarre dai ricoveri ordinari (che essendo nella popolazione considerata numericamente ben più consistenti, ossia 1.148.521, richiedono, a parità di precisione delle stime, campioni più numerosi).

Per la determinazione della dimensione ottimale di ambo i campioni, data la natura della variabile di interesse (la tipologia di ricovero) si è fatto riferimento alla metodologia per il calcolo di n nel caso di inferenza su frequenze relative (cfr., ad es., Delvecchio 2015), che per popolazioni molto grandi, come nel nostro caso, si riduce alla formula:

$$n = \frac{p(1-p) \cdot z_{\alpha/2}^2}{(d/2)^2} + 1 .$$

Detta formula evidenzia che per il calcolo di n occorre:

- prefissare il livello di significatività del test z , ossia α ;
- prefissare $d/2$, cioè lo scostamento massimo della frequenza relativa nel campione (f) da quella nella popolazione (p): insomma, la precisione della stima;
- calcolare (o stimare) la varianza della popolazione, ossia $p(1 - p)$.

Normalmente non è possibile conoscere il valore di p (e dunque della varianza) prima di aver estratto il campione, per cui si massimizza ponendo $f=0,5$. Nel nostro caso, invece i dati sono noti (il campione viene infatti estratto solo per motivi computazionali), e dunque sarà $p(1 - p) = 0,876 \cdot 0,124 = 0,0109$. La semiampiezza $d/2$ è stata fissata a livelli minimi (per garantire la massima precisione delle stime), ossia a 0,005, mentre α è stato posto pari a 0,05.

La formula descritta in precedenza fornisce, sulla base di tali dati, una numerosità minima pari a 37.173 ricoveri ordinari; da tale ammontare, che deve rappresentare il 60% del totale dei ricoveri ordinari da estrarre, è immediato ottenere la numerosità del campione di ricoveri ordinari da estrarre, ossia $n=37.173/0,6=61.955$. Tale ammontare è pari al 5,4% della numerosità della popolazione dei ricoveri ordinari. Dato che esso è, comunque, il valore MINIMO per ottenere la precisione di stima richiesta in caso di varianza massima, si è ritenuto opportuno estrarre dalla popolazione dei ricoveri ordinari un campione casuale pari approssimativamente al 5,6% (la procedura di estrazione casuale presente in SPSS, infatti, garantisce percentuali solo approssimate), che a posteriori si è rivelato essere pari a circa 64.460 casi. È stato dunque necessario estrarre approssimativamente lo stesso numero di casi anche dalla popolazione dei 162.195 ricoveri in regime diurno (di cui tale numero rappresenta circa il 39,66%). In definitiva, il campione bilanciato ammonta a 128.925 ricoveri, equidistribuiti in modo quasi perfetto tra modalità ordinaria e modalità diurna, come mostra la tabella 6.

Tabella 6 *Distribuzione per anno del campione estratto, secondo il regime di ricovero.*

Anno	Regime Ricovero		Totale
	Ordinario	Day Hospital	
	Numero ricoveri	Numero ricoveri	Numero ricoveri
2009	11.590	11.590	23.180
2010	12.225	12.225	24.450
2011	11.309	11.310	22.619
2012	10.138	10.138	20.276
2013	9.793	9.793	19.586
2014	9.407	9.407	18.814
Totale	64.462	64.463	128.925

Fonte: elaborazioni proprie su dati Ares

4.2 Analisi di Classificazione

Come ampiamente spiegato, l'obiettivo di questo lavoro è valutare l'appropriatezza organizzativa all'interno delle strutture sanitarie, la cui mancanza interferisce con l'appropriatezza diagnostica e clinica per il regime di ricovero ospedaliero. Per questo motivo nell'analisi di classificazione compiuta si è concentrata l'attenzione su quelli eventi (caratteristiche dei pazienti o della struttura di ricovero) che influenzano in maniera significativa la scelta delle modalità di ricovero (la quale dovrebbe dipendere esclusivamente da motivi clinici e diagnostici).

È stato possibile procedere con la costruzione del modello di segmentazione, in quanto abbiamo verificato che tutte le condizioni di applicabilità di tale modello.

Appare pienamente soddisfatta la condizione che la popolazione esibisca una elevata numerosità; è stato superato, come già specificato, il limite che la distribuzione della variabile risposta risultasse fortemente asimmetrica; la scelta delle variabili esplicative è stata dettata dagli obiettivi di questo lavoro di individuare le relazioni tra la scelta del regime di ricovero e le caratteristiche di tipo clinico o diagnostico e altri fattori come alcune caratteristiche socio demografiche dei pazienti, caratteristiche territoriali o legate alla tipologia dei differenti istituti di ricovero.

Come set di training, dal nuovo dataset bilanciato è stato estratto un campione casuale semplice pari a circa il 60% dello stesso (per cui, ovviamente, il set di verifica ammonta al restante 40% del dataset).

È stata quindi applicata una serie di analisi di segmentazione al nuovo dataset bilanciato, utilizzando la modalità di ospedalizzazione come variabile risposta e, come predittive, tutte le altre informazioni disponibili; sono stati altresì sperimentati gli algoritmi CHAID, EXHAUSTIVE CHAID, QUEST e CRT (anche noto come CaRT).

Per ogni algoritmo è stato scelto un livello di profondità differente per ottimizzare i risultati a secondo delle differenti caratteristiche peculiari di ognuno di essi; infine, abbiamo scelto di uniformare il numero minimo di casi nei nodi genitori e nei nodi figli (rispettivamente, 100 e 50 unità).

Infine la scelta è ricaduta sul modello fornito dall'algoritmo CRT⁶, nonostante si tratti di una tecnica molto dispersiva in termini di risultati (dato l'algoritmo binario di suddivisione dei nodi), perché fornisce "predizioni" migliori del QUEST e comparabili con quelle di CHAID (queste ultime leggermente migliori nella fase esplorativa), pur se ad un maggior costo computazionale rispetto al primo, ma in tempi incomparabilmente più brevi degli altri due metodi.

In definitiva, per l'algoritmo CRT è stato utilizzato un criterio di arresto di tipo esplorativo (profondità massima=9 livelli) con opzione di *pruning* dell'albero, ossia lo sfoltimento "a posteriori" delle suddivisioni meno rilevanti rispetto alla complessiva capacità predittiva, allo scopo di ridurre sia la dimensione dell'albero e sia la sua complessità, che si potrebbe risolvere in una scarsa possibilità di generalizzazione dei risultati.

In questo modello sono stati identificati 93 nodi, di cui 47 non ulteriormente divisibili. Il risultato dell'analisi è senz'altro interessante, posto che il modello identificato raggiunge delle buone percentuali di corretta classificazione pari a ben oltre l'80% dei casi, sia nel campione di training che in quello di test (tabella 7).

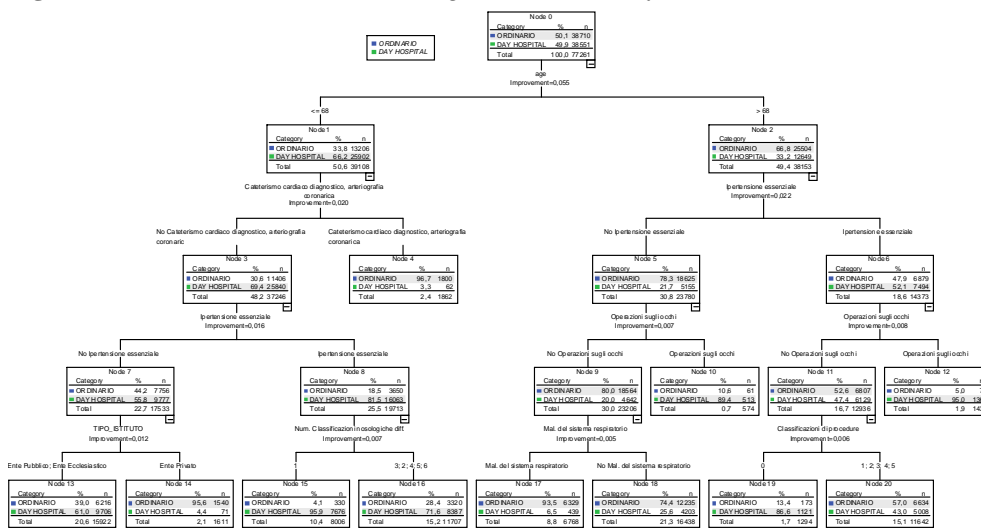
Inoltre il modello si ritiene di discreta utilità in quanto evidenzia una soddisfacente "ripartizione" nella classificazione dei casi osservati per le due modalità rispetto a quelli previsti. I ricoveri ordinari infatti sono stati classificati correttamente in circa il 79% dei casi, e quelli diurni in una percentuale di circa l'84% dei casi.

Tabella 7 Campione di Test e di Training: matrice di confusione derivante da analisi di classificazione dei ricoveri (campione bilanciato).

Campione	Ricoveri osservati	Ricoveri previsti		Totale osservato	% previsioni corrette
		Ordinario	Day Hospital		
Training	Ordinario	30.762	7.948	38.710	79,47
	Day Hospital	6.082	32.469	38551	84,22
	Totale previsto	36.844	40417	77.261	
	% complessiva	47,69	52,31		81,84
Test	Ordinario	20309	5443	25.752	78,86
	Day Hospital	4289	21623	25.912	83,45
	Totale previsto	24.598	27.066	51.664	
	% complessiva	47,61	52,39		81,16

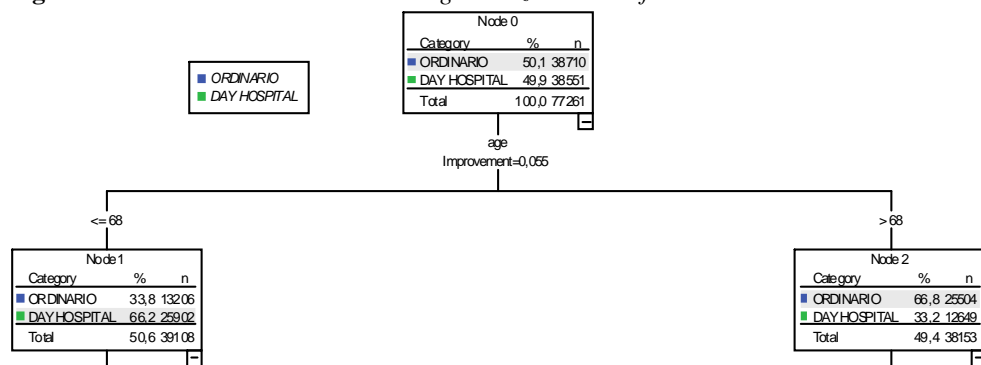
⁶ Per economia editoriale non è possibile riportare i risultati degli altri modelli testati.

Figura 1. *Primi 4 livelli dell'albero di segmentazione identificato.*



Il grafico dell'albero di classificazione, è difficile da rappresentare nella sua complessità, e dunque per ragioni di praticità e chiarezza si è scelto di riportarne qui solo i primi quattro livelli (Fig. 1). Tuttavia, persino questo sottoinsieme dell'albero di segmentazione risulta di difficile osservazione nel dettaglio. Per tal motivo, nel seguito esso viene descritto in riferimento ai singoli rami.

Figura 2. *Primo livello dell'albero di segmentazione identificato.*



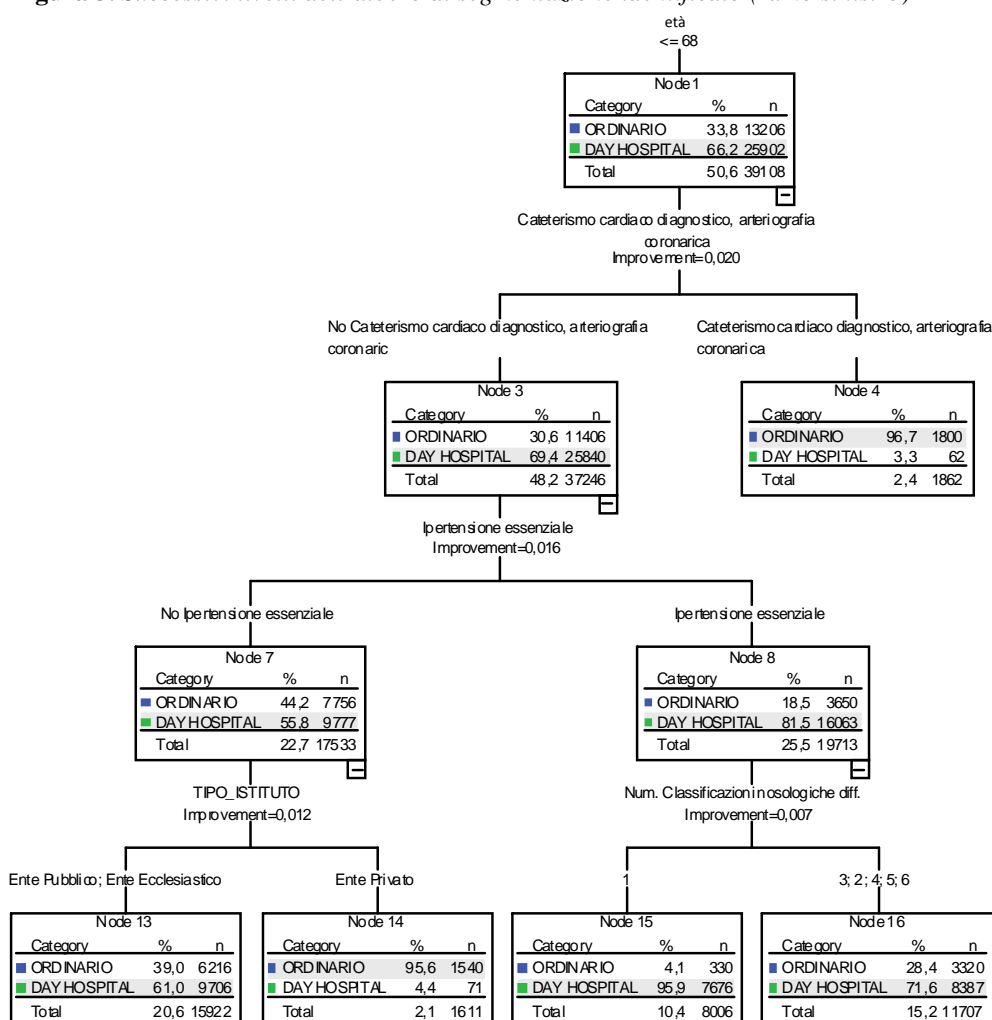
Innanzitutto, va sottolineato che nel primo livello la variabile esplicativa più significativamente legata alla variabile risposta “regime di ricovero” è l’età (Fig. 2). Il nostro campione si divide quasi perfettamente a metà tra il nodo uno (ricoveri di pazienti di età inferiore o uguale a 68 anni) e il nodo due (ricoveri di pazienti di età superiore a 68 anni). Nel primo caso viene preferito il regime diurno in circa il 66% dei casi, nel secondo si sceglie il ricovero ordinario in circa il 67% dei casi; si ri-

cordi, peraltro, che una relazione inversamente proporzionale tra età e scelta del regime DH è stata identificata già nella prima analisi descrittiva.

Nel secondo livello dell'albero intervengono variabili di tipo clinico e diagnostico. Il nodo dei pazienti meno anziani (Fig. 3) si divide in: soggetti che non sono stati sottoposti alla procedura cateterismo cardiaco diagnostico, arteriografia coronarica (nodo 3, circa il 48%), per i quali viene scelto il regime Day hospital in circa il 70% dei casi, e soggetti che invece sono stati sottoposti a tale procedura (nodo 4, circa il 2%), per i quali appare quasi assoluta la scelta del ricovero ordinario.

Nel terzo livello dell'albero di segmentazione interviene un altro importante fattore clinico-diagnostico: l'ipertensione essenziale.

Figura 3. Successivi livelli dell'albero di segmentazione identificato (ramo sinistro)



Dai nodi 7 e 8 si evince però un meccanismo particolare, che sarà chiarito nel seguito: questi soggetti meno anziani (e *non* sottoposti a cateterismo cardiaco arteriografia coronarica) se non ipertesi vengono trattati in regime diurno in circa il 55% dei casi, ma se ipertesi l'opzione Day Hospital sale a oltre l'81%.

Proseguendo nell'analisi di questo ramo dell'albero, nel quarto livello, oltre ai fattori di tipo diagnostico e clinico subentra il fattore "tipo Istituto", ossia la *forma giuridica della struttura di ricovero*. I nodi 13 e 14 evidenziano che, in assenza di ipertensione essenziale, gli istituti pubblici ed ecclesiastici utilizzano il regime diurno nel 61% dei casi mentre quelli privati in solo poco più del 4% dei casi.

Dai nodi 15 e 16 si evince che, in presenza di ipertensione essenziale, i medesimi vengono trattati in regime diurno in percentuale sensibilmente minore quando vengono diagnosticate anche altre patologie non appartenenti alla classificazione nosologica cardiovascolare (i ricoveri con una sola classificazione nosologica vengono avviati al DH in circa 96% dei casi; con più di una, in circa 72% dei casi).

Spostando l'attenzione sul lato destro dell'albero (Fig. 4), si nota che la quota dei ricoveri dei soggetti di età maggiore di 68 anni si suddivide a sua volta, in percentuali pari a circa il 30% e circa il 20% del campione complessivo, nei nodi 5 e 6, rispettivamente ricoveri di pazienti non affetti da ipertensione essenziale e soggetti ipertesi.

I ricoveri dei pazienti non ipertesi sono di tipo ordinario in circa il 78% dei casi, ma l'ipertensione non sembra, di primo acchito, discriminante nella scelta del regime di ricovero: i soggetti ipertesi vengono infatti trattati all'incirca indifferentemente in regime ordinario o diurno, anche se con una leggera preferenza per la seconda modalità (Ord. 48% vs DH 52%).

Subito di seguito, possiamo evincere dai nodi 9 e 10 (anziani non ipertesi) che per coloro che subiscono un intervento agli occhi viene preferito il ricovero ordinario nell'80% dei casi (nodo 9), mentre una tendenza opposta viene riscontrata per gli stessi soggetti che però hanno subito interventi agli occhi, per i quali viene scelto il ricovero diurno in circa il 90% dei casi (nodo 10).

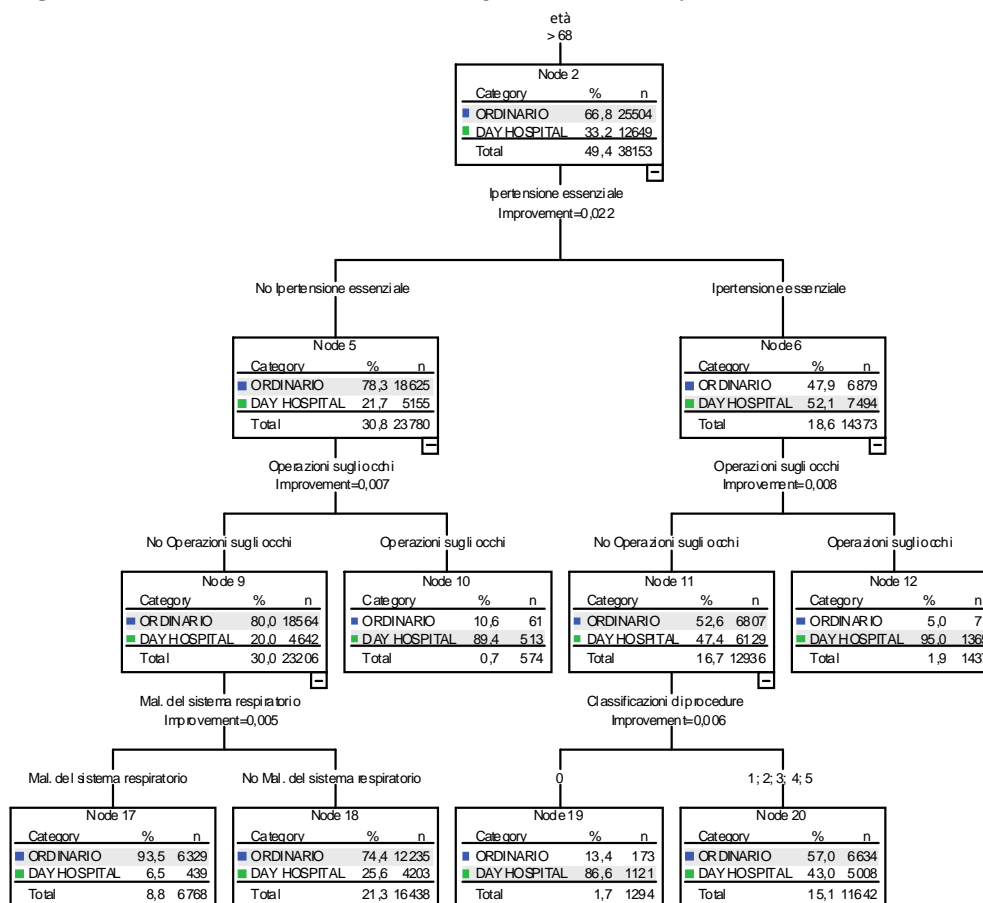
Proseguendo nell'osservazione del ramo, nei nodi 11 e 12 vengono classificati i ricoveri degli anziani ipertesi che, rispettivamente, non hanno subito un intervento agli occhi (i quali vengono trattati in quasi egual misura nei due diversi regimi di ricovero, con leggera prevalenza della tipologia ordinaria, circa 53% contro circa 47%), e quelli che hanno invece subito tale intervento (nodo 12), i quali sono ricoverati in regime diurno in ben il 95% dei casi.

Al quarto livello di questo ramo di segmentazione, nei nodi 17 e 18 vengono classificati i ricoveri dei soggetti più anziani non ipertesi (e che non hanno subito opera-

zioni agli occhi). Questi soggetti vengono trattati prevalentemente in regime ordinario quando sono affetti da malattie del sistema respiratorio in circa il 93% dei casi (nodo 17), mentre in circa il 74% dei casi quando non lo sono (nodo 18).

Per i pazienti più anziani ipertesi (ancora senza operazioni agli occhi), viene scelto il ricovero diurno in circa 87% dei casi quando non si rilevano classificazioni di procedure o interventi (nodo 19) e in circa il 43% dei casi quando vengono riscontrate una o più classificazioni di procedure (nodo 20).

Figura 4. Successivi livelli dell'albero di segmentazione identificato (ramo destro)



Dunque, entro il terzo/quarto livello di segmentazione una diagnosi appartenente al CCS di ipertensione essenziale sembra avere un effetto “protettivo” contro l’opzione di ricovero ordinario, anche in presenza di interventi (relativi o meno al sistema cardiovascolare). Questa apparente incongruenza è dovuta, come viene chiarito osservando i livelli successivi, alla constatazione che chi soffre di una pa-

tologia cronica è tenuto sotto controllo medico in misura maggiore di soggetti apparentemente sani, e quindi sono più facilmente gestibili eventuali stati acuti della malattia (o l'insorgere di altre patologie acute) che normalmente richiedono un immediato ricovero ordinario.

I successivi livelli dell'albero di classificazione cercano di rendere sempre più precisa l'attribuzione dei casi all'una o all'altra modalità di ricovero, ma la loro descrizione si fa sempre più complessa, e appare meno utile rispetto ai primi 4 livelli (anche a causa della sempre minore importanza e capacità discriminante dei relativi nodi). Ai fini della definizione delle regole di decisione, che in seguito servono per identificare le situazioni "presumibilmente anomale", la descrizione puntuale di tali livelli di classificazione risulta pressoché superflua. Basti qui sapere che la variabile "tipo di Istituto" ricorre in più nodi, e sempre connotando una più frequente opzione per il regime Day Hospital nelle strutture pubbliche, una quasi assoluta preferenza per il ricovero ordinario nelle strutture private e una situazione intermedia nelle strutture sanitarie rette da enti ecclesiastici.

Di seguito (Tab. 8) è riportato, in ordine di importanza, l'elenco delle variabili indipendenti valutate nel modello (anche quando poi escluse in fase di *pruning*).

Tabella 8. *Importanza normalizzata delle variabili indipendenti nell'analisi effettuata*

Variabili indipendenti	Importanza normalizzata
<i>Ipertensione essenziale</i>	100,00
<i>Malattie del sistema respiratorio</i>	85,34
<i>Età</i>	61,33
<i>Operazioni sugli occhi</i>	53,14
<i>Cateterismo cardiaco diagnostico, arteriografia coronarica</i>	52,34
<i>Aritmie cardiache</i>	52,03
<i>Numero di classificazioni nosologiche diverse</i>	46,49
<i>Numero di differenti CCS di interventi e procedure</i>	37,38
<i>Tipo Istituto di ricovero</i>	36,00
<i>Aterosclerosi coronarica e altre malattie ischemiche cardiache</i>	31,88
<i>Lesioni e avvelenamenti</i>	29,84
<i>Numero di differenti Classificazioni di interventi e procedure</i>	28,34
<i>Insufficienza cardiaca congestizia, non da ipertensione</i>	25,55
<i>Altre procedure terapeutiche cardiovascolari non di sala operatoria</i>	25,18
<i>Infarto miocardico acuto</i>	22,60
<i>Procedure diagnostiche e terapeutiche varie</i>	18,61
<i>Vasculopatie cerebrali acute</i>	17,30
<i>ASL di ricovero</i>	15,84
<i>Provincia di residenza</i>	15,83

Variabili indipendenti	Importanza normalizzata
<i>Ipertensione con complicazioni e ipertensione secondaria</i>	14,39
<i>Operazioni sull'apparato tegumentario</i>	14,19
<i>Operazioni su naso, bocca e faringe</i>	13,23
<i>Malattie del sistema genito-urinario</i>	12,33
<i>Inserzione, revisione, sostituzione, rimozione di pacemaker cardiaco o defibrillatore automatico</i>	12,32
<i>Provincia di nascita</i>	11,98
<i>Operazioni sul sistema endocrino</i>	11,78
<i>Malattie del sistema nervoso e degli organi di senso</i>	10,99
<i>Circolazione extracorporea ausiliaria per procedure a cuore aperto</i>	10,45
<i>Altre e mal definite vasculopatie cerebrali</i>	9,54
<i>Varici degli arti inferiori</i>	9,04
<i>Stripping di vene varicose dell'arto inferiore</i>	9,01
<i>Operazioni sull'apparato gastro-intestinale</i>	8,86
<i>Operazioni sul sistema nervoso</i>	8,81
<i>Operazioni sul sistema sull'apparato cardiovascolare</i>	8,80
<i>Disturbi della conduzione</i>	8,52
<i>Operazioni sull'apparato respiratorio</i>	8,40
<i>Dolore toracico non specifico</i>	8,07
<i>Reparto di ricovero</i>	7,59
<i>Malattia mentale</i>	7,38
<i>Malattia cardiopolmonare</i>	6,85
<i>Bypass dell'arteria coronarica (cabg)</i>	6,28
<i>Procedure sulle valvole cardiache</i>	6,21
<i>Operazioni sull'apparato urinario</i>	5,73
<i>Arresto cardiaco e fibrillazione ventricolare</i>	5,71
<i>Ischemia cerebrale transitoria</i>	4,93
<i>Istituto di ricovero</i>	4,36
<i>Altre procedure di sala operatoria sul cuore</i>	4,30
<i>Malattie delle valvole cardiache</i>	4,26
<i>Malattie del sistema muscoloscheletrico e del tessuto connettivo</i>	4,09
<i>Altre malattie delle vene e dei vasi linfatici</i>	3,21
<i>Aterosclerosi delle arterie periferiche e viscerali</i>	2,84
<i>Mese di ricovero</i>	2,70
<i>Altre procedure di sala operatoria su vasi, esclusi capo e collo</i>	2,23
<i>Aneurismi aortici, delle arterie periferiche e viscerali</i>	2,20
<i>Malattie del sangue e degli organi ematopoietici</i>	2,13
<i>Endoarteriectomia, vasi del capo e del collo</i>	2,05
<i>Malattie del sistema digestivo</i>	2,03
<i>Embolia o trombosi dell'aorta e delle arterie periferiche</i>	1,97
<i>Codici residui e non classificati + codici E</i>	1,79

Variabili indipendenti	Importanza normalizzata
<i>Bypass vascolare periferico</i>	1,65
<i>Formazione, revisione e rimozione di fistola arteriovenosa o cannula intervasale per dialisi</i>	1,61
<i>Sintomi, segni, cond. Indefinite, fattori influenzanti lo stato di salute</i>	1,60
<i>Resezione dell'aorta, sostituzione o anastomosi</i>	1,57
<i>Flebite, tromboflebite e tromboembolia</i>	1,49
<i>Altro cateterismo vascolare, escluso cuore</i>	1,45
<i>Altre procedure di sala operatoria su vasi del capo e del collo</i>	1,42
<i>Peri-,endo-,miocarditi, cardiomiopatie (no da tubercolosi o da MST)</i>	1,29
<i>Operazioni sull'apparato muscolo-scheletrico</i>	1,15
<i>Embolectomia e endoarteriectomia su vasi degli arti inferiori</i>	1,14
<i>Mal. Endocrine-nutrizionali-metaboliche e disordini immun.</i>	0,95
<i>Condizioni originate nel periodo perinatale</i>	0,84
<i>Tipo di procedura</i>	0,65
<i>Postumi delle malattie cerebrovascolari</i>	0,64
<i>Interventi ostetrici</i>	0,53
<i>Operazioni sul sistema sanguigno e linfatico</i>	0,46
<i>Operazioni sugli organi genitali maschili</i>	0,44
<i>Altre procedure diagnostiche cardiovascolari</i>	0,43
<i>Altre e mal definite cardiopatie</i>	0,41
<i>Mal. Della pelle e del tessuto sottocutaneo</i>	0,39
<i>Neoplasie</i>	0,36
<i>Genere</i>	0,36
<i>Cittadinanza</i>	0,26
<i>Anomalie congenite</i>	0,26
<i>Malattie infettive e parassitarie</i>	0,21
<i>Complicanze della gravidanza del parto e del puerperio</i>	0,19
<i>Cittadinanza</i>	0,17
<i>Altre malattie circolatorie</i>	0,12
<i>Occlusione o stenosi delle arterie precerebrali</i>	0,11
<i>Operazioni sugli organi genitali femminili</i>	0,07
<i>Altri bypass e anastomosi vascolari, escluso cuore</i>	0,06
<i>Operazioni sull'orecchio</i>	0,05
<i>Trombolisi coronarica</i>	0,01

A sorpresa, la variabile più importante non è l'età dei ricoverati (pur essendo questa la prima variabile a intervenire nella segmentazione) ma il CCS *Iperensione essenziale*, che presenta la maggiore importanza contribuendo per oltre il 9% al miglioramento complessivo rispetto alla pura classificazione casuale (paritetica fra le due modalità); rilevante è anche il contributo delle *Malattie del sistema respira-*

torio, pari a quasi l'8% (oltre l'85% dell'effetto del fattore più importante), mentre l'Età apporta un miglioramento pari a circa il 6% (61% in termini relativi), le Operazioni sugli occhi, il Cateterismo cardiaco diagnostico, arteriografia coronarica e le Aritmie cardiache un miglioramento pari a circa il 5% (importanza relativa, 52%-53%), seguite dal Numero di Classificazioni nosologiche diverse (circa il 4%) dal Numero di differenti CCS di interventi e procedure e dal Tipo Istituto di ricovero (circa il 3% ciascuno).

Gli altri fattori prognostici (effettivi o potenziali) sono localizzati sotto la soglia di importanza del 3% (ossia a meno del 35% di importanza relativa).

Si vuole ora verificare se e quanto il modello di segmentazione descritto fino a questo punto, definito in base a un campione stratificato uniforme, sia valido per l'intera popolazione dalla quale detto campione è stato estratto.

Per effettuare tale *validazione finale* del modello le regole di classificazione estrapolate dall'analisi fin qui condotta sono state applicate all'intera popolazione dei ricoveri con diagnosi di tipo cardiocircolatorio in Puglia, assegnando una "pre-dizione" della modalità di ricovero in base alle caratteristiche (personali o clinico-diagnostiche) dei ricoverati e a quelle delle strutture ospitanti

Tabella 9. *Matrice di confusione della popolazione complessiva dei ricoveri, ottenuta applicando le regole di classificazione derivanti dall'analisi effettuata sul campione bilanciato.*

Ricoveri osservati	Ricoveri previsti		Totale osservato	% previsioni corrette
	Ordinario	Day Hospital		
Ordinario	906.830	241.691	1.148.521	79,0
Day Hospital	24.480	137.715	162.195	84,9
Totale previsto	931.310	379.406	1.310.716	
% complessiva	71,1	28,9		79,7

Dalla tabella 9 è possibile apprezzare immediatamente la bontà del modello identificato anche quando applicato alla totalità della popolazione, ma anche i suoi difetti. Esso infatti si discosta in misura ridottissima dai risultati campionari, in particolare per quanto riguarda la capacità di discriminare i ricoveri in regime diurno, in questo caso si registra addirittura un miglioramento, seppur lieve.

Questo risultato assume un importante significato per l'intero studio condotto fino a questo punto: scendendo più nello specifico si potrà ritenere di poter valutare l'appropriatezza organizzativa del sistema con una buona approssimazione dei ri-

sultati. Dunque, per il complesso delle diagnosi si registra, in media, una tipologia di ricovero attinente a quello previsto nel 79% dei casi di degenza ordinaria e in quasi l'85% dei ricoveri diurni effettuati.

Tuttavia, si assiste a una sovrastima complessiva dei ricoveri in regime Day Hospital (che risultano previsti in quasi il 29% dei casi, contro un valore effettivo del 12,4%), e ciò può essere dovuto a diversi motivi: tra cui il fatto che l'influenza del *Tipo di Istituto*, riconosciuta per alcuni nodi gerarchicamente più o meno rilevanti, potrebbe essere ben più diffusa.

5. Conclusioni

Obiettivo principale di questo lavoro, più volte dichiarato, è cercare di valutare con opportune tecniche statistiche multivariate la qualità dei servizi ospedalieri, con particolare attenzione all'appropriato utilizzo dello strumento "day hospital". Un utilizzo appropriato di tale strumento come alternativa al ricovero ordinario, quando possibile (ossia quando il quadro clinico del paziente lo consente), comporta infatti una notevole riduzione dei costi legati all'assistenza sanitaria.

In questa ottica, il presente lavoro rileva e analizza le abitudini di ricovero negli Istituti pubblici e privati pugliesi nel periodo 2009-2014.

La prima parte di questo lavoro è stata dedicata ad una breve definizione del concetto di efficienza sanitaria, alla descrizione del piano di ricerca, inoltre è stata effettuata una prima esplorazione della struttura del dataset. È stata analizzata la distribuzione dei ricoveri secondo il regime e per anno e da tale studio è emersa un'importante asimmetria della variabile regime di ricovero per le due sue modalità (ricovero ordinario 87,6% e Day hospital 12,4%).

Si è entrati nel dettaglio dell'analisi descrivendo la distribuzione percentuale dei ricoveri secondo il regime e per diagnosi connesse: è emersa la scelta predominante del regime di ricovero Ordinario sia per le diagnosi di tipo cardiocircolatorio che per le altre, rispettivamente circa l'88% e il 73% dei ricoveri totali; proseguendo è stata studiata la distribuzione percentuale dei ricoveri secondo le diagnosi ad essi connesse, per regime ed è stata riscontrata la presenza di un maggior numero di ricoveri sia ordinari che day hospital per le diagnosi classificate come non cardiocircolatorie, rispettivamente si registra circa il 67% e l'84% sul totale dei ricoveri.

L'ipotesi di base del nostro lavoro è che la scelta tra i diversi tipi di ricovero dipenda solo da motivi clinici e diagnostici. Se invece intervengono altri elementi bisogna studiare in che modo tali elementi influiscano sulle scelte, in modo da poter determinare delle politiche adeguate. Dunque, nella seconda parte di questo lavoro,

abbiamo arricchito la prima analisi descrittiva studiando la distribuzione percentuale dei ricoveri secondo il regime di ricovero per quelle che sono le principali caratteristiche socio demografiche, territoriali e cliniche del nostro campione nell'arco temporale considerato (2009/2014).

Si è riscontrato un legame inversamente proporzionale tra le percentuali dei ricoveri in regime Day Hospital e l'età dei pazienti inoltre, i pazienti nati nella provincia di Bari e ivi residenti sono risultati essere quelli maggiormente trattati in regime diurno (circa il 17% dei casi), stessa tendenza è stata riscontrata per la Asl.

In questa fase esplorativa, è stato osservato un legame tra la scelta del regime di ricovero e il Tipo di Istituto di ricovero: il Day hospital è utilizzato prevalentemente in quelli pubblici, seguono quelli privati ed in fine quelli ecclesiastici.

Sono emersi inoltre dei legami tra la scelta del regime di ricovero e alcune caratteristiche diagnostiche e cliniche dei pazienti che si possono ritenere appropriati: in generale, viene preferito il ricovero ordinario al complicarsi del quadro clinico diagnostico del paziente.

Dopo questa ampia e dettagliata analisi descrittiva si è cercato di perseguire il proposito di indagare l'adeguatezza organizzativa delle procedure di assistenza sanitaria e le determinanti dei tassi più bassi di utilizzo del regime di ricovero Day Hospital rispetto al tasso complessivo attraverso un'analisi di classificazione utilizzando differenti metodologie.

Dopo aver costruito un nuovo dataset bilanciato rispetto alle due modalità del regime di ricovero, è stata applicata un'analisi di segmentazione, utilizzando la modalità di ospedalizzazione come variabile risposta e tutte le altre informazioni disponibili come predittive.

I risultati dell'analisi sono apprezzabili, ammontando ad appena al 18% la stima di errata classificazione, ossia la percentuale dei casi erroneamente classificati nel campione di apprendimento. Il modello esibisce una soddisfacente "ripartizione" nella classificazione dei casi osservati per le due modalità rispetto a quelli previsti. I ricoveri ordinari sono stati classificati correttamente in circa il 79% dei casi e quelli diurni in una percentuale di circa l'84% dei casi, inoltre l'algoritmo raggiunge delle buone percentuali di corretta classificazione sia nel campione di training che in quello di test pari a oltre 80% dei casi.

In definitiva l'albero di segmentazione ha gerarchizzato le variabili predittive considerate classificando come più importanti alcuni fattori di tipo diagnostico e clinico (Ipertensione essenziale, Malattie del sistema respiratorio), l'età del paziente al momento del ricovero, alcuni interventi (Operazioni sugli occhi, Cateterismo cardiaco diagnostico, arteriografia coronarica), il numero di diagnosi differenti, il

tipo di Istituto di ricovero (Pubblico, privato ed ecclesiastico) e in ordine altri fattori di tipo diagnostico e clinico.

La validazione di tale modello sull'intera popolazione, avvenuta applicando le regole di classificazione estrapolate dall'analisi fin qui condotta, ha restituito risultati confortanti. I risultati ottenuti sulla totalità della popolazione si discostano in misura ridottissima da quelli campionari, in particolare per quanto riguarda la capacità di discriminare i ricoveri in regime diurno, in questo caso si registra addirittura un miglioramento, seppur lieve. Dunque, per il complesso delle diagnosi si registra, in media, una tipologia di ricovero attinente a quello previsto nel 79% dei casi di degenza ordinaria e in quasi l'85% dei ricoveri diurni effettuati.

Il proposito è di sviluppare questo studio utilizzando la regressione categoriale come confermativa dell'analisi di classificazione e come strumentale per identificare le variabili da inserire nelle successive analisi.

Dunque l'intenzione è sviluppare un'analisi multicriteri al fine di costruire un modello parsimonioso capace di perfezionare i risultati ottenuti fino a questo momento. Per cercare di identificare la soluzione più stabile e affidabile, saranno testati differenti algoritmi di reti neurali artificiali modificandone di volta in volta la struttura, e verificando tutte le possibili combinazioni di funzioni nascoste, funzioni di uscita, ottimizzazioni e diversi strati nascosti.

Questo studio potrebbe offrire interessanti spunti di riflessione per i dirigenti del settore sociale e sanitario, come "passo preliminare" per i loro interventi identificando i punti di debolezza del sistema e gestire proficuamente tutti i fattori di tipo non clinico o diagnostico che di fatto sembrerebbero discriminanti nella scelta del tipo di regime di ricovero.

Riferimenti bibliografici

- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J. (1984). *Classification and Regression Trees*, Chapman & Hall, New York-London.
- Brito P. (2016). *Statistical Computing for Data Science*, Invited Talk to the International Conference on Data Science & Social Research, Naples, Italy, February 2016.
- Cowen M.E., Dusseau D.J., Toth B.G. et al. (1998). Casemix adjustment of managed care claims data using the clinical classifications for health policy research method. *Medical Care*, 36: 1108-1113.
- Cutillo P. (1974). "La segmentazione", Atti della XXII Riunione scientifica della Società Italiana di Economia, Demografia e Statistica (Siena 30 nov.-2 dic.

- 1972), *Rivista Italiana di Economia, Demografia e Statistica*, vol, XXVIII, n. 4: 215-225.
- d'Ovidio F.D., Mancarella R., Toma E. (2015). *Assessing the Health Care Efficiency by using Classification Analyses*, Book of Abstracts, International Conference "IES 2015 - Statistical methods for service assessment" (University of Bari, 8-9 July), pag. 13.
- De Leeuw J. (1977) *Normalized cone regression*, Department of Data Theory, University of Leiden, The Netherlands.
- De Leeuw J., Young F.W., Takane Y. (1976) Additive Structure in Qualitative Data: an Alternative Least Squares Method with Optimal Scaling Features. *Psychometrika*, **41**: 471-504.
- Delvecchio F. (2015) *Statistica per l'analisi dei fenomeni sociali*, CLEUP, Padova.
- Elixhauser A, Steiner C.A. (1999). *Hospital inpatient statistics, 1996*. Healthcare Cost and Utilization Project (HCUP) Research Note. Rockville, MD: Agency for Health Care Policy and Research;. AHCPR Pub. No. 99-0034.
- Fabbris L. (1997). *Statistica multivariata. Analisi esplorativa dei dati*, McGraw-Hill, Milano.
- Gillo M.W. (1972) MAID, a Honeywell 600 program for an automatic survey analysis. *Behavioral Science*, **17**: 251-252.
- Gillo M.W., Shelley M.W. (1974) Predictive modelling of multivariable and multivariate data. *Journal of the American Statistical Association*, **69**: 646-653.
- Guttman L.A. (1941) The quantification of a class of attributes: A theory and method of scale construction.. In: P. Horst (ed.), *The prediction of personal adjustment*. Social Science Research Council, New York.
- Kass G. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, **29.2**: 119-127.
- King, G., Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis* **9** (2): 137-163.
- Morgan J. N., Messenger R. C. (1973) *THAID, a Sequential Analysis Program for the Analysis of Nominal Scale Dependent Variables*, Institute for Social Research, University of Michigan, Ann Arbor (Mich.).
- Sonquist J. A (1970) *Multivariate Model Building. The Validation of a Search Strategy*, Institute for Social Research, The University of Michigan, Ann Arbor (Mich.).
- Sonquist J. A., Morgan J.N. (1964) *The Detection of Interaction Effects*, Institute for Social Research, The University of Michigan, Ann Arbor (Mich.).

Total Factor Productivity Index

Domenico Summo*

Università degli Studi di Bari "A. Moro"

Riassunto: Il lavoro di ricerca ha riguardato l'implementazione di un approccio metodologico innovativo rispetto all'impostazione classica per la misura della produttività. Si tratta di porre in essere indici di produttività di Malmquist che, a differenza di quelli del tipo *Laspeyres*, *Paasche*, *Fisher*, *Tornquist*, *Solow*, considerano l'ipotesi di inefficienza delle strutture produttive a confronto e pertanto consentono di scindere i cambiamenti dovuti alla efficienza tecnica e ai cambiamenti dovuti al progresso tecnologico, noti, nella terminologia anglo-sassone, come *catching-up effect* e *tecnology schift effect*; questi indici, tuttavia, richiedono la definizione di una metodologia prodromica alla loro generazione: la *Data Envelopment Analysis (DEA)*.

Keywords: Total Factor Productivity Index, Data Envelopment Analysis.

1. Introduzione

In termini generici, la produttività¹ può essere intesa come l'espressione concettuale di una capacità o attitudine a produrre di una organizzazione produttiva, ovvero a

* Corresponding author: domenico.summo@uniba.it

¹ Cfr. (per es.) OWYONG (2000), pag. 19: "At its most basic level, productivity is based on the economics of the firm. It is measured as the ratio of output to input. Historically, productivity is often expressed as the ratio of output to the most limited or critical input, with all the other inputs held constant"; pag. 29: "... With the increasing recognition that productivity growth is the key to sustained economic expansion, measuring productivity is becoming important to economists and policy makers alike. The accurate measurement of productivity growth plays an important role in providing the information economists need to put forth better policy recommendations and for policy makers to make the right decisions". Lo studioso Peter Druker, nel 1974, affermava che "Senza obiettivi in tema di produttività un'impresa è senza direzione; senza un sistema di misurazione della produttività un'impresa è senza controllo"; cfr. DRUKER (1974).

conseguire un risultato di natura fondamentale diversa dai mezzi adoperati nel processo produttivo e generalmente di qualità in qualche modo superiore. La sua misura è generata dal confronto tra l'ammontare della produzione ottenuta dal processo produttivo e le risorse (o fattori produttivi) impiegate nello stesso ovvero da un rapporto tra output ed input. Per quanto possibile, i termini del rapporto dovrebbero esprimere quantità fisiche². Evidentemente, più che la produttività *sic et simpliciter*, interessa analizzare la dinamica temporale della stessa per modo che, con riferimento ad una struttura produttiva oggetto di analisi (impresa, gruppo, settore, sistema economico), il riscontro di differenti variazioni temporali tra input ed output possa segnalare la direzione della dinamica della produttività³; pertanto ai fini delle analisi, sono di notevole interesse, più che gli indicatori di produttività, gli indici della stessa; comunque la definizione dei primi rimane prodromica alla costruzione dei secondi.

Occorre distinguere la produttività globale o totale (*total factor productivity*)⁴ da quella parziale (*partial productivity*). Nelle indagini relative alla prima, vengono considerati tutti gli output e tutti gli input caratterizzanti il processo di produzione, in quelle relative alla seconda, la produzione viene rapportata ad un solo input; si parla in tal caso di produttività del fattore lavoro, capitale, materie prime, ..., con possibilità di ulteriore penetrazione analitica.

² Il condizionale è d'obbligo se si considera la sostanziale eterogeneità di output ed input che, in genere, è crescente nel passaggio da analisi condotte su strutture produttive semplici (singola impresa) a quelle condotte su altre più complesse (per es. intero sistema economico) e che implica il ricorso a valori monetari come soluzione aggregativa. In ogni caso, i problemi di aggregazione sono sempre presenti nelle analisi di produttività globale, data la conclamata eterogeneità degli input considerati. Peraltro si complicano ulteriormente nelle ipotesi di strutture caratterizzate da produzione congiunta, data la conseguente necessità di aggregare i diversi output.

³ Talvolta l'interesse per l'analisi della produttività ricade sulla sua dinamica spaziale invece che su quella temporale. Allora l'obiettivo diventa quello di valutare come varia la produttività tra due strutture produttive (semplici o complesse che siano) con riferimento ad un dato tempo. In pratica, considerate due o più strutture produttive, l'obiettivo è di confrontare le rispettive produttività registrate in un determinato periodo t . In presenza di più di due strutture produttive, sorge il problema di verificare la consistenza interna del sistema dei confronti bilaterali, vale a dire la possibilità che gli indici di produttività soddisfino la proprietà della transitività, per modo che (per es.) se una struttura produttiva A risulta più produttiva di quella B del 15% e questa ultima più produttiva di quella C del 15%, per transitività dovrebbe risultare anche la struttura A più produttiva di quella C nella misura del 32,25% (essendo $1,15 \cdot 1,15 = 1,3225$). Una soluzione al problema deriverebbe dall'adozione del cosiddetto metodo EKS (dal nome degli studiosi Eletto-Koves e Szulk). Cfr. con (tra gli altri) COELLI, PRASADA RAO., BATTESE (1998), pag. 84 e succ.. Nel presente lavoro ci si occupa esclusivamente della dinamica temporale della produttività.

⁴ Cfr. (per es.) OWYONG (2000), pag. 20: "Total factor productivity is the combined productivity of all inputs, and hence avoids the problems faced by measures based on just one factor. One does not have to choose any factor on which to base productivity growth, since all factors are included."

Se sotto il profilo concettuale la trattazione della produttività non presenta particolari problemi, le difficoltà sorgono quando l'analisi della stessa passa dal livello strettamente concettuale a quello applicativo. Il superamento di tali difficoltà sembra ormai risiedere nell'assemblaggio di un modello metodologico i cui elementi sono mutuabili dalla economia politica e dalle discipline aziendali e statistico-matematiche. Una opportuna combinazione di questi contributi consente di modellare metodologie adeguate alle finalità dell'analisi di produttività che si vuole condurre, fronteggiando i problemi insiti nella stessa.

Atteso che, come si vedrà in seguito, l'analisi della produttività globale fornisce indicazioni più utili e congrue di quanto possa fare quella della produttività parziale, questi problemi sono minimi quando (ipotesi piuttosto accademica) i processi produttivi si caratterizzano per essere mono-input e mono-output e conseguentemente la distinzione tra produttività parziale e globale scompare. Si complicano in presenza (ipotesi realistica) di una pluralità di input e/o output al punto che si rende necessaria una loro aggregazione resa possibile in prima istanza dall'impiego dei rispettivi prezzi e dall'assunzione secondo la quale i processi produttivi osservati operino in condizioni di efficienza tecnica ed allocativa, vale a dire che output ed input osservati riflettano comportamenti ottimizzanti da parte degli operatori traducibili fondamentalmente nella massimizzazione dei ricavi e minimizzazione dei costi⁵. Questa semplice aggregazione non tiene conto, tuttavia, della probabile curvatura degli isoquanti e pertanto gli indici di produttività che ne derivano potrebbero fornire informazioni distorte. Una strada alternativa sarebbe quella di sfruttare le proprietà parametriche delle funzioni di produzione ricavando un indice di produttività per via econometrica; tuttavia, le difficoltà operative poste dalla complessità della procedura farebbero propendere per la costruzione di numeri indici di produttività basati indirettamente su funzioni di produzione e considerati come approssimazione dei risultati ottenuti per via parametrica. Un'altra soluzione sarebbe quella di considerare le variazioni di produttività come variazioni di output non spiegate da variazioni di input, ovvero di misurarle in quanto residuo ascrivibile al progresso tecnico (residuo di Solow). In ogni caso, la congruità delle indicazioni fornite da indici di produttività globale così concepiti sarebbe inficiata dalle assunzioni di efficienza tecnica ed allocativa. Questi problemi possono essere supe-

⁵ Occorre evidenziare che gli input ed output espressi in valore monetario devono essere previamente deflazionati, per modo che siano valorizzati con moneta avente lo stesso potere d'acquisto: l'evidente obiettivo è di consentire all'indicatore di produttività di cogliere solo variazioni quantitative. Un simile problema sussiste anche nelle analisi della dinamica spaziale della produttività, cioè la differenza spaziale (invece che temporale) del potere d'acquisto.

rati ricorrendo all'impiego di funzioni di distanza per la rappresentazione delle tecnologie produttive. Si tratta di particolari funzioni che non richiedono alcuna assunzione specifica in merito al comportamento degli operatori, superano qualunque problema di carattere aggregativo degli input ed output e rilassano l'ipotesi estrema di processi produttivi tecnicamente efficienti, consentendo di poter distinguere le variazioni di produttività ascrivibili alla dinamica del livello di efficienza e quelle ascrivibili alla dinamica del progresso tecnico. Ne derivano degli indici di produttività (di Malmquist) molto flessibili ed adatti ad analisi su qualsiasi tipologia di struttura produttiva (*profit*, *non-profit*, produzione congiunta e non, ...).

In definitiva emerge la considerazione che la ricerca metodologica ha aggiornato la "cassetta degli attrezzi" dello statistico-economico o, quanto meno, l'ha integrata con approcci metodologici estremamente flessibile ed in grado di far luce anche sulle determinanti fondamentali della dinamica temporale della produttività.

2. Total Factor Productivity Index (TFPI): generalità

Un indicatore di produttività⁶ è definito dal rapporto tra il risultato dell'attività produttiva e l'ammontare dei fattori in essa impiegati. Concretamente trattasi di un valore numerico che scaturisce dal rapporto tra due entità numeriche: a numeratore l'output, a denominatore l'input. Tali entità, a loro volta, sono generate da un procedimento di aggregazione di input e/o output, quando ce n'è più d'uno, oppure quando il livello d'analisi supera la singola struttura produttiva. Se nella elaborazione vengono coinvolti tutti i fattori produttivi⁷, si ottiene un indicatore di produttività totale, (*Total Factor Productivity*, *TFP*), altrimenti un indicatore di *produttività parziale* in relazione al fattore produttivo considerato. Tuttavia un indicatore di produttività totale assolve una funzione segnaletica di maggior rilievo se considerato in termini relativi piuttosto che in termini assoluti. Per esempio, in sede empirica, verificare che la produttività di una struttura produttiva sia pari a 4,5 può essere meno significativo di verificare che quella stessa organizzazione denoti un indice di produttività pari a 1,5 posta uguale ad 1 la produttività di una diversa e contemporanea struttura produttiva oppure quella sua stessa relativa ad un periodo di tem-

⁶ L'obiettivo dell'indagine sulla dinamica temporale della produttività di una determinata struttura produttiva (impresa, gruppo, distretto, settore, sistema economico) può essere sintetizzato nella opportunità di monitorare nel tempo la capacità di produrre livelli superiori di output impiegando livelli inferiori di input, ovvero l'economicità della produzione.

⁷ Talvolta l'analisi può ridursi al coinvolgimento degli input più importanti. Se la tecnologia produttiva fosse anche *multi-output*, ossia tale da produrre congiuntamente più output, l'analisi deve tener conto, in qualche modo, anche di questi ultimi.

po precedente, ossia, genericamente, relativa ad un valore scelto come base o di riferimento, *benchmark*)⁸. Pertanto, selezionato un valore base, un indice di produttività totale è genericamente definibile come segue:

$$\begin{aligned}TFPI = YI/XI &= (Y^{t+1}/Y^t)/(X^{t+1}/X^t) = \\ &= (Y^{t+1}/X^{t+1})/(Y^t/X^t) = TFP^{t+1}/TFP^t\end{aligned}\quad [2.1]$$

dove Y e X denotano rispettivamente i valori delle aggregazioni output ed input, mentre YI e XI i rispettivi indici. Quando l'analisi della produttività verte sulla dinamica temporale, disponendo almeno di *time-series data*, gli apici $t+1$ e t denotano rispettivamente i periodi di tempo cui si riferiscono i valori di Y e X , essendo t il periodo base, $t+1$ quello da confrontare. Quando l'interesse cade sulla dinamica spaziale, disponendo di *cross-sectional data*, gli stessi apici denotano due differenti e contemporanee strutture produttive⁹. In ogni caso, dalla [2.1] emerge che, in linea di principio, l'indice della TFP è un indice del "prodotto medio" (*average product*), ovvero il rapporto tra due indici: quello YI relativo all'output e quello XI relativo ai

⁸ Quando si adottano indici di produttività generati mediante rapporti tra funzioni di distanza, come si vedrà più avanti a proposito degli indici Malmquist di produttività, il problema delle difficoltà insite nell'aggregazione della pluralità degli input e/o output, resta sostanzialmente superato proprio dal ricorso alle funzioni di distanza. A tali difficoltà si aggiungono anche quelle relative alla valutazione del contributo fornito nell'attività produttiva da certi input quali il capitale fisico e/o finanziario. Si tratta di difficoltà che determinano una maggior propensione all'impiego di indicatori parziali, atteso che quelli totali, talvolta, potrebbero essere inficiati nella loro attendibilità da opzioni metodologiche inevitabilmente non obiettive e quindi probabilmente arbitrarie. Se ciò è vero, cioè se la scelta tra indicatori parziali e totali può essere pilotata dalla possibilità/facilità di utilizzo dei primi piuttosto che dei secondi (fatte, altresì, le debite considerazioni relative al *budget* di spesa stanziato per le indagini), non è men vero che anche il trascurare alcuni input potrebbe rendere distorte le segnalazioni fornite dagli indicatori parziali. Quello che in definitiva si viene a delineare è un vero e proprio *trade-off*: le difficoltà di definire con congrua attendibilità l'impiego di taluni input, senza delle quali sarebbe logico optare aprioristicamente per gli indicatori totali, lascerebbe propendere per l'adozione di indicatori parziali. Così facendo, però, si rischia di inficiare l'attendibilità complessiva dell'indagine, considerando assunzioni che potrebbero non essere affatto realistiche, vale a dire, sinteticamente, che la dinamica della produttività possa essere ben spiegata senza analizzare congiuntamente il contributo di tutti i fattori produttivi. Questo *trade-off* è più forte quando le indagini vertono su strutture produttive caratterizzate dalla introduzione sempre più massiccia di capitale fisico e/o finanziario. Insomma le iniezioni di capitale nei processi produttivi e il ruolo crescente di altri fattori non tradizionalmente riconducibili a quelli primari e l'opportunità-necessità di tenerne conto in nome di una maggiore attendibilità delle indagini, imporrebbe di far entrare dalla finestra quello che esigenze di semplicità e facile, veloce ed economica reperibilità dei dati vorrebbero si facesse uscire dalla porta: l'analisi della dinamica della produttività totale dei fattori. Nel presente lavoro l'attenzione ricade proprio su quest'ultima.

⁹ Nel primo caso, l'obiettivo è quello di valutare, relativamente ad una data struttura produttiva, la sua *performance* (nella fattispecie, la TFP) del periodo $t+1$ rispetto a quella del periodo t . Nel secondo caso, l'obiettivo è quello di valutare, relativamente ad un dato periodo di tempo, la *performance* della struttura produttiva $t+1$ rispetto ad una diversa t . Nel seguito, ove non diversamente specificato, si fa riferimento solo al primo caso, sintetizzabile come valutazione della dinamica temporale.

fattori impiegati nell'attività produttiva. A questi indici si può pervenire previo ricorso a procedimenti di aggregazione degli input, per tener conto dell'azione congiunta di tutti gli input che si ritiene siano coinvolti nella generazione dell'output e, nel caso di tecnologia *multi-output*, anche dei vari output congiuntamente prodotti.

Spesso, il più delle volte per semplicità, si fa e si è fatto ricorso ad una aggregazione di tipo contabile. Mediante fissazione di un sistema di prezzi ponderanti, le quantità fisiche vengono riassunte nei noti *TFPI* di Laspeyres (*LPI*), Paasche (*PPI*), Fisher (*FPI*)¹⁰:

$$TFPI = \frac{\sum_j^M p_j^t y_j^{t+1} / \sum_j^M p_j^t y_j^t}{\sum_i^N c_i^t x_i^{t+1} / \sum_i^N c_i^t x_i^t} = \frac{LOQI}{LIQI} = LPI \quad [2.2]$$

$$TFPI = \frac{\sum_j^M p_j^{t+1} y_j^{t+1} / \sum_j^M p_j^{t+1} y_j^t}{\sum_i^N c_i^{t+1} x_i^{t+1} / \sum_i^N c_i^{t+1} x_i^t} = \frac{POQI}{PIQI} = PPI \quad [2.3]$$

$$TFPI = \sqrt{(LPI)(PPI)} = \frac{\sqrt{(LOQI)(POQI)}}{\sqrt{(LIQI)(PIQI)}} = \frac{FOQI}{FIQI} = FPI \quad [2.4]$$

dove y e x denotano rispettivamente gli output e gli input, p e c i rispettivi prezzi, M e N le rispettive numerosità, gli apici t e $t+1$ i soliti periodi di tempo, i pedici j e i due indici di riferimento ai singoli output ed input¹¹. La [2.2] può anche essere riformulata come segue:

$$TFPI = \sum_j^M \frac{y_j^{t+1}}{y_j^t} / \sum_i^N \frac{x_i^{t+1}}{x_i^t} = LPI \quad [2.5]$$

dove:

¹⁰ Per dettagli sul *framework* metodologico di questi indici si rinvia ad un qualsiasi testo di statistica economica. Cfr. (per es.) VITALI (1999), pag. 269-272; GUARINI, TASSINARI (2000), pag. 19-26. Nella [2.2], *LOQI* sta per *Laspeyres Output Quantity Index*, *LIQI* per *Laspeyres Input Quantity Index*, mentre nella [2.3], *POQI* sta per *Paasche Output Quantity Index*, *PIQI* per *Paasche Input Quantity Index*. Analogamente vale per *FOQI* e *FIQI*.

¹¹ Nelle applicazioni empiriche, il ricorso preferenziale all'indice di Fisher è giustificabile con la conseguente riduzione dell'arbitrio insito nella scelta tra l'indice di Laspeyres e quello di Paasche. Questi ultimi, infatti, generano risultati sensibilmente diversi, essendo diverso il riferimento temporale dei prezzi-pesi usati in sede di aggregazione degli output ed input. Tuttavia l'evidenza empirica rivela un maggior ricorso all'indice di Laspeyres.

$$\hat{\gamma}_j^t = p_j^t y_j^t / \mathbf{p}^t \mathbf{y}^t, \quad \%_i^t = c_i^t x_i^t / \mathbf{c}^t \mathbf{x}^t, \quad i = 1, \dots, N, \quad j = 1, \dots, M$$

$$\mathbf{p}^t \mathbf{y}^t = \sum_j^M p_j^t y_j^t, \quad \mathbf{c}^t \mathbf{x}^t = \sum_i^N c_i^t x_i^t, \quad t = 1, \dots, T$$

con $\hat{\gamma}_j^t$ [$\%_i^t$] che denota la quota di valore del singolo output y_j^t [input x_i^t] sul valore totale $\mathbf{p}^t \mathbf{y}^t$ [$\mathbf{c}^t \mathbf{x}^t$]¹², essendo $\mathbf{y}^t = (y_1^t, \dots, y_M^t)$ [$\mathbf{x}^t = (x_1^t, \dots, x_N^t)$] il vettore output [input] e $\mathbf{p}^t = (p_1^t, \dots, p_M^t)$ [$\mathbf{c}^t = (c_1^t, \dots, c_N^t)$] quello dei relativi prezzi.

Mutatis mutandis, analoghe riformulazioni valgono per gli *PPI* e *FPI*.

Alternativamente ma, soprattutto, atteso che siffatta aggregazione, prescindendo dalle caratteristiche tecnologiche del processo produttivo, può comportare importanti distorsioni¹³ nella valutazione della dinamica della produttività, si può optare per un processo aggregativo il cui risultato sia una *proxy* di quello derivante dal diretto coinvolgimento di specifiche funzioni di produzione. Nei lavori empirici si è fatto prevalentemente ricorso alle forme funzionali *Cobb-Douglas* e *Translog*¹⁴, con conseguente generazione dei due noti indici di Konüs-Byushgens e Törnqvist-Theil, rispettivamente.

Questi indici consentono di riflettere nel loro *framework* le caratteristiche tecnologiche del processo di produzione il cui funzionamento si vuole sia “mimato” da una selezionata¹⁵ funzione di produzione (*fdp*) che, nel ruolo di *tecnologia astratta*, si assume simili coerentemente la *tecnologia reale*¹⁶ del processo di trasformatio-

¹² Si tratta della cosiddetta *output [input] value share* per l'output [input].

¹³ Ipotizzando, per semplicità, una tecnologia produttiva *single-output & two-input*, questa imprecisione è dovuta al fatto che l'aggregazione contabile degli input ignora la forma degli isoquanti e pertanto l'elasticità di sostituzione. Ne deriva che, essendo la curvatura degli isoquanti in relazione inversa rispetto alla sostituibilità dei fattori, l'errore di stima commesso dall'indice di produttività così costruito, sarà massimo nell'ipotesi in cui l'isoquante approssimi un angolo retto (bassa o, addirittura, assoluta insostituibilità) e sostanzialmente nullo in quella in cui l'isoquante approssimi un andamento rettilineo (elevata o, addirittura, perfetta sostituibilità), atteso che l'aggregazione contabile presuppone una tecnologia caratterizzata da perfetta sostituibilità. Sorge quindi l'interesse verso *TFPI* che tengano conto di tecnologie produttive più realistiche e quindi della possibilità, molto più concreta, che gli isoquanti siano caratterizzati da curvatura. Sul punto, *cf.* FRIGERO (1996), pag. 49-50.

¹⁴ Caratterizzata da un maggior livello di flessibilità, quest'ultima forma funzionale è il risultato di un filone di ricerca mirante a definire approssimazioni lineari della funzione *CES* (*Constant Elasticity of Substitution*) e che ha condotto alla formulazione di funzioni di produzione *VES* (*Variable Elasticity of Substitution*) delle quali la *Translog* è una formulazione generale. Matematicamente è nota per configurarsi come approssimazione di una qualunque funzione produttiva mediante un polinomio di Taylor del secondo ordine. *Cfr.* (per es.) VITALI (1999), pag. 260-261; GUARINI, TASSINARI (2000), pag. 312-316. Per i primigeni riferimenti alle forme funzionali citate, *cf.* COBB, DOUGLAS (1928); CHRISTENSEN, JORGENSON, LAU (1971); CHRISTENSEN, JORGENSON, LAU (1973).

¹⁵ In base a proprietà matematiche che descrivono le caratteristiche della *tecnologia astratta*, tra le quali vi è proprio il livello di produttività totale rilevato da un parametro tecnologico.

¹⁶ Normalmente, nelle indagini empiriche statistiche-economiche sono disponibili i dati (in forma di: *time-series data*, *cross-sectional data*, *panel data*) ma non si conosce il *reale* modello matematico che

ne input-output. In caso di tecnologia *multi-output*, si ipotizza la produzione di un unico tipo di output, e quindi, dovendosi esprimere globalmente la produzione in funzione dell'impiego dei fattori produttivi, si aggregano gli output, normalmente con procedimento contabile, e si seleziona una opportuna *fdp* per rappresentare l'attività trasformazione degli input nella produzione vista come un output scalare¹⁷.

In effetti la funzione di produzione consente di studiare le caratteristiche della *tecnologia astratta*¹⁸, compresa l'evoluzione nei modi di produrre generata dal cambiamento tecnologico¹⁹. Tra le molteplici applicazioni empiriche cui la *fdp* si presta, vi è proprio quella relativa all'analisi della produttività totale. La *fdp* assolve un compito di aggregazione-trasformazione dei fattori produttivi, il cui risultato esprime l'ammontare massimo dell'output Y ottenibile dall'impiego di una certa combinazione di input X . Nella ipotesi *single-output* la *fdp* rappresenta la frontiera (*production frontier*) dell'insieme delle possibili combinazioni produttive input-output. Precisamente è il suo sottoinsieme efficiente.

Considerando l'ipotesi *single-output* o comunque quella di un output scalare risultante da una qualche aggregazione nell'ipotesi *multi-output*²⁰, alla definizione di una generica *fdp* si può pervenire come di seguito indicato. Sia:

li ha generati come "eco" del funzionamento del processo di trasformazione input-output. Di qui l'esigenza di selezionare un modello *astratto* di funzionamento che sia il più possibile aderente alla realtà che si intende simulare, "mimare". Nell'approccio marginalista alla teoria della produzione, tale modello è proprio la *fdp*. Anche quando il modello reale è noto, permane il carattere astratto della *fdp* come conseguenza della necessità di sintetizzare la complessità tecnologica della trasformazione input-output in caratteristiche rilevanti, prescindendo dai *reali* aspetti fisico-chimici, dalle reali procedure aziendali.

¹⁷ Normalmente, specie nelle indagini a livello di sistema o settore economico, per esprimere la produzione come funzione degli input, si usa il valore della produzione oppure il valore aggiunto. In effetti misure in termini fisici sono possibili solo per le produttività parziali (per es. quantità per addetto, quantità per ora lavorata), di non interesse nel presente lavoro. Quando si ricorre al valore della produzione, tra gli input si annoverano tutti i fattori coinvolti nel processo di produzione, inclusi quelli intermedi; quando si ricorre al valore aggiunto, solo il lavoro, il capitale, la terra. Sulla problema della esistenza di una *fdp* del valore aggiunto, *cfr.* FRIGERO (1996), pag. 43-44.

¹⁸ Per una utile trattazione di queste caratteristiche, *cfr.* (per es.) COELLI, PRASADA RAO, BATTESE (1998), pag. 12-37.

¹⁹ Nel presente lavoro, i termini *tecnico* e *tecnologico* sono usati in modo equivalente. Una dinamica migliorativa [peggiorativa] nei modi di produrre, viene normalmente definita come progresso [regresso] tecnologico. In termini generici si è soliti riferirsi a tale dinamica con la locuzione '*technical change*' (cambiamento tecnologico) che può connotarsi, appunto, come progresso o regresso. È utile precisare che spesso, in letteratura, il riferimento a tale dinamica avviene col la sola locuzione '*progresso tecnico*' usandola, ove non diversamente precisato, in senso neutro: ossia come sinonimo di cambiamento. Succede anche nel presente lavoro.

²⁰ La necessità di pervenire comunque ad uno scalare nell'ipotesi *multi-output*, deriva da una ovvietà matematica: per definizione la *fdp* può gestire solo l'ipotesi di produzione singola. Questo problema può essere risolto, come meglio evidenziato nel successivo paragrafo a proposito degli indici di Malmquist, mediante ricorso ad una sorta di generalizzazione della *fdp*, ossia alla funzione di distanza che

- $t = 1, \dots, T$, un indicatore dei periodi di tempo cui riferire l'attività produttiva;
- $y^t \in \mathfrak{R}_+$, il singolo output in cui viene trasformato il vettore input $\mathbf{x}^t = (x_1^t, \dots, x_N^t) \in \mathfrak{R}_+^N$;
- $L^t(y^t) \subseteq \mathfrak{R}_+^N$, l'insieme dei fabbisogni di input (*input requirement set*), ossia l'insieme di tutti i vettori input $\mathbf{x}^t \in \mathfrak{R}_+^N$ il cui impiego consente la produzione del livello $y^t \in \mathfrak{R}_+$ dell'output;
- $P^t(\mathbf{x}^t) \subseteq \mathfrak{R}_+^{M=1}$, l'insieme dei livelli di output producibili (*output producible set*), ossia l'insieme dei livelli $y^t \in \mathfrak{R}_+$ dell'output ottenibili dall'impiego del vettore $\mathbf{x}^t \in \mathfrak{R}_+^N$;
- Z^t , la tecnologia produttiva (*production technology*) che gestisce la trasformazione di vettori input nell'output, definendo l'insieme delle possibili combinazioni produttive input-output (*technology set* o *graph set*).

In termini più formali, sia:

$$Z^t = \left\{ (\mathbf{x}^t, y^t) \in \mathfrak{R}_+^{N+1} \mid \mathbf{x}^t \text{ consente di produrre } y^t \right\} \quad [2.6]$$

$$L^t(y^t) = \left\{ \mathbf{x}^t \in \mathfrak{R}_+^N \mid y^t \in P^t(\mathbf{x}^t) \right\} \quad [2.7]$$

$$P^t(\mathbf{x}^t) = \left\{ y^t \in \mathfrak{R}_+ \mid \mathbf{x}^t \in L^t(y^t) \right\} \quad [2.8]$$

Una combinazione produttiva input-output è tecnologicamente possibile se $\mathbf{x}^t \in L^t(y^t)$ ovvero $y^t \in P^t(\mathbf{x}^t)$. Conseguentemente si può definire l'insieme di tutte le possibili combinazioni input-output (*technology set*), $(\mathbf{x}^t, y^t) \in \mathfrak{R}_+^{N+1}$:

$$\begin{aligned} Z^t &= \left\{ (\mathbf{x}^t, y^t) \in \mathfrak{R}_+^{N+1} \mid y^t \in P^t(\mathbf{x}^t), \mathbf{x}^t \in \mathfrak{R}_+^N \right\} = \\ &= \left\{ (\mathbf{x}^t, y^t) \in \mathfrak{R}_+^{N+1} \mid \mathbf{x}^t \in L^t(y^t), y^t \in \mathfrak{R}_+ \right\}, \end{aligned} \quad [2.9]$$

e quindi:

$$L^t(y^t) = \left\{ \mathbf{x}^t \in \mathfrak{R}_+^N \mid (\mathbf{x}^t, y^t) \in Z^t \right\} \quad [2.10]$$

$$P^t(\mathbf{x}^t) = \left\{ y^t \in \mathfrak{R}_+ \mid (\mathbf{x}^t, y^t) \in Z^t \right\} \quad [2.11]$$

A questo punto si può definire una generica *fdp* $f : \mathfrak{R}_+^N \rightarrow \mathfrak{R}_+$ come segue:

$$\forall \mathbf{x}^t \in \mathfrak{R}_+^N : f(\mathbf{x}^t) = \max \left\{ y^t \in \mathfrak{R}_+ \mid (\mathbf{x}^t, y^t) \in Z^t \right\} \quad [2.12]$$

costituisce un *tool* idoneo alla rappresentazione della fenomenologia produttiva complessa: *multi-input & multi-output*.

risultando che:

$$Z^t = \left\{ (\mathbf{x}^t, y^t) \in \mathfrak{R}_+^{N+1} \mid y^t \leq f(\mathbf{x}^t) \right\} \quad [2.13]$$

Naturalmente è necessario che la *fdp* sia ben definita, ovvero che il problema di massimo ammetta sempre una soluzione $\forall \mathbf{x}^t \in \mathfrak{R}_+^N$. Ciò è garantito dal fatto che l'insieme $\left\{ y^t \in \mathfrak{R}_+ \mid (\mathbf{x}^t, y^t) \in Z^t \right\}$ sia chiuso e limitato superiormente²¹. La *fdp* così definita esprime l'insieme di tutte le combinazioni input-output tecnicamente efficienti che lo stato della tecnologia consente alla struttura produttiva di attuare, ossia esprime il livello massimo dell'output $y^t \in \mathfrak{R}_+$ ottenibile da un determinato vettore input $\mathbf{x}^t \in \mathfrak{R}_+^N$. Si capisce il motivo per cui la *fdp* è definibile come *production frontier*²² di Z^t . Spesso si usa indifferentemente la locuzione 'efficient frontier'.

Assumendo un cambiamento tecnologico di tipo neutrale²³, la generica *fdp* potrebbe essere riscritta come segue²⁴:

$$y^t = A^t f(\mathbf{x}^t) \quad [2.14]$$

²¹ Questa proprietà è a sua volta garantita dal fatto che il *technology set* gode, tra le altre, delle proprietà di: a) chiusura, b) limitazione superiore degli input, c) possibilità di inazione, d) *free disposability* degli input. In virtù di quest'ultima, sono attivabili combinazioni produttive input-output che a parità di livello dell'output, impiegano maggiori quantità di input. Anche la *fdp* gode di varie proprietà, in parte corrispondenti a quelle del *technology set*, in parte proprie. Per dettagli su queste proprietà, cfr. (per es.) TANI (1986), pag. 39-43; IMPICCIATORE (1998), pag. 157-160.

²² Cfr. (per es.) IMPICCIATORE (1998), pag. 164: "La funzione di produzione costituisce l'artificio di cui la teoria economica si avvale, tradizionalmente, per rappresentare sinteticamente il complesso delle tecniche efficienti." In verità, con riferimento alla frontiera del *technology set* intesa come *fdp*, più rigorosamente si dovrebbe parlare di luogo geometrico di combinazioni input-output che sono, con certezza, solo output-efficienti. Non si può escludere che le stesse possano essere input-inefficienti. In effetti l'output potrebbe essere ottenuto con un minor impiego degli input. Si immagini una tecnologia produttiva di tipo *single-output & single-input* e ad una sua rappresentativa *fdp* che denoti tratti paralleli all'asse dell'input.

²³ In termini teorici, si ha un progresso tecnico di tipo neutrale (secondo Hicks) quando l'accumulazione delle conoscenze tecnico-scientifico-organizzative introdotte nel processo produttivo non determina alcuna variazione nella forma del diagramma della *fdp* e quindi del sistema degli isoprodotti, verificandosi solo un suo spostamento verso il basso (progresso) o verso l'alto (regresso): le produttività marginali dei fattori produttivi variano in modo proporzionale. Si ha un progresso tecnico non neutrale quando, invece, risulta variata la forma del diagramma della *fdp* e quindi degli isoquanti. In questo caso il progresso tecnico consente di ridurre in modo differenziato l'impiego di fattori: si è soliti parlare di progresso tecnico risparmiante lavoro (*labor saving*) o capitale (*capital saving*). Una formulazione della nozione di progresso tecnico neutrale secondo Harrod, invece, è fornita in FRIGERO (1996), pag. 75-76: "E' infatti definita come progresso tecnico "neutrale", in un modello in cui i fattori siano solo capitale e lavoro, la crescita della produttività del lavoro che avviene a parità di rapporto capitale/prodotto e dato il tasso di profitto...". Sulla scorta della nozione di neutralità secondo Harrod, Momigliano ha derivato una classificazione del progresso tecnico riportata in FRIGERO (1979), pag. 51-52.

²⁴ Convienne ribadire che t è un apice e non un esponente ed indica che la variabile è funzione del tempo.

dove A^t è un parametro moltiplicativo funzione di t (tempo) che esprime gli effetti del progresso tecnico²⁵ sulla produzione di y^t , atteggiandosi come indicatore di *TFP*. Si noti che, così configurato, il progresso tecnico è tale per cui la struttura della tecnologia $f(\mathbf{x}^t)$ non cambia nel tempo, ma può subire spostamenti (*shift*) proprio attraverso il parametro A^t . Ne risulta che seguendo le variazioni di tale parametro, si valuta la dinamica temporale di *TFP*, ossia si può pervenire a valori di *TFPI*.

Pertanto, definito come segue un generico indicatore *TFP*²⁶:

$$TFP^t = \frac{y^t}{f(\mathbf{x}^t)} = \frac{\text{output effettivo}}{\text{output potenziale}} = A^t \quad [2.15]$$

il conseguente indice sarebbe:

$$TFPI = \frac{YI}{XI} = \frac{y^{t+1}/y^t}{f(\mathbf{x}^{t+1})/f(\mathbf{x}^t)} = \frac{y^{t+1}/f(\mathbf{x}^{t+1})}{y^t/f(\mathbf{x}^t)} = \frac{A^{t+1}}{A^t} \quad [2.16]$$

da cui si evince che *TFPI* consiste, in definitiva, in rapporti tra *fdp* aggreganti, ed è interpretabile come rapporto tra “produttività medie”. L’indicatore *TFP*^t è pari ad uno quando l’aggregazione degli input è impiegata nel miglior modo reso possibile dalla tecnologia disponibile. Una variazione nel livello di y^t per una ragione qualsiasi, determina una variazione nel livello di *TFP*^t segnalando una variazione del livello di produttività.

Atteso che la *fdp* rappresenta la frontiera della *production technology*, una sua traslazione verso l’alto [basso] segnala una espansione [contrazione] del *technology set*, ossia un progresso [regresso] tecnologico e per ciò stesso un incremento [decremento] della produttività totale²⁷. E’ in questo senso che la *fdp* è concepibile

²⁵ Si osservi che, ove si assumi un progresso tecnico non-neutrale, in luogo di un parametro moltiplicativo, ci sarebbe un argomento nella relazione funzionale tra output ed input. In questo senso gli effetti del progresso tecnico determinano modifiche nella struttura della *fdp*.

²⁶ Coerentemente con la interpretazione della [2.1], tale indicatore può essere considerato come espressivo del *average product*.

²⁷ Più avanti, si evidenzia come il rilassamento della assunzione di efficienza tecnica, introduca un ulteriore fattore determinante della dinamica della produttività totale, ossia l’approssimarsi della struttura produttiva (rappresentata dalle coordinate input-output) alla frontiera di produzione, segnaletico di un recupero di efficienza tecnica. Anche in questo secondo caso si può parlare di incremento di produttività con la differenza che, nel primo l’incremento è ascrivibile al cambiamento migliorativo della tecnologia (progresso tecnico), ossia ad spostamento verso l’alto della *fdp*, e nel secondo alla eliminazione o riduzione del *gap* di efficienza tecnica, riconducibile alla incapacità del *management* di sfruttare le potenzialità della tecnologia disponibile evitando sprechi di risorse; secondo la prevalente terminologia anglosassone, si tratta rispettivamente dei seguenti due eventi caratterizzanti la produzione: *technology shift effect* e *catching up effect*. Peraltro entrambi gli eventi possono coesistere e più o meno compensarsi.

come un *tool* aggregatore di fattori produttivi, il cui risultato è uno scalare segnale-tico del massimo livello di output ed è in questo senso che si può pensare alla dinamica della *TFP* come alla dinamica di rapporti tra *fdp* aggreganti, ovvero come ad un *TFPI*. Infatti, nell'ipotesi di tecnologia *single-input & single-output* la [2.16] diventa²⁸:

$$TFPI = \frac{YI}{XI} = \frac{y^{t+1}/y^t}{x^{t+1}/x^t} = \frac{y^{t+1}/x^{t+1}}{y^t/x^t} = \frac{A^{t+1}}{A^t} \quad [2.17]$$

ossia un rapporto tra *average productivity*²⁹. Peraltro, si può pensare alla dinamica della *TFP* anche in termini di tassi di variazione:

$$\frac{\Delta TFP}{TFP^t} = \frac{y^{t+1} - y^t}{y^t} - \frac{f(\mathbf{x}^{t+1}) - f(\mathbf{x}^t)}{f(\mathbf{x}^t)} = \frac{A^{t+1} - A^t}{A^t} \quad [2.18]$$

ossia, la variazione percentuale della *TFP* è data dalla differenza tra la variazione percentuale dell'output e quella degli input.

In definitiva la variazione della *TFP* può essere concepita come variazione dell'output al netto di quella dell'input, *ex* [2.18]³⁰, ovvero come *shift* della *production frontier* tra $t+1$ e t , *ex* [2.16]-[2.17]. Il dato essenziale è che, in ipotesi di efficienza tecnica³¹ (*technical efficiency*), espansione [contrazione] della *TFP* e progresso [regresso] tecnologico sono sinonimi.

²⁸ E' evidente l'analogia concettuale, oltre che formulistica, tra la [2.16], la [2.17] e la [2.1]. Si sottolinea che le [2.1], [2.16] e la successiva [2.18], a differenza delle [2.2]-[2.5] e di ogni altra concreta formulazione di *TFPI*, prescindono dalle specifiche e concrete soluzioni al problema dell'aggregazione: sono, più che altro, generiche formulazioni dei concetti teorici relativi alla dinamica della *TFP*.

²⁹ Che il *TFPI* sia un rapporto tra *average productivity* è del tutto evidente nell'ipotesi *single-output & single-input*. Che sia interpretabile come tale anche in ipotesi *multi-output* è coerente con il ricorso a *fdp* aggreganti che conduce pur sempre ad uno scalare *proxy* della totalità degli input impiegati. Tuttavia che il rapporto tra produttività medie esprima lo *shift* tecnologico è "esatto" solo se la tecnologia di produzione esibisce rendimenti di scala costanti; diversamente lo si può intendere solo come approssimazione. L'errore può essere smorzato considerando la media geometrica dei due *shift* valutati in corrispondenza delle due osservazioni in t e $t+1$.

³⁰ E' spontanea una interpretazione della [2.18] come una discretizzazione del modello (*residuo*) di Solow cui si fa cenno più avanti.

³¹ Il concetto di efficienza tecnica è stato formalizzato da Koopmans nei seguenti termini: una struttura produttiva è tecnicamente efficiente se, e solo se, *i*) un incremento in qualsiasi output richiede un decremento in almeno un altro output oppure un incremento in almeno un input (*output-oriented approach*) ovvero *ii*) un decremento in qualsiasi input richiede un incremento in almeno un altro input oppure un decremento in almeno un output (*input-oriented approach*). Talvolta si parla di efficienza di tipo Pareto-Koopmans.

In seguito Farrell ha introdotto una misura di efficienza tecnica conferendo contenuto empirico al coefficiente di utilizzazione delle risorse definito da Debreu ed atteggiandosi come prototipo empirico per la misura orientata di efficienza teoricamente definita da Koopmans. Tuttavia la misura di effi-

Evidentemente l'attendibilità segnaletica delle [2.15]-[2.16] dipenderà dalla coerenza della forma funzionale selezionata con le effettive caratteristiche tecnologiche della produzione, ossia dalla coerenza tra queste ultime e i parametri matematici della *tecnologia astratta* $f(\mathbf{x}^t)$ simulante quella *reale*. E' altresì importante che le effettive variazioni di quest'ultima restino coerenti con la prima. A questo punto le difficoltà che affettano la definizione e stima della produttività e della sua dinamica si traducono nella stima statistica dei parametri della forma funzionale selezionata, tra questi essendocene uno dedicato alla livello della produttività totale e, ancor prima, nelle difficoltà proprie della stessa selezione. Si tratta di applicare tecniche econometriche rigorose e complesse. Spesso tale rigore viene sacrificato in nome di una maggior semplicità richiesta anche dalle risorse a disposizione per condurre le indagini e dall'ammontare delle quali dipende, in buona sostanza, la possibilità di reperire le necessarie informazioni per giungere ad una stima dei parametri tra cui A^t .

Per queste ragioni si preferisce optare per l'impiego di *TFPI*, quali approssimazioni dei risultati che dovrebbero essere generati dalla stima econometrica dei parametri. Ecco che i problemi si spostano sulla selezione di un indice che sia coerente con il tipo di *fdp* che si vuole simulare l'effettiva tecnologia produttiva. Un *TFPI* > 1 segnalerà che il processo produttivo ha sperimentato un miglioramento della sua

cienza tecnica proposta da Farrell (misura radiale) non coincide con quella concettuale di Koopmans, in quanto la metodologia non-parametrica suggerita da Farrell per la stima della *production frontier* può comportare la presenza di differenze residuali (*input slacks* e/o *output slacks*) nei vincoli tecnici dei problemi di PL (Programmazione Lineare).

Alla luce di questo distinguo, quando si fa riferimento ad una struttura produttiva tecnicamente efficiente, occorrerebbe distinguere se la qualificazione deriva da una interpretazione/stima dell'efficienza secondo Farrell o secondo Koopmans, soprattutto se si usa la locuzione '*efficient frontier*'. Vale la pena di precisare che la *efficient frontier* secondo Koopmans è un sottoinsieme (*efficient subset*) di quella secondo Farrell (isoquanto) e che per *cross-sectional data* di grandi dimensioni, tendono a coincidere. Su questo aspetto si ritorna nel successivo paragrafo.

Nel prosieguo del presente lavoro, ove non diversamente specificato, con la locuzione '*production frontier*' usata *sic et simpliciter*, si fa riferimento all'insieme di combinazioni produttive efficienti secondo Farrell. Inoltre, sino al termine del paragrafo e salvo diversa specificazione, l'efficienza tecnica continua ad essere considerata di tipo *output-oriented*, come fatto implicitamente sinora. In vista di quanto esposto in seguito, torna utile precisare che Farrell ha distinto tre concetti di efficienza:

- a) efficienza tecnica (*technical efficiency*): che riflette, come noto, l'abilità di una struttura produttiva ad ottenere (impiegare) il massimo (minimo) livello di output (input), dato un vettore input (output);
- b) efficienza allocativa (*price o allocative efficiency*): che riflette l'abilità di una struttura produttiva nell'impiegare (produrre) gli input (output) in ottime proporzioni, dati i rispettivi prezzi e la tecnologia produttiva;
- c) efficienza economica (*overall o economic efficiency*): che riflette la combinazione delle precedenti due abilità, ovvero la capacità di minimizzare (massimizzare) i costi (ricavi): questa minimizzazione (massimizzazione) può non essere raggiunta proprio a causa di inefficienza tecnica o allocativa.

Cfr. KOOPMANS. (1951); DEBREU G. (1951); FARRELL (1957).

produttività che, assumendo l'ipotesi di efficienza tecnica, è interpretabile come cambiamento migliorativo della tecnologia.

Ad ogni funzione di produzione selezionata corrisponde un indice di produttività diverso. Così alla funzione *Coob-Douglas* corrisponde, come già anticipato, l'aggregatore di Konüs-Byushgens secondo il quale gli input sono oggetto della aggregazione di seguito definita:

$$XI = \frac{f(\mathbf{x}^{t+1})}{f(\mathbf{x}^t)} = \prod_{i=1}^N \left(\frac{x_i^{t+1}}{x_i^t} \right)^{\%_i^t} = KBI \quad [2.19]$$

dove $\%_i^t = c_i^t x_i^t / \mathbf{c}^t \mathbf{x}^t$ denota le quote di costo totale dei singoli input³². Conseguentemente il generico *TFPI* ex [2.16] può essere specificamente riformulato come segue:

$$TFPI = YI/XI = (y^{t+1} / y^t) / KBPI \quad [2.20]$$

Invece alla funzione *Translog* corrisponde l'aggregatore di Törnqvist-Theil³³ secondo il quale gli input sono oggetto della seguente altra soluzione aggregativa:

$$XI = \frac{f(\mathbf{x}^{t+1})}{f(\mathbf{x}^t)} = \prod_{i=1}^N \left(\frac{x_i^{t+1}}{x_i^t} \right)^{(\%_i^{t+1} + \%_i^t)/2} = TTI \quad [2.21]$$

dove:

$$(\%_i^{t+1} + \%_i^t)/2 = \left((c_i^{t+1} x_i^{t+1} / \mathbf{c}^{t+1} \mathbf{x}^{t+1}) + (c_i^t x_i^t / \mathbf{c}^t \mathbf{x}^t) \right) / 2.$$

Conseguentemente si avrebbe che:

$$TFPI = YI/XI = (y^{t+1} / y^t) / TTI \quad [2.22]$$

Nei lavori empirici, in ipotesi di produzione *multi-output*, si ricorre più frequentemente ad un *TFPI* caratterizzato dalla aggregazione anche degli output con un

³² Si consideri che in un'ottica di equilibrio di lungo periodo di concorrenza perfetta (con rendimenti di scala costanti), il valore della produzione risulta pari alla somma dei valori dei fattori produttivi.

³³ Com'è noto, l'indice di Törnqvist-Theil (d'ora in poi, solo Törnqvist) è una approssimazione discreta dell'indice di Divisia, cui si accenna più avanti, essendo quest'ultimo valido nel caso continuo in cui si ipotizzano variazioni dei dati continue nel tempo, mentre quelli di cui normalmente si dispone in indagini statistiche-economiche sono di tipo discreto, da cui l'opportunità di addivenire ad una approssimazione discreta quale è appunto l'indice di Törnqvist.

procedimento analogo a quello degli input, *ex* [2.21]. Si tratta della seguente generalizzazione del *TFPI* definito *ex* [2.21]-[2.22], nota semplicemente come *Törnqvist Productivity Index (TPI)*:

$$TFPI = YI/XI = \prod_{j=1}^M \left(\frac{y_j^{t+1}}{y_j^t} \right)^{(\hat{\gamma}_j^{t+1} + \hat{\gamma}_j^t)/2} \bigg/ \prod_{i=1}^N \left(\frac{x_i^{t+1}}{x_i^t} \right)^{(\%_{00}^{t+1} + \%_{00}^t)/2} = TOQI/TIQI = TPI \quad [2.23]$$

dove³⁴:

$$\begin{aligned} (\hat{\gamma}_j^{t+1} + \hat{\gamma}_j^t)/2 &= \left((p_j^{t+1} y_j^{t+1} / \mathbf{p}^{t+1} \mathbf{y}^{t+1}) + (p_j^t y_j^t / \mathbf{p}^t \mathbf{y}^t) \right) / 2, \\ (\%_{00}^{t+1} + \%_{00}^t)/2 &= \left((c_i^{t+1} x_i^{t+1} / \mathbf{c}^{t+1} \mathbf{x}^{t+1}) + (c_i^t x_i^t / \mathbf{c}^t \mathbf{x}^t) \right) / 2. \end{aligned}$$

Questo indice, che tra i *TFPI* è senz'altro il più noto ed è generalmente più usato di quello Fisher, a sua volta preferibile rispetto a quelli Paasche e Laspeyres, oltre a rispondere positivamente ai *tests* di Fisher, presenta una importante relazione con un altro indice i cui aspetti teorici vengono sintetizzati nel prossimo paragrafo: l'indice Malmquist di produttività³⁵.

Spesso si è soliti far riferimento al *TPI* nella sua forma logaritmica:

$$\begin{aligned} \ln TPI = \ln YI - \ln XI &= \frac{1}{2} \sum_j^M (\hat{\gamma}_j^{t+1} + \hat{\gamma}_j^t) (\ln y_j^{t+1} - \ln y_j^t) - \\ &+ \frac{1}{2} \sum_i^N (\%_{00}^{t+1} + \%_{00}^t) (\ln x_i^{t+1} - \ln x_i^t) \end{aligned} \quad [2.24]$$

ossia nella differenza tra l'indice degli output e quello degli input entrambi nella forma logaritmica della struttura dell'indice Törnqvist di quantità.

La dinamica temporale della produttività totale dei fattori, oltre che con il ricorso ai numeri indici suddetti, può essere valutata anche ricorrendo all'indice (o *residuo* di) Solow (*SPI*) oppure al *metodo delle variazioni relative*, che genera un indice talvolta ricondotto a Kendrick (*KPI*), entrambi riconducibili al *growth accounting approach*. Con riferimento al primo, l'obiettivo di Solow è quello di fornire un “*ele-*

³⁴ Nella [2.23] *TOQI* sta per *Törnqvist Output Quantity Index*, *TIQI* per *Törnqvist Input Quantity Index*.

³⁵ Per una panoramica sulla *performance* di vari (numeri) indici rispetto ai *tests* di Fisher o, più in generale, rispetto all'approccio assiomatico, *cf.* (per es.) CECCHI, GALEOTTI (1990), pag. 25-33, 43-49; DIEWERT (1992), pag. 211-248; COELLI, PRASADA RAO., BATTESE (1998), pag. 79-81; GUARINI, TASSINARI (2000), pag. 27-28; DIEWERT, NAKAMURA (2002); McLELLAN (2004), pag. 4-6. Per i primigeni riferimenti a questo ambito della letteratura, *cf.* FISHER (1922); FRISCH (1936).

mentary way of segregating variations in output per head due to technical change from those due to changes in availability of capital per head', concependo il progresso tecnico come "a shorthand for any kind of shift in the production function"³⁶.

Piuttosto che procedere alla costruzione di indici di quantità di output ed input, come nel caso dei precedenti *TFPI*, si procede direttamente alla valutazione della dinamica della produttività intesa come variazione dell'output non spiegata dalla variazione degli input, ovvero un *residuo* segnalato dallo spostamento della *fdp*. Considerata la generica *fdp*, ex [2.14], in cui A^t ne coglie lo *shift* (*technical change*) nel tempo, l'obiettivo di Solow è giungere ad una sua stima, assumendo che:

- la $f(\mathbf{x}^t)$ sia lineare omogenea, cioè di grado 1, non sussistendo economie di scala (assumendo, quindi, rendimenti di scala costanti)³⁷:

$$f(k\mathbf{x}^t) = k^{n-1} f(\mathbf{x}^t), \quad n = \text{grado di omogeneità} \quad [2.25]$$

- si operi in regime di concorrenza perfetta, con la conseguenza che il comportamento degli operatori economici, funzionale alla massimizzazione del profitto, porta verso un sostanziale livellamento tra il prezzo in termini reali di ogni input e la sua produttività marginale, ossia ad una situazione in cui ciascun input viene retribuito per la propria produttività marginale; in definitiva si assume una efficienza tecnico-allocaativa:

$$f(k\mathbf{x}^t) = k^{n-1} f(\mathbf{x}^t), \quad n = \text{grado di omogeneità} \quad [2.26]$$

$$\partial y^t / \partial x_i^t = c_i^t / p^t$$

Dai precedenti due punti, consegue che:

$$y^t = \sum_i^N x_i^t \frac{\partial y^t}{\partial x_i^t} = \sum_i^N x_i^t \frac{c_i^t}{p^t}, \quad p^t y^t = \sum_i^N c_i^t x_i^t = \mathbf{c}^t \mathbf{x}^t \quad [2.27]$$

Applicando una differenziazione logaritmica rispetto a t nella [2.14], si giunge alla seguente equazione differenziale:

$$\frac{1}{y^t} \frac{dy^t}{dt} = \frac{1}{A^t} \frac{dA^t}{dt} + \sum_i^N \alpha_i^t \left(\frac{1}{x_i^t} \frac{dx_i^t}{dt} \right). \quad [2.28]$$

³⁶ SOLOW R. A. (1957), pag. 312. L'idea alla base dell'indice di Solow è che le variazioni di produttività siano definibili come variazioni della produzione non spiegabili da variazioni intervenute nella disponibilità dei fattori: la variazione di produttività è identificata con una sorta di *residuo* che si forma grazie al cambiamento tecnologico.

³⁷ L'originale modello di Solow assume una *fdp Cobb-Douglas* di grado 1. Questa assunzione comporta che il *residuo* ingloba anche eventuali rendimenti variabili di scala.

Infatti:

$$\begin{aligned}
 \frac{1}{y^t} \frac{dy^t}{dt} &= \frac{1}{A^t} \frac{dA^t}{dt} + \sum_i^N \frac{1}{f(\mathbf{x}^t)} \frac{\partial f}{\partial x_i^t} \frac{dx_i^t}{dt} = \\
 &= \frac{1}{A^t} \frac{dA^t}{dt} + \sum_i^N \frac{A^t}{y^t} \frac{\partial f}{\partial x_i^t} \frac{dx_i^t}{dt} \frac{x_i^t}{x_i^t} = \\
 &= \frac{1}{A^t} \frac{dA^t}{dt} + \sum_i^N \left(A^t \frac{\partial f}{\partial x_i^t} \right) \left(\frac{x_i^t}{y^t} \right) \frac{1}{x_i^t} \frac{dx_i^t}{dt} = \\
 &= \frac{1}{A^t} \frac{dA^t}{dt} + \sum_i^N \left(\frac{\partial y^t}{\partial x_i^t} \right) \left(\frac{x_i^t}{y^t} \right) \frac{1}{x_i^t} \frac{dx_i^t}{dt}
 \end{aligned} \tag{2.29}$$

dove:

$$\left(A^t \frac{\partial f}{\partial x_i^t} \right) = \left(\frac{\partial y^t}{\partial x_i^t} \right), \quad \%_{0,i}^t = \left(\frac{\partial y^t}{\partial x_i^t} \right) \left(\frac{x_i^t}{y^t} \right) = \frac{c_i^t x_i^t}{p^t y^t} = \frac{c_i^t x_i^t}{\mathbf{c}^t \mathbf{x}^t}$$

con $\%_{0,i}^t$ che esprime l'elasticità dell'output rispetto all' i -esimo input.

Riordinando la [2.29] si ottiene:

$$\frac{1}{A^t} \frac{dA^t}{dt} = \frac{1}{y^t} \frac{dy^t}{dt} - \sum_i^N \%_{0,i}^t \left(\frac{1}{x_i^t} \frac{dx_i^t}{dt} \right) \tag{2.30}$$

risultando che il tasso di variazione della produttività totale (*technical change*) è dato dalla differenza tra il tasso di variazione dell'output e la media ponderata dei tassi di variazione degli input, usando pesi pari alle rispettive elasticità $\%_{0,i}^t$, la cui somma è pari ad uno.

Per rendere empirica la [2.30], Solow assume che le derivate rispetto a t possano essere approssimate da variazioni discrete in analogia alla [2.18]. Pertanto, passando alle differenze discrete, la [2.30] diventa:

$$\frac{\Delta A^t}{A^t} = \frac{A^{t+1} - A^t}{A^t} = \frac{\Delta y^t}{y^t} - \sum_i^N \%_{0,i}^t \left(\frac{\Delta x_i^t}{x_i^t} \right) \tag{2.31}$$

da cui si passa all'indice:

$$TFPI = 1 + \frac{\Delta A^t}{A^t} = 1 + \frac{\Delta y^t}{y^t} - \sum_i^N \%_{0,i}^t \left(\frac{\Delta x_i^t}{x_i^t} \right) = SPI \tag{2.32}$$

che esprime la variazione della *TFP* nel passaggio dal periodo t a quello $t+1$, ricordando sempre che, in virtù dell'assunzione dell'ipotesi di efficienza tecnica, va intesa come cambiamento tecnologico³⁸.

Vale la pena di evidenziare che, nel continuo, la [2.32] ha il suo equivalente nell'indice *TFP* di Divisia definito come segue:

$$TFPI = \left(\frac{y^{t+1}}{y^t} \right) / \exp \int_t^{t+1} \sum_i^N \%_i^t \left(\frac{1}{x_i^t} \frac{dx_i^t}{dt} \right) = DPI \quad [2.33]$$

la cui formulazione logaritmica è:

$$\ln \left(\frac{TFPI^{t+1}}{TFPI^t} \right) = \ln \left(\frac{y^{t+1}}{y^t} \right) - \int_t^{t+1} \sum_i^N \%_i^t \left(\frac{1}{x_i^t} \frac{dx_i^t}{dt} \right) = \ln DPI \quad [2.34]$$

In pratica *DPI* è un indice curvilineo continuo definito da un integrale di linea, non suscettibile di applicazione empirica, essendo necessaria la disponibilità di dati nel continuo³⁹. Se nella [2.30] si sostituiscono i tassi di crescita continui con differenze discrete nei logaritmi e, per ciascun input, la *value share* continua con la media aritmetica dei suoi due valori riferiti a t e $t+1$, si perviene alla forma logaritmica del *TFPI* definito dalla [2.22]⁴⁰, ossia si perviene ad una sua approssimazione

³⁸ In realtà, secondo autorevoli studiosi, il modello condurrebbe ad una stima distorta del cambiamento tecnologico, proprio in quanto il relativo indice verrebbe generato per *residuo* (da cui il nome) trascinandosi dentro non solo elementi desiderati, riconducibili al miglioramento ed accumulazione delle conoscenze scientifico-tecnico, bensì anche elementi di distorsione ascrivibili ad imprecisioni e/o omissioni che ne andrebbero ad inficiare la purezza. Tra questi vanno senz'altro menzionate: a) l'omissione di fattori produttivi, b) ciclicità del *residuo*, c) sostituibilità tra fattori, d) tipologia dei mercati, e) rendimenti di scala (l'elencazione non è esaustiva). In proposito è emblematico citare l'espressione con cui Abramovitz si riferisce al *residuo*: "*measure of our ignorance*", cfr. ABRA-MOVITZ (1956), pag. 8. Per approfondimenti sul punto, cfr. (per es.) FRIGERO (1996), pag. 60-66; HULTEN (2000), pag. 10-13.

³⁹ Cfr. (per es.) HULTEN (1973). Il *DPI* non è suscettibile di applicazione empirica, in quanto il concreto calcolo integrale di linea di cui alle [2.33]-[2.34] necessita di dati continui tra i due estremi di integrazione.

⁴⁰ Cfr. JORGENSON, GRILICHES (1967). Le ricerche di Diewert hanno consentito di dimostrare che l'approssimazione logaritmica di tipo Törnqvist all'indice di Divisia definita da Jorgenson e Griliches nel 1967 può definirsi un *exact TFPI* se la *fdp* di cui alla [2.14] è di tipo *Translog*. In pratica, in tale ipotesi, la versione log del *TFPI* di cui alla [2.22] non è una semplice approssimazione, ma un indice esatto. Sicché, partendo da una *fdp Translog*, passando per le [2.27]-[2.30], si giunge alla approssimazione logaritmica di tipo Törnqvist, ossia al *TFPI* versione log di cui alla [2.22]. Inoltre, considerato la flessibilità della *Translog* linearmente omogenea, ossia il suo essere una approssimazione di Taylor al secondo ordine di una qualsiasi *fdp* linearmente omogenea, il *TPI* assurge al rango di "*superlative index number*", usando una felice espressione di Diewert con la quale si vuole evidenziare che l'impiego del *framework* Törnqvist nella costruzione dei numeri indici costituisce una scelta 'esatta' anche quando la *tecnologia reale* non è di tipo *Translog*. Il 'grado di esattezza' dipende dalla capacità della *fdp Translog* di 'mimarla'. Cfr. DIEWERT (1976).

che, nell'ipotesi di tecnologia *multi-input & multi-output*, ha il suo corrispondente nel *TFPI* definito dalla [2.24].

Prossimi a quelli forniti dall'indice *SPI*, ex [2.32], sono i risultati di quello definito talvolta come *indice Kendrick*⁴¹:

$$TFPI = \frac{y^{t+1}/y^t}{\sum_i^N \%_i^t (x_i^{t+1}/x_i^t)} = KPI \cong SPI \quad [2.35]$$

Nelle indagini empiriche in cui si considera il valore aggiunto come output e lavoro e capitale come input, allora può essere utile la seguente approssimazione:

$$TFPI = \frac{y^{t+1}/y^t}{\%_L^t (x_L^{t+1}/x_L^t) + \%_K^t (x_K^{t+1}/x_K^t)} = KPI \cong \frac{1}{S} \cong SPI \quad [2.36]$$

dove S è il prodotto tra y^t/y^{t+1} e x_L^{t+1}/x_L^t .

3. Frontier approach versus non-frontier approach

Dalle generiche formulazioni di cui alle [2.16] e [2.18], emerge che l'assunzione di efficienza tecnica implica una sostanziale coincidenza tra la dinamica della *TFP* e dinamica della tecnologia produttiva. In tutti *TFPI* visti sinora, tale assunzione è implicita: la costruzione degli indici presuppone che le strutture produttive oggetto di indagine siano osservate sulla frontiera del *technology set*. Del resto si è visto che la costruzione di ciascuno di essi coinvolge una *fdp*⁴². Va altresì evidenziato che gli stessi indici assumono anche una efficienza allocativa (*allocative efficiency*) delle strutture produttive osservate. In effetti l'aggregazione degli input, coerente col funzionamento di una *tecnologia astratta*, si risolve concretamente nel coinvolgimento delle *input value share* assumendo la sussistenza delle condizioni di equilibrio concorrenziale, minimizzazione dei costi (*cost-minimizing*), uguaglianza tra

⁴¹ Cfr. (per es.) KENDRICK (1956). Il *metodo delle variazioni relative* è stato impiegato anche da altri autorevoli studiosi, fra cui Abramovitz, Graziani, Stigler. E' stato impiegato anche nella prima Relazione del Comitato nazionale della Produttività: *Misura della produttività in Italia*, giugno 1963. Gli indici *SPI* e *KPI* sono stati applicati a livello macroeconomico, facendo normalmente ricorso al valore aggiunto per l'output ed al lavoro e capitale per gli input. Meno diffuso è stato il ricorso al valore della produzione per l'output ed al conseguente coinvolgimento dei beni intermedi tra gli input. Per una notevole applicazione al settore bancario, cfr. MARTELLI (1984).

⁴² Tenuto conto del funzionamento dell'aggregatore contabile, si può pensare ad una *fdp* sottesa agli indici di Laspeyres e di Paasche, caratterizzata da una elasticità infinita dei fattori.

produttività marginale di ciascun input e rispettivo prezzo reale. *Mutatis mutandis*, lo stesso dicasi per l'aggregazione degli output nel caso di produzione *multi-output*⁴³. Si tratta di una assunzione restrittiva, non sussistendo la quale l'aggregazione non avrebbe fondamento economico. In estrema sintesi, si può affermare che alla base del *framework* metodologico dei *TFPI* sinora richiamati, vi è l'assunzione dell'ipotesi di comportamento economico ottimale (*optimal behaviour*) delle strutture produttive, intese come imprese di tipo *profit* volte alla massimizzazione del profitto (*profit-maximizing*)⁴⁴. Ne derivano inevitabilmente alcune considerazioni⁴⁵.

Innanzitutto, occorre valutare la coerenza dell'assunzione di efficienza tecnico-allocaativa con la realtà economica effettuale. Di primo acchito, non vi sarebbe alcunché di eccezionale, considerata l'impostazione classica della teoria microeconomia, secondo la quale, in un mercato concorrenziale degli input ed output, strutture produttive tecnicamente ed allocativamente inefficienti verrebbero espulse dal mercato. Il problema è che, pur essendo teoricamente vero, ciò non necessariamente si realizza istantaneamente. Per effetto di frizioni o imperfezioni fisiologiche di tali mercati, le strutture inefficienti potrebbero continuare ad operare anche per diversi anni. Nel frattempo potrebbero aver luogo virtuosi aggiustamenti manageriali che potrebbero far recuperare il *gap* di efficienza con la ovvia conseguenza che le strutture produttive condannate dalla teoria microeconomia classica all'espulsione dal mercato, potrebbero in effetti conquistarsi una stabile permanenza. Se ciò è logico, allora occorre riconoscere, da un lato, la possibilità che posizioni di inefficienza possono coesistere con l'effettiva presenza in un mercato anche concorrenziale, dall'altro l'opportunità di isolare, nella dinamica della produttività, l'effetto di strategie manageriali che modificano la posizione in termini di efficienza tecnica della struttura produttiva osservata, posizione che non può aprioristicamente coincidere con una *full efficiency* come, al lume di una rigida interpretazione della classica teoria microeconomica, vorrebbe la logica sottostante al *framework* dei *TFPI* su esposti. D'altra parte gli effettivi comportamenti economici delle imprese potrebbero non essere tutti di tipo *optimal behaviour*. Alcune imprese potrebbero ottimizzare ma nel rispetto di determinati vincoli, altre potrebbero non ottimizzare affatto.

⁴³ Nel caso di tecnologia anche *multi-output*, vale un parallelo coinvolgimento delle *output value share* di tipo *revenue-maximizing*.

⁴⁴ In definitiva le imprese opererebbero in situazione di tendenziale efficienza economica. *Cfr. supra*, nota 26.

⁴⁵ Va altresì evidenziato che, in tutti i *TFPI* sinora richiamati, l'impiego delle *value share* poggia sull'assunzione di una tecnologia produttiva linearmente omogenea.

L'indagine empirica potrebbe riguardare operatori del *public sector* o del *not-for-profit sector*: in tali casi, parlare di *optimal behaviour* non avrebbe senso economico. Inoltre le informazioni sui prezzi potrebbero essere incomplete, distorte, quando non indisponibili⁴⁶. La conseguenza è che le *value share* aggreganti possono essere distorte, generando a loro volta stime distorte (*biased*)⁴⁷ della dinamica della *TFP*. La distorsione può essere più grave se la struttura produttiva osservata non opera sulla frontiera del *technology set*, sperimentando una inefficienza tecnica⁴⁸. In questo caso la dinamica della *TFP* non può essere intesa, *sic et simpliciter*, come cambiamento tecnologico, scandito⁴⁹ dallo spostamento della *fdp* nel tempo. Quest'ultimo non può e non deve essere considerato come fattore assorbente della dinamica temporale della *TFP*⁵⁰, atteso che potrebbero essersi verificati miglioramenti o peggioramenti nella gestione della efficienza tecnica da parte del *management*. In definitiva se la struttura produttiva osservata operasse al di sotto della *fdp* (ma pur sempre nel *technology set*), i *TFPI* esposti sinora fornirebbero stime distorte del progresso tecnico, dovendo necessariamente isolare dalle stime, un fattore ascrivibile alla dinamica della efficienza tecnica.

L'opportunità di scomporre un *TFPI* in due componenti principali, l'uno segnaletico del cambiamento migliorativo [progresso] o peggiorativo [regresso] dello stato tecnologico, l'altro segnaletico del miglioramento o peggioramento dello stato di (in)efficienza tecnica, assume preminente rilievo in sede di *economic policy*. Infatti con riferimento ad indagini empiriche condotte a livello di sistema economico, non è affatto irrilevante conoscere se un *trend* negativo della *TFP* è ascrivibile maggiormente ad uno piuttosto che all'altro dei due fattori.

⁴⁶ L'indisponibilità, oltre che da *deficit* informativo, può essere fisiologica del settore di attività della struttura produttiva osservata: è il caso non solo delle realtà *not-for-profit*, bensì anche di realtà produttive caratterizzate da produzione congiunta di *good & bad output*. I *bad output* essendo generati congiuntamente ed inevitabilmente ai *good* ma non desiderati (*undesirable*), normalmente non sono commerciabili e pertanto non sono prezzati non avendo mercato. Per questo motivo, spesso si fa riferimento ad essi come a dei *non marketable output*. Ove nella costruzione dei *TFPI* si volesse tener conto anche di questa particolare tipologia di output, sarebbe necessario stimare dei prezzi ombra (*shadow prices*). Sul punto, *cfr.* (per es.) PITTMAN (1983).

⁴⁷ La distorsione della stima deriverebbe dal fatto che gli *input [output] shares* non siano effettivamente *cost-minimizing [revenue-minimizing]*, ossia non siano coerenti con l'assunzione di *optimal behaviour* su cui poggia la giustificazione economica dei *TFPI* del tipo di quelli visti sinora.

⁴⁸ "Note that there are two possible sources of inefficiency: the first is technical inefficiency (i.e., production below the frontier), and the second is allocative inefficiency. Allocative inefficiency would be reflected in the shares used to aggregate inputs.", *cfr.* FÀRE, GROSSKOPF, NORRIS, ZHANG (1994).

⁴⁹ Nel caso *single-output*.

⁵⁰ La sostanziale coincidenza tra dinamica della *TFP* e dinamica della tecnologia produttiva permarrrebbe ove la dinamica della efficienza tecnica fosse piatta.

Se, per esempio, il *trend* fosse sfavorevole a causa di una dinamica negativa della efficienza tecnica, una politica di sostegno alla crescita della *TFP* si rivelerebbe inefficace se diretta a sostenere l'innovazione tecnologica; efficace sarebbe, invece, una politica di rimozione di barriere istituzionali alla diffusione della innovazione, se fosse proprio la loro presenza ad incidere sullo stato generale di inefficienza.

Ridefinendo il livello delle problematiche, ragionamenti analoghi sono pertinenti con riferimento ad indagini condotte a livello di settore economico. Evidentemente l'opportunità di scomposizione di un *TFPI* interessa anche in sede di *policy committee*. Qui, la possibilità di individuare vuoti di efficienza ascrivibili alla condotta del *management*⁵¹ diventa cruciale per la *performance* economica dell'impresa.

Da un punto di vista formale, il rilassamento dell'ipotesi di efficienza tecnica può essere descritto facendo ricorso al concetto di funzione di distanza, ossia ad una particolare funzione che consente di valutare la distanza (euclidea) nello spazio input-output tra la struttura produttiva osservata e la posizione potenziale che essa stessa sperimenterebbe se operasse sulla *production frontier*. Pertanto, se la struttura produttiva opera al di sotto della *fdp*, allora la [2.14] deve essere ridefinita come disuguaglianza:

$$y^t < A^t f(\mathbf{x}^t) \quad [3.1]$$

Naturalmente tale distanza varia in funzione del posizionamento della struttura osservata nel *production technology* rispetto al suo *subset* efficiente (secondo Farrell), la *production frontier* appunto. In forza di ciò la funzione di distanza rappresenta un *tool* diverso e, per certi aspetti, più potente della *fdp*, che si rivela utile proprio ai fini dello studio e della rappresentazione delle tecnologie produttive. Più potente perché consente di aderire maggiormente alla realtà effettuale, consentendo di assecondare ipotesi tecnologiche di *technical inefficiency* e, inoltre, di tipo *multi-input & multi-output*⁵². Precisamente la funzione di distanza relativa alla [3.1] è orientata

⁵¹ Per esempio, l'esistenza di sacche di inefficienza può consistere anche nella mancanza di uno spirito di appartenenza e/o realizzazione nelle risorse umane gestite dal *management*. L'evidenza empirica ha dimostrato che è possibile colmare tali sacche agendo proprio sul *management*, pur mantenendo le stesse risorse umane.

⁵² Cfr. (per es.) TANI (1986), pag. 49: "... i recenti sviluppi delle teorie comportamentistiche dell'impresa, e in particolare la teoria della X-inefficienza introdotta da Leibenstein, hanno messo in luce possibili comportamenti definibili come razionali e pur compatibili con situazioni anche tecnicamente inefficienti, conducendo a prendere sempre più in considerazione, anche nell'analisi empirica ed econometria, la possibilità di osservare situazioni produttive che non si collocano sulla frontiera dell'insieme di produzione. Da questo punto di vista l'analisi in termini di funzione di distanza, nonostante le sue difficoltà concettuali ed analitiche, offre maggiore spazio a considerazioni relative a situazioni di inefficienza ed alla trattazione della produzione congiunta, ...".

all'output, è cioè una *output-oriented distance function* secondo Shephard⁵³ ed è reciproca della misura di efficienza tecnica secondo Farrell nell'orientamento all'output (*output technical efficiency measure*)⁵⁴ e può essere definita come segue:

$$\begin{aligned} D_o^t(\mathbf{x}^t, y^t) &= \min \left\{ [\ : \left(\mathbf{x}^t, \frac{y^t}{[\]} \right) \in Z^t \right\} = \\ &= \left(\max \left\{ [\ : (\mathbf{x}^t, [\] y^t) \in Z^t \right\} \right)^{-1} = \left(E_o^t(\mathbf{x}^t, y^t) \right)^{-1} \end{aligned} \quad [3.2]$$

dove $[\]$ è una espansione del livello di y^t (parallela al suo asse⁵⁵), condizionata dal vettore \mathbf{x}^t . Considerato che lo stare sulla frontiera significa, a parità di \mathbf{x}^t impiegato, produrre il massimo livello di output, è evidente che $D_o^t(\mathbf{x}^t, y^t) < 1$ se e solo se $y^t < A^t f(\mathbf{x}^t)$ e d'altra parte $D_o^t(\mathbf{x}^t, y^t) = 1$ se e solo se $y^t = A^t f(\mathbf{x}^t)$, mentre $D_o^t(\mathbf{x}^t, y^t) > 1$ se e solo se $y^t \notin Z^t$. Il primo caso è quello coerente con la [3.1], il secondo con la [2.14], il terzo caso presuppone che la produzione di y^t non sia compatibile con alcuna delle possibili combinazioni input-output appartenenti al *production technology* Z^t . La relazione tra *fdp* e funzione di distanza può essere meglio evidenziata come segue:

$$D_o^t(\mathbf{x}^t, y^t) = \min \left\{ [\ : \frac{y^t}{[\]} \leq A^t f(\mathbf{x}^t) \right\} = \frac{y^t}{A^t f(\mathbf{x}^t)} \quad [3.3]$$

con conseguente definizione del livello massimo di output nei periodi t e $t+1$:

$$\frac{y^t}{D_o^t(\mathbf{x}^t, y^t)} = A^t f(\mathbf{x}^t), \quad \frac{y^{t+1}}{D_o^{t+1}(\mathbf{x}^{t+1}, y^{t+1})} = A^{t+1} f(\mathbf{x}^{t+1}) \quad [3.4]$$

⁵³ Per i primigeni riferimenti, *cfr.* SHEPHARD (1970); FÄRE (1988).

⁵⁴ Nel paragrafo successivo, in chiave prodromica alla esposizione degli indici di Malmquist, si ritorna con maggior dettaglio sulla funzione di distanza, generalizzata ad una ipotesi di tecnologia produttiva di tipo *multi-input & multi-output*, e si definisce anche la *input-oriented distance function* introdotta da Shepard. *Cfr.* SHEPHARD (1953). Conviene ribadire che, per il momento, si sta considerando uno scenario di tecnologia produttiva di tipo *single-output & multi-input*. In tale scenario, lo si ripete, la *production frontier* coincide con la *fdp*.

⁵⁵ Nell'ipotesi *multi-input & multi-output*, l'espansione degli output è di tipo radiale, ovvero equiproportionale, atteso che lascia invariate le proporzioni tra i vari output. Nell'ipotesi qui considerata, trattandosi di output scalare, se gli input fossero due, ossia se il *technology set* fosse rappresentato dallo spazio input-output 3D, l'espansione dell'output si configurerebbe come proiezione verticale, parallela al suo asse, della combinazione input-output osservata. La massima espansione possibile, proietterebbe quest'ultima sulla superficie della *production frontier*, costituita da tutte le combinazioni input-output tecnologicamente efficienti secondo Farrell. Se la frontiera fosse di tipo *smooth-surface* (per es., una *Cobb-Douglas*) sussisterebbe anche una efficienza secondo Koopmans. In tale ipotesi il punto di proiezione costituirebbe una combinazione input-output efficiente anche secondo Koopmans.

dove se l'output osservato y^t giace sulla frontiera, coerentemente con la [2.14], $D_o^t(\mathbf{x}^t, y^t) = 1$; se invece dista dalla stessa rispetto alla sua corrispondente proiezione, coerentemente con la [2.21], allora $D_o^t(\mathbf{x}^t, y^t) < 1$, il cui preciso valore *ex post* [2.22] dà contezza di tale distanza, ossia del reciproco della corrispondente *output technical efficiency measure*⁵⁶.

Ritornando al generico *TFPI*, tenendo presente la [2.15] ed in virtù della [3.4], la [2.16] può essere ridefinita come segue:

$$TFPI = \frac{y^{t+1}/f(\mathbf{x}^{t+1})}{y^t/f(\mathbf{x}^t)} = \frac{A^{t+1} D_o^{t+1}(\mathbf{x}^{t+1}, y^{t+1})}{A^t D_o^t(\mathbf{x}^t, y^t)} \quad [3.5]$$

dalla quale si evince chiaramente che, rilassando la restrittiva ipotesi di efficienza tecnica delle strutture produttive osservate, la dinamica temporale della *TFP* non è detto che continui ad esser spiegata in modo esaustivo dagli *shift* della *production frontier*, data la presenza di un rapporto di funzioni di distanza che esprime la dinamica della efficienza tecnica⁵⁷. *Mutatis mutandis*, stesse considerazioni valgono per la [2.18] che può essere ridefinita come segue:

$$\begin{aligned} \frac{\Delta TFP}{TFP^t} &= \frac{y^{t+1} - y^t}{y^t} - \frac{f(\mathbf{x}^{t+1}) - f(\mathbf{x}^t)}{f(\mathbf{x}^t)} = \\ &= \frac{A^{t+1} - A^t}{A^t} + \frac{D_o^{t+1}(\mathbf{x}^{t+1}, y^{t+1}) - D_o^t(\mathbf{x}^t, y^t)}{D_o^t(\mathbf{x}^t, y^t)} \end{aligned} \quad [3.6]$$

dove la differenza tra la variazione percentuale dell'output e quella degli input non determina necessariamente la variazione percentuale del progresso tecnologico, data l'addizionale presenza della variazione percentuale della efficienza tecnica.

Se consideriamo la [3.5], ci si rende conto che esistono inevitabilmente due effettivi valori per il cambiamento tecnologico, uno in corrispondenza del vettore \mathbf{x}^t , l'altro in corrispondenza del vettore \mathbf{x}^{t+1} .

Tuttavia se si assume, proprio come nella [2.14], un progresso tecnico di tipo Hicks-neutrale ed una omogeneità lineare della tecnologia, i due valori sono identici. In altre ipotesi sono normalmente differenti, per cui si può considerare una me-

⁵⁶ Naturalmente, minore [maggiore] è il valore della *output-oriented distance function* [*output technical efficiency measure*] rispetto alla soglia unitaria, più grave è lo stato di inefficienza.

⁵⁷ La sostanziale coincidenza tra indice della *TFP* ed indice del progresso tecnico ritorna ove, *ex post*, la struttura produttiva osservata si riveli efficiente nei due tempi t e $t+1$, ovvero ove la sua inefficienza sia statica.

dia geometrica tra essi⁵⁸. In pratica possiamo riscrivere l'indice del progresso tecnico come segue:

$$\frac{A^{t+1}}{A^t} = AI(\mathbf{x}^t) = \frac{y^t / f^t(\mathbf{x}^t)}{y^t / f^{t+1}(\mathbf{x}^t)} \quad [3.7]$$

oppure:

$$\frac{A^{t+1}}{A^t} = AI(\mathbf{x}^{t+1}) \frac{y^{t+1} / f^t(\mathbf{x}^{t+1})}{y^{t+1} / f^{t+1}(\mathbf{x}^{t+1})} \quad [3.8]$$

oppure come media geometrica di entrambi:

$$\frac{A^{t+1}}{A^t} = AI(\mathbf{x}^t, \mathbf{x}^{t+1}) = \left[\frac{f^{t+1}(\mathbf{x}^t)}{f^t(\mathbf{x}^t)} \frac{f^{t+1}(\mathbf{x}^{t+1})}{f^t(\mathbf{x}^{t+1})} \right]^{1/2} \quad [3.9]$$

Per quanto su specificato, a proposito della omogeneità di grado 1 e della neutralità del progresso tecnico, risulta:

$$AI = AI(\mathbf{x}^t) = AI(\mathbf{x}^{t+1}) = AI(\mathbf{x}^t, \mathbf{x}^{t+1}) \quad [3.10]$$

Inoltre, coerentemente con le [3.7]-[3.9], tenuto conto della [3.2], possiamo riscrivere l'indice dell'efficienza tecnica della [3.5] come segue:

$$\begin{aligned} EI &= \frac{D_o^{t+1}(\mathbf{x}^{t+1}, y^{t+1})}{D_o^t(\mathbf{x}^t, y^t)} = \frac{y^{t+1}}{f^{t+1}(\mathbf{x}^{t+1})} \bigg/ \frac{y^t}{f^t(\mathbf{x}^t)} = \\ &= \frac{f^t(\mathbf{x}^t)}{y^t} \bigg/ \frac{f^{t+1}(\mathbf{x}^{t+1})}{y^{t+1}} = \frac{E_o^t(\mathbf{x}^t, y^t)}{E_o^{t+1}(\mathbf{x}^{t+1}, y^{t+1})} = EI \end{aligned} \quad [3.11]$$

In definitiva, sia pure in linea di principio, è legittimo pensare ad una scomposizione di un generico *TFPI*. Vale a dire, sussistendo una tecnologia produttiva caratterizzata da rendimenti di scala costanti e con una sua dinamica temporale di tipo Hicks-neutrale, un generico *TFPI*, inteso come indice di *average productivity*, ovvero come indice del progresso tecnico, rilassata l'ipotesi restrittiva della efficienza tecnica, è teoricamente scomponibile in due sue componenti principali: un indice di efficienza tecnica (*catching-up effect*) ed un indice del cambiamento tecnologico (*technological shift effect*):

⁵⁸ E' quanto si fa quando la stima della dinamica del progresso tecnico avviene nell'ambito di quella degli indici Malmquist della *TFP*.

$$TFPI = \frac{y^t}{f(\mathbf{x}^t)} \bigg/ \frac{y^{t+1}}{f(\mathbf{x}^{t+1})} = EI \cdot AI \quad [3.12]$$

dalla quale, tuttavia, considerate le [3.7]- [3.9] e la [3.11], emerge che una scomposizione così concepita è concretamente percorribile se è dato conoscere le tecnologie produttive su cui è (implicitamente) basata la costruzione del generico *TFPI*. Se, come normalmente, non è dato conoscerle, si può comunque pensare di stimarle⁵⁹.

Coerente con la [3.12] è la scomposizione che può essere definita teoricamente facendo ricorso al *framework* degli indici di quantità di Malmquist, e precisamente a quelli produttività (*Malmquist Productivity Index, MPI*). Il passaggio dal livello teorico a quello empirico è reso possibile mediante definizione e applicazione di appositi stimatori di realtà numeriche inosservabili. Si tratta degli stimatori *DEA* (*Data Envelopment Analysis*) che consentono di definire stimatori non-parametrici (*MPI-DEA* appunto) degli indici teorici *MPI*. In effetti questi ultimi sono definiti teoricamente come rapporti di funzioni di distanza, e lo stesso vale per le loro componenti principali. Tutte le funzioni di distanza che si rendono necessarie, tuttavia, sono definite in modo teorico, ossia rispetto *production frontier* delle reali e sconosciute *production technology* coinvolte nella valutazione. La loro applicazione empirica è resa possibile, dalla implementazione degli stimatori *DEA* che, previa disponibilità di *balanced panel data*, consentono di stimare in modo non-parametrico tutto quanto è necessario ma sconosciuto in quanto inosservabile: 1) reale *technology set*, e quindi 2) reale *input requirement set* o 3) reale *output producible set*. La stima del primo rende disponibile quella della reale *production frontier*. Disporre di una stima di quest'ultima significa disporre anche di una stima della reale e sconosciuta frontiera degli insiemi 2) e 3). Disponendo di una stima di tutte le *production frontier* coinvolte nell'applicazione empirica, si perviene alle conseguenti stime delle reali e sconosciute *distance function*, quindi ad una stima degli sconosciuti valori dei teorici (reali) *MPI*⁶⁰.

⁵⁹ Si supponga di voler valutare la dinamica della *TFP* di una data struttura produttiva (per es., una impresa di un determinato settore) in un determinato intervallo osservazionale (per es., dal periodo/anno t a quello $t+3$). Per usare un qualsiasi *TFPI* per ciascuna coppia di periodi adiacenti (t e $t+1$, $t+1$ e $t+2$, $t+2$ e $t+3$) occorrono le informazioni relative alle quantità ed ai prezzi output ed input della struttura in valutazione e per ciascun periodo. Ove si volesse scomporre tale indice, occorrerebbe poter osservare (cioè conoscere) la reale tecnologia produttiva relativa a ciascun periodo. Non conoscendola, si può pensare di stimarla con un qualche *framework* metodologico.

⁶⁰ Se i reali insiemi 1)-3) sono inosservabili, lo sono anche le rispettive frontiere e, per conseguenza, le funzioni di distanza ad esse relative: in definitiva i valori degli indici teorici *MPI* non sono calcolabili. Se un qualche *framework* metodologico consente di generare una qualche stima di quanto non è suscettibile di osservazione, si possono conseguentemente generare tutte le altre stime prodromiche a quella definitiva di tali indici e loro componenti. Nella metodologia *Malmquist-DEA*, il *framework*

In chiusura di paragrafo conviene tentare una schematizzazione tra possibili approcci metodologici allo studio della dinamica temporale della produttività. In tal senso si può procedere assumendo come criterio discretivo proprio la formulazione di ipotesi in merito alla efficienza tecnica ed ulteriormente il *framework* metodologico che presiede alla stima dei valori di tale dinamica.

Sicché, è possibile definire il seguente schema distinguendo tra tecniche metodologiche di tipo:

- *Non-frontier approach*:

- Non-parametriche

Si tratta proprio dei *TFPI* già sinteticamente esposti, siano essi basati propriamente sull'aggregatore contabile, quali quelli di Paasche, Laspeyres, Fisher, o basati implicitamente su funzioni di produzione, quali quelli di Konüs, Törnqvist, o riconducibili propriamente al *growth accounting approach*, quali quelli di Solow, Kendrick.

- Parametriche

Si tratta di quelle tecniche basate direttamente sull'impiego delle *fdp* e che si risolvono nella stima econometrica dei loro parametri, tra cui quello necessario per esprimere le variazioni di produttività totale, ossia i cambiamenti intervenuti nella tecnologia (progresso/regresso tecnico)⁶¹.

metodologico che consente il *by-pass* dell'inosservabilità è costituito proprio dagli stimatori *DEA*, definiti a loro volta, mediante problemi di Programmazione Lineare, come esposto nel paragrafo 4.

⁶¹ Genericamente, questo tipo di tecniche conduce alla individuazione di una equazione (con parametri stimati econometricamente), la cui derivata rispetto al tempo esprime la dinamica temporale della produttività totale in forma di funzione del tempo che, assunta l'ipotesi di efficienza tecnica, coincide con quella del cambiamento tecnologico. A parte la necessità di individuare una congrua forma funzionale della *fdp* e di procedere alla stima dei relativi parametri, vi è anche l'inconveniente di non poter trattare ipotesi tecnologiche di tipo *multi-output*. In sede empirica, per superare quest'ultimo inconveniente, ove non si intenda optare per una *fdp* il cui output scalare sia generato per aggregazione (per es., con un indice Laspeyres) dei molteplici output, si può fare (e si è fatto) ricorso alla teoria della dualità tra *fdp* e funzione di costo (*cost function*), in forza del fatto che quest'ultima può gestire la pluralità degli output. Più raramente il ricorso alla dualità si è risolto nell'impiego della *revenue function* o *profit function*. In pratica si possono stimare i parametri di una funzione di costo o ricavo o profitto in luogo di quelli di una *fdp* che non è suscettibile di definizione in ipotesi *multi-output*, ipotesi in cui la funzione di costo ha la sua corrispondente rappresentazione primale nella *input-oriented distance function*, la funzione di ricavo nella *output-oriented distance function*, la funzione di profitto nella *directional distance function*. In sede empirica è diffuso il ricorso alla forma funzionale *Translog* per sfruttare la sua flessibilità. Per una primigenia trattazione sulla teoria della dualità in economia, cfr. SHEPHARD (1970); cfr. (anche) TANI (1986), pag. 95-126; FÄRE, PRIMONT (1995); CHAMBERS, CHUNG, FÄRE (1998). Nel tipo di tecniche in questione, in sede empirica, la stima della *production frontier*, primale o duale che sia, è di tipo stocastico, mentre in passato, si è fatto prevalente ricorso al tipo deterministico. Sulla differenza tra frontiera stocastica e deterministica si fa sinteticamente cenno più avanti in paragrafo.

- *Frontier approach*:

- Non-parametriche

Si tratta proprio degli indici Malmquist di produttività (*MPI*). La reciprocità sul piano teorico tra funzioni di distanza e misure di efficienza tecnica secondo Farrell, sussiste naturalmente anche in sede empirica, ossia in sede di stima di quanto è vero ed inosservabile, ed è proprio questo che, in definitiva, rende possibile la stima non-parametrica degli indici teorici *MPI* mediante gli stimatori *MPI-DEA*⁶². Si è soliti riferirsi a questi ultimi usando la locuzione: ‘metodologia *Malmquist-DEA*’⁶³.

- Parametriche

Queste ultime fanno ricorso all’impiego di *fdp* o, più in generale di frontiere, di tipo stocastico per la stima dei parametri, tra cui quelli necessari per esprimere le variazioni di efficienza e di progresso tecnico⁶⁴.

Sino ai primi anni ’80, l’approccio tradizionale allo studio sulla dinamica della produttività è stato caratterizzato dall’utilizzo di metodologie *non-frontier approach* e, nelle indagini a livello di intero sistema economico o di settore di produzione, si è

⁶² Esiste un’altra metodologia non-parametrica basata sempre sulla implementazione di problemi di PL: si tratta della *Free Disposable Hull (FDH)* che a differenza di quella *DEA* si caratterizza per il rilassamento del vincolo di convessità. Questa metodologia è stata introdotta da DEPRINS, SIMAR, TULKENS (1984); *cfr.* (anche) TULKENS (1993). Tuttavia la letteratura internazionale dimostra un generalizzato ricorso alla *DEA*. Per una applicazione empirica della metodologia *FDH*, *cfr.* (per es.) DESTEFANIS (1996).

⁶³ Vale la pena di precisare che non sono mancati tentativi di procedere ad una generazione di *MPI* con stima delle relative funzioni di distanza definite rispetto a frontiere stocastiche, ossia ad una generazione parametrica di tali indici. *Cfr.* (per es.) COELLI, PRASADA RAO, BATTESE (1998), pag. 233-242.

⁶⁴ Genericamente, questo tipo di tecniche conduce alla individuazione di una equazione (con parametri stimati econometricamente), la cui derivata rispetto al tempo esprime la dinamica temporale della produttività totale in forma di funzione del tempo in cui sono distinguibili due componenti additive: una relativa al cambiamento tecnologico, l’altra relativa al cambiamento nell’efficienza. Per un esempio di concreta applicazione di questo tipo di tecniche, *cfr.* FECHER, PESTIEAU (1991), in cui si fa ricorso ad una *fdp* di tipo *Cobb-Douglas*. A proposito della gestione dell’ipotesi *multi-output*, valgono anche per questo tipo di tecniche considerazioni analoghe a quelle svolte in nota a proposito di quelle *non-frontier approach* di tipo parametrico. Pertanto è necessario che, della *production frontier*, si faccia ricorso ad una sua rappresentazione duale (per es., una *cost function*) o primale (*distance function*), a meno che non si intenda optare per una *fdp* il cui output scalare sia una qualche aggregazione dei molteplici output. Per una applicazione di una *cost function* di tipo *Translog*, *cfr.* (per es.) BAUER (1990).

La primogenitura di questo filone metodologico (*frontier approach* di tipo parametrico) va riconosciuta ad un lavoro di Nishimizu e Page anche se la frontiera usata è di tipo (*Translog*) deterministico e non stocastico, *cfr.* NISHIMIZU, PAGE (1982). Nel lavoro appena citato, la dinamica della produttività è vista per la prima volta come risultante di due contributi distinti ed additivi: quello relativo al cambiamento tecnologico, quello relativo al cambiamento nell’efficienza. Da questo punto di vista, ossia degli obiettivi finali, il *framework* metodologico ivi impiegato può essere considerato un precursore della metodologia *Malmquist-DEA*.

fatto particolarmente ricorso agli indici di Solow, Kendrick, oltre che all'uso di indici di produttività parziale; ad altri livelli si è usato maggiormente l'indice di Laspeyres e quello di Törnqvist. In generale è stata l'intera categoria delle metodologie *non-frontier* di tipo non-parametrico ha trovato maggior impiego, per le ovvie ragioni di semplicità ed economicità che caratterizzano il loro impiego. Infatti i dati necessari per l'utilizzo di questa metodologia tradizionale sono organizzati in *time series* (aggregati al livello rispetto al quale si intende condurre l'indagine)⁶⁵.

Naturalmente, escludendo l'ipotesi di inefficienza tecnica (cioè essendo *non-frontier approach*) queste metodologie conducono ad una "consapevole" confusione (non solo terminologica) tra variazioni di produttività e cambiamento tecnologico. Pertanto, per esempio, l'indice di Solow esprime cambiamenti tecnologici intesi come variazioni della produttività totale generando, almeno in presenza di effettive posizioni di inefficienza, una forzata equivalenza concettuale tra i due effetti. *Mutatis mutandis*, ciò vale per tutte le metodologie *non-frontier approach*.

Tuttavia, questo inconveniente non può essere propriamente interpretato come conseguenza di una *defaillance* metodologica delle tecniche in questione, piuttosto come costo in termini di precisione che va pagato quando non si è nella disponibilità di un *panel data* che consenta di stimare la frontiera di produzione. Allora emerge, piuttosto, che in genere è la disponibilità o meno di dati in forma di *panel* che pilota l'opzione rispettivamente verso metodologie *frontier* o *non-frontier approach* e che le ultime dovrebbero atteggiarsi come una sorta di ripiego in presenza di sole *time series data*.

A far tempo da primi anni '80 hanno cominciato a diffondersi indagini empiriche condotti sia sulla efficienza che sulla produttività facendo ricorso a tecniche di tipo *frontier approach*. In particolare, in maniera pressoché generalizzata, si è trattato di quella non-parametrica. Pertanto la metodologia *DEA* è stata prevalentemente usata nelle indagini sulla efficienza (soprattutto tecnica) e ha consentito di rendere operativi gli indici *Malmquist* di produttività definiti mediante da funzioni di distanza e rimasti noti solo sul piano teorico, data la impossibilità di procedere concretamente al loro diretto calcolo, o quanto meno ad una loro stima.

Vale la pena di evidenziare che se da un lato i *TFPI* sono estremamente semplici da calcolare e, a differenza dell'approccio *non-frontier* parametrico, richiedono solo due osservazioni relative alla stessa struttura produttiva in due periodi differenti, dall'altro i dati necessari sono quelli di quantità degli output ed input e dei rispettivi prezzi.

⁶⁵ Inoltre i *TFPI* sono stati usati anche nelle indagini sui confronti multilaterali (per es., tra stati, regioni, aziende) usando *cross-sectional data*.

Le tecniche *frontier-approach*, disponibilità di *panel data* permettendo, consentono di scomporre le variazioni della produttività totale in cambiamenti di efficienza e cambiamenti tecnologici senza necessitare di dati relativi ai prezzi degli input ed output⁶⁶, tuttavia le osservazioni necessarie non sono, appunto, solo due. Disponendo solo di una *cross-sectional data* relativa al periodo t , queste tecniche consentono di stimare solo la *production frontier* di tale periodo nonché le deviazioni dalla stessa delle strutture produttive osservate.

Tuttavia c'è una differenza. Nell'approccio parametrico (stocastico) le deviazioni dalla *production frontier* hanno anche una origine stocastica, infatti la deviazione totale dalla frontiera va ascritta ad una prima componente erratica riconducibile direttamente alla inefficienza e ad una seconda componente erratica governata dal *caso*, ossia da ogni altro fattore casuale di disturbo (errori di misurazione, eventi atmosferici, accidentali, sindacali ...)⁶⁷. Nell'approccio non-parametrico, le deviazioni dalla frontiera⁶⁸ sono ascritte alla sola inefficienza tecnica e, pertanto, affinché le stesse siano correttamente segnaletiche di posizionamenti inefficienti, è

⁶⁶ A meno che, ovviamente, non si opti per una rappresentazione duale della tecnologia di produzione (per es. ove si opti per l'impiego di una frontiera di costo, occorrono i prezzi degli input).

⁶⁷ In letteratura la prima componente viene solitamente definita *one-sided error* ad indicare un errore unidirezionale che è responsabile del posizionamento inefficiente. Ove si ignori la seconda componente erratica, la cui distribuzione generalmente si assume sia una *Normal* (con media zero), la frontiera stimata si caratterizzerebbe come non più stocastica, bensì deterministica (variante ormai trascurata). Un notevole inconveniente nella adozione di frontiere stocastiche risiede nella sostanziale arbitrarietà con cui si fissa la distribuzione dell'errore *one-sided* (in genere si opta tra *Exponential*, *Half-Normal*, *Truncated-Normal*, *Gamma*). In effetti le applicazioni pratiche hanno dimostrato che la stima della frontiera stocastica è sensibile alla selezione della distribuzione. Peraltro resta irrisolto l'analogo problema della arbitrarietà nella selezione della forma funzionale della rappresentazione tecnologica che (sia nel *frontier* che nel *non-frontier approach*) ricade quasi sempre, dopo la predominanza della *Cobb-Douglas*, sulla *Translog*.

Come già indicato, la disponibilità di un'unica *cross-sectional data* consente solo la stima della efficienza tecnica. Tuttavia va precisato che in caso di utilizzo di rappresentazioni duali, le stime di efficienza non riguardano esclusivamente l'aspetto tecnico, bensì quello economico, sicché, per esempio, ove si utilizzi una frontiera di costo, si otterrebbero stime relative alla inefficienza economica di costo. Queste possono essere scomposte in stime di efficienza allocativa e tecnica solo se la rappresentazione primale implicata dalla frontiera di costo è una *fdp* suscettibile di esplicita derivazione: è il caso della forma funzionale *Cobb-Douglas*, non lo è per la *Translog*. Una *fdp Translog*, pur essendo implicata da una duale funzione di costo, non può essere matematicamente derivata. Soluzioni econometriche a questo problema, peraltro molto complesse, sono stata proposte in KOPP R. J., DIEWERT (1982); ZIESCHANG (1983). Solo la disponibilità di un *panel data* consente, in sede di stima parametrica della frontiera stocastica, di pervenire ad informazioni sulla dinamica della efficienza tecnica e della tecnologia, espressa in tassi di variazione. In presenza di un *panel data* la stima della inefficienza specifica per ciascuna struttura produttiva è più agevole di quella in presenza di una sola *cross-sectional data*. Tuttavia, per poterne trarre vantaggio è opportuno formulare una qualche assunzione circa il tipo di dinamica (stabile o variabile) della inefficienza.

⁶⁸ Queste deviazioni sono espresse proprio dalle stime dei valori delle funzioni di distanza definite rispetto a frontiere non-parametriche stimate mediante stimatori *DEA*.

necessario assumere che i dati osservati non siano perturbati da fattori non riconducibili al *management*⁶⁹.

4 - Considerazioni finali

Dal punto di vista empirico i dati necessari alla implementazione della metodologia *Malmquist-DEA* devono essere organizzati in forma di *panel data*, ossia si deve disporre di una *cross-sectional data* per ciascuno dei t anni del periodo di osservazione: in altri termini, per ciascuna delle k unità produttive del campione, occorrono i dati relativi al vettore degli output ed a quello degli input osservati nei t anni; ciò presuppone che il *panel data* sia bilanciato. Non sono necessari i dati relativi ai prezzi dei singoli input ed output; pertanto, una prima rilevante differenza rispetto al tradizionale approccio dei numeri indici *TFP*, consiste proprio nel fatto che i vettori relativi ai prezzi non sono necessari, in quanto il tradizionale sistema di aggregazione mediante prezzi viene sostituito da una implicita aggregazione derivante dall'impiego delle funzioni di distanza (*Shephard distance function*).

Disponendo di un *balanced panel data*, l'applicazione della metodologia *Malmquist-DEA*, per ciascuna delle K strutture produttive (denominate, in gergo *DEA*, come *DMU*, acronimo di *Decisional Making Unit*) genera un indice di produttività per ogni coppia di anni adiacenti (per esempio, gli anni t e $t+1$). Il fatto che la metodologia prescindendo dalla disponibilità di prezzi è di rilevanza essenziale se si considera che spesso, in sede di analisi empirica, i prezzi di cui si dispone possono essere non corretti, e/o incompleti quando non mancanti del tutto. Inoltre vi sono settori in cui i prezzi sono di difficile rilevazione (per esempio, quello bancario) o non esistono (per esempio i settori *non-profit*).

Si tratta pur sempre di indici di produttività totale dei fattori (*TFPI*) o, meglio, di stimatori della dinamica temporale della *TFP*, ma a differenza di quelli di tipo *non-frontier approach*, sono suscettibili di essere scomposti nelle loro componenti principali, ossia: dinamica della efficienza tecnica *EI* (*catching-up effect* o *technical efficiency change*) e dinamica della tecnologia produttiva *TI* (*technological shift effect* o *technological change*).

⁶⁹ In letteratura sono state proposte alcune varianti di una impostazione stocastica della metodologia *DEA* che non sembra aver trovato un ruolo rilevante in sede empirica. Per dettagli metodologici, *cfr.* (per es.) BANKER R. D., MAINDIRATTA A. (1992), *Maximum Likelihood Estimation of Monotone and Concave Production Frontiers*, *Journal of Productivity Analysis*, vol. 3, pag. 401-415; BANKER R. D. (1993), *Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation*, *Management Science*, vol. 39, n° 10, pag. 1265-1273; RETZLAFF-ROBERT, MOREY (1993); RESTI (2000); FETHI, JACKSON, WEYMAN-JONES (2001).

Con riferimento a ciascuna unità produttiva osservata, le due componenti segnalano il ruolo che nella definizione della propria dinamica della *TFP*, giocano i cambiamenti temporali intervenuti rispettivamente nel proprio livello di efficienza tecnica (efficienza secondo Farrell) e nello stato della comune tecnologia produttiva (progresso/regresso tecnologico).

Bibliografia

- Abramovitz M. (1956). *Resource and Output Trends in the United States since 1870*, American Economic Review, vol. 46, n° 2.
- Aigner D. J., Chu S. F. (1968). *On Estimating the Industry Production Function*, American Economic Review, n° 13, pag. 568-598.
- Aigner D. J., Lovell C. A. K., Schmidt P. (1977). *Formulation and Estimation on Stochastic Frontier Production Function Models*, Journal of Econometrics, n° 6, pag. 21-37.
- Althin R., Fre R., Grosskopf S. (1996). *Profitability and Productivity Change: an Application to Swedish Pharmacies*, Annals of Operations Research, vol. 66, pag. 219-230.
- Balk B. M. (1993). *Malmquist Productivity Index and Fisher Ideal Indexes: Comment*, Economic Journal, vol. 103.
- Banker R. D., Maindiratta A. (1992). *Maximum Likelihood Estimation of Monotone and Concave Production Frontiers*, Journal of Productivity Analysis, vol. 3, pag. 401-415.
- Banker R. D. (1993). *Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation*, Management Science, vol. 39, n° 10, pag. 1265-1273.
- Bauer P. W. (1990). *Decomposing TFP Growth in the Presence of Cost Inefficiency, Nonconstant Returns to Scale and Technological Progress*, Journal of Productivity Analysis, vol. 1, n° 4, pag. 287-301.
- Berg S. A., Førsund F. R., Jansen E. S. (1992). *Malmquist Indices of Productivity Growth during the Deregulation of Norwegian Banking, 1980-89* – in Scandinavian Journal of Economics, vol. 94, Supplement.
- Bjurek H. (1996). *The Malmquist Total Factor Productivity Index*, Scandinavian Journal of Economics, vol. 98, n° 2, pag. 303-313.
- Caves D. W., Christensen L. R., Diewert W. E. (1982). *The Economic Theory of Index Number and the Measurement of Input, Output, and Productivity*, Econometrica, vol. 50, n° 6.
- Cecchi C., Galeotti G. (1990). *I numeri indici. Teoria e Tecnica*, Cacucci, Bari.

- Chambers R., F re R., Grosskopf S. (1994). *Efficiency, Quantity Indexes, and Productivity Indexes: a Synthesis*, Bulletin of Economic Research, vol. 46, n° 1.
- Chambers R., Chung Y., F re R. (1998). *Profit, Directional Distance functions, and Nerlovian Efficiency*, Journal of Optimization Theory and Applications, vol. 98, n° 2, pag. 351-364.
- Charnes A., Cooper W. W., Rhodes E. (1978). *Measuring the Efficiency of Decision Making Units*, European Journal of Operational Research, vol. 2, n° 6, pag. 429-444.
- Christensen L. R., Jorgenson D. W., Lau L. J. (1971). *Coniugate Duality and the Transcendental Logarithmic Function*, Econometrica, vol. 39, pag. 255-256.
- Christensen L. R., Jorgenson D. W., Lau L. J. (1973). *Transcendental Logarithmic Production Frontiers*, Review of Economics and Statistics, vol. 55, pag. 28-45.
- Cobb C. W., Douglas P. H. (1928). *A Theory of Production*, American Economic Review, vol. 18, (Suppl. di marzo), pag. 139-165.
- Coelli T., Prasada Rao D. S., Battese G. E. (1998). *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers, London.
- Coelli T. J. (1998). *A Multi-stage Methodology for the Solution of Oriented DEA Models*, CEPA Working Papers, n° 1, Department of Econometrics, University of New England, Armidale, Australia.
- Debreu G. (1951). *The Coefficient of Resource Utilization*, Econometrica, vol. 19, n° 3, pag. 273-292.
- Deprins D., Simar L., Tulkens H. (1984). *Measuring Labour Efficiency in Post Offices*, in MARCHAND M., PESTEAU P. (Eds.), *The performance of Public Enterprises: Normative, Positive and Empirical Issues*, Amsterdam, North Holland, pag. 243-267
- Destefanis S. (1996). *Un'analisi non parametrica dell'efficienza per le aziende di credito italiane*, Rivista di Politica Economica, vol. 86, n° 11/12, pag. 385-414.
- Diewert W. E. (1976). *Exact and Superlative Index Numbers*, Journal of Econometrics, vol. 4, pag. 115-145.
- Diewert W. E. (1992). *Fisher Ideal Output, Input and Productivity Indexes Revisited*, Journal of Productivity Analysis, n° 3, vol. 3.
- Diewert W. E., Nakamura A. O. (2002). *The Measurement of Aggregate Total Factor Productivity Growth*, Working Paper pubblicato in: DIEWERT W. E., NAKAMURA A. O., HECKMAN J. J., LEAMER E. E. (Eds.) (2005). *Handbook of Econometrics Methods*, Amsterdam, North Holland.
- Druker P. (1974). *Management Task, Responsibilities*, Harper & Row, New York.
- F re R. (1988). *Fundamentals of Production Theory*, Springer-Verlag, New York.
- F re R., Grosskopf S. (1990). *The Fisher Ideal Index and the Indirect Malmquist Productivity Index: A Comparison*, New Zeland Economic Paper, n° 24.

- F re R., Grosskopf S. (1992). *Malmquist Productivity Index and Fisher Ideal Indexes*, Economic Journal, vol. 102, n° 4.
- F re R., Grosskopf S., Norris M., Zhang Z. (1994). *Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries*, The American Economic Review, vol. 84, n° 1, pag. 66-83.
- F re R., Grosskopf S., Lovell C. A. K. (1994). *Production Frontier*, Cambridge University Press, Cambridge.
- F re R., Grosskopf S., Roos P. (1996). *On Two Definition of Productivity*, Economics Letters, vol. 53.
- F re R., Grifell-Tatjé E., Grosskopf S., Lovell C. A. K. (1997). *Biased Technical Change and the Malmquist Productivity Index*, Scandinavian Journal of Economics, vol. 99, n° 1, pag. 119-127.
- F re R., Førsund F. R., Grosskopf S., Hayes K., Hesmati A. (2001). *A note on decomposing the Malmquist productivity index by means of subvector homoteticity*, Economic Theory, vol. 17, pag. 239-245.
- F re R., Primont D. (1995). *Multi-Output Production and Duality: Theory and Applications*, Kluwer Academic Publishers, Boston.
- Farrell M. J. (1957). *The Measurement of Productive Efficiency*, Journal of Royal Statistical Society, Series A, vol. 120, n° 3, pag. 253-290.
- Fecher F., Pestieau P., (1991). *Efficiency and Competition in Financial Services*, in: FRIED H. O., LOVELL C. A. K., SCHMIDT S. S. (Eds.) (1993), *The Measurement of Productive Efficiency*, Oxford University Press, New York, pag. 374-385.
- Fethi M. D., Jackson P. M., Weyman-Jones T. G. (2001). *An Empirical Study of Stochastic DEA and Financial Performance: the Case of the Turkish Commercial Banking Industry*, INFORMS International Hawaii Conference, 17-20/06/2001, Maui, Hawaii, USA.
- Fisher I. (1922). *The making of Index Number*, Houghton-Mifflin, Boston.
- Fisher F. M., Shell K. (1972). *The Economic Theory of Price Indexes*, Academic Press, New York.
- Førsund F. R. (2002). *On the Circularity of the Malmquist Productivity Index – Working Paper*, Department of Economics, University of Oslo, and the Frisch Centre, Oslo.
- Frigero P. (1979). *Per uno studio della produttività: contributi teorici e metodi di ricerca*, Bollettino CERIS-CNR, n° 3.
- Frigero P. (1996). *Metodi e temi di analisi della produttività*, Utet, Torino.
- Frisch R. (1936). *Annual Survey of general Economic Theory: The Problem of Index Numbers*, Econometrica, vol. 50, pag. 1393-1414.
- Grosskopf S. (1993). *Efficiency and productivity*, in FRIED H. O., LOVELL C. A. K., SCHMIDT S. S. (Eds.), *The Measurement of Productive Efficiency*, Oxford University Press, New York, pag. 160-191.

- Grosskopf S. (2002). *Some Remarks on Productivity and its Decompositions*, Working Paper 01/2002, pag. 10 (presentato il 21-22/09/2001 al Exclusive Workshop on DEA, University of Southern Denmark, Odense).
- Guarini R., Tassinari F. (2000). *Statistica economica. Problemi e metodi di analisi*, Il Mulino, Bologna.
- Hulten C. R. (2000). *Total Factor Productivity: A Short Biography*, Working Paper n° 7471, National Bureau of Economic Research, University of Maryland, Washington.
- Hulten C. R. (1973). *Divisia Index Numbers*, *Econometrica*, vol. 41, n° 6, pag. 1017-1025.
- Impicciatore G. (1998). *Introduzione alla moderna microeconomia*, Cedam, Padova.
- Jorgenson D. W., Griliches Z. (1967). *The Explanation of Productivity Change*, *Review of Economic Studies*, vol. 34, n° 3, 249-283.
- Jondrow J., Lovell C. A. K., Materov I. S. Schmidt P. (1982). *On Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model*, *Journal of Econometrics*, n° 19, pag. 233-238.
- Kendrick J. W. (1956). *Productivity trends: Capital and Labour*, *Review of Economics and Statistics*, vol. 39, n°3, pag. 248-257.
- Koopmans T. C. (1951). *An Analysis of Production as an Efficient Combination of Activities*, in KOOPMANS T. C. (Ed.), *Activity Analysis of Production and Allocation*, Cowles Commission for Research in Economics, Monograph n° 13, Wiley & Sons, New York.
- Kopp R. J., Diewert W. E. (1982). *The Decomposition of Frontier Cost Function Deviations into Measures of Technical and Allocative Efficiency*, *Journal of Econometrics*, vol. 19, pag. 319-331.
- Lovell C. A. K. (1993). *Production Frontiers and Productive Efficiency*, in: FRIED H. O., LOVELL C. A. K., SCHMIDT S. S. (Eds.), *The Measurement of Productive Efficiency*, Oxford University Press, New York.
- Malmquist S. (1953). *Index Numbers and Indifference Surfaces*, *Trabajos de Estadística*, vol. 4, pag 209-242.
- Martelli G. (1984). *La produttività nelle aziende di credito*, Giuffrè, Milano.
- Mc Lellan N. (2004). *Measuring Productivity using the Index Number Approach: An Introduction*, Working Paper n° 4, New Zeland Treasury.
- Meeusen W., Van den Broeck J. (1977). *Efficiency Estimation from Cobb-Douglas Production Functions With Composed Error*, *International Economic Review*, n° 18, pag. 435-444.
- Nishimizu M., Page J. M. (1982). *Total Factor Productivity Growth, Technological Progress and Technical Efficiency Change: Dimensions of Productivity Change in Yugoslavia 1965-78*, *Economic Journal*, vol. 92, pag. 920-936.

- Owyong D. T. (2000). *Productivity Growth: Theory and Measurement*, APO Productivity Journal, pag. 19-29.
http://www.apo-tokyo.org/productivity/016_prod.htm
- Pittman R. (1983). *Multilateral productivity Comparisons with Undesirable Output*, Economic Journal, vol. 93, n° 372, pag. 883-891.
- Resti A. (2000). *Efficiency Measurement for Multi-Product Industries: A Comparison of Classic and Recent Techniques Based on Simulated Data*, European Journal of Operation Research, vol. 121, pag. 559-578.
- Retzlaff-Robert D. L., Morey R. C. (1993). *A Goal Programming Method of Stochastic Allocative DEA*, European Journal of Operational Research, vol. 71, n°3, pag. 379-397.
- Samuelson P. A., Swamy S. (1974). *Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis*, American Economic Review, vol. 64, pag. 566-593.
- Shephard R. W. (1953). *Cost and Production Functions*, Princeton University Press, Princeton, NJ.
- Shephard R. W. (1970). *Theory of Cost and Production Functions*, Princeton University Press, Princeton, NJ.
- Solow R. A. (1957). *Technical change and the Aggregate Production Function*, Review of Economics and statistics, vol. 39.
- Tani P. (1986). *Analisi microeconomia della produzione*, Utet, Torino.
- Tulkens H. (Ed.). *The performance of Public Enterprises: Normative, Positive and Empirical Issues*, Amsterdam, North Holland, pag. 243-267.
- Tulkens H. (1993). *On FDH Efficiency Analysis: Some Methodological Issues and Application to Retail Banking, Courts and Urban Transit*, Journal of Productivity Analysis, vol. 4, n° 1, pag. 183-210.
- Vitali O. (1999). *Statistica economica*, Cacucci, Bari.
- Zieschang K. D. (1983). *A Note on the Decomposition of Cost Efficiency into Technical and Allocative Components*, Journal of Econometrics, vol. 23, pag. 401-405.



UNIVERSITÀ
DEGLI STUDI DI BARI
ALDO MORO

DIPARTIMENTO DI
SCIENZE ECONOMICHE E
METODI MATEMATICI

PDF finito di comporre
il 29 dicembre 2016

ISBN 978-88-88793-99-3



9 788888 793993