



SAPIENZA  
UNIVERSITÀ DI ROMA

## Modelli di mortalità con indicatori socioeconomici

Facoltà di Ingegneria dell'informazione, informatica e statistica  
Corso di Laurea Magistrale in Scienze Attuariali e Finanziarie

Candidato

Marco Maddaloni

Matricola 1604100

Relatore

Prof.ssa Susanna Levantesi

Correlatore

Prof.ssa Giulia Rotundo

Anno Accademico 2018/2019

Tesi discussa il 22 Marzo 2019  
di fronte a una commissione esaminatrice composta da:

Prof. Fabio Grasso (presidente)

Prof. Massimo Franchi

Prof. Paolo Giordani

Prof.ssa Susanna Levantesi

Prof.ssa Filonema Racioppi

Prof.ssa Giulia Rotundo

Prof. Claudio Russo

Prof.ssa Giulia Zacchia

---

**Modelli di mortalità con indicatori socioeconomici**

Tesi di Laurea Magistrale. Sapienza – Università di Roma

© 2019 Marco Maddaloni. Tutti i diritti riservati

Questa tesi è stata composta con  $\text{\LaTeX}$  e la classe Sapthesis.

Email dell'autore: [marcomaddaloni.sefa@libero.it](mailto:marcomaddaloni.sefa@libero.it)

# Indice

<b>1</b>	<b>Introduzione</b>	<b>1</b>
<b>2</b>	<b>Analisi storica</b>	<b>4</b>
2.1	Sviluppo storico dei modelli . . . . .	4
2.2	Mortalità e fluttuazioni macroeconomiche . . . . .	11
2.3	Aspettativa di vita differente per uomini e donne . . . . .	22
<b>3</b>	<b>Indicatori socioeconomici e mortalità in Italia</b>	<b>24</b>
3.1	Dati . . . . .	24
3.2	Modello di Lee-Carter . . . . .	35
3.3	Modello di Niu-Melenberg . . . . .	39
3.4	Relazione tra indicatori socioeconomici e mortalità . . . . .	46
3.4.1	Correlazione . . . . .	46
3.4.2	Test di stazionarietà . . . . .	50
3.4.3	Test di cointegrazione . . . . .	52
<b>4</b>	<b>Modelli proposti</b>	<b>54</b>
4.1	Modelli OLS . . . . .	54
4.2	Modello con Indice di sviluppo umano depurato . . . . .	58
4.3	Modello con Spesa sanitaria pubblica sul PIL . . . . .	60
4.4	Modello con Inflazione . . . . .	62
4.5	Modello con Stipendi . . . . .	64
4.6	Misure di fitting . . . . .	66

---

<b>5</b>	<b>Stima dei tassi di mortalità futuri</b>	<b>70</b>
5.1	Forecasting . . . . .	70
5.2	Test Out-of-sample . . . . .	78
<b>6</b>	<b>Conclusione</b>	<b>81</b>
	<b>Bibliografia</b>	<b>85</b>
	<b>Sitografia</b>	<b>89</b>

# Capitolo 1

## Introduzione

Nell'ultimo secolo l'aspettativa di vita è cresciuta enormemente in tutti i Paesi del mondo raggiungendo, per alcune Nazioni, la soglia degli 80 anni.

L'aspetto impressionante di questa crescita è il suo picco in epoca recente. Si è passati da un'aspettativa di vita media di 40/45 anni nella metà del 18esimo secolo ai 70 anni agli inizi del 21esimo secolo.

L'allungamento della aspettativa vita, come affermato anche da molti accademici, quali ad esempio, S.S.Yang et al. (2010), Y.Deng et al. (2012) e V.D'amato et al. (2014) riflette una miglior qualità della vita rispetto alla generazioni precedenti. Il miglioramento di alcuni fattori socioeconomici come educazione, occupazione, stipendi e nutrizione hanno condotto la popolazione ad uno stile di vita più sano.

L'aumento dell'aspettativa di vita è legato anche ad altri fattori, quali ad esempio:

- introduzione e sviluppo dei farmaci;
- sviluppo del sistema sanitario;
- sviluppo tecnologico.

Lo sviluppo tecnologico esponenziale dell'ultimo secolo ha permesso alla medicina di raggiungere risultati che fino a qualche decennio fa sarebbero sembrati impossibili. La disponibilità di macchinari sempre più avanzati, i vaccini e gli investimenti nella ricerca hanno permesso di debellare moltissime malattie che nei secoli passati risultavano devastanti per intere popolazioni.

L'educazione ha portato ad una maggiore consapevolezza ed informazione rispetto

ad uno stile di vita più sano, ad esempio, come testimoniano i dati OCSE, in Italia il consumo di litri di alcolici pro capite dal 1974 ad oggi si è ridotto del 70%.

Il mondo lavorativo è stato totalmente rivoluzionato nell'ultimo mezzo secolo. Il numero di impiegati in lavori deterioranti per la salute dell'uomo è diminuito drasticamente, in favore del crescente numero di lavoratori nel mondo dei servizi e dell'informatica.

Tutti questi fattori, sommati ad altri, hanno condotto ad un incredibile aumento dell'aspettativa di vita, la quale però comporta di conseguenza delle complicazioni. Il numero degli anziani e pensionati presenti nel nostro paese è cresciuto enormemente, il che, incide sulla spesa sanitaria e sulla spesa previdenziale. Negli ultimi decenni questo fenomeno ha dunque richiamato l'attenzione degli "Holders del longevity risk" quali imprese di assicurazioni o fondi pensionistici. Queste imprese necessitano una previsione il più affidabile possibile sui tassi di mortalità futuri.

La previsione dei tassi di mortalità futuri è un problema a cui gli studiosi cercano di dare risposta da secoli e, dato l'interesse crescente di molte istituzioni economiche, nell'ultimo trentennio il numero di modelli proposti è stato molto elevato.

I metodi di previsione dei tassi di mortalità possono essere divisi in tre categorie:

- Aspettative;
- Esplicativi;
- Estrapolativi.

Il primo si basa sulle opinioni degli esperti, il secondo cerca di collegare la dinamica della mortalità con alcuni risk factors ed infine ci sono i metodi estrapolativi, i quali assumono che la mortalità continuerà a seguire i trend passati.

La maggior parte dei modelli per la mortalità sono estrapolativi. In generale questi modelli si concentrano sui dati storici riguardanti la mortalità e tentano di estrapolarne alcuni fattori latenti.

In questa tesi verranno utilizzati il modello di Lee-Carter, il modello Niu-Melenberg e alcuni nuovi modelli da noi proposti. Questi ultimi, come il modello di Niu-Melenberg, possono essere visti come una combinazione di un approccio estrapolativo ed esplicativo. Esplicativo in quanto utilizzeranno alcune grandezze socioeconomiche come indicatori che cattureranno la correlazione fra i trend della mortalità e quelli della

crescita economica ed estrapolativo in quanto i modelli cattureranno il trend della mortalità sulla base di dati storici e da questi trend si stimeranno i tassi futuri di mortalità.

Nell'applicazione dei modelli utilizzeremo i dati italiani sulla mortalità dal 1990 al 2016 per la fascia di età che va dai 40 ai 90 anni per uomini, donne ed entrambi i sessi aggregati.

Non prenderemo in considerazione i tassi di mortalità fra i 0 e i 39 anni in quanto questi sono legati alla mortalità infantile ed alle morti accidentali che potrebbero avere effetti distorsivi per la nostra analisi. Allo stesso modo non vengono prese in considerazione le età più estreme i cui andamenti spesso sono poco prevedibili.

Nella tesi verranno introdotte delle grandezze socioeconomiche e verrà osservata la loro capacità di catturare informazioni rispetto ai trend della mortalità.

Su questi verranno selezionati solo quelli che presenteranno maggiore correlazione con il fattore latente del modello Lee-Carter e con i tassi di mortalità. Verrà poi studiata la loro stazionarietà e la presenza di cointegrazione con il fattore latente del modello di Lee-Carter  $k_t$ .

Basandoci sui risultati ottenuti vengono introdotti dei modelli simili a quelli di Niu-Melenberg dove però il PIL pro capite viene sostituito da questi nuovi indici. Con lo scopo di confrontare i modelli usati viene studiata la bontà di fitting di questi, osservandone il BIC, MAPE e MAD e verrà effettuato un test out-of-sample per il periodo 2008-2016.

Infine, questi modelli verranno poi usati per effettuare delle previsioni sui tassi di mortalità futuri per il ventennio tra il 2017 ed il 2036.

## Capitolo 2

# Analisi storica

### 2.1 Sviluppo storico dei modelli

La stima dei tassi di mortalità futuri è da sempre un argomento di particolare interesse da parte degli accademici di tutto il mondo. Un discreto numero di studi risale al XV secolo. Nel 1693 il famoso astronomo E.Halley(1693) sviluppo la prima tavola di mortalità.

Age	Number	Age	Number	Age	Number
1	1000	29	539	57	272
2	855	30	531	58	262
3	798	31	523	59	252
4	760	32	515	60	242
5	732	33	507	61	232
6	710	34	499	62	222
7	692	35	490	63	212
8	680	36	481	64	202
9	670	37	472	65	192
10	661	38	463	66	182
11	653	39	454	67	172
12	646	40	445	68	162
13	640	41	436	69	152
14	634	42	427	70	142
15	628	43	417	71	131
16	622	44	407	72	120
17	616	45	397	73	109
18	610	46	387	74	98
19	604	47	377	75	88
20	598	48	367	76	78
21	592	49	357	77	68
22	586	50	346	78	58
23	579	51	335	79	50
24	573	52	324	80	41
25	567	53	313	81	34
26	560	54	302	82	28
27	553	55	292	83	23
28	546	56	282	84	20

**Figura 2.1.** Tavola di mortalità della città di Breslau, di E.Halley, An estimate of the degrees of the mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslaw with an attempt to ascertain the price of annuities upon lives(1693).

I dati della tavola vennero raccolti da C. Neumann per la città di Breslau, l'attuale Breslavia in Polonia, che all'epoca era parte dell'impero Asburgico.

Nella tavola mostrata nella figura 2.1 il numero dei sopravviventi di età  $k+1$  veniva calcolato sottraendo il numero degli individui deceduti di età  $k$  ai sopravviventi di età  $k$ :

$$P_{k+1} = P_k - D_k$$

E. Halley utilizzò come popolazione di partenza 1238 persone, che corrispondeva al numero medio di nascite per la città di Breslau.

Si aveva che  $l_0=1238$  e  $l_1=890$  e dunque si calcolava  $L_0$  come:

$$L_0 = l_0 - \frac{1}{2}D_0 = 1238 - \frac{348}{2} = 1064$$

Essendo  $L_0 = 1064$  sembra dunque che E. Halley abbia effettuato una proporzione a 1000 per convenienza.

Successivamente A. De Moivre (1725) propose la prima formula matematica per modellare la mortalità:

$$l_x = k\left(1 - \frac{x}{86}\right) \text{ per } 12 \leq x \leq 86$$

dove  $l_x$  è il numero degli individui ancora in vita all'età  $x$  partendo da una popolazione di dimensione  $l_0$ ,  $k$  è una costante mentre 86 rappresenta l'età massima raggiungibile dagli individui.

Agli inizi del XIX secolo B. Gompertz (1825) propose una nuova legge demografica basata sulla forza di mortalità:

$$\mu_x = \alpha \exp(\beta x)$$

nella formula  $\alpha$  e  $\beta$  sono parametri positivi ed  $x$  rappresenta l'età.

Oltre 30 anni dopo venne introdotta la funzione di Gompertz-Makeham (1860). Il modello proposto prendeva in considerazione una componente di mortalità per causa accidentale ed una componente dovuta all'invecchiamento. La forza di mortalità veniva dunque definita come:

$$\mu_x = \alpha + \beta c^x$$

La forza di mortalità è un concetto tuttora utilizzato in demografia; viene definita anche tasso istantaneo di mortalità e definisce la probabilità che un individuo di età

$x$  muoia fra  $x$  e  $x+dx$ ; si ottiene calcolando il limite del tasso di mortalità quando  $n$  tende a zero:

$$\mu_x = \lim_{n \rightarrow 0} n \bar{m}_x$$

Attraverso l'integrazione di questa forza di mortalità è possibile ottenere anche altre grandezze derivate, come ad esempio la funzione di sopravvivenza:

$$S_x = S_0 \exp\left[-\int_0^x \mu(a) da\right]$$

Negli anni a seguire lo studio dei tassi di mortalità, delle tavole di mortalità e della demografia si è di molto evoluto portando alla formulazione di modelli sempre più complessi e completi.

In epoca recente il numero dei modelli proposti è stato molto elevato, dato l'interesse di molti istituti assicurativi e pensionistici.

La complessità dei modelli è cresciuta per cercare di ottenere un miglior fitting per età elevate e per la mortalità infantile.

L.Heligman e J.H.Pollard(1980) proposero un nuovo modello in cui la mortalità veniva definita da tre termini per un totale di 8 parametri che dovevano essere stimati:

$$m(x) = A^{(x+B)^C} + De^{E(\ln x - \ln F)^2} + \frac{GH^x}{(1 + GH^x)}$$

la prima componente descrive la mortalità infantile, la seconda il crescere della mortalità per l'invecchiamento ed infine il terzo descrive la mortalità accidentale.

Un decennio dopo è stato introdotto il modello di Lee-Carter(1992). Questo è un modello puramente estrapolativo, molto semplice che ha posto le basi per una nuova forma di modellizzazione per i tassi di mortalità:

$$\ln(m_{x,t}) = \alpha_x + \beta_x k_t + \epsilon_{x,t}$$

Nell'ultimo trentennio questa modellizzazione è stata usata da molti accademici come punto di partenza per l'introduzione di nuovi modelli che rappresentavano un perfezionamento del modello di Lee-Carter.

Fra i maggiori studi ricordiamo:

- N. Brouhns et al. (2002);
- A.E.Renshaw e S.Haberman (2006);

- A.J.G.Cairns et al. (2006);
- R.Plat (2009);
- C.O'Hare e Y.Li (2012);
- G.Niu e B.Melenberg (2014).

N.Brouhns, M.Denuit e J.Vermunt proposero di modellizzare i tassi centrali di mortalità con il modello di Lee-Carter, ipotizzando però che i decessi fossero distribuiti come una Poisson:

$$D_{x,t} \sim \text{Poisson}(E_{x,t}m_{x,t})$$

Questa ipotesi risulta essere più realistica del SVD (singular value decomposition) proposta nel modello di Lee-Carter.

Anche altri studiosi successivamente accettarono le ipotesi poste da N. Brouhns et al., ad esempio A.E.Renshaw e S.Haberman, i quali introdussero un modello che prendeva in considerazione anche un fattore di coorte:

$$\ln(m_{x,t}) = \beta_x^1 + \beta_x^2 k_t + \beta_x^3 \gamma_{t-x} + \epsilon_{x,t}$$

$\gamma_{t-x}$  è un effetto coorte funzione dell'anno di nascita  $t-x$ . Il modello proposto mostrò un migliore fitting dei dati storici rispetto al modello di Lee-Carter anche se presentava dei limiti di robustezza.

Nel 2006 A.J.G.Cairns, D.Blake e K.Dowd introdussero un modello in cui veniva considerato un doppio effetto dal tempo:

$$\log\left(\frac{q_{x,t}}{1 - q_{x,t}}\right) = k_t^1 + (x - \bar{x})k_t^2 + \epsilon_{x,t}$$

dove:

$$q_{x,t} \approx 1 - e^{-m_{x,t}}$$

con  $q_{x,t}$  che rappresenta la probabilità per l'individuo di età  $x$  nell'anno  $t$  di morire fra l'età  $x$  ed  $x+1$ .  $(k_t^1, k_t^2)$  si ipotizza che sia una passeggiata aleatoria con drift bivariata, mentre  $\bar{x}$  è l'età media nel campione.

R.Plat nel 2009 prese in considerazione sia l'effetto coorte che l'age-period, introducendo un modello in cui i parametri time-varying erano cinque:

$$\ln(m_{x,t}) = b_x^1 + k_t^1 + (\bar{x} - x)k_t^2 + (\bar{x} - x)^+ k_t^3 + \gamma_{t-x} + \epsilon_{x,t}$$

dove  $(\bar{x} - x)^+ = \max(\bar{x} - x, 0)$  e  $\bar{x}$  è la media delle età considerate..

C.O'Hare e Y.Li nel 2012 proposero un estensione del modello di Plat:

$$\ln(m_{x,t}) = b_x^1 + k_t^1 + (\bar{x} - x)k_t^2 + ((\bar{x} - x)^+ + ([\bar{x} - x]^+)^2)k_t^3 + \gamma_{t-x} + \epsilon_{x,t}$$

In questo modello vi sono più fattori latenti che permettono di catturare in maniera migliore i cambiamenti storici osservati dei tassi di mortalità per differenti età mentre  $\gamma_{t-x}$  modella l'effetto coorte.

Negli ultimi anni alcune delle modifiche al modello di Lee-Carter prevedevano l'introduzione di fattori macroeconomici, ambientali o biomedici per osservarne l'impatto sui tassi di mortalità.

I modelli introdotti che prendono in considerazione anche grandezze osservate non sono stati molti, uno fra i più famosi è sicuramente quello proposto da G.Niu e B.Melenberg. Tal modello è una forma estesa del modello di Lee-Carter in cui viene incorporato anche il PIL pro capite:

$$\ln(m_{x,t}) = \alpha_x + \beta_x k_t + \gamma_x g_t + \epsilon_{x,t}$$

$\alpha_x$ ,  $\beta_x$  e  $k_t$  continuano ad avere il medesimo significato del modello di Lee-Carter;  $g_t$  è il logaritmo del PIL pro capite e  $\gamma_x$  è un parametro che definisce la sensibilità della mortalità al fattore osservato.

Negli ultimi anni infine sono stati introdotti anche modelli molto più complessi che prevedono un approccio multi-popolazione in cui è prevista la divisione dei tassi di mortalità in più popolazioni. Gli  $m_{i,x,t}$  sono dunque i tassi di mortalità della popolazione  $i$  nell'anno  $t$  all'età  $x$ .

La condizione di base nella modellizzazione multi-popolazione è la condizione di "coerenza", cioè, il rapporto fra i tassi di due popolazioni differenti,  $m_{i,x,t}/m_{j,x,t}$  con  $i \neq j$ , deve rimanere stabile.

Un esempio è lo studio di A.J.G.Cairns et al.(2016) con dati relativi alla popolazione maschile Danese dai 55 ai 94 anni fra il 1985 ed il 2012. La popolazione veniva divisa in dieci sottogruppi utilizzando un indice rappresentativo della ricchezza e del reddito:

$$\log(m_{i,x,t}) = \beta_0^{(i)} + k_1^{(i)}(t) + (x - \bar{x})k_2^{(i)}(t)$$

$k_1^{(i)}(t)$  cattura i cambiamenti nel livello della mortalità, mentre  $k_2^{(i)}(t)$  cambia in funzione alla curva della log-mortalità relativa alla baseline.  $\beta_0^{(i)}(x)$  è un parametro che dipende dall'età.

N.Li e R.D.Lee(2005) introdussero un estensione multi-popolazione del modello di Lee-Carter. J.T.Boonen e H.Li(2017), seguendo l'idea introdotta da G.Niu e B.Melenberg, estesero il modello di Li-Lee introducendo il PIL:

$$\log(m_{i,x,t}) = a_{i,x} + \sum_{j=1}^J B_{j,x} K_{j,t} + \sum_{l=1}^L \gamma_{l,x} g_{l,t} + b_{i,x} k_{i,t} + \epsilon_{i,x,t}$$

$$\epsilon_{i,x,t} \sim N(0, \sigma_{\epsilon_{i,x}}^2) \quad i.i.d$$

$a_{i,x}$ ,  $b_{i,x}$  e  $k_{i,t}$  sono i parametri specifici della popolazione,  $K_{j,t}$  è il j-esimo trend,  $g_{l,t}$  è la l-esima componente principale del PIL della I popolazione;  $B_{j,x}$  e  $\gamma_{l,x}$  sono i rispettivi carichi. Il modello include L componenti principali con  $L < I$ , con I numero delle popolazioni. Il modello mostrava un buon fitting in-sample ed una buona capacità di previsione out-of-sample.

Interessante risulta infine essere lo studio effettuato da M.Seklecka et al.(2017) in cui viene studiato l'impatto delle fluttuazioni del PIL sulla mortalità di nove Paesi europei, i quali, vengono divisi in due gruppi in base alla loro crescita economica.

Lo studio presentato fornisce una chiave di lettura molto precisa e attenta permettendo di confrontare fra loro vari modelli introdotti negli anni (Lee-Carter, O'Hare-Li, Niu-Melenberg) con il modello da loro proposto, che risulta essere un estensione del O'Hare-Li:

$$\ln(m_{x,t}) = b_x^1 + k_t^1 + c g_x (\bar{x} - x) k_t^2 + (\bar{x} - x)^+ k_t^3 + ([\bar{x} - x]^+)^2 k_t^4 + \gamma_{t-x} + \epsilon_{x,t}$$

$k_t^{(2)}$  riflette l'osservazione storica del miglioramento dei tassi per età differenti;  $k_t^{(3)}$  cattura gli effetti delle età più giovani e  $k_t^{(4)}$  la presenza di non-linearità. Questo modello, in paragone con il Lee-Carter mostrava una migliore capacità predittiva, osservata attraverso le misure dell'errore MAPE e MAD.

Nella tesi non utilizzeremo un approccio multi-popolazione, ma la relazione fra mortalità e fattori socioeconomici verrà studiata usando come benchmark il modello di Lee-Carter ed il modello di Niu-Melenberg.

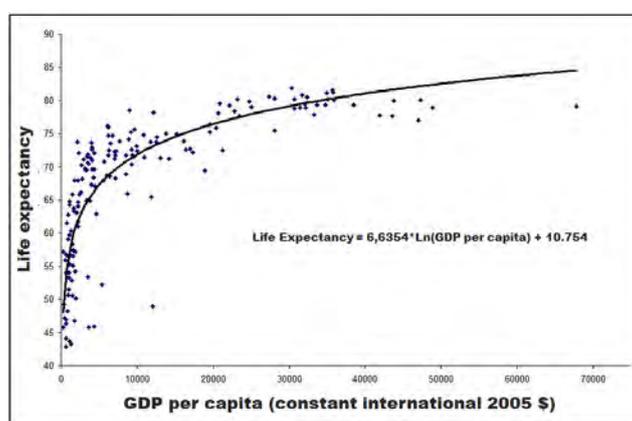
I nuovi modelli introdotti saranno delle varianti del modello Niu-Melenberg, dove,

invece di usare il PIL come fattore macroeconomico utilizzeremo altre grandezze socioeconomiche che risultano essere altamente correlate con il fattore latente del Lee-Carter.

## 2.2 Mortalità e fluttuazioni macroeconomiche

La triplice relazione fra crescita economica, salute e mortalità è stata vastamente studiata nel corso degli anni. Nonostante questo, il dibattito resta ancora aperto, in quanto i risultati sono stati spesso contrastanti.

S.H.Preston(1975) introdusse la famosa curva di Preston, mostrata nella figura 2.2. La curva mostrava empiricamente come Nazioni con un PIL pro capite più elevato avessero, in media, anche una maggiore aspettativa di vita. Questa relazione era più forte per Paesi più poveri e più debole per i Paesi più ricchi.



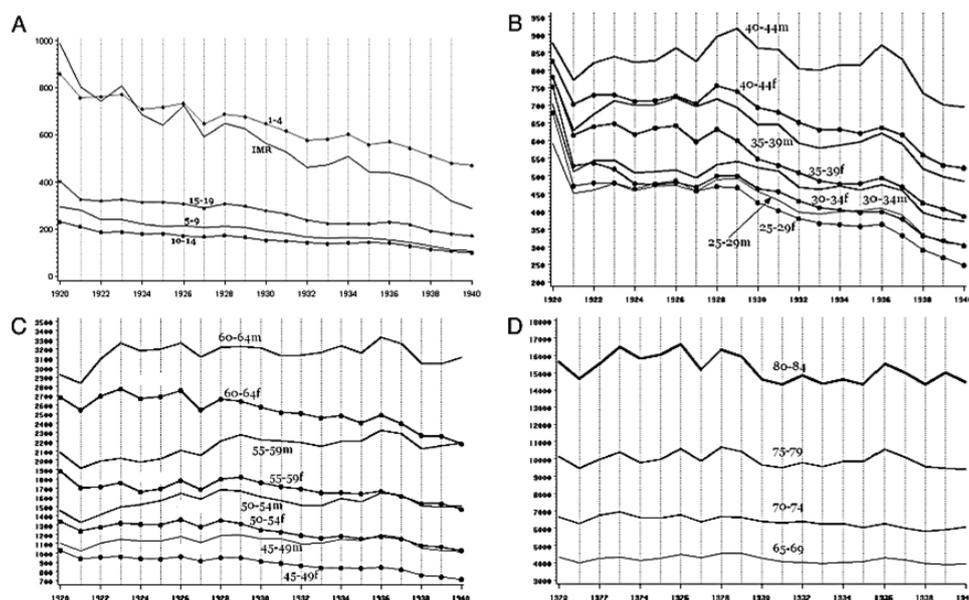
**Figura 2.2.** Curva di Preston, di S.H.Preston, *The Changing Relation between Mortality and Level of Economic Development* (1975).

S.L.Ettner(1996) documentò come un reddito più elevato migliorasse significativamente la salute fisica e psicologica degli individui. A sostegno di queste ipotesi M.H.Brenner(2005) osservò che la crescita economica si rispecchia non solo in un aumento dei salari, ma, stimola anche investimenti in nuove medicine, ospedali, servizi, ricerca, i quali portano conseguenzialmente ad un generale aumento dello stato di salute della popolazione e dunque un aumento dell’aspettativa di vita.

R.Swift(2011) propose uno studio sull’esistenza di cointegrazione fra salute e PIL per 13 Paesi con dati relativi agli ultimi due secoli. Questo studio confutò una relazione tra PIL ed aspettativa di vita per la maggior parte dei Paesi analizzati. D.French e C.O’Hare(2014) trovarono una relazione tra il fattore latente del modello di Lee-Carter ed alcuni indicatori economici e della salute. Il loro studio “ Forecasting

death rates using exogenous determinants” mostrò come l’inclusione nei modelli di alcune variabili esogene alla salute, come per esempio il PIL o la spesa sanitaria conducevano a migliori risultati nella previsione dei tassi di mortalità per i Paesi studiati. Gli accademici però non sono tutti d’accordo rispetto alla relazione tra queste grandezze economiche e l’aspettativa di vita. J.A.Tapia Granados effettuò vari studi per confutare come la crescita economica comporti in realtà un aumento dei tassi di mortalità.

J.A.Tapia Granados e A.V.D.Roux(2009) pubblicarono un articolo che mirava ad osservare l’andamento della mortalità nel ventennio fra il 1920 ed il 1940 negli Stati Uniti. Lo studio mostrò come i tassi di mortalità decrescevano per quasi tutte le età durante i quattro anni della Grande Depressione e crescevano durante i periodi di grandi espansione, come nel 1923,1926,1929 e 1936-37. L’unica eccezione a questo andamento della mortalità risultavano essere le morti per suicidio che rappresentavano però meno del 2% del totale. Negli anni successivi alla prima guerra mondiale gli Stati Uniti si trovavano in un periodo di recessione; nel 1921 la disoccupazione era all’11,3%. Gli anni immediatamente successivi furono segnati da una rapida crescita economica che si stabilizzò fino alla fine degli anni 20 e che portò ad una diminuzione del tasso di disoccupazione attorno al 5%. Il periodo di stabilità economica si concluse con il famoso “Giovedì Nero” nel 24 Ottobre del 1929 in cui la crisi del New York Stock Exchange generò un crollo della borsa dei valori causando effetti devastanti e recessivi per tutti i Paesi industrializzati. Durante la Grande Depressione il commercio internazionale diminuì notevolmente e con esso anche il reddito dei lavoratori. Nel 1932 la disoccupazione raggiunse il suo picco storico del 22,9%. La crisi negli Stati Uniti durò fino alla metà del 1933. Nel biennio fra il 1934-36 vi fu di nuovo una forte crescita economica che fu seguita da una recessione nel 1938. La disoccupazione calò fra il 1933 ed il 1937 per poi ricrescere nel 1938 per la recessione economica. Confrontando l’andamento economico del paese con l’aspettativa di vita, si osserva che, nonostante l’aspettativa di vita nel periodo osservato sia generalmente cresciuta, gli anni di maggiore crescita economica sono stati anni di decrescita per l’aspettativa di vita. Durante la Grande Depressione invece l’aspettativa di vita passò da 57.1 a 63.3 anni.



**Figura 2.3.** Andamenti della mortalità per gruppi di età e sesso negli Stati Uniti tra il 1920 ed il 1940. IMR, mortalità infantile per 10.000 nascite, tassi di mortalità per 100.000 morti, di J.A.Tapia Granados e A.V.D.Roux, *Life and death during the Great Depression* (2009).

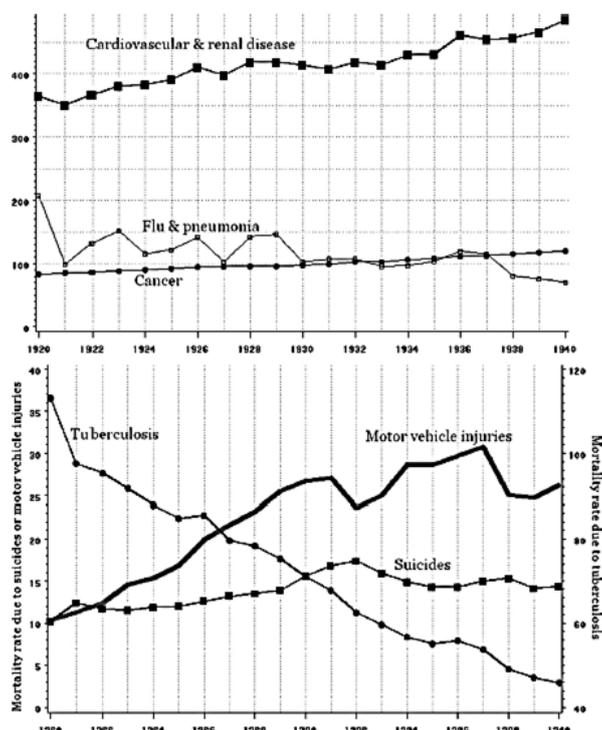
Osservando singolarmente i gruppi di età, mostrati nella figura 2.3, si constatò come i tassi di mortalità infantile e per età inferiori a 20 anni erano decresciuti fra il 1920 ed il 1930 ma aumentavano negli anni di grande crescita economica (1923,1926,1928-29 e 1934-36). La mortalità per uomini e donne dai 25 ai 44 anni presentava un andamento stazionario fra il 1920 ed il 1928 per poi crescere nel biennio 1928-29, anni di forte crescita economica. Successivamente in tutti i gruppi demografici la mortalità ha avuto un andamento decrescente negli anni della Grande Depressione per poi crescere negli anni di sviluppo economico. Per le età più avanzate, 45-64 e 65-84 anni, la mortalità presentava dei picchi negli anni di crescita economica.

Per una maggiore comprensione degli andamenti dei tassi di mortalità vennero prese in considerazione singolarmente le sei maggiori cause di morte:

- Malattie cardiovascolari e renali;
- Polmoniti ed influenza;
- Tubercolosi;

- Cancro;
- Incidenti stradali;
- Suicidi.

Le cause di morte sopraelencate fra il 1920 ed il 1940 erano responsabili di due terzi delle morti.



**Figura 2.4.** Tassi di mortalità per le maggiori cause di morte negli Stati Uniti tra il 1920 ed il 1940, di J.A.Tapia Granados e A.V.D.Roux, *Life and death during the Great Depression* (2009).

Dalla figura 2.4 si osserva che la maggiore causa di morte erano le malattie cardiovascolari e renali. Queste fra il 1920 ed il 1930 crebbero per poi stabilizzarsi durante la crisi fra il 1930 ed il 1932. Le morti per influenza e polmoniti decrebbero nel ventennio, con dei minimi negli anni di Grande Depressione. Le morti per tubercolosi calarono fortemente durante il ventennio in maniera costante. Le morti per cancro aumentarono leggermente senza presentare però picchi dovuti agli anni di crisi o di espansione economica. Per quanto riguarda le morti per incidenti stradali si osservava

un andamento generalmente crescente dovuto alla diffusione delle auto; l'andamento di queste morti è legato molto all'economia in quanto le auto nel ventennio fra il 1920 ed il 1940 erano un bene di lusso ed il numero degli individui che potevano permettersi questo bene naturalmente cresceva negli anni di espansione economica per decrescere negli anni di recessione. Infine, come già detto, le morti per suicidi sono le uniche eccezioni in cui i picchi si presentarono negli anni di crisi per decrescere negli anni di espansione.

La disoccupazione e la crisi economica conducono molte persone in uno stato di povertà e sfiducia che spesso sfocia in problematiche psicologiche quali per esempio isolamento e depressione che da sempre sono catalizzatori per i suicidi.

Osservando poi la correlazione fra l'andamento dell'aspettativa di vita e cambiamenti nella condizione economica si osservò che l'aspettativa di vita era negativamente correlata con la crescita del PIL e positivamente correlata con la crescita della disoccupazione.

		Correlations with	
		GDP growth	Annual change in the unemployment rate
Annual gain in life expectancy at birth	Total	-0.54*	0.64**
	Males	-0.59**	0.69***
	Females	-0.48*	0.58**
	Whites	-0.51*	0.61**
	Non whites	-0.66**	0.72***
	Males	-0.67***	0.73***
	Females	-0.63**	0.68***
	Annual percentage decline in age-specific mortality at ages	Infant mortality (< 1)	-0.40†
1-4		-0.29	0.43†
20-24		-0.54*	0.67**
30-34		-0.40†	0.57*
40-44		-0.53*	0.64**
50-54		-0.56**	0.66**
60-64		-0.39†	0.49*
70-74		-0.28	0.35
80-84		-0.31	0.35

\*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; †,  $P < 0.1$

**Figura 2.5.** Correlazione fra salute, misurata come crescita annuale dell'aspettativa di vita alla nascita, con la crescita annuale del PIL e le variazioni del tasso di disoccupazione negli Stati Uniti fra il 1920 ed il 1940, di J.A.Tapia Granados e A.V.D.Roux, *Life and death during the Great Depression* (2009).

Come mostra la figura 2.5, la relazione fra crescita del PIL e miglioramenti nella salute risultava essere statisticamente significativa per 10 delle 16 categorie osservate mentre, per la relazione tra disoccupazione e miglioramenti nella salute era significativa per 13 delle 16 categorie osservate.

Alcuni risultati simili ottenuti precedentemente da J.A.Tapia Granados(2005) furono però criticati. M.H. Brenner(2005), per esempio, contestò le ipotesi poste da J.A.Tapia Granados osservando che la relazione fra la crescita del PIL e la decrescita della mortalità su un orizzonte medio-lungo è stata più volte studiata e documentata; osservò però che un innalzamento della mortalità negli anni di espansione economica può essere legata ad alcuni fattori precedenti alla crisi, le quali presentano i loro effetti con un lag-temporale.

E.Neumayer(2005) osservò che nonostante sia ovvio che i periodi recessivi e la conseguenziale disoccupazione possa danneggiare psicologicamente gli individui e ridurre la qualità della vita, vi sono varie ricerche scientifiche che attestano una crescita della mortalità negli anni di espansione ed una decrescita negli anni di recessione. Giustificò questi andamenti “inaspettati” notando che in periodi di espansione economica le ore lavorative, lo stress, il consumo di sostanze nocive per la salute ed il numero di incidenti sul lavoro crescono inevitabilmente causando un aumento della mortalità sull'immediato ma non sul lungo periodo.

J.A.Tapia Granados contestò le critiche relative al lag temporale osservando che, in relazione al periodo della Grande Depressione, queste non potevano essere critiche soddisfacenti, in quanto il picco di mortalità avvenuto nel 1936, quattro anni dopo la crisi economica, si presentava anche per bambini con età inferiore ai quattro anni; inoltre, l'aumento di morti causato da incidenti stradali non poteva essere legato ai precedenti anni di recessione economica.

J.A.Tapia Granados e A.V.D.Roux osservarono come la crescita economica sia legata a molti fattori che incidono negativamente sulla salute umana e che possono condurre a malattie croniche:

- Aumento del consumo di alcol;
- Aumento del consumo di tabacco;
- Riduzione delle ore di sonno;
- Aumento dello stress;
- Aumento inquinamento.

In conclusione dunque, dai risultati mostrati, si può osservare come la crescita del PIL e decrescita della disoccupazione in un contesto di lunga durata possono migliorare la

ricchezza individuale, la qualità della vita, la nutrizione ed altri fattori che incidono favorevolmente sulla aspettativa di vita, ma, osservando gli effetti a breve termine si nota che una rapida crescita economica può porre sotto stress una popolazione che non è capace ad assorbire in maniera sana la rapida crescita e dunque questo può condurre a dei picchi di mortalità in anni in cui il benessere economico del Paese è in crescita.

K.Hanewald(2011) tentò di spiegare le dinamiche della mortalità studiandone la relazione con le fluttuazione macroeconomiche e le cause di morte.

K.Hanewald usò i dati dal 1950 al 2006 di sei Paesi appartenenti all'OECD: Australia, Canada, Giappone, Paesi Bassi, Regno unito, Stati Uniti.

Sui tassi di mortalità di questi Paesi implementò il modello di Lee-Carter per entrambi i sessi e per differenti gruppi di età (0-99,25-44,65-84, 85-99).

La correlazione fra i fattori latenti  $k_t$  stimati e gli indici macroeconomici risultò significativa per il 20,8% dei casi. I risultati vengono mostrati nella figura 2.6.

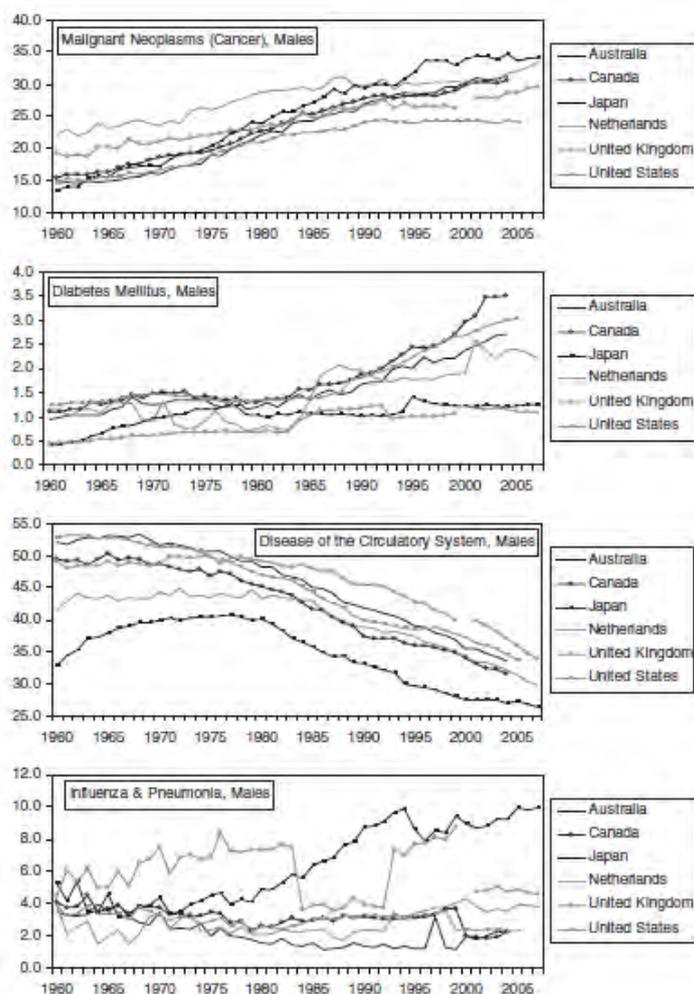
Age Group	Australia	Canada	Japan	Netherlands	United Kingdom	United States
Real GDP Growth Rates (1950–2006)						
Males						
25–44	0.228 <sup>+</sup>	0.217	-0.451 <sup>***</sup>	0.093	0.294 <sup>*</sup>	0.194
45–64	0.338 <sup>*</sup>	0.447 <sup>***</sup>	-0.215	0.329 <sup>*</sup>	0.134	0.323 <sup>*</sup>
65–84	0.350 <sup>**</sup>	0.267 <sup>*</sup>	0.101	0.263 <sup>+</sup>	0.041	0.281 <sup>*</sup>
85–99	0.272 <sup>*</sup>	0.075	0.089	-0.027	-0.025	0.261 <sup>+</sup>
0–99	0.365 <sup>**</sup>	0.391 <sup>**</sup>	-0.114	0.218	0.073	0.309 <sup>*</sup>
Females						
25–44	0.259 <sup>+</sup>	-0.145	-0.528 <sup>***</sup>	-0.036	0.230 <sup>+</sup>	0.171
45–64	0.182	0.316 <sup>*</sup>	-0.233 <sup>+</sup>	0.076	0.115	0.158
65–84	0.250 <sup>+</sup>	0.018	0.158	0.106	0.083	0.250 <sup>+</sup>
85–99	0.265 <sup>*</sup>	0.047	0.191	-0.003	0.043	0.331 <sup>*</sup>
0–99	0.281 <sup>*</sup>	0.177	-0.051	0.085	0.076	0.254 <sup>+</sup>
Unemployment Rate Changes (1960–2006)						
Males						
25–44	-0.327 <sup>*</sup>	-0.065	0.387 <sup>**</sup>	-0.076	-0.289 <sup>+</sup>	-0.182
45–64	-0.124	-0.072	-0.013	0.062	0.054	-0.249 <sup>+</sup>
65–84	-0.248 <sup>+</sup>	0.128	-0.155	0.128	0.075	-0.156
85–99	-0.269 <sup>+</sup>	0.058	-0.155	-0.002	0.053	-0.124
0–99	-0.239	0.053	-0.037	-0.010	0.058	-0.226
Females						
25–44	-0.162	-0.004	0.113	-0.030	-0.223	-0.349 <sup>*</sup>
45–64	-0.152	-0.216	-0.062	-0.191	-0.023	-0.193
65–84	-0.264 <sup>+</sup>	0.118	-0.169	-0.254	0.012	-0.268 <sup>+</sup>
85–99	-0.350 <sup>*</sup>	-0.053	-0.238	-0.079	0.015	-0.309 <sup>*</sup>
0–99	-0.297 <sup>*</sup>	-0.018	-0.147	-0.225	-0.006	-0.305 <sup>*</sup>

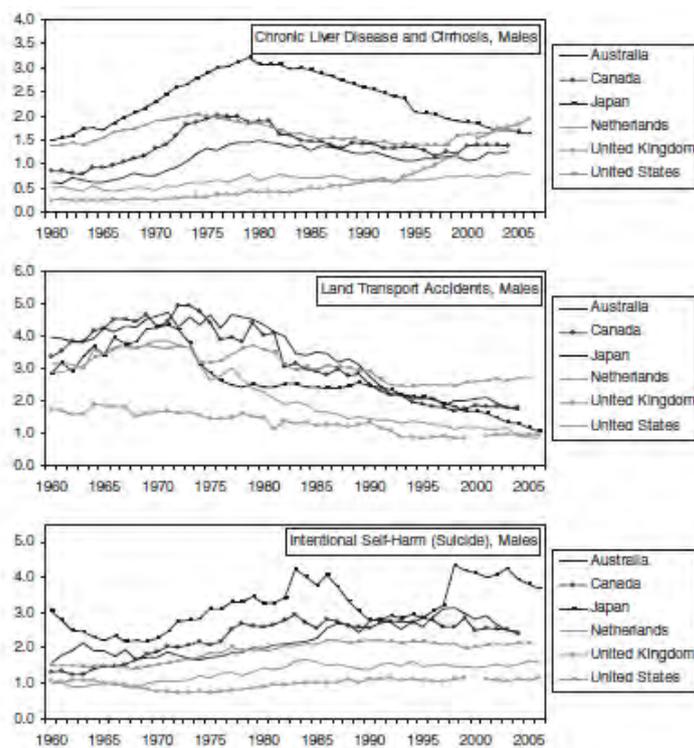
Note: Unemployment rates for the Netherlands are available for 1973–2006 only. <sup>+</sup>  $p < 0.10$ , <sup>\*</sup>  $p < 0.05$ , <sup>\*\*</sup>  $p < 0.01$ , <sup>\*\*\*</sup>  $p < 0.001$ .

**Figura 2.6.** Correlazione per ogni Paese, sesso e gruppo di età fra la prima differenza del  $k_t$  con il PIL e la disoccupazione fra il 1950 ed il 2006, di K.Hanewald, Explaining mortality dynamics: the role of macroeconomic fluctuations and cause of deaths trends (2011).

L'analisi venne poi estesa per le sette maggiori cause di morte:

- Cancro;
- Diabete;
- Malattie cardiovascolari;
- Influenza;
- Malattie del fegato;
- Incidenti stradali;
- Suicidi.





**Figura 2.7.** Percentuale delle morti per le maggiori cause di morte degli uomini per sei differenti Paesi fra il 1960 ed il 2006, di K.Hanewald, *Explaining mortality dynamics: the role of macroeconomic fluctuations and cause of deaths trends* (2011).

I grafici nella figura 2.7 fanno riferimento ai trend per gli uomini, ma i risultati per le donne sono molto simili. Le morti per malattie cardiovascolari e per le malattie al fegato decrescono a partire dal 1980, mentre quelle per incidenti stradali calarono dal 1970; le morti per diabete e cancro presentano un trend crescente; le morti per influenza e suicidi presentano andamenti molto differenti tra i Paesi.

Dai grafici si può dunque osservare come le percentuali per causa di morte siano molto cambiate dal 1960 al 2006. Nel 1960 il 35-55% delle cause di morte era dovuto a malattie cardiovascolari, nel 2006 solo il 25-40% delle morti era causato da queste malattie; sono cresciute le percentuali per morti di cancro passando dal 15-20% delle morti del 1960 al 20-35% del 2006.

La figura 2.8 mostra la correlazione fra le variazioni percentuali nelle specifiche cause di morte e tassi di crescita del PIL. Si osserva che sia per uomini che per donne si ha in generale una correlazione positiva fra la crescita dei tassi del PIL con le morti per

incidenti stradali e con malattie cardiovascolari. I risultati mostravano una differente relazione fra le specifiche cause di morte e le fluttuazioni macroeconomiche.

Cause	Australia	Canada	Japan	Netherlands	United Kingdom	United States
Males						
Cancer	0.226	0.141	0.283 <sup>+</sup>	0.113	0.077	0.079
Diabetes	0.029	0.088	0.276 <sup>+</sup>	-0.082	0.095	0.181
Circulatory disease	0.417 <sup>**</sup>	0.196	0.331 <sup>*</sup>	0.142	-0.031	0.311 <sup>*</sup>
Influenza	0.051	-0.096	-0.037	0.112	0.065	0.211
Liver disease	0.133	0.379 <sup>*</sup>	0.547 <sup>***</sup>	0.144	0.252 <sup>+</sup>	0.359 <sup>*</sup>
Transport accident	0.488 <sup>***</sup>	0.606 <sup>***</sup>	0.581 <sup>***</sup>	0.413 <sup>**</sup>	0.314 <sup>*</sup>	0.692 <sup>***</sup>
Suicide	0.025	-0.089	-0.361 <sup>*</sup>	-0.027	-0.697	-0.198
Females						
Cancer	-0.099	-0.003	0.181	-0.110	0.127	0.050
Diabetes	0.046	-0.172	0.370 <sup>*</sup>	0.031	0.110	0.118
Circulatory disease	0.390 <sup>**</sup>	-0.072	0.376 <sup>*</sup>	0.083	0.038	0.334 <sup>*</sup>
Influenza	0.093	-0.220	-0.019	0.145	0.068	0.237
Liver disease	0.046	0.211	0.333 <sup>*</sup>	-0.145	0.254 <sup>+</sup>	0.347 <sup>*</sup>
Transport accident	0.442 <sup>**</sup>	0.483 <sup>***</sup>	0.581 <sup>***</sup>	0.338 <sup>*</sup>	0.231	0.676 <sup>***</sup>
Suicide	0.032	0.191	-0.212	0.224	0.009	0.034

Note: "Circulatory disease" denotes the mortality rate from diseases of the circulatory system and "Transport accident" denotes land transport accidents. <sup>+</sup>  $p < 0.10$ , <sup>\*</sup>  $p < 0.05$ , <sup>\*\*</sup>  $p < 0.01$ , <sup>\*\*\*</sup>  $p < 0.001$ .

**Figura 2.8.** Correlazione tra cambio percentuale dei tassi di mortalità delle specifiche cause di morte e la crescita del PIL tra il 1960 ed il 2006 per uomini, donne e sei differenti Paesi, di K.Hanewald, *Explaining mortality dynamics: the role of macroeconomic fluctuations and cause of deaths trends* (2011).

Testando la cointegrazione fra le specifiche cause di morte e le fluttuazioni macroeconomiche con il fattore latente del Lee-Carter si ottennero dei risultati molto eterogenei tra loro.

In generale il numero di cointegrazioni significative risultò troppo alto per rifiutare l'ipotesi di una relazione casuale fra il fattore latente e le fluttuazioni macroeconomiche o le cause di morte, però, i risultati furono troppo differenti per permettere di tracciare una conclusione unica per le differenti età e Paesi.

K.Hanewald mostrò che non esiste una soluzione generale ed unica "Broad-Brush solution" che permetta di modellizzare la mortalità, in quanto questa risulta essere infatti molto differente fra sessi, gruppi di età e Paesi.

La relazione che lega la mortalità con le fluttuazioni macroeconomiche tende a modificarsi nel tempo in base alle percentuali di composizione delle cause di morte, in quanto queste reagiscono in maniera differente alle oscillazioni economiche.

In conclusione K.Hanewald osservò che gli studi futuri dovrebbero concentrarsi su alcuni fattori, come i miglioramenti nella medicina, fattori ambientali o educazionali, stili di vita ed osservarne la loro evoluzione nel tempo ed il loro impatto sulla morta-

lità. Sulla base di queste osservazioni decidiamo dunque di introdurre dei modelli che prendano in considerazione delle grandezze socioeconomiche direttamente impattanti sulla salute degli individui e dunque sulla loro aspettativa di vita.

## 2.3 Aspettativa di vita differente per uomini e donne

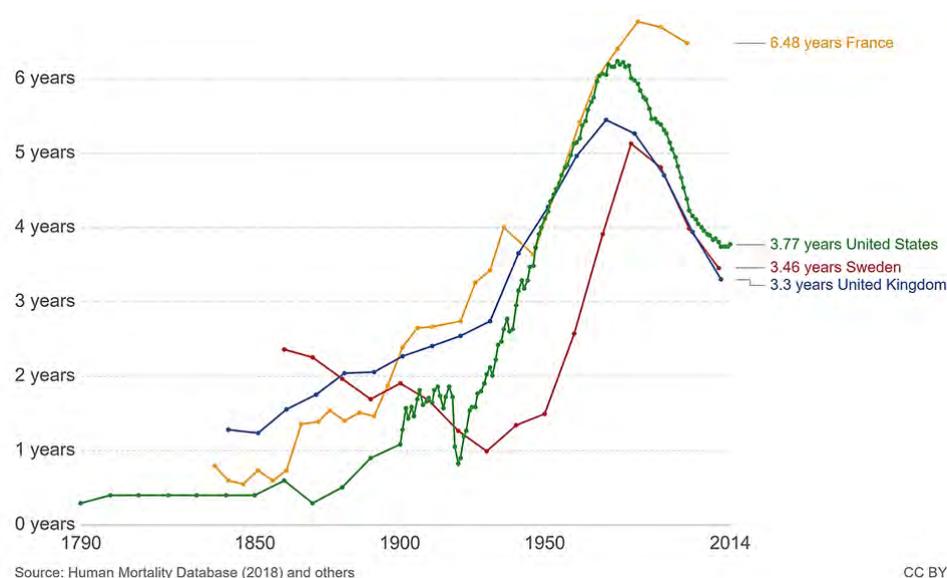
Il fatto che le donne in media vivano più degli uomini è un risultato empirico demograficamente accettato.

L'istat ha stimato che per l'anno 2017 la speranza di vita alla nascita per gli uomini era di 80.6 anni mentre per le donne era di 84.9 anni, dunque con una differenza di 4.3 anni.

Una differenza fra uomini e donne in termini di aspettativa di vita vi è sempre stata, ma questo rapporto si è molto trasformato nel tempo.

Da uno studio effettuato da E.Ortiz-Ospina e D.Beltekian (2018) si mostra come la differenza in questa aspettativa di vita sia evoluta nel tempo dal 1800 al 2014 per Svezia, Regno Unito, Francia e Stati Uniti.

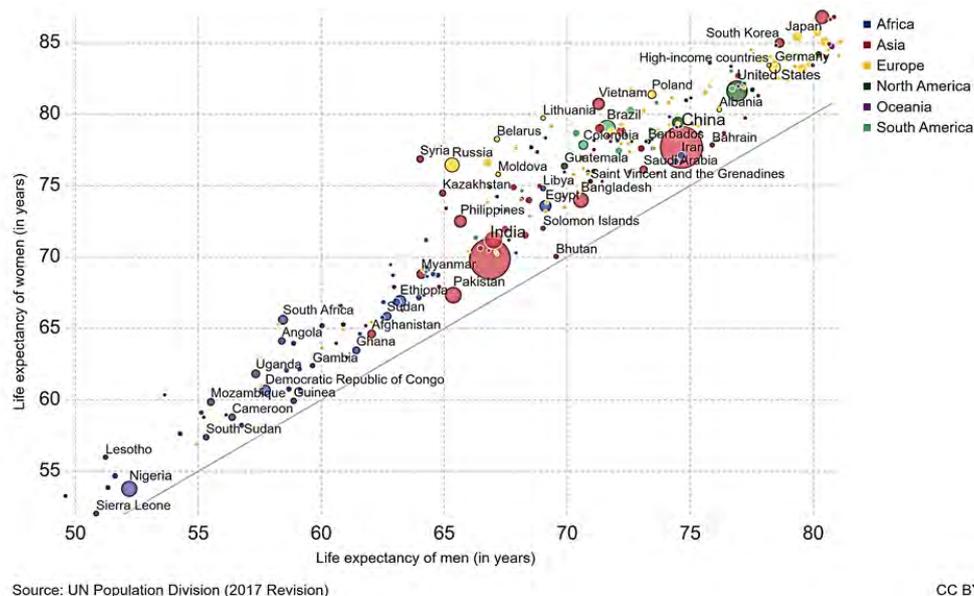
La figura 2.9 mostra che una differenza nell'aspettativa di vita fra uomini e donne già era presente nei primi nell'ottocento ma, in epoca recente, questa differenza si è evoluta raggiungendo una differenza fra i 4 ed i 6 anni per i Paesi osservati da E.Ortiz-Ospina e D.Beltekian.



**Figura 2.9.** Differenza tra l'aspettativa di vita degli uomini e delle donne di 45 anni di età per Francia(giallo), Stati Uniti(verde), Svezia(rosso) e Regno Unito(blu) dal 1790 al 2014, di E.Ortiz-Ospina e D.Beltekian, Why do women live longer than men? (2014).

I due accademici mostrarono anche come nei Paesi di tutto il mondo l'aspettativa di vita delle donne sia superiore di quella degli uomini.

La figura 2.10 mostra la differenza nell'aspettativa di vita fra uomini e donne nel 2015 per vari Paesi e continenti:



**Figura 2.10.** Confronto fra aspettativa di vita alla nascita di donne e uomini nel 2015 per differenti Paesi, di E.Ortiz-Ospina e D.Beltekian, *Why do women live longer than men?* (2014).

Il motivo di questa differenza non è stato ancora definito con chiarezza, le ipotesi poste sono:

- Comportamentali, per esempio il maggiore consumo di sigarette da parte degli uomini;
- Lavorative, attività più rischiose o dannose per il fisico per gli uomini;
- Genetica, presenza di due cromosomi X sembrerebbe poter essere uno dei fattori principali, nel caso di una mutazione genetica nel cromosoma X le donne possono compensare con il secondo cromosoma X;
- Ormonale, gli estrogeni facilitano l'eliminazione del colesterolo LDL.

## Capitolo 3

# Indicatori socioeconomici e mortalità in Italia

### 3.1 Dati

Nella tesi prenderemo in considerazione i dati Istat sulla mortalità per la popolazione italiana di età compresa tra i 40 ed i 90 anni per il periodo dal 1990 al 2016. Nelle modellizzazioni studieremo per entrambi i sessi e per il totale della popolazione la relazione tra alcuni indicatori socioeconomici ed i tassi di mortalità.

I dati relativi a queste grandezze socioeconomiche sono stati raccolti dall'OECD (Organisation for Economic Co-operation and Development), dall'IMF (International Monetary Fund), dall'Istat e dall'Human Development Reports.

Le grandezze socioeconomiche studiate sono:

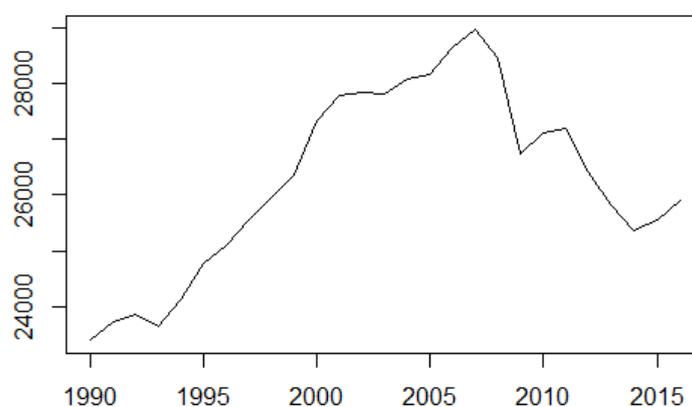
- PIL pro capite;
- indice di sviluppo umano depurato
- tasso di disoccupazione;
- debito pubblico su PIL;
- spesa sanitaria pubblica sul PIL;
- inflazione;
- stipendi italiani lordi.

Il PIL, o prodotto interno lordo, è una grandezza macroeconomica che misura il valore di mercato aggregato di tutte le merci finite e di tutti i servizi prodotti nei

confini di una nazione in un dato periodo di tempo.

Il termine prodotto è riferita ai beni e servizi che hanno un valore in un processo di scambio. Si definisce interno in quanto l'indice prende in considerazione solo le attività economiche svolte all'interno del paese, escludendo beni e servizi prodotti dagli operatori nazionali, imprese e lavoratori all'estero mentre sono inclusi i prodotti realizzati da lavoratori ed imprese esteri all'interno del Paese. Il termine lordo indica che il valore della produzione è al lordo degli ammortamenti, ossia non tiene conto del deprezzamento dello stock di capitale fisico intervenuto nel periodo.

Il PIL pro capite è una misura macroeconomica che consente di operare confronti tra aree di dimensione demografica differente e viene calcolato rapportando il PIL, espresso ai prezzi di mercato, alle dimensioni della popolazione. Nella tesi quando parleremo di PIL faremo riferimento sempre al PIL pro capite.



**Figura 3.1.** Pil pro capite italiano dal 1990 al 2016, Fonte Istat.

Utilizzando il modello di Niu-Melenberg sui recenti dati italiani, mostrati nella figura 3.1, bisogna prendere in considerazione che il PIL italiano dal 2008 ad oggi ha subito una forte decrescita, causata dalla crisi economica. Questo andamento del PIL potrebbe comportare una diminuzione della capacità predittiva del modello.

L'HDI, Human Development Index, è un indice sviluppato dall'indiano Amartya Sen

e dal Pakistano Mahbub ul Haq(1990). Questa grandezza socioeconomica dal 1993 viene utilizzata assieme al PIL dall'Organizzazione delle Nazioni Unite per valutare la qualità della vita nei Paesi membri.

Fino al 2009 l'indice veniva calcolato come media aritmetica dell'indice di aspettativa di vita, di un indice del PIL pro capite e dell'indice di educazione, che a sua volta era composto da livello di istruzione degli adulti e dall'indice lordo di iscrizione scolastiche.

$$\text{Indice di aspettativa di vita} = \frac{AV-25}{85-25}$$

$$\text{Indice PIL pro capite} = \frac{(\log(PIL_{pc}) - \log(100))}{(\log(40000) - \log(100))}$$

$$\text{Indice di educazione} = \frac{2}{3}LIA + \frac{1}{3}ILLI$$

$$\text{Livello di istruzione degli adulti(LIA)} = \frac{(TIA-0)}{(100-0)}$$

$$\text{Indice lordo di iscrizioni scolastiche(ILLI)} = \frac{(TCLI-0)}{(100-0)}$$

dove AV è l'aspettativa di vita, TIA è il tasso di istruzione degli adulti, TCLI è il tasso combinato di iscrizioni scolastiche ed infine PILpc è il PIL pro capite alla parità del potere d'acquisto in dollari.

Dal 2010 l'Indice di Sviluppo Umano, viene definito come la media geometrica dell'indice di aspettativa di vita, indice del livello di istruzione e indice di reddito.

$$HDI = \sqrt[3]{I_{Health} * I_{Education} * I_{Income}}$$

dove:

$$I_{Health} = \text{indice di aspettativa di vita(IAV)} = \frac{AV-20}{85-20}$$

$$I_{Education} = \text{indice di istruzione(II)} = \frac{IAMI + IAPI}{2}$$

$$\text{indice anni medi di istruzione(IAMI)} = \frac{AMI}{15}$$

$$\text{indice anni previsti di istruzione(IAPI)} = \frac{API}{18}$$

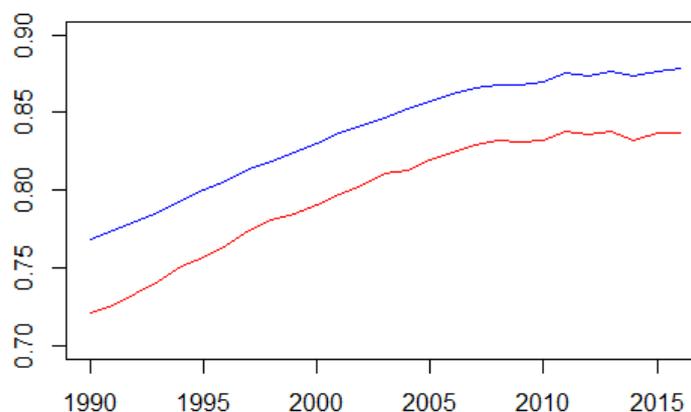
$$I_{Income} = \text{indice di reddito(IR)} = \frac{(\log(RNL_{pc}) - \log(100))}{(\log(75000) - \log(100))}$$

dove  $AMI$  denota gli anni medi di istruzione,  $API$  gli anni previsti di istruzione,  $RNL_{pc}$  è il reddito nazionale lordo pro capite e  $\log(x)$  è il logaritmo naturale.

Essendo l'HDI una combinazione anche dell'aspettativa di vita potremmo ottenere risultati alterati da questo legame con i tassi di mortalità. Per depurare questo indice da questa componente bisogna calcolare l'indice dell'aspettativa di vita ed eliminare il suo apporto. L'indice depurato è una media geometrica di due indici:

$$\text{indice di sviluppo umano depurato} = \sqrt{I_{Education} * I_{Income}}$$

Osservando graficamente nella figura 3.2 i due indici a confronto si nota che l'eliminazione della componente legata all'aspettativa di vita non ha modificato l'andamento della serie, la quale continua a risultare monotona crescente, con una crescita che si è arrestata negli anni recenti.

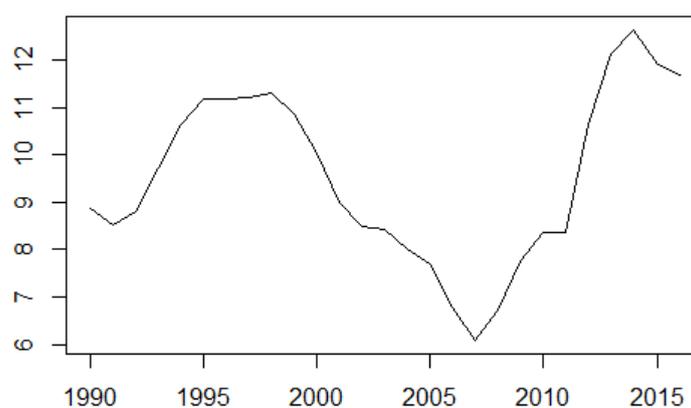


**Figura 3.2.** Indice di sviluppo umano(blu) ed indice di sviluppo umano depurato(rosso) in Italia dal 1990 al 2016, Fonte Human development reports.

La relazione tra tasso di disoccupazione e aspettativa di vita è stata argomento di dibattito per molti accademici. Alcuni articoli mostrarono empiricamente, per esempio J.A.Tapia Granados e A.V.D.Roux(2009), come storicamente si sia osservato più volte che la diminuzione del tasso di disoccupazione portasse ad un aumento dei

tassi di mortalità. Questi risultati vennero però contestati sulla base di un'assenza di un lag-temporale e sulla necessità di prendere in considerazione anche la durata della disoccupazione. Osservando periodi in cui vi è una veloce alternazione fra periodi di forte crescita e recessione, come gli Stati Uniti fra il 1920 ed il 1940, nonostante si presentino periodi con un elevatissimo tasso di disoccupazione, questi danneggiano in maniera limitata la popolazione in quanto non perdurano nel tempo. Il lavoro è una condizione necessaria sia a livello economico che psicologico per un individuo. Periodi di elevata disoccupazione possono inizialmente ridurre i tassi di mortalità riducendosi il numero di malattie croniche ed incidenti connessi all'attività lavorativa, però, a lungo termine, un tasso di disoccupazione elevato riduce la qualità della vita per molti individui, danneggiandone conseguenzialmente lo stato di salute. Il tasso di disoccupazione viene calcolato rapportando il numero di persone che sono in cerca di lavoro con il numero di persone che hanno un lavoro.

$$\text{tasso di disoccupazione} = \frac{\text{persone in cerca di lavoro}}{\text{forza lavoro}}$$



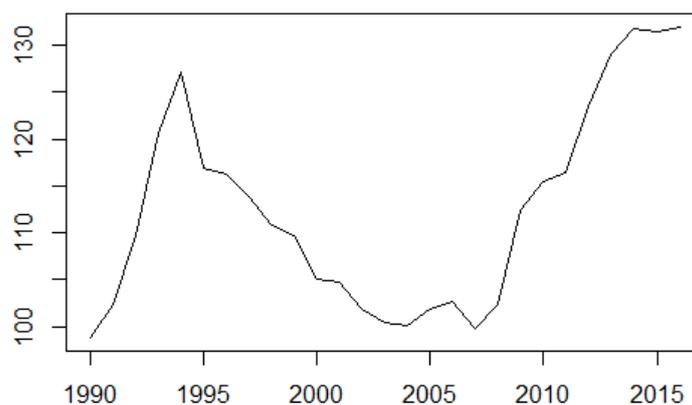
**Figura 3.3.** Tasso di disoccupazione in Italia dal 1990 al 2016, Fonte OECD.

Come si mostra nella figura 3.3 ad oggi l'Italia presenta un elevato problema di disoccupazione ed in particolar modo di disoccupazione giovanile. La politica in

quest'ambito ha provato a fornire risposte anche se tutt'oggi il mercato italiano non sembra in grado di inglobare la grande mole di giovani che non riescono a trovare un lavoro. Dai primi anni duemila al 2008 il tasso di disoccupazione è calato fino a toccare la soglia del 6,7%. Successivamente però, la crisi mondiale e la recessione economica, hanno ridotto enormemente il numero di posti di lavoro ed vi è stata un'impennata del tasso di disoccupazione fino al 12%.

Il debito pubblico è il debito dello Stato nei confronti di altri soggetti economici nazionali o esteri.

Il debito pubblico rapportato al PIL è un importante indice di solidità economica e finanziaria di un Paese. Il rapporto con il PIL permette di identificare in che misura il Paese è indebitato rispetto alla propria ricchezza interna. Questa misura di debito è molto legato alla crescita del paese ma anche alla politica, che può essere più o meno restrittiva. La figura 3.4 mostra l'andamento del debito pubblico sul PIL dal 1990 al 2016.



**Figura 3.4.** Debito pubblico su PIL in Italia dal 1990 al 2016, Fonte IMF.

Il debito pubblico/PIL presenta un primo picco fra il 1990 ed il 1995, successivamente, si osserva una forte decrescita determinata da una diminuzione del debito pubblico ed una crescita del PIL. Nel 2008, con l'avvento della grande Crisi Economica, il PIL

ha iniziato a calare e il debito Pubblico a risalire. Il rapporto è passato così in soli 8 anni dal 100% al 130%.

Mortalità e spesa sanitaria presentano una doppia relazione in quanto, un aumento della aspettativa di vita conduce ad un invecchiamento generale della popolazione e dunque una conseguenziale crescita della spesa sanitaria; inoltre un aumento della spesa sanitaria dovrebbe condurre ad un miglioramento delle strutture ospedaliere o della qualità dei servizi offerti, permettendo così cure migliori e più accessibili impattando dunque in maniera positiva sulla salute e l'aspettativa di vita degli individui. Il sistema sanitario italiano è costituito su tre pilastri:

- Sanità pubblica;
- Sanità integrativa;
- Sanità privata.

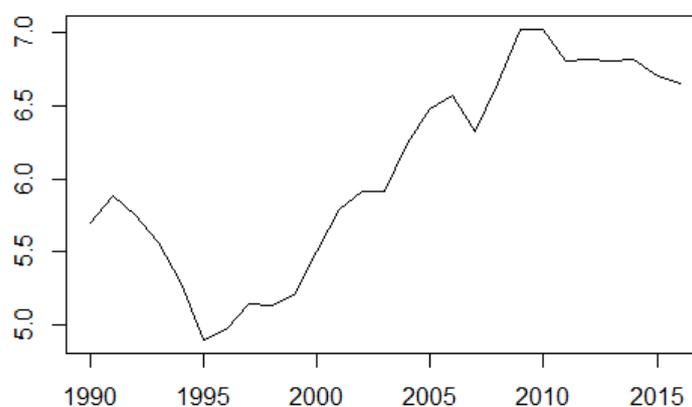
Con sanità privata si fa riferimento ai servizi offerti da strutture ed enti private. La sanità integrativa è una copertura che si pone a metà fra pubblico e privato ed ha un ruolo complementare e/o supplementare rispetto al sistema sanitario Nazionale. La sanità pubblica fa riferimento alla copertura offerta dallo stato italiano ai propri cittadini. Il sistema sanitario pubblico italiano è un sistema di carattere “universalistico”, sulla base di quanto definito dall'art. 32 della Costituzione Italiana che sancisce il diritto alla salute di tutti gli individui.

La sanità pubblica viene finanziata dallo Stato attraverso la fiscalità generale e le entrate percepite attraverso ticket sanitari o prestazioni a pagamento. D.French e C.O'Hare (2014) proposero di prendere in considerazione nella modellizzazione della mortalità anche fattori impattanti direttamente sulla salute degli individui come la spesa sanitaria. Lo studio di questa grandezza socioeconomica presenta però dei limiti legati ai differenti sistemi sanitari presenti nei vari Paesi. Confrontando ad esempio il sistema sanitario universalistico italiano con quello statunitense si osserva che il ruolo della spesa sanitaria privata è totalmente differente.

Il sistema statunitense si basa in maniera maggiore sulla sanità privata. In due sistemi sanitari di questo tipo si ha che periodi recessivi o espansivi hanno conseguenze differenti. Il sistema sanitario statunitense è molto più legato alla ricchezza individuale ed un aumento del tasso di disoccupazione ridurrebbe inevitabilmente

il numero di individui che possono far fronte alle spese mediche. Se considerassimo invece, il miglioramento della qualità della vita in funzione di un aumento della spesa sanitaria pubblica americana, osserveremmo dei risultati alterati in quanto una maggiore spesa sanitaria pubblica impatta solamente su chi usufruisce della sanità pubblica. Dunque, in un sistema come quello statunitense sarebbe maggiormente corretto osservare l'aumento della spesa sanitaria totale e non solamente quella pubblica. Nel caso italiano, basandoci su un sistema "universalistico", prenderemo in considerazione la spesa sanitaria pubblica e non la spesa sanitaria totale in quanto la prima rappresenta maggiormente l'impegno dello stato nella sanità italiana.

I cittadini italiani fanno perlopiù riferimento al sistema pubblico, anche se, ritardi e inefficienze negli anni hanno portato ad una crescita dell'utilizzo di altri strumenti sanitari come la sanità integrativa o le polizze private.



**Figura 3.5.** Spesa sanitaria pubblica sul PIL in Italia dal 1990 al 2016, Fonte OECD.

La figura 3.5 mostra nei primi anni, fra il 1990 al 1995, un calo della spesa sanitaria pubblica sul PIL, con la perdita quasi di un punto percentuale; successivamente si osserva una crescita della spesa sanitaria dal 1995 al 2010 per poi stabilizzarsi negli ultimi anni. Questa crescita mostra dunque un maggiore investimento nel sistema sanitario rispetto agli anni precedenti.

Questo andamento può essere dunque spiegato da più fattori:

- Un invecchiamento della popolazione con conseguente aumento della spesa sanitaria;
- Uno sviluppo tecnologico che comporta inevitabilmente un aumento dei costi per macchinari all'avanguardia e la necessità dunque di un personale sempre più specializzato;
- Un aumento della spesa pubblica in ambito sanitario per sanare le inefficienze ed offrire un servizio migliore.

In macroeconomia l'inflazione è l'aumento prolungato del livello medio generale dei prezzi di beni e servizi in un dato periodo di tempo, che genera una diminuzione del potere d'acquisto della moneta. Nella figura 3.6 si osserva l'inflazione Italiana per il periodo che va dal 1990 al 2016. L'inflazione mostra un andamento decrescente che passa dal 6% allo 0,1%.



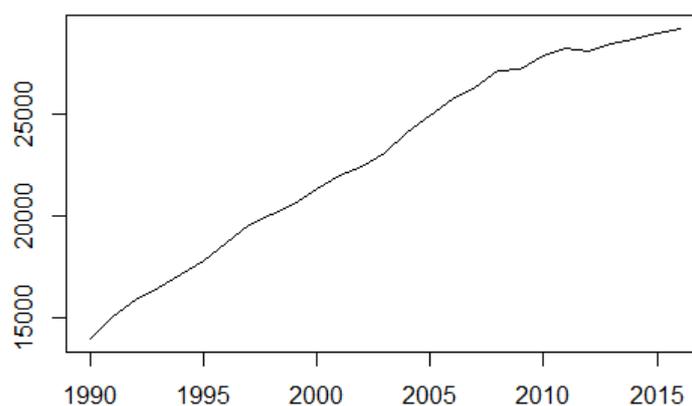
**Figura 3.6.** Inflazione in Italia dal 1990 al 2016, Fonte OECD.

L'inflazione presenta un trend decrescente già a partire dalla fine degli anni 70. Negli ultimi 35/40 anni si è infatti passati da un'inflazione attorno all'20% a valori molto prossimi allo zero. Il medesimo trend si è osservato per tutte le economie avanzate.

Nel perseguimento della stabilità dei prezzi, la BCE si prefigge di mantenere il tasso d'inflazione su livelli inferiori ma prossimi al 2 per cento su un orizzonte di medio periodo. Una situazione, come quella attuale, in cui i tassi di inflazione sono troppo prossimi allo zero rischia di tradursi in una situazione di deflazione. La deflazione viene definita, in macroeconomia, come una diminuzione generale dei prezzi che genera un incremento del potere d'acquisto della moneta.

La riduzione dei prezzi genera aspettative di un'ulteriore flessione, famiglie e imprese dunque reagiscono posticipando gli acquisti non indispensabili. Questo atteggiamento tende ad accentuare il declino dell'attività economica generando una reazione a catena che comporta la chiusura di molte imprese ed il conseguenziale aumento dei tassi di disoccupazione. In risposta a questo periodo di crisi le imprese sono costrette ad un continuo abbassamento dei prezzi innescando dunque un'inevitabile spirale deflazionistica.

Nella figura 3.7 viene mostrato come i salari dagli anni 90 ad oggi siano passati da 13.000 ai 29.000 euro attuali.



**Figura 3.7.** Stipendi italiani dal 1990 al 2016, Fonte OECD.

L'effetto del cambio di moneta sugli stipendi italiani non porta ad effetti distorsivi. L'aumento dei salari naturalmente porta un miglioramento della qualità della vita

per le famiglie italiane. Una maggiore disponibilità economica infatti rende beni e servizi molto più fruibili, migliorando in generale il benessere degli individui e dunque andando ad impattare su molti fattori che conducono ad un generale allungamento dell'aspettativa di vita.

In passato gli stipendi sono già stati proposti come una grandezza socioeconomica che possa impattare lo stato di salute degli individui (S.Gaille e M.Sherris(2010) e A.Roy(2012)). Nel 2016 A.J.G.Cairns et al. (2016) hanno proposto un modello multipopolazione in cui gli individui vengono divisi in dieci gruppi in base alla ricchezza e allo stipendio. La divisione in gruppi si basa sull'ipotesi che la ricchezza degli individui e i loro stipendi influiscono sui tassi di mortalità.

## 3.2 Modello di Lee-Carter

Il modello di Lee-Carter è un modello estrapolativo introdotto da Ronald D.Lee e Lawrence R.Carter in un articolo pubblicato sul Journal of the American Statistical Association nel 1992. Il modello utilizza come dati di partenza una matrice di tassi centrali storici di mortalità  $m_{x,t}$  ed ha lo scopo di stimare una matrice di tassi futuri di mortalità. I tassi centrali di mortalità all'età  $x$  ed anno  $t$  sono calcolati rapportando il numero dei decessi degli individui di età  $x$  nell'anno  $t$  con gli esposti al rischio di età  $x$  nell'anno  $t$ :

$$m_{x,t} = D_{x,t}/E_{x,t}$$

Il modello di Lee-Carter è definito attraverso la seguente equazione:

$$\ln(m_{x,t}) = \alpha_x + \beta_x k_t + \epsilon_{x,t}$$

oppure in forma alternativa con:

$$m_{x,t} = \exp(\alpha_x + \beta_x k_t + \epsilon_{x,t})$$

$\alpha_x$  è la media aritmetica rispetto al tempo del logaritmo di ciascun tasso specifico di mortalità. Questo parametro presenta naturalmente un andamento crescente, dovuto all'aumento della probabilità di sperimentare l'evento morte ad età sempre più elevate.  $k_t$  è un indice di mortalità nel tempo, fornisce informazioni su come il fenomeno della mortalità si sia evoluto nel tempo.  $\beta_x$  è un parametro di sensibilità, ci da informazioni di come i tassi centrali di mortalità all'età  $x$  cambino in funzione di  $k_t$ . Infine  $\epsilon_{x,t}$  rappresenta le variazioni che non sono catturate dalla parte sistematica del modello. L'errore si ipotizza omoschedastico con media 0 e varianza  $\hat{\sigma}_\epsilon^2$ . Per la stima dei parametri del modello è necessario introdurre due vincoli:

$$\sum_t k_t = 0$$

$$\sum_x \beta_x = 1$$

L'equazione di Lee-Carter può essere stimata usando la decomposizione ai valori singolari (SVD). Per far sì che le morti stimate combacino con il numero di morti

osservate nel dato anno  $t$  si impone che  $D(t) = \sum_x D_{x,t}$ . R.D.Lee e L.R.Carter proposero di correggere  $k_t$  dopo il primo step di stima in questa maniera:

$$D(t) = \sum_x (E_{x,t} \exp(\hat{\alpha}_x + \hat{\beta}_x k_t))$$

dove  $\hat{\alpha}_x$  e  $\hat{\beta}_x$  sono parametri stimati attraverso SVD. Con SVD in algebra lineare si intende una particolare fattorizzazione di una matrice basata su autovalori ed autovettori.

In questo modello per quanto riguarda la previsione della mortalità, questa si basa sui parametri stimati e la proiezione futura di  $k_t$ . Generalmente il fattore latente  $k_t$  viene modellato attraverso un  $\text{arima}(0,1,0)$ , cioè, un passeggiata aleatoria con drift:

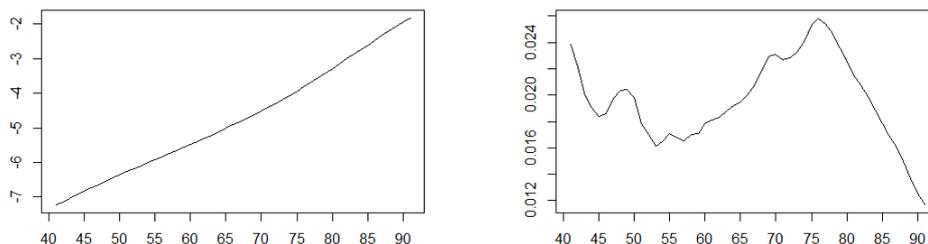
$$k_t = \theta + k_{t-1} + \delta_t$$

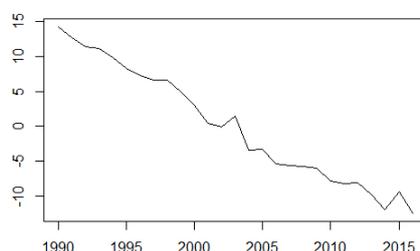
dove  $\theta$  è il termine di drift e  $\delta_t$  il white noise.

Il Lee-Carter è un modello estrapolativo in quanto non tende in alcun modo ad inglobare informazioni mediche, sociali o comportamentali, ma tutte le considerazioni sono basate su come il fenomeno si è evoluto nel tempo.

Nella nostra analisi stimeremo i parametri del modello di Lee-Carter usando le funzioni presenti nel pacchetto “LifeMetrics” introdotto da G.Coughlan et al. (2007). Nell’applicazione del modello si useranno le matrici degli esposti al rischio e dei tassi di mortalità italiani per ogni età dai 40 ai 90 anni per il periodo dal 1990 al 2016. Uomini e donne saranno trattati sia separatamente che congiuntamente.

La figura 3.8 mostra i risultati ottenuti per la stima dei parametri del modello Lee-Carter per le donne Italiane di età fra i 40 ed i 90 anni:





**Figura 3.8.** Parametri  $\alpha_x$ ,  $\beta_x$  e  $k_t$  del modello di Lee-Carter stimati sui tassi di mortalità italiani delle donne dai 40 ai 90 anni dal 1990 al 2016.

Il  $k_t$  mostra un trend decrescente, il quale descrive l'andamento dei tassi di mortalità nel periodo osservato.  $\alpha_x$  presenta una crescita costante che descrive dunque l'aumentare del rischio di morire in funzione dell'età.  $\beta_x$  invece presenta un andamento oscillatorio che tende a decrescere per le età più avanzate. La decrescita del  $\beta_x$  è legata alla dimensione campionaria sempre minore, la quale porta ad una difficoltà del fattore latente a spiegare l'andamento dei tassi di mortalità.

Se applicassimo il modello per la totalità delle età osserveremmo per l' $\alpha_x$  una leggera decrescita nei primi anni, dovuta probabilmente alla mortalità infantile ed un picco attorno ai 20/30 anni di età dovuto verosimilmente agli incidenti stradali in giovane età.

I grafici dei parametri sia per gli uomini che per unisex presentano andamenti molto simili a quelli osservati per le donne.

Il modello di Lee-Carter è uno dei modelli estrapolativi più noti ed ampiamente accettati all'interno della comunità scientifica; nonostante ciò il modello presenta però dei limiti:

- non è in grado di spiegare cosa causi i trend;
- non è in grado di affermare se questi trend continueranno.

Dal 1992, anno della sua pubblicazione, ad oggi molti accademici hanno presentato vari modelli che partendo dal modello di Lee-Carter tentavano di ottenere risultati migliori e di superare i limiti di questo modello.

Molti di questi modelli hanno tentato di unire un approccio esplicativo, introdu-

cendo variabili che potessero spiegare i trend dei tassi di mortalità, con l'approccio estrapolativo tipico del modello di Lee-Carter.

### 3.3 Modello di Niu-Melenberg

Il modello di Niu-Melenberg è un modello stocastico per la mortalità che fu introdotto da Geng Niu e Bertrand Melenberg nel 2014.

Il modello è una delle estensioni più famose e accettate del modello di Lee-Carter, in quanto, partendo da quest'ultimo ha introdotto anche un fattore macroeconomico osservabile, il logaritmo del PIL pro capite:

$$\ln(m_{x,t}) = \alpha_x + \beta_x k_t + \gamma_x g_t + \epsilon_{x,t}$$

I parametri del modello si dividono, come nel Lee-Carter, in parametri invariabili nel tempo, dunque  $\alpha_x$ ,  $\beta_x$  e  $\gamma_x$  ed il parametro “time-varying”  $k_t$ .  $\epsilon_{x,t}$  è il termine di errore, il quale risulta essere incorrelato con  $k_t$  e  $g_t$ . I parametri hanno interpretazione simile a quella dei parametri del Lee-Carter.  $\alpha_x$  è la media aritmetica rispetto al tempo del logaritmo di ciascun tasso specifico di mortalità. Per far sì che  $\alpha_x$  abbia questo andamento bisogna sottrarre al fattore osservato la propria media prima della stima dei parametri.  $k_t$  è un indice che fornisce informazioni su come il fenomeno della mortalità si sia evoluto nel tempo.  $\beta_x$  è un parametro di sensibilità, ci fornisce informazioni su come i tassi centrali di mortalità cambino in funzione di  $k_t$ . L'unico nuovo parametro da stimare è il  $\gamma_x$ , il quale misura la sensibilità dei tassi di mortalità all'età  $x$  al  $g_t$ , cioè al PIL pro capite misurato in forma logaritmica. Se nel modello di Niu-Melenberg si imponesse  $\gamma_x = 0$  si otterrebbe il modello di Lee-Carter. Ipotizzando che  $\mu_{x,t}$  definisca la parte sistematica del  $\ln(m_{x,t})$  e supponendo dati un set di parametri:

$$\theta = (\alpha_1, \dots, \alpha_X, \beta_1, \dots, \beta_X, k_1, \dots, k_T, \gamma_1, \dots, \gamma_X)$$

I parametri non possono essere identificati senza vincoli addizionali in quanto per ogni scalare  $c$ , e ed  $d \neq 0$  si ha che:

$$\begin{aligned} \mu_{x,t} &= \alpha_x + \beta_x k_t + \gamma_x g_t \\ &= \alpha_x + \beta_x (k_t - e g_t) + (\gamma_x + e \beta_x) g_t \\ &= (\alpha_x - \beta_x c) + \frac{\beta_x}{d} \{d(k_t - e g_t + c)\} + (\gamma_x + e \beta_x) g_t \\ &= \tilde{\alpha}_x + \tilde{\beta}_x \tilde{k}_t + \tilde{\gamma}_x g_t \end{aligned}$$

dove,

$$\tilde{\alpha}_x = \alpha_x - \beta_x c$$

$$\tilde{\beta}_x = \frac{\beta_x}{d}$$

$$\tilde{k}_t = d(k_t - eg_t + c)$$

$$\tilde{\gamma}_x = \gamma_x + e\beta_x$$

I vincoli necessari per la stima dei parametri sono i seguenti:

$$\sum_t k_t = 0$$

$$\sum_x \beta_x = 1$$

$$\text{cov}(k_t, g_t) = 0$$

$$k = (k_1, \dots, k_t) \neq 0$$

Il terzo vincolo fa riferimento alla covarianza campionaria. Il quarto è un vincolo ex-post. Non è incluso nella fase di stima ma è necessario per identificare  $\beta_x$ .

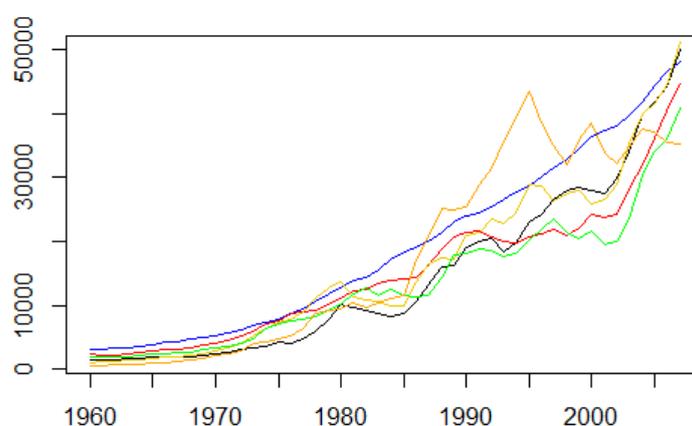
Questi vincoli permettono di identificare i parametri univocamente, come viene dimostrato dal seguente teorema:

dato  $\mu = (\mu_{x,t}, x = 1, \dots, X, t = 1, \dots, T)$ , dove  $\mu = \mu(\theta)$  rispetta  $\mu_{x,t} = \alpha_x + \beta_x k_t + \gamma_x g_t$  per alcune  $\theta$ . Dunque la parametrizzazione di  $\theta^o$ , dove  $\sum_{x=1}^X \beta_x^o = 1$ ,  $\sum_{t=1}^T k_t^o = \theta$ ,  $\text{cov}(k_t^o, g_t) = \theta$  e  $k^o \neq \theta$  soddisfa le seguenti ipotesi:

- $\theta^o$  è una funzione di  $\theta$ .
- $\mu$  è una funzione di  $\theta$  attraverso  $\theta^o$ .
- La parametrizzazione di  $\mu$  è esattamente identificata mediante  $\theta^o$ . Se  $\theta^1 \neq \theta^2$  sono due parametri che soddisfanno i vincoli, avremmo che  $\mu(\theta^1) \neq \mu(\theta^2)$ .

Nel 2014, anno dell'introduzione del modello, la relazione fra crescita economica ed aspettativa di vita era stata già ampiamente studiata, con risultati in parte contraddittori. G.Niu e B.Melenberg furono fra i primi a non limitarsi ad uno studio della relazione tra il fattore latente del modello di Lee-Carter e gli indicatori macroeconomici, ma, introdussero un modello stocastico per la mortalità che include un fattore latente ed un fattore osservato, ed ha lo scopo di interpretare e prevedere i tassi di mortalità. La pubblicazione nella quale venne introdotto il modello è intitolata "Mortality Decrease and economic Growth". In questo testo, utilizzando i dati di sei Paesi industrializzati (Stati Uniti, Regno Unito, Paesi Bassi, Canada,

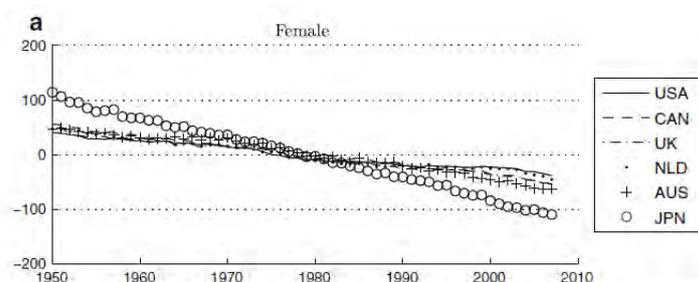
Australia e Giappone) per il periodo dal 1950 al 2007, venne studiata la correlazione e la relazione a lungo termine fra la mortalità ed il PIL. Il PIL dei Paesi osservati nella seconda metà del XX secolo presentava un andamento fortemente crescente. La figura 3.9 mostra l'andamento del PIL pro capite dei Paesi studiati da G.Niu e B.Melenberg fra il 1960 ed il 2007:



**Figura 3.9.** PIL pro capite Stati Uniti(blu),Regno Unito(nero), Paesi Bassi(giallo), Canada(rosso), Australia(verde) e Giappone(arancione) dal 1960 al 2007, Fonte World Bank Open Data.

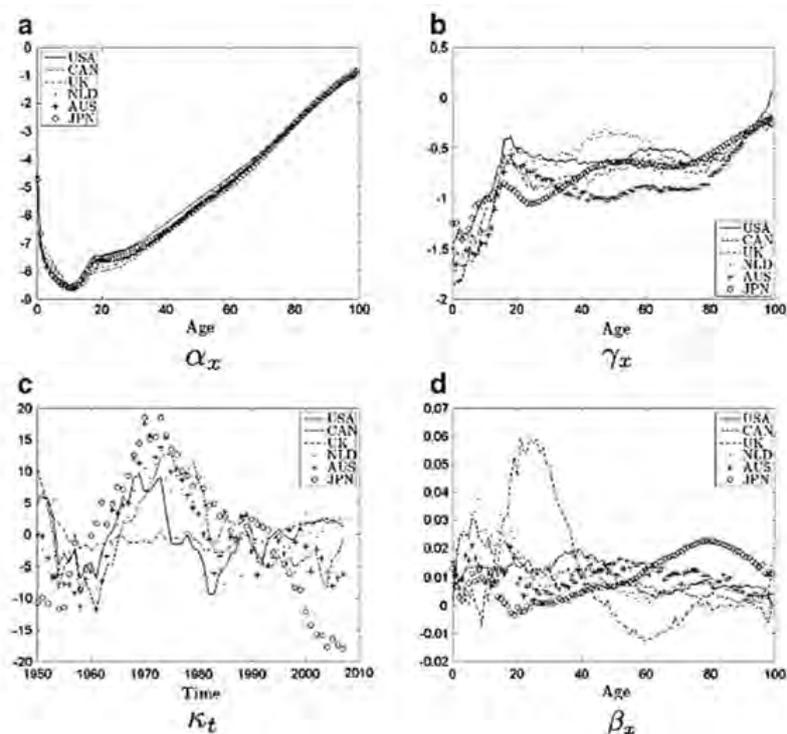
Per lo studio dell'andamento dei tassi di mortalità gli studiosi applicarono il modello di Lee-Carter ai dati dei Paesi osservati. Il  $k_t$ , come mostrato nella figura 3.10, presentava per tutti i Paesi un andamento decrescente e la varianza spiegata dal modello risultava essere molto elevata. Nel paper veniva poi testata la stazionarietà delle serie, per verificare il rispetto della condizione di cointegrazione la quale prevede che le serie cointegrate debbano essere dello stesso ordine. I risultati ottenuti dal Phillips-Perron e Kwiatkowski-Phillips-Schmidt-Shin mostravano che le serie erano integrate di ordine(1). Sulla base dei test di stazionarietà venne effettuato l'Engle-Granger test per la cointegrazione e come analisi supplementare il Johansen test. I risultati furono misti ma in generale mostrarono una possibile cointegrazione fra

l'indice economico e la mortalità.



**Figura 3.10.** Fattore latente  $k_t$  per le donne di Stati Uniti, Canada, Regno Unito, Paesi Bassi, Australia e Giappone dal 1950 al 2007, di G.Niu e B.Melenberg, Trends in Mortality Decrease and Economic Growth(2014).

In base ai risultati ottenuti venne così introdotto il modello di Niu-Melenberg. I valori stimati per i parametri possono essere osservati nella figura 3.11:



**Figura 3.11.** Parametri  $\alpha_x, \gamma_x, \beta_x$  e  $k_t$  del modello originale di Niu-Melenberg, di G.Niu e B.Melnberg, Trends in Mortality Decrease and Economic Growth (2014).

$\alpha_x$  presenta un andamento crescente molto simile a quello che si osservava per il modello di Lee-Carter.  $\beta_x$  varia molto fra i vari Paesi.  $\gamma_x$  è negativa e presenta un andamento molto simile nei vari Paesi, valori più alti nei primi anni di età per poi stabilizzarsi fino agli 80 anni con valori che vanno, approssimativamente, da -0,5 al -1. Dopo gli 80 anni si presenta una leggera crescita, dovuta probabilmente alla riduzione della popolazione esposta e dunque un aumento della variabilità dell'evento morte. Infine però la maggiore differenza si presenta nel fattore latente  $k_t$ . Nel modello Lee-Carter questo presentava un andamento decrescente che rappresentava il trend della decrescita della mortalità. In questo modello  $k_t$  non presenta questo trend e non sembrerebbe essere in grado di catturare l'effetto del tempo sui tassi di mortalità.

Dopo aver implementato il modello ne venne dunque calcolata la capacità di fitting e confrontata con il Lee-Carter. La misura utilizzata è stata il BIC definito come:

$$BIC = N \log(\hat{\sigma}_\epsilon^2) + v \log(N)$$

dove  $\hat{\sigma}_\epsilon^2$  è la varianza stimata del termine di errore,  $v$  è la differenza fra il numero di parametri e quello dei vincoli ed  $N$  è il numero delle osservazioni. I risultati mostrarono un miglior fitting del modello introdotto rispetto al Lee-Carter nella quasi totalità dei Paesi.

Successivamente venne anche confrontata la capacità predittiva del modello attraverso un test out-of-sample. Usando i dati dal 1950 al 1999 stimarono i tassi di mortalità fra il 2000 e il 2007.

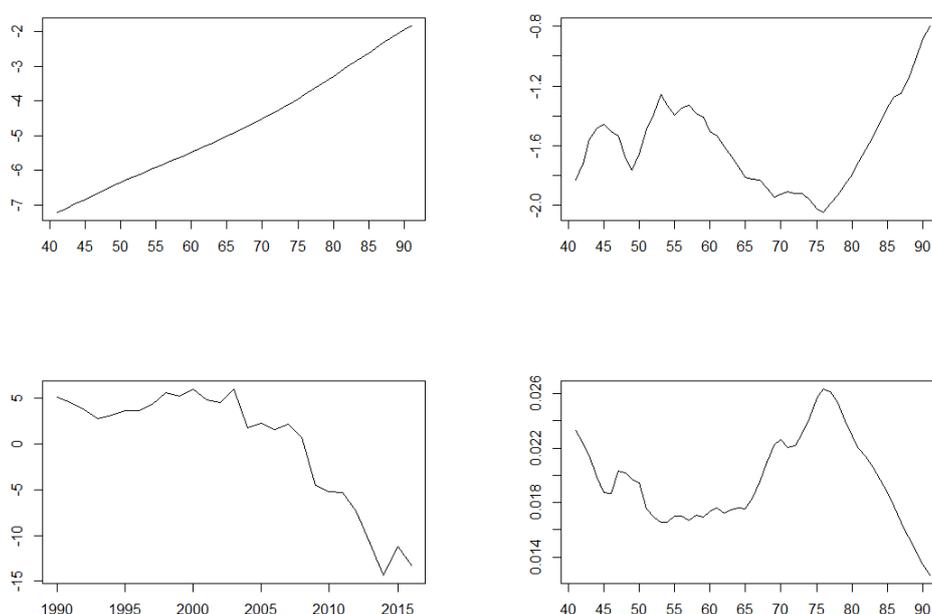
La previsione dei tassi di mortalità futuri venne effettuata proiettando i valori futuri del  $g_t$  e del fattore latente per quanto riguarda il modello di Niu-Melenberg e proiettando il fattore latente per quanto riguardava il modello di Lee-Carter. Le previsioni sono state comparate usando il root mean squared forecast error (RMSFE):

$$RMSFE = \sum_t \sum_i (y_{it} - \hat{y}_{it})^2$$

dove  $y_{it}$  sono i tassi di mortalità osservati all'età  $i$  nell'anno  $t$  mentre  $\hat{y}_{it}$  sono i corrispondenti tassi stimati. RMSFE mostrò nella maggior parte dei casi una maggiore capacità predittiva del Niu-Melenberg rispetto al Lee-Carter. Infine il modello introdotto venne utilizzato per stimare i tassi futuri di mortalità dal 2008 al 2037. I

risultati mostravano una crescita dell'aspettativa di vita per tutti i Paesi, con un'età massima raggiunta dalle donne del Giappone.

Stimando i valori dei parametri del modello di Niu-Melenberg sui nostri dati si utilizzerà il logaritmo del PIL pro capite Italiano dal 1990 al 2016 e si prenderanno in considerazione solo gli individui di età dai 40 ai 90 anni. I valori stimati dei parametri per le donne vengono osservati nella figura 3.12:



**Figura 3.12.** Parametri  $\alpha_x$ ,  $\gamma_x$ ,  $k_t$  e  $\beta_x$  del modello di Niu-Melenberg stimati sui tassi di mortalità italiani per donne di età dai 40 ai 90 anni dal 1990 al 2016.

Rappresentando graficamente i parametri si osserva che  $\alpha_x$  ha un andamento monotono crescente; prendendo in considerazione semplicemente le età fra i 40 e i 90 anni non si osserva la decrescita iniziale dovuto alla mortalità infantile e il picco attorno ai 20 anni dovuto agli incidenti stradali.  $\gamma_x$  è sempre negativa ha un andamento oscillatorio per la fascia di età che va dai 40 ai 60 anni per poi decrescere per le età fra i 60 ed i 75 anni ed infine crescere per le età più elevate.  $\beta_x$  ha un andamento oscillatorio con un picco verso gli 80 anni per poi calare fortemente nelle età più avanzate.

$k_t$  presenta un andamento costante per i primi 15 anni per poi decrescere negli ultimi

10 anni. Nel modello originale il  $k_t$  non presentava un andamento definito. Nel nostro caso si osserva invece un trend decrescente del fattore latente che parte negli anni della crisi, ciò potrebbe rappresentare un aumento dell'effetto esplicativo del  $k_t$  negli anni in cui il PIL modifica il suo andamento perdendo la sua capacità predittiva.

I risultati nel caso Italiano dunque mostrano dei risultati che differiscono in parte da quelli osservati nel modello originale ma che comunque si presentano attendibili.

Successivamente nella tesi introdurremo nuovi modelli e confronteremo attraverso differenti misure di fitting la loro capacità nel modellizzare e prevedere i tassi futuri di mortalità.

## 3.4 Relazione tra indicatori socioeconomici e mortalità

### 3.4.1 Correlazione

Partendo dagli indicatori introdotti vengono selezionati solo quelli che mostrano una maggiore correlazione con il fattore latente del Lee-Carter e la mortalità. La correlazione viene misurata attraverso l'indice di correlazione lineare definito anche indice di Pearson.

Questo indice, date due variabili statistiche X e Y, viene definito come la covarianza divisa per il prodotto delle deviazioni standard delle due variabili:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

I risultati vengono mostrati nella tabella 3.1 per uomini, donne ed unisex.

	$k_t$ uomini	$k_t$ donne	$k_t$ unisex
PIL	-0.5558	-0.6132	-0.5881
Tasso di disoccupazione	-0.0219	0.0192	0.0020
Indice di sviluppo umano depurato	-0.9575	-0.9707	-0.9659
Debito pubblico/ PIL	-0.3952	-0.3438	-0.3665
Spesa sanitaria sul PIL	-0.8537	-0.8429	-0.8495
Inflazione	0.8098	0.8159	0.8139
Stipendi	-0.9756	-0.9841	-0.9813

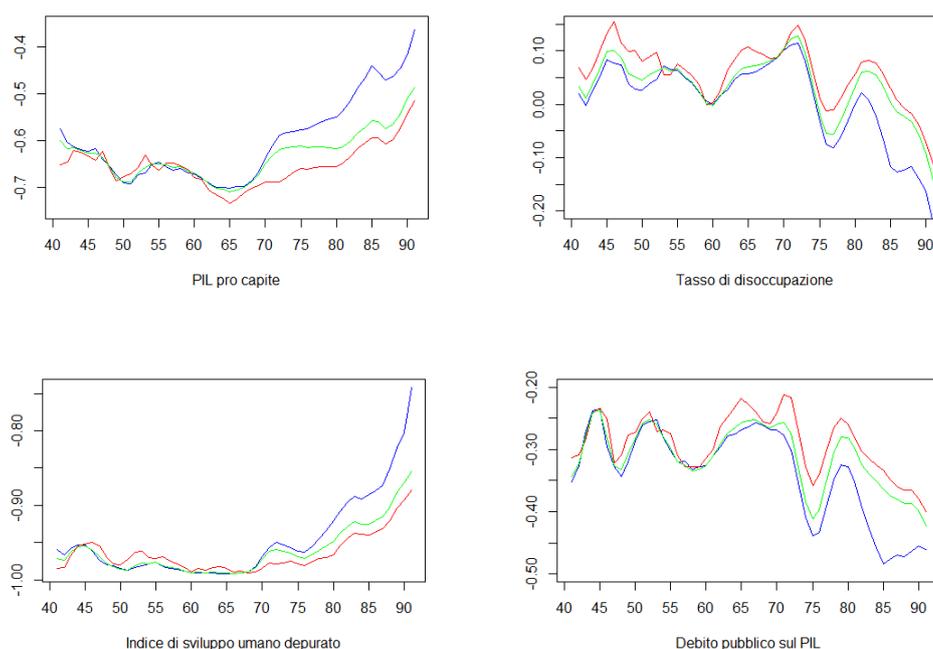
**Tabella 3.1.** Correlazione fra gli indicatori socioeconomici e il fattore latente del modello di Lee-Carter per uomini, donne ed unisex.

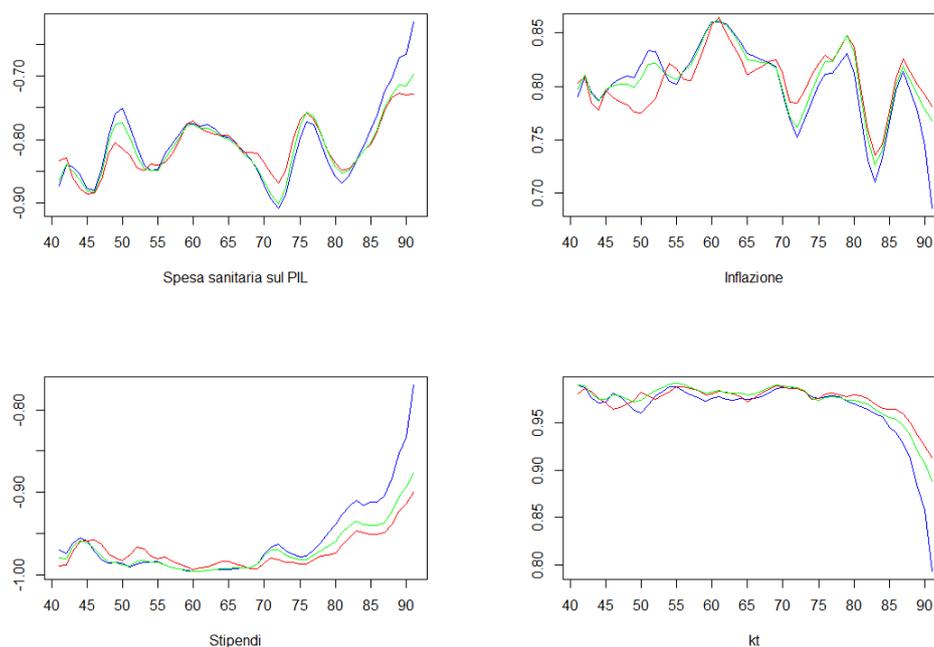
Osservando i risultati ottenuti testando la correlazione fra gli indicatori ed il fattore latente per uomini, donne ed unisex si nota che la correlazione varia molto fra le grandezze socioeconomiche proposte:

- Il PIL presenta una correlazione negativa con il fattore latente che varia leggermente fra i sessi, nelle donne la correlazione negativa è più forte e si aggira attorno al -0,61, negli uomini la correlazione è uguale a -0.5558;

- L'indice di sviluppo umano depurato, gli stipendi e la spesa sanitaria presentano una correlazione negativa molto forte con il fattore latente del Lee-Carter, che per i primi due supera il -0.95;
- L'inflazione è l'unico indice socioeconomico che presenta una correlazione positiva con il  $k_t$ , il cui valore in tutti e tre i casi si aggira attorno lo 0.80;
- Il debito pubblico sul PIL presenta una correlazione negativa che non sembrerebbe particolarmente significativa rispetto agli altri risultati ottenuti;
- Per il tasso di disoccupazione si osserva una correlazione con il fattore latente quasi inesistente; per gli uomini è negativa e per le donne positiva ma in entrambi i casi è molto prossima allo zero.

Dai risultati ottenuti sembrerebbe che il debito pubblico italiano sul PIL ed il tasso di disoccupazione non rappresentino due indici socioeconomici molto legati con il fattore latente e dunque con la mortalità. Prima di scartarli dalla nostra analisi effettuiamo un'analisi supplementare osservando direttamente la correlazione con il tasso centrale di mortalità per ogni età dai 40 ai 90 anni. I risultati vengono mostrati nella figura 3.13:





**Figura 3.13.** Correlazione fra gli indicatori socioeconomici e i tassi di mortalità per ogni età dai 40 ai 90 anni per uomini(blu), donne(rosso) ed unisex(verde).

In tutti i grafici osservati si nota come la correlazione diminuisca fortemente per le età più elevate, in particolar modo per gli uomini. Questo andamento deriva dalla riduzione della popolazione che comporta un aumento di effetti distorsivi legati alla natura stocastica del verificarsi dell'evento morte. La mortalità per gli uomini sembra essere molto più soggetta a questo effetto distorsivo in quanto il numero di sopravvissuti alle età più elevate è minore rispetto a quello delle donne, avendo gli uomini in media un'aspettativa di vita minore.

Come benchmark viene anche testata la correlazione del fattore latente del Lee-Carter con i tassi di mortalità. Il grafico mostra un'elevatissima correlazione, molto vicina a 1, per tutte le età fino agli 85 anni per poi calare fortemente per le età più estreme. I risultati non si discostano molto da quelli osservati precedentemente, sembrerebbe opportuno scartare dalle nostre analisi la disoccupazione ed il debito pubblico, non mostrando questi un fortissimo legame con i tassi di mortalità.

Gli effetti delle fluttuazioni dei tassi di disoccupazione sulla mortalità in passato erano già state studiate da alcuni accademici, tra i quali per esempio J.A.Tapia

Granados(2005) e M.H.Brenner(2005). Questi ultimi, anche se con risultati in parte contraddittori, mostrarono una significativa correlazione con i tassi di mortalità. Nel nostro caso si osserva invece una correlazione con il fattore latente quasi inesistente. La differente correlazione osservata rispetto alcuni studi precedenti può essere giustificata osservando che le fluttuazioni dei tassi di disoccupazione:

- tendono a manifestare i propri effetti con un ritardo temporale di qualche anno;
- hanno effetti differenti in base al periodo storico e al Paese;
- hanno effetto sulla salute sono se perdurano nel tempo.

Sulla base dei nostri risultati concordiamo con la letteratura scientifica già proposta sull'argomento, la quale sostiene che non esista una chiara e definita relazione fra disoccupazione e mortalità.

### 3.4.2 Test di stazionarietà

Dopo averne testato la correlazione ed aver scartato gli indicatori che risultavano meno correlati procediamo con un'analisi della relazione sul lungo termine fra le grandezze socioeconomiche e la mortalità.

Prima di tutto va testata la stazionarietà delle serie e dell'indice di mortalità  $k_t$ . Il primo test utilizzato è il Phillips-Perron Test(1988), il quale presenta come ipotesi nulla la presenza di radice unitaria. Il test si basa sulla seguente relazione:

$$y_t = \beta' D_t + \phi y_{t-1} + u_t$$

in cui  $D_t$  è un vettore che contiene termini deterministici,  $\beta'$  è un parametro da stimare e  $u_t$  è un processo  $I(0)$  che può essere eteroschedastico. I risultati per il test di Phillips-Perron, assieme ai test successivi sono mostrati nella tabella 3.2.

I fattori latenti del modello Lee-Carter per uomini, donne e unisex rifiutano l'ipotesi nulla al 5%. Tutte le altre serie storiche invece non rifiutano l'ipotesi nulla. Il rifiuto dell'ipotesi nulla da parte del  $k_t$  anche se potrebbe sembrare un esito anomalo in realtà è un risultato molto simile a quello ottenuto nei test effettuati per il Giappone da G.Niu e B.Melenberg (2014).

Come analisi supplementare effettuiamo il KPSS test(1992) con l'ipotesi nulla che le serie siano trend stazionarie. Il KPSS è un test che è stato introdotto da D.Kwiatkowski et al. (1992) e si basa sul seguente modello:

$$y_t = \beta' D_t + \mu_t + u_t$$

$$\mu_t = \mu_{t-1} + \epsilon_t$$

in cui  $D_t$  è la parte deterministica,  $\beta'$  è un parametro da stimare,  $\mu_t$  è una passeggiata aleatoria,  $u_t$  è un processo  $I(0)$  che può essere eteroschedastico ed infine si ha che  $\epsilon_t \sim WN(0, \sigma_\epsilon^2)$ .

L'ipotesi nulla del KPSS test viene rifiutata nella totalità dei casi:

- al 5% per il  $k_t$  delle donne ed unisex e per la spesa sanitaria pubblica e l'inflazione;
- al 1% per il  $k_t$  degli uomini, il PIL, l'indice di sviluppo umano depurato e gli stipendi.

Infine, assumendo che tutte le serie presentino unit root, si effettua una differenziazione per tutte le serie per tentare di rendere le serie stazionarie:

$$\Delta Y_t = Y_t - Y_{t-1}$$

Dopo aver differenziato le serie ripetiamo nuovamente il pp.test. I risultati sono mostrati nella tabella 3.2:

	pp.test	kpss.test	pp.test prima differenza
$k_t$ uomini	-3.7263*	0.1060**	-9.0453**
$k_t$ donne	-3.7897*	0.2133*	-12.1300**
$k_t$ unisex	-3.7737*	0.1705*	-10.4803**
PIL	-0.6569	0.3413**	-3.5560+
indice di sviluppo umano depurato	1.4473	0.3475**	-7.3455**
Spesa sanitaria sul PIL	-1.8774	0.1701*	-3.3745+
Inflazione	-2.6711	0.1576*	-6.2279**
Stipendi	-1.3054	0.3348**	-5.6395**

**Tabella 3.2.** Phillips-Perron e Kwiatkowski, Phillips, Schmidt e Shin tests per stazionarietà.

+p<0.10; \*p<0.05; \*\*p<0.01.

Dopo la differenziazione tutte le serie rifiutano l'ipotesi nulla del test di Phillips-Perron all'1%, ad eccezione del PIL e della spesa sanitaria pubblica che lo rifiutano al 10%. Dai risultati ottenuti assumiamo che tutte le serie storiche sono integrate di ordine 1, I(1). Una serie temporale  $Y_t$  si dice integrata di ordine 1, se è non stazionaria, ma la serie temporale definita dalle differenze prime è stazionaria.

### 3.4.3 Test di cointegrazione

Basandoci sui risultati ottenuti precedentemente occorre effettuare una differenziazione delle serie originali per renderle stazionarie. Effettuando però una manipolazione delle serie originali si incorre nel rischio di perdita di informazioni rilevanti contenuti nella serie. Una strategia comune per studiare la relazione sul lungo periodo tra le serie storiche è quello di studiarne la cointegrazione. Generalmente le combinazioni lineari di variabili integrate di ordine  $d$  risultano integrate dello stesso ordine; in alcuni casi particolari però esistono combinazioni lineari in cui l'ordine di integrazione è inferiore a quello delle serie di partenza.

Se  $x_t$  è  $I(d)$  e  $y_t$  è  $I(b)$  allora  $z_t = x_t + \lambda y_t$  è  $I(c)$ , dove:

$$\begin{cases} c = \max(d, b) & \text{se } d \neq b \\ c \leq \max(d, b) & \text{se } d = b \end{cases}$$

Quando la disuguaglianza vale in senso stretto si ha che le due serie sono cointegrate. La cointegrazione può esistere solo tra serie che mostrano lo stesso ordine di integrazione, come si evince dalla definizione precedente, nel caso di combinazione lineare di serie di ordine di integrazione differente si avrà che l'ordine di integrazione della combinazione sarà il massimo fra gli ordini di integrazione delle due serie. Una combinazione lineare tra una serie  $I(1)$  ed una serie  $I(0)$  è una serie  $I(1)$ .

Nel nostro caso abbiamo delle serie integrate di ordine 1. Per verificare la cointegrazione fra il fattore latente  $k_t$  e gli indici socioeconomici proposti si è utilizzato il Johansen test (1988). La procedura proposta da Johansen permette di identificare il rango di cointegrazione. Questa metodologia utilizza due test, il trace test, che fa riferimento alle seguenti ipotesi:

$$\begin{cases} H_0 : \text{rango}(\Pi) = r_0 \\ H_1 : r_0 < \text{rango}(\Pi) < k \end{cases}$$

oppure il test del massimo autovalore:

$$\begin{cases} H_0 : \text{rango}(\Pi) = r_0 \\ H_1 : \text{rango}(\Pi) = r_0 + 1 \end{cases}$$

Nel nostro caso si è preferito l'utilizzo del primo test.

I risultati sono mostrati nella tabella 3.3:

		$k_t$ uomini	$k_t$ donne	$k_t$ unisex
PIL	$r \leq 1$	4.43	4.06	4.27
	$r = 0$	27.24**	22.30*	25.44**
indice di sviluppo umano depurato	$r \leq 1$	10.09**	5.51	7.04
	$r = 0$	29.63**	23.91*	25.98**
Spesa sanitaria sul PIL	$r \leq 1$	9.25*	10.53*	10.04*
	$r = 0$	31.33**	29.06**	30.94**
Inflazione	$r \leq 1$	9.74*	9.27*	9.57*
	$r = 0$	31.74**	28.03**	30.46**
Stipendi	$r \leq 1$	10.06*	6.63	7.32
	$r = 0$	27.43**	21.61*	23.44*

**Tabella 3.3.** Johansen test per cointegrazione fra gli indicatori socioeconomici ed il fattore latente del Lee-Carter per ogni sesso. + $p < 0.10$ ; \* $p < 0.05$ ; \*\* $p < 0.01$ .

L'ipotesi nulla di non presenza di cointegrazione ( $r=0$ ) fra i fattori latenti e gli indicatori viene rifiutata in tutti i casi all'1% o 5%. L'ipotesi nulla  $r \leq 1$  conduce invece a risultati più eterogenei. Viene rifiutata al 5% per l'inflazione e la spesa sanitaria. Per il PIL non viene rifiutata in nessuno dei tre i casi. Per l'indice di sviluppo umano depurato e gli stipendi viene rifiutata solo con il  $k_t$  degli uomini.

Dai risultati ottenuti non appare evidente una cointegrazione fra gli indicatori socioeconomici proposti ed i  $k_t$ , nonostante ciò non viene rifiutata una possibile relazione di lungo termine fra i tassi di mortalità e le grandezze socioeconomiche.

## Capitolo 4

# Modelli proposti

### 4.1 Modelli OLS

Storicamente sono state molte le proposte di indici macroeconomici da prendere in considerazione nello studio dei tassi di mortalità futuri.

Il dibattito più combattuto è stato sempre quello legato all'andamento del PIL, utilizzato come indice descrittivo della situazione economica del Paese.

La ricchezza di una Nazione però non incide direttamente sul miglioramento della salute degli individui ma generalmente questa impatta su un insieme di grandezze socioeconomiche:

- Migliore sistema sanitario;
- Tassi di disoccupazione più bassi;
- Più alto livello di scolarizzazione;
- Stipendi più elevati con conseguente aumento del benessere individuale.

Questi indici generalmente migliorano la qualità della vita degli individui, riducendo così i tassi di mortalità.

In questa tesi si è cercato di studiare l'impatto di queste grandezze singolarmente per osservarne in che maniera e quanto queste influiscano sull'aumento dell'aspettativa di vita. Sono stati introdotti dei modelli OLS per i tassi di mortalità per ogni età dai 40 ai 90 anni in cui vengono utilizzate come variabili esplicative gli indicatori socioeconomici che mostravano un'elevata correlazione con la mortalità.

È stato inoltre utilizzato un modello in cui si utilizza come variabile esplicativa il

fattore latente della mortalità  $k_t$  come benchmark.

I modelli introdotti sono:

$$\ln(m_{x,t}) = \beta_{0,x}^g + \beta_{1,x}^g g_t + \epsilon_{x,t}^g$$

$$\ln(m_{x,t}) = \beta_{0,x}^{ind} + \beta_{1,x}^{ind} ind_t + \epsilon_{x,t}^{ind}$$

$$\ln(m_{x,t}) = \beta_{0,x}^{sp} + \beta_{1,x}^{sp} sp_t + \epsilon_{x,t}^{sp}$$

$$\ln(m_{x,t}) = \beta_{0,x}^{infl} + \beta_{1,x}^{infl} infl_t + \epsilon_{x,t}^{infl}$$

$$\ln(m_{x,t}) = \beta_{0,x}^w + \beta_{1,x}^w w_t + \epsilon_{x,t}^w$$

$$\ln(m_{x,t}) = \beta_{0,x}^k + \beta_{1,x}^k k_t + \epsilon_{x,t}^k$$

Per studiare la bontà di adattamento dei modelli viene introdotto il coefficiente di determinazione  $R^2$  il cui valore varia tra 0 ed 1 e misura la frazione della varianza della variabile dipendente espressa dalla regressione. Questa si può calcolare rapportando la devianza spiegata dal modello (ESS) alla devianza totale (TSS) oppure sottraendo ad 1 il rapporto tra la devianza residua (RSS) e la devianza totale (TSS):

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

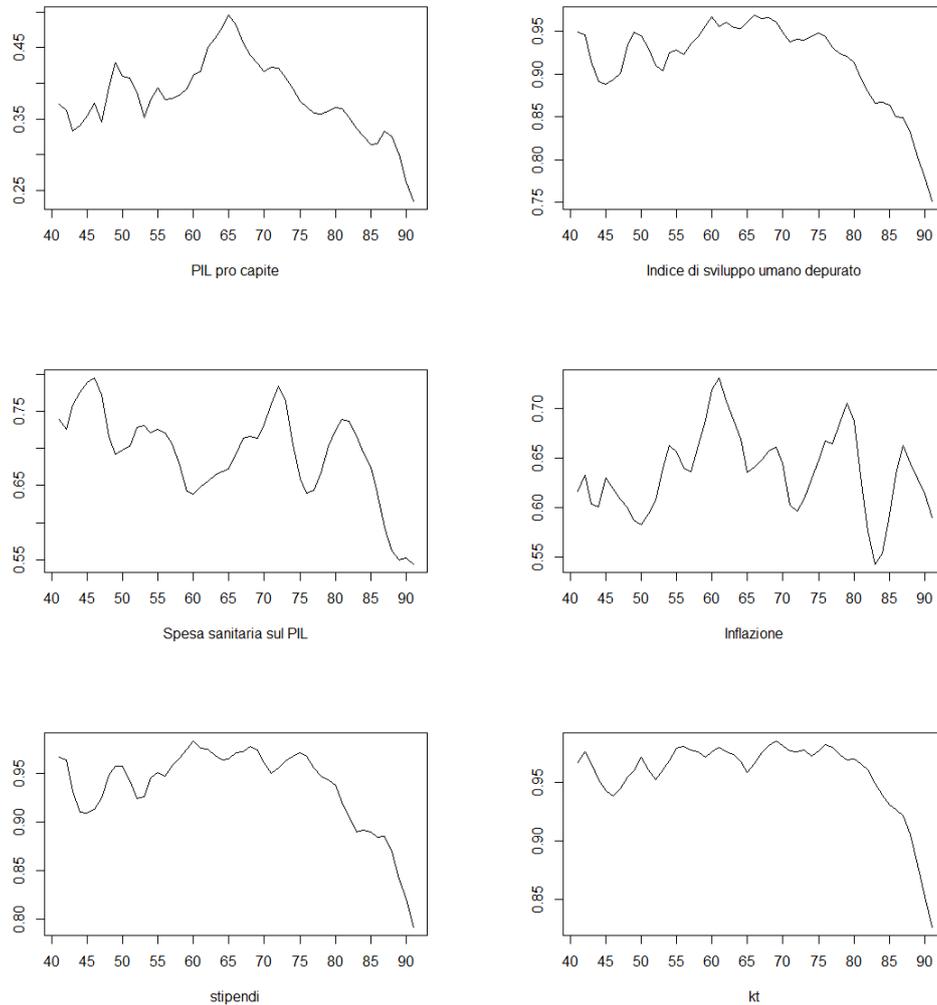
$$ESS = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

$$TSS = \sum_{i=1}^n (y_i - \bar{y})^2$$

$$RSS = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

dove  $y_i$  sono i dati osservati,  $\bar{y}$  è la loro media e  $\hat{y}_i$  sono i valori stimati dal modello.

Nella figura 4.1 si osserva graficamente l' $R^2$  per ogni età dai 40 ai 90 anni per i modelli OLS introdotti per i tassi di mortalità delle donne.



**Figura 4.1.**  $R^2$  dei modelli OLS per i tassi di mortalità delle donne dai 40 ai 90 anni con PIL, indice di sviluppo umano depurato, spesa sanitaria sul PIL, Inflazione, stipendi ed il fattore latente del Lee-Carter per il periodo dal 1990 al 2016.

Dai grafici si nota come i modelli con il fattore latente, gli stipendi e l'indice di sviluppo umano depurato riescano a catturare con grande fedeltà l'andamento dei tassi di mortalità. Il valore del  $R^2$  rimane elevato fino alle età più estreme dove cala drasticamente.

I modelli con spesa sanitaria sul PIL ed inflazione presentano un valore di  $R^2$  che oscilla fra le varie età, rimanendo però su livelli sufficientemente elevati. Anche il modello con il PIL mostra un andamento oscillatorio che si muove fra 0,35 e 0,45.

Gli indicatori mostrano in generale un buon fitting ad eccezione del modello con il PIL, il cui risultato si discosta da quelli ottenuti da G.Niu e B.Melenberg(2014). Nel loro studio infatti il coefficiente di determinazione aveva in media un valore superiore allo 0.90 per ogni Paese.

Sulla base dei risultati osservati si ipotizza dunque che sia possibile introdurre dei modelli stocastici per la mortalità in cui vengano incluse le grandezze socioeconomiche osservate. Proponiamo dunque dei modelli affini al modello di Niu-Melenberg in cui il logaritmo del PIL pro capite viene sostituito dagli indicatori studiati finora.

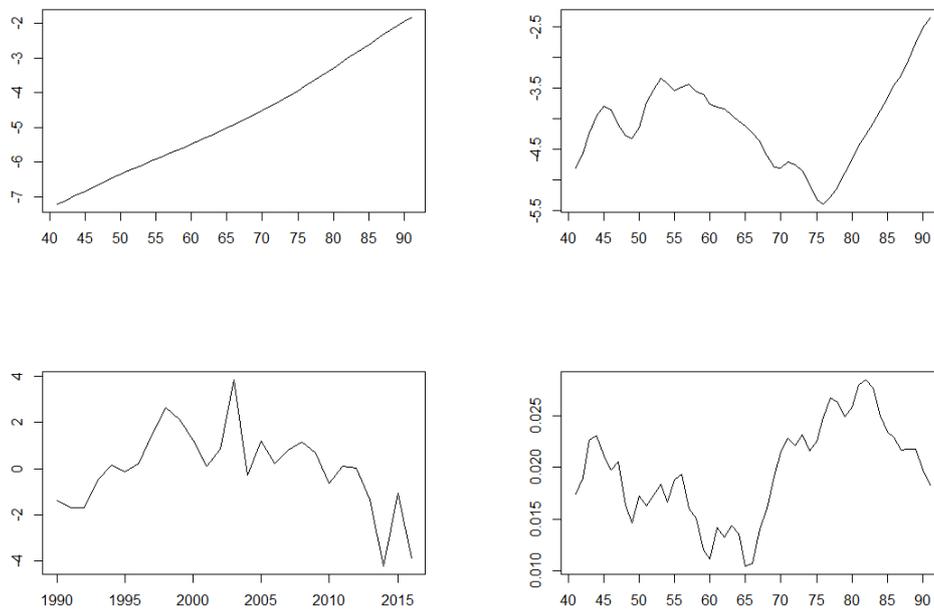
I valori dei parametri stimati risulteranno differenti tra i vari modelli proposti, ad eccezione del parametro  $\alpha_x$ . Per una migliore interpretazione di questo parametro, nella fase di stima dei parametri, alle grandezze osservate viene sottratta la loro media, in modo che  $\alpha_x$  sia la media aritmetica rispetto al tempo del logaritmo di ciascun tasso specifico di mortalità.

## 4.2 Modello con Indice di sviluppo umano depurato

Per primo introduciamo il modello con l'indice di sviluppo umano depurato:

$$\ln(m_{x,t}) = \alpha_x^{ind} + \beta_x^{ind} k_t^{ind} + \gamma_x^{ind} ind_t + \epsilon_{x,t}^{ind}$$

Dai grafici nella figura 4.2 osserviamo i valori stimati dei parametri del modello con indice di sviluppo umano depurato per le donne italiane.



**Figura 4.2.** Parametri  $\alpha_x^{ind}$ ,  $\gamma_x^{ind}$ ,  $k_t^{ind}$  e  $\beta_x^{ind}$  del modello con Indice di sviluppo umano depurato stimati sui tassi di mortalità italiani delle donne di età dai 40 ai 90 anni dal 1990 al 2016.

$\alpha_x^{ind}$  come già detto continua ad avere il suo andamento monotono crescente.

$\gamma_x^{ind}$  è un parametro che misura la sensibilità dei tassi di mortalità al fattore osservato per ogni età  $x$ . I valori del parametro sono negativi per la relazione negativa fra l'indice di sviluppo umano depurato ed i tassi di mortalità. L'andamento del parametro mostra un forte legame fra mortalità e l'indice dai 40 ai 75 anni. Questo legame tende a decrescere per le età più estreme dove la correlazione negativa fra l'indice e

la mortalità tende a decrescere.

$k_t^{ind}$  non presenta un chiaro trend decrescente come avveniva nel modello di Lee-Carter. L'andamento lascia presumere che il trend dei tassi di mortalità ora sia perlopiù catturato dal fattore osservato e che dunque il fattore latente non sia in grado di catturare l'andamento dei tassi di mortalità.

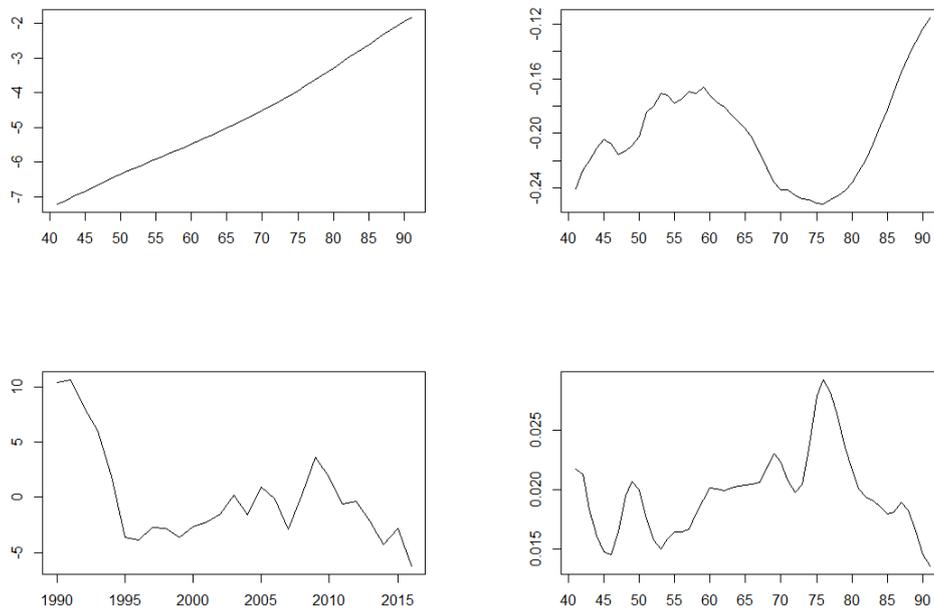
$\beta_x^{ind}$  cattura la sensibilità del logaritmo dei tassi di mortalità ad ogni età  $x$  alle variazioni di  $k_t^{ind}$ . I suoi valori sono positivi ed oscillano fra 0.010 e 0.025, picco raggiunto fra gli 80 ed 85 anni.

### 4.3 Modello con Spesa sanitaria pubblica sul PIL

Per la spesa sanitaria il modello introdotto è:

$$\ln(m_{x,t}) = \alpha_x^{sp} + \beta_x^{sp} k_t^{sp} + \gamma_x^{sp} sp_t + \epsilon_{x,t}^{sp}$$

Dai grafici nella figura 4.3 osserviamo i valori stimati dei parametri del modello con spesa sanitaria pubblica sul PIL per le donne italiane.



**Figura 4.3.** Parametri  $\alpha_x^{sp}$ ,  $\gamma_x^{sp}$ ,  $k_t^{sp}$  e  $\beta_x^{sp}$  del modello con Spesa sanitaria pubblica sul PIL stimati sui tassi di mortalità italiani delle donne dai 40 ai 90 anni dal 1990 al 2016.

Data la relazione negativa fra spesa sanitaria e mortalità i valori del parametro  $\gamma_x^{sp}$  sono negativi. Come nel caso dell'indice di sviluppo umano depurato anche qui i valori del parametro tendono ad avvicinarsi allo zero fra i 40 ed i 60 anni per poi decrescere, e dunque aumentare la sensibilità, fra i 60 ed i 75 anni. Questo legame tende a diventare più debole per le età più estreme dove la correlazione negativa fra la spesa sanitaria e la mortalità tende a decrescere.

$k_t^{sp}$  fra il 1990-1995 presenta un chiaro trend decrescente per poi stabilizzarsi e

rimanere stazionario fra il 1995 ed il 2016. L'andamento sembrerebbe mostrare che il fattore latente per i primi 5 anni è in grado di catturare il trend della mortalità. Questo potrebbe essere giustificato dall'andamento della spesa sanitaria per quel periodo, la quale presenta un primo andamento decrescente per poi presentare un trend crescente; dunque il  $k_t^{sp}$  sembrerebbe catturare il trend della mortalità nei primi anni, in cui l'andamento della spesa sanitaria era poco chiaro, per poi stabilizzarsi negli anni successivi. L'andamento costante dal 1995 al 2016 lascia presumere che il trend dei tassi di mortalità, in quel arco temporale, sia perlopiù catturato dal fattore osservato e che dunque il fattore latente non sia in grado di catturare l'andamento dei tassi di mortalità.

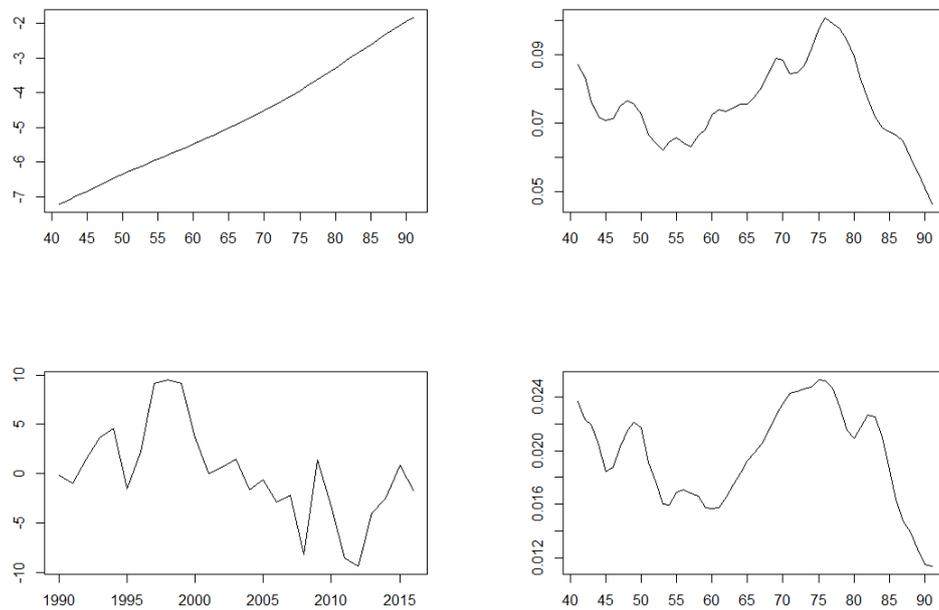
$\beta_x^{sp}$  ha valori positivi per tutte le età e questi oscillano tra 0.10 e 0.20 con un picco di valore 0.28 agli 80 anni.

## 4.4 Modello con Inflazione

Per l'inflazione il modello introdotto è:

$$\ln(m_{x,t}) = \alpha_x^{infl} + \beta_x^{infl} k_t^{infl} + \gamma_x^{infl} infl_t + \epsilon_{x,t}^{infl}$$

Dai grafici nella figura 4.4 osserviamo i valori stimati dei parametri del modello con inflazione per le donne italiane.



**Figura 4.4.** Parametri  $\alpha_x^{infl}$ ,  $\gamma_x^{infl}$ ,  $k_t^{infl}$  e  $\beta_x^{infl}$  del modello con Inflazione stimati sui tassi di mortalità italiani delle donne dai 40 ai 90 anni dal 1990 al 2016.

L'inflazione è l'unica grandezza socioeconomica da noi studiata correlata positivamente con i tassi di mortalità. Sulla base di questa correlazione positiva il parametro  $\gamma_x^{infl}$  mostra valori positivi. Come avveniva nei casi precedenti la sensibilità catturata dal parametro tende a calare per le età fra i 40 ed i 60 anni per poi crescere e raggiungere il suo picco attorno ai 75 anni e successivamente calare nuovamente per le età più estreme.

$k_t^{infl}$  non presenta un chiaro trend e dunque non sembra in grado di catturare l'effetto

temporale sui tassi di mortalità. Partendo da valori prossimi allo zero tende ad oscillare tra -10 e 10 tornando ad avvicinarsi allo zero per gli anni finali.

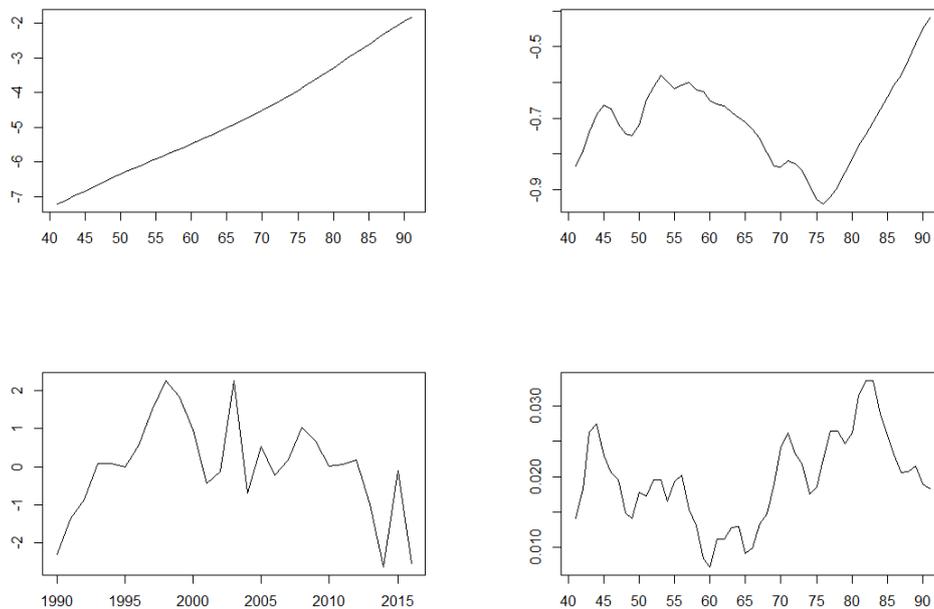
$\beta_x^{infl}$  ha un andamento decrescente dai 40 ai 60 anni per poi crescere nuovamente fino 0.24 agli 80 e calare drasticamente per le età più elevate.

## 4.5 Modello con Stipendi

Come avveniva per il PIL, non si introduce nel modello il valore osservato degli stipendi ma il suo logaritmo naturale. Il modello introdotto è:

$$\ln(m_{x,t}) = \alpha_x^w + \beta_x^w k_t^w + \gamma_x^w w_t + \epsilon_{x,t}^w$$

Dai grafici nella figura 4.5 osserviamo i valori stimati dei parametri del modello con stipendi per le donne italiane.



**Figura 4.5.** Parametri  $\alpha_x^w$ ,  $\gamma_x^w$ ,  $k_t^w$  e  $\beta_x^w$  del modello con Stipendi stimati sui tassi di mortalità italiani delle donne dai 40 ai 90 anni dal 1990 al 2016.

$\gamma_x^w$  presenta valori negativi ed un andamento molto simile a quello osservato per l'indice di sviluppo umano depurato;  $\gamma_x^w$  tende a crescere, quindi diminuisce la sensibilità del modello al fattore osservato, dai 40 fino ai 60 anni per poi decrescere nuovamente fino i 75 anni e dunque nelle età più estreme crescere fino ad avvicinarsi allo zero. Questo andamento del  $\gamma_x$  viene riscontrato per tutti i modelli introdotti, se considerassimo il modulo di questo parametro osserveremmo che il suo valore

tende a decrescere nell'intervallo fra i 40 ed i 60 anni, in cui dunque la mortalità sembrerebbe meno legata con i fattori osservati, per poi crescere tra i 60 ed i 75 anni ed infine calare nuovamente alle età più estreme in cui i tassi di mortalità presentano un andamento molto più casuale e la correlazione fra la mortalità e gli indici è molto più debole.

$k_t^w$  non presenta un chiaro trend lasciando presumere che il trend dei tassi di mortalità sia perlopiù catturato dal fattore osservato.

$\beta_x^w$  tende a decrescere con un andamento oscillatorio tra i 45 ed i 60 anni per poi crescere nuovamente dai 60 agli 85 anni. La sensibilità al fattore latente sembrerebbe crescere per le età in cui la sensibilità verso il fattore osservato si riduce ( $\gamma_x^w$  cresce). Nonostante ciò non è possibile affermare che questo effetto sia legato ad un aumento dell'effetto del fattore latente  $k_t^w$ , non mostrando questi un chiaro trend.

## 4.6 Misure di fitting

I risultati ottenuti dalla stima dei parametri dei nuovi modelli introdotti sembrerebbero attendibili. Per comprendere però se l'utilizzo di indici alternativi permetta di raggiungere risultati migliori bisognerà applicare delle misure di fitting ai modelli. La prima metodologia introdotta è il Bayesian information Criterion. Il BIC è un criterio per la selezione di modelli che prende in considerazione una penalizzazione in base al numero di parametri nel modello. Matematicamente viene definito come:

$$BIC = L(\psi) - \frac{1}{2}k \ln(n)$$

dove  $L(\psi)$  è il valore massimizzato della funzione di log-verosimiglianza,  $n$  è il numero delle osservazioni e  $k$  è il numero dei parametri stimati dal modello.

La tabella 4.1 mostra il BIC calcolato per i modelli studiati:

	Uomini	Donne	Unisex
Lee-Carter	-11579.1	-10678.06	-15000.28
Niu-Melenberg	-11117.36	-10829.12	-14938.75
Indice di sviluppo umano depurato	-10935.04	-10737.88	-14722.63
Spesa sanitaria sul PIL	-11341.76	-10721.69	-14583.06
Inflazione	-11550.10	-10793.06	-14873.94
Stipendi	-11008.52	-10761.89	-14712.27

**Tabella 4.1.** BIC dei modelli studiati per uomini, donne ed unisex.

Per gli uomini il criterio sembrerebbe prediligere il modello con l'indice di sviluppo umano depurato mentre per unisex il risultato migliore è ottenuto dal modello con la spesa sanitaria. Per quanto riguarda le donne il criterio di Schwarz sembrerebbe preferire il modello di Lee-Carter agli altri.

Quest'ultimo risultato potrebbe sembrare anomalo e contrastare dunque la letteratura citata precedentemente come K.Hanewald(2011), D.French e C.O'Hare (2014), G.Niu e B.Melenberg (2014) in cui si mostrava una relazione esistente fra le grandezze socioeconomiche e la mortalità.

Risultati simili per le donne sono stati ottenuti da M.Seklecka et al.(2017) in uno studio in cui venivano presi in considerazione nove Paesi Europei per il periodo dal 1960 al 2012.

Il BIC veniva usato per confrontare il fitting di 4 modelli:

- Lee-Carter;
- Niu-Melenberg;
- O'Hare-Li;
- Modello proposto da M.Seklecka et al.

Il criterio prediligeva il modello di Lee-Carter per i Paesi con minore popolazione, Austria, Belgio ed Olanda.

Sulla base dei risultati ottenuti da M.Seklecka et al. è possibile ipotizzare che l'esito anomalo del BIC sia legato ad un errore dovuto dalla non elevata profondità dei dati storici osservati e dal limitato range di età. Per confutare dunque che i modelli introdotti forniscano un miglior fitting per tutti i sessi introduciamo due misure di errore:

- Il MAPE, mean absolute percentage error

$$MAPE = \frac{1}{NM} \sum_{x,t} \frac{\|\hat{m}_{x,t} - m_{x,t}\|}{m_{x,t}}$$

- IL MAD, mean absolute deviation

$$MAD = \sum_{x,t} \frac{\|\hat{m}_{x,t} - m_{x,t}\|}{NM}$$

in cui  $m_{x,t}$  sono i valori osservati dei tassi di mortalità e  $\hat{m}_{x,t}$  la loro stima, M è il numero di età osservate ed N è il numero degli anni presi in considerazione.

I risultati delle due misure di errore sono mostrate rispettivamente nelle tabelle 4.2 e 4.3.

I valori del MAPE per gli uomini prediligono il modello che utilizza l'indice di sviluppo umano depurato, come avveniva utilizzando il BIC. Per le donne i valori sono molto vicini tra loro e quello con il minore errore è il modello con la spesa sanitaria. Per unisex viene prediletto il modello con l'indice di sviluppo umano depurato.

	Uomini	Donne	Unisex
Lee-Carter	0.0276	0.0238	0.0241
Niu-Melenberg	0.0238	0.0229	0.0226
Indice di sviluppo umano depurato	0.0232	0.0229	0.0222
Spesa sanitaria sul PIL	0.0260	0.0226	0.0224
Inflazione	0.0265	0.0231	0.0231
Stipendi	0.0237	0.0230	0.0223

**Tabella 4.2.** MAPE dei modelli studiati per uomini, donne ed unisex.

Anche per il MAD i risultati mostrano come il modello con l'indice di sviluppo umano depurato risulti essere il migliore nel fitting dei dati per gli uomini. Per le donne ed unisex i risultati migliori si ottengono con il modello originale di Niu-Melenberg.

	Uomini	Donne	Unisex
Lee-Carter	0.000422	0.000162	0.000264
Niu-Melenberg	0.000319	0.000146	0.000210
Indice di sviluppo umano depurato	0.000308	0.000151	0.000219
Spesa sanitaria sul PIL	0.000411	0.000154	0.000234
Inflazione	0.000394	0.000156	0.000233
Stipendi	0.000327	0.000152	0.000225

**Tabella 4.3.** MAD dei modelli studiati per uomini, donne ed unisex.

Dai risultati ottenuti solo per gli uomini viene eletto un unico modello come migliore nella capacità di fitting. Per le donne ed unisex della popolazione i risultati sono misti, mostrando però un buon fitting per il modello di Niu-Melenberg e per il modello con spesa sanitaria sul PIL. Il risultato ottenuto dal BIC che eleggeva come modello migliore per le donne il Lee-Carter non è stato assolutamente confermato dalle misure di errore, come avveniva in M.Seklecka et al. (2017).

I risultati ottenuti dunque sembrano concordare con le ipotesi di esistenza di relazione

fra grandezze le socioeconomiche e la mortalità.

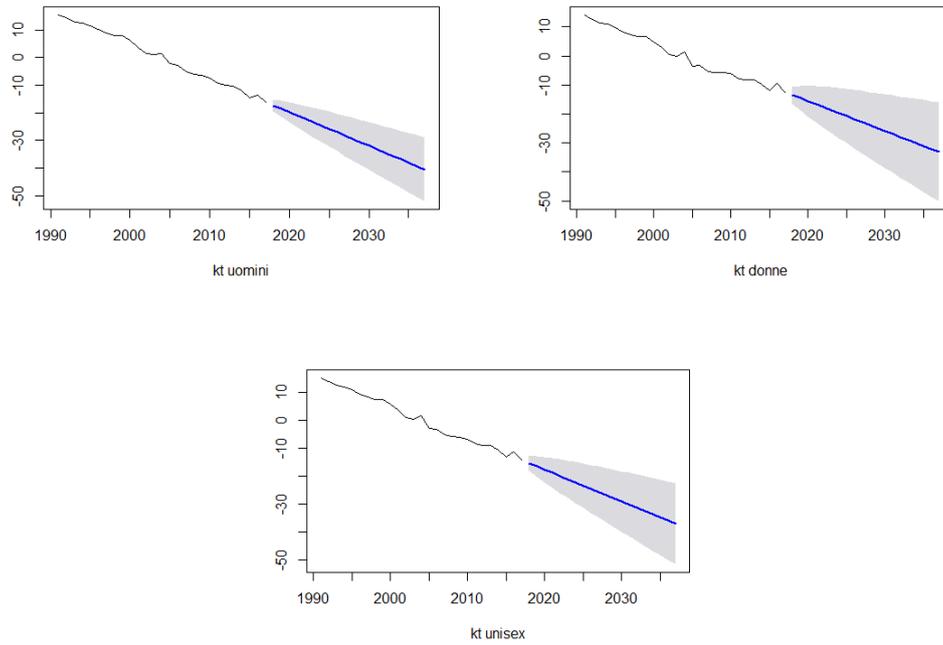
## Capitolo 5

# Stima dei tassi di mortalità futuri

### 5.1 Forecasting

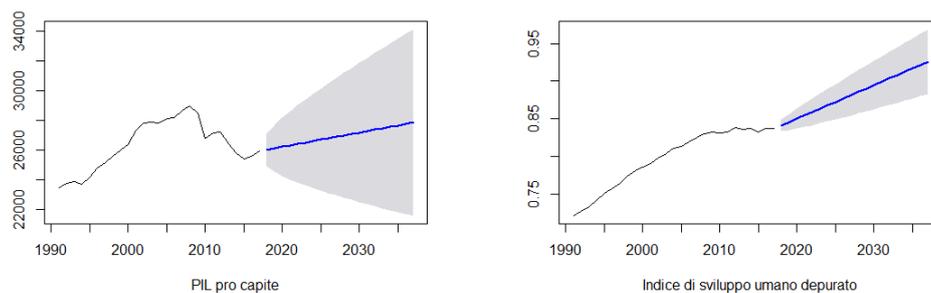
I modelli introdotti, come quelli di Niu-Melenberg e Lee-Carter, hanno come fine ultimo quello di stimare i tassi futuri di mortalità. La stima dei tassi di mortalità futuri viene effettuata stimando i valori futuri dei fattori latenti e degli indicatori socioeconomici. Stimando i tassi di mortalità futuri per il periodo dal 2017 al 2036 si osserva che per tutti i sessi ed i modelli i tassi presentano andamenti decrescenti. L'andamento decrescente dei tassi di mortalità implica una crescita dell'aspettativa di vita. In tutti i casi i tassi di mortalità delle donne risultano essere più bassi di quelle degli uomini e i due andamenti non si incrociano mai. Questo è un risultato empirico demograficamente accettato in quanto le donne in media vivono più degli uomini in tutto il mondo.

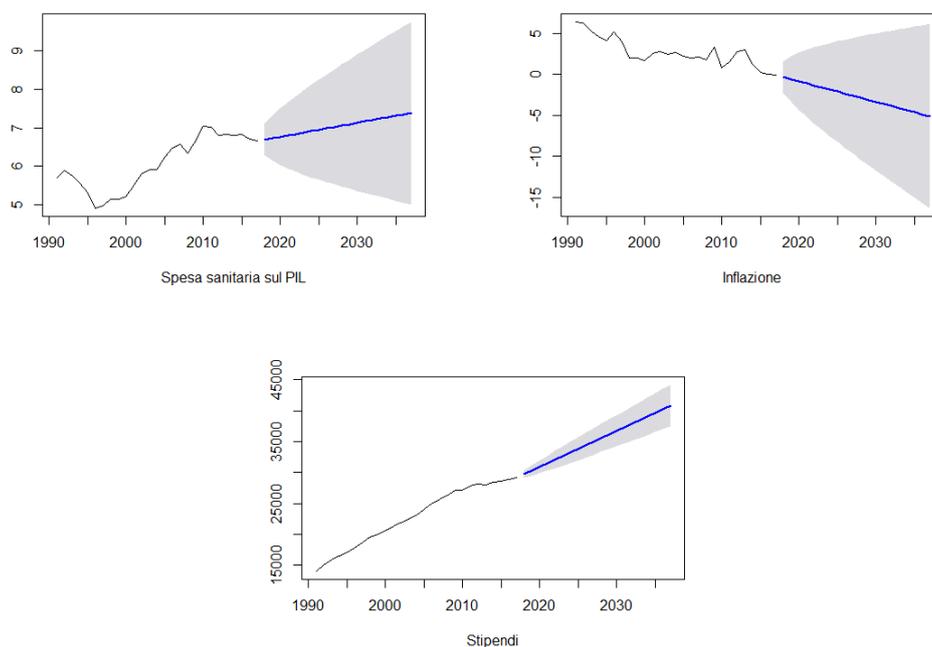
Stimando i valori futuri del fattore latente di Lee-Carter e degli indicatori socioeconomici con un intervallo di confidenza al 95% si nota come gli intervalli siano piuttosto differenti fra di loro. La figura 5.1 mostra gli andamenti stimati per il periodo dal 2017 al 2036 per i  $k_t$  del modello di Lee-Carter per tutti i sessi. Il fattore latente si presenta decrescente per tutti i sessi ma per le donne presenta degli intervalli di confidenza molto più ampi.



**Figura 5.1.** Andamenti stimati dei  $k_t$  del modello di Lee-Carter per il periodo dal 2017 al 2036 con intervalli di confidenza al 95%.

Gli andamenti degli indicatori socioeconomici sono mostrati nella figura 5.2 e presentano andamenti molto eterogenei. Per l'indice di sviluppo umano depurato e gli stipendi viene stimato una crescita con un intervallo molto stretto, mentre, per spesa sanitaria, inflazione e PIL pro capite gli intervalli di confidenza sono molto più larghi e comprendono anche la possibilità che il valore degli indicatori decresca.





**Figura 5.2.** Andamenti stimati degli indicatori socioeconomici studiati per il periodo dal 2017 al 2036 con intervalli di confidenza al 95%.

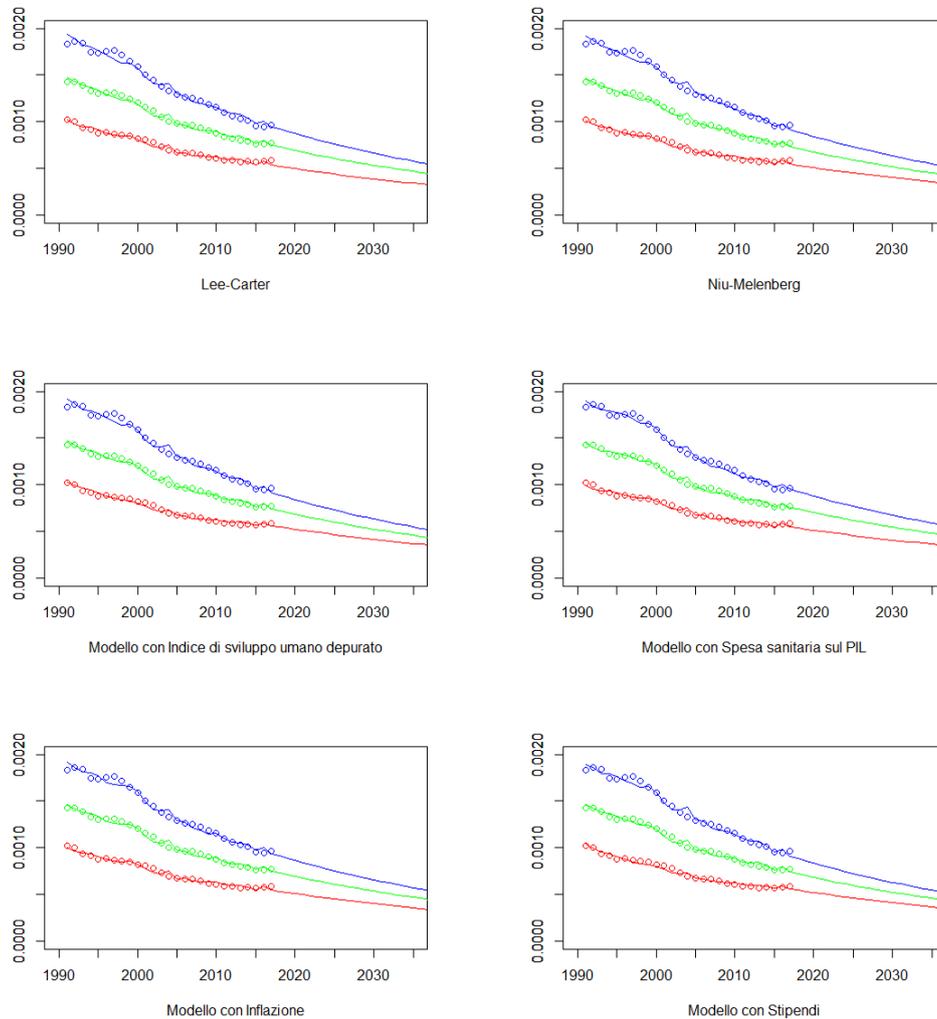
Questi intervalli così eterogenei sono il risultato di andamenti molto meno prevedibili rispetto all'indice di sviluppo umano depurato e gli stipendi, i quali sono cresciuti costantemente negli ultimi 27 anni e si stima continuino a crescere ancora.

Dalla stima dei fattori latenti ed osservati abbiamo calcolato i tassi di mortalità futuri. Prendiamo in riferimento due età in particolare per osservarne l'andamento futuro stimato dei tassi:

- 40 anni, mortalità ad età lavorativa;
- 80 anni, mortalità per età più estreme.

La figura 5.3 mostra che i tassi di mortalità stimati per individui di 40 anni di età per tutti i modelli hanno andamenti molto simili. I tassi di mortalità per gli uomini si attestano più elevati. Tutti i tassi di mortalità presentano una decrescita che risulta però essere più o meno forte.

La differenza maggiore che riscontriamo in questi modelli nella stima dei tassi di mortalità futuri è legata però al livello di confidenza.

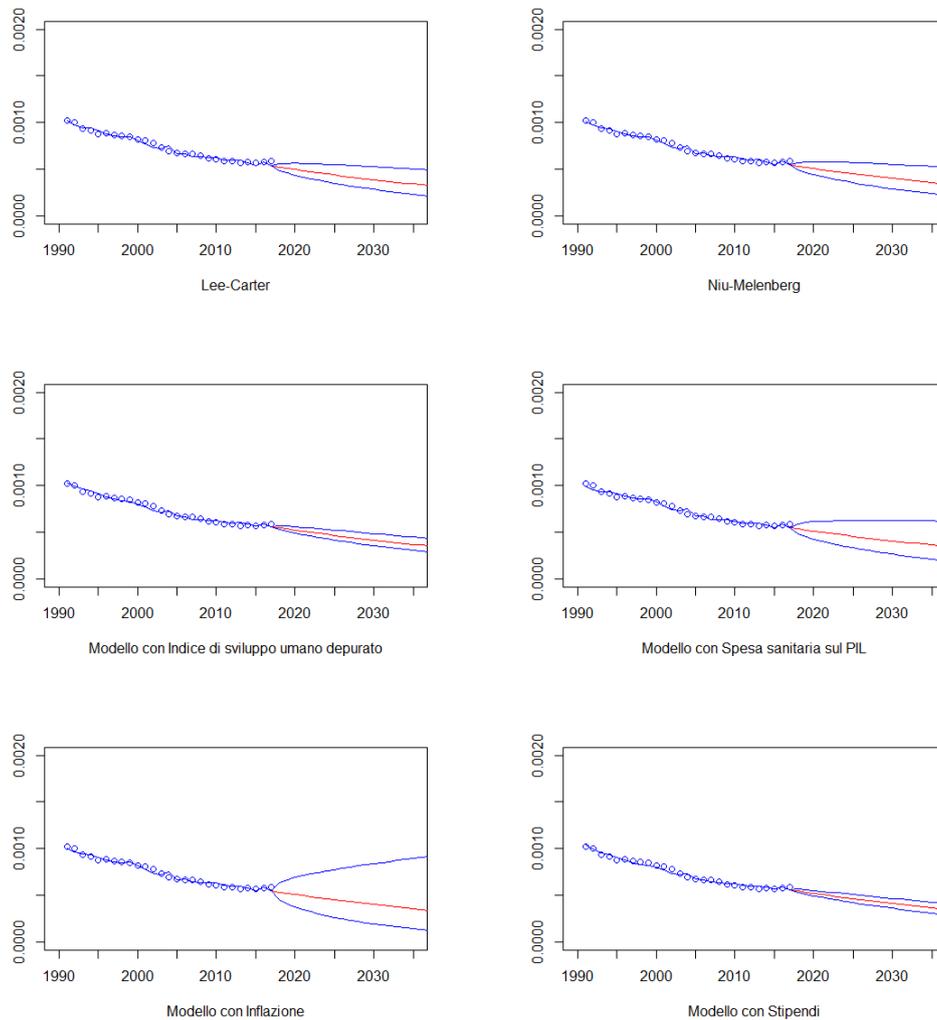


**Figura 5.3.** Stima tassi di mortalità dal 2017 al 2036 per uomini(blu), donne(rosse) ed unisex(verde) di 40 anni.

Nella figura 5.4, concentrandoci sui tassi di mortalità delle donne con intervalli di confidenza al 95%, si ottengono dei risultati molto differenti fra i vari modelli proposti. Questi grafici ci mostrano come l'aspettativa dei tassi di mortalità nei vari modelli sia molto simile, ma il loro livello di confidenza sia molto differente.

Questo risultato si ottiene in quanto la stima futura degli indicatori socioeconomici e dei fattori latenti presenta degli intervalli di confidenza molto differenti. Questa eterogeneità nella larghezza degli intervalli è legata alla capacità di prevedere con

maggiore o minore precisione i valori futuri di questi indici.

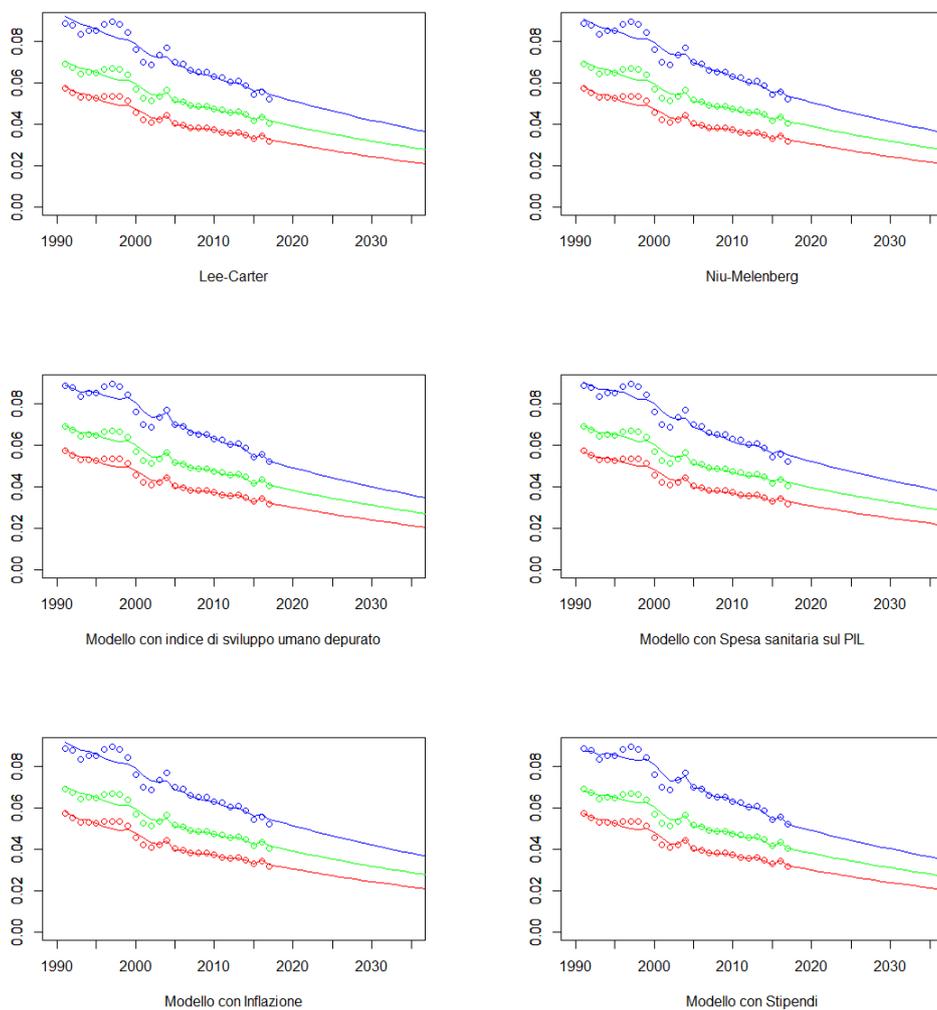


**Figura 5.4.** Stima tassi di mortalità per donne di 40 anni dal 2017 al 2036 con intervalli di confidenza al 95%.

La stima effettuata con i modelli con l'indice di sviluppo umano depurato e gli stipendi presentano degli intervalli di confidenza molto stretti, mentre, per quanto riguarda i modelli con l'inflazione e la spesa sanitaria questi intervalli risultano essere molto ampi. Il risultato di queste stime future dei tassi di mortalità rispecchia l'andamento stimato dei fattori osservati, gli indici che presentavano intervalli di confidenza più ampi, spesa sanitaria ed inflazione, presentano anche degli intervalli di

confidenza più ampi per la stima dei tassi di mortalità. Dai risultati sembrerebbe che gli intervalli di confidenza siano asimmetrici rispetto al valore medio, il risultato è ottenuto in quanto il calcolo della stima dei tassi di mortalità futuri prevede l'utilizzo dell'esponenziale, il quale non è una funzione lineare. Una funzione  $F$  è lineare se  $F(aX + bY) = a * F(X) + b * F(Y)$  per ogni  $a, b$  e per ogni  $X, Y$ .

Stimando i tassi di mortalità per individui di 80 anni si ottengono i grafici mostrati nella figura 5.5:



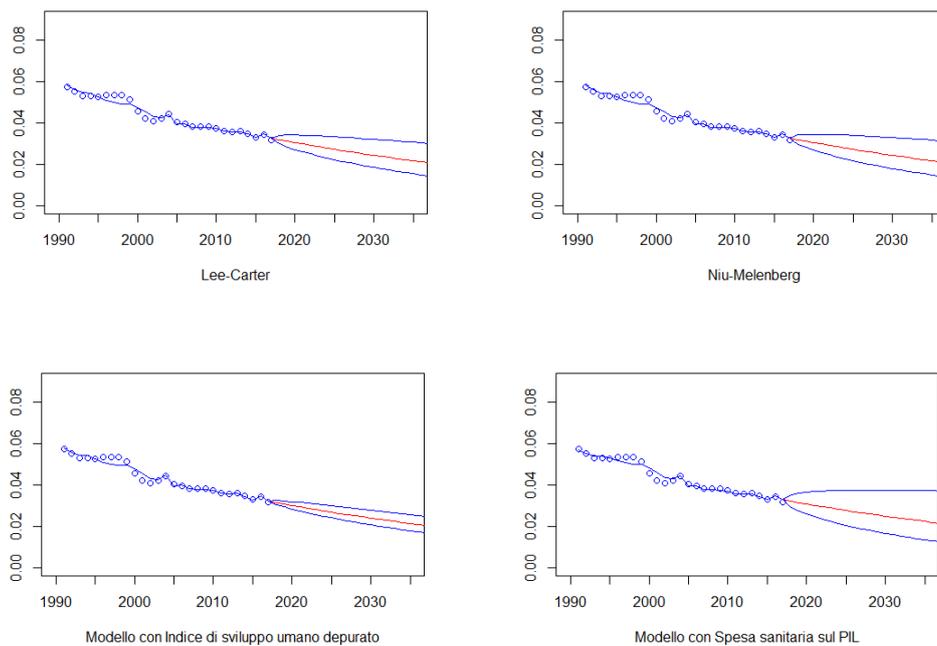
**Figura 5.5.** Stima tassi di mortalità dal 2017 al 2036 per uomini(blu), donne(rosse) ed unisex(verde) di 80 anni.

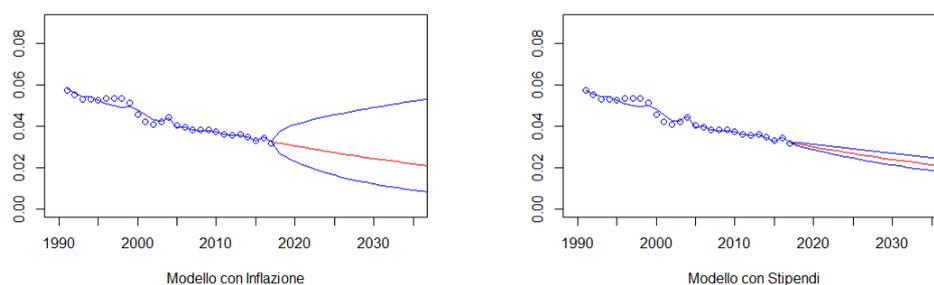
In tutti i modelli e per tutti i sessi anche alle età più avanzate l'andamento dei tassi della mortalità rimangono molto simili.

Una differenza però evidente con i dati osservati precedentemente è che per le età più avanzate la mortalità ha un andamento molto più stocastico e come si può osservare dal grafico i valori osservati e quelli stimati per il periodo 1990-2016 divergono per alcuni anni. Questo fenomeno si osserva in particolar modo per il primo periodo, fra il 1990 ai primi anni 2000, in cui l'andamento dei tassi di mortalità si presentava più stocastico di quanto non sia stato in epoca recente, dove osserviamo invece una costante decrescita e nessun andamento oscillatorio.

La figura 5.6 mostra la stima dei tassi di mortalità per le donne di 80 anni per il periodo dal 2017 al 2036 con un intervallo di confidenza al 95%.

Dai grafici si osserva come il risultato ottenuto precedentemente per i tassi di mortalità di 40 anni persista anche per le età più elevate. Si osserva un evidente decrescita dei tassi di mortalità che per ogni sesso e modello risulta molto simile, però, osservando gli intervalli di confidenza si nota come il grado di certezza collegato a questa stima cambi molto tra i modelli usati.





**Figura 5.6.** Stima tassi di mortalità per donne di 80 anni dal 2017 al 2036 con intervalli di confidenza al 95%.

Il modello che utilizza l'indice di sviluppo umano depurato e quello che utilizza gli stipendi sembrano poter stimare con un maggiore grado di fedeltà i tassi di mortalità futuri per ogni età. Dai risultati dunque sembrerebbe che gli ultimi due modelli citati riescano a prevedere con un maggiore grado di affidabilità i tassi futuri. Per confrontare però la capacità predittiva dei modelli ci riconduciamo a dei test out-of-sample.

## 5.2 Test Out-of-sample

Per testare la capacità predittiva dei modelli dividiamo i nostri dati in due gruppi:

- in-sample data, la sezione di dati utilizzata per la stima dei parametri dei modelli;
- out-of-sample data, la sezione di dati utilizzata per testare la capacità di previsione.

Il primo gruppo va dal 1990 al 2007 mentre il secondo dal 2008 al 2016. Usiamo i dati del primo gruppo per fare una stima dei valori dei tassi di mortalità futuri e successivamente ne testeremo la performance di previsione attraverso un test Out-of-sample.

Per la previsione dei tassi di mortalità con il modello di Lee-Carter bisogna stimare i parametri del modello per il periodo che va dal 1990 al 2007 e successivamente viene effettuata una stima dei valori futuri del fattore latente dal 2008 al 2016. Questi valori stimati verranno introdotti nel modello di Lee-Carter e si otterrà una stima dei tassi di mortalità.

Allo stesso modo nel modello di Niu-Melenberg e nei modelli a lui affini, da noi proposti, bisogna ripetere la stima dei parametri basandoci sui dati relativi al periodo dal 1990 al 2007 e successivamente stimare i valori dei fattori osservati e dei fattori latenti per il periodo che va dal 2008 al 2016.

Per testare la capacità di previsione dei modelli utilizziamo delle misure di errore:

- MAPE;
- MAD.

I risultati ottenuti sono presentati rispettivamente nelle tabelle 5.1 e 5.2.

Per il MAPE il modello di Niu-Melenberg presenta l'errore minore per unisex. Per le donne l'errore minore è per il modello che utilizzava gli stipendi come indicatore ed infine per gli uomini si prediligeva il modello con l'indice di sviluppo umano depurato. Dai risultati finora ottenuti sembra dunque che sia nel caso delle misure di fitting che per il test out-of-sample il modello con l'indice di sviluppo umano depurato è il migliore per gli uomini mentre per le donne ed unisex i risultati sono misti.

	Uomini	Donne	Unisex
Lee-Carter	0.0600	0.0591	0.0539
Niu-Melenberg	0.0532	0.0506	0.0459
Indice di sviluppo umano depurato	0.0526	0.0542	0.0481
Spesa sanitaria sul PIL	0.0561	0.0589	0.0518
Inflazione	0.0562	0.0559	0.0502
Stipendi	0.0548	0.0502	0.0478

**Tabella 5.1.** MAPE test Out-of-sample

Effettuiamo dunque un'ulteriore misurazione dell'errore out-of-sample attraverso la Mean Absolute Deviation:

	Uomini	Donne	Unisex
Lee-Carter	0.000849	0.000580	0.000647
Niu-Melenberg	0.000807	0.000412	0.000466
Indice di sviluppo umano depurato	0.000793	0.000497	0.000525
Spesa sanitaria sul PIL	0.000840	0.000616	0.000581
Inflazione	0.000803	0.000511	0.000582
Stipendi	0.000841	0.000461	0.000493

**Tabella 5.2.** MAD test Out-of-sample.

I risultati ottenuti da questa misura confermano in parte ciò che era stato osservato precedentemente:

- Il modello che utilizza l'indice di sviluppo umano depurato come fattore osservato conduce al minore errore per gli uomini;
- Il minore errore per unisex si osserva per il modello originale di Niu-Melenberg;
- Per le donne il modello con errore minore è il Niu-Melenberg originale, sembra dunque che non sia possibile stabilire con chiarezza quale sia il modello migliore per le donne.

Per quanto riguarda i risultati ottenuti bisogna considerare che i dati prendono in considerazione una profondità storica non molto elevata ed inoltre l'intervallo out-of-sample analizzato ha subito delle perturbazioni legate alla crisi mondiale che prese avvio nel 2007 negli Stati Uniti d'America. Inoltre l'eterogeneità dei risultati è legata alle differenti misure di errore utilizzate per definire quale sia il modello migliore.

Come veniva osservato da K.Hanewald(2011) non esiste un "broad-brush solution", una modellizzazione unica che possa essere applicata alla mortalità in linea generale, in quanto questa varia molto in funzione del sesso, delle età e del Paese.

Tutti gli indicatori socioeconomici osservati presentano un'elevata correlazione con il fattore latente del modello di Lee-Carter e con i tassi di mortalità. La capacità di previsione di questi differenti modelli è però molto legata anche alla capacità di previsione delle serie storiche sulle quali si basano. Avendo preso in considerazione il periodo 2008-2016 ovviamente la stima delle grandezze socioeconomiche risulta alterata dalla crisi mondiale del 2007-08, per esempio l'andamento del PIL è cambiato dalla crisi passando da crescente a decrescente. Bisogna dunque prendere in considerazione la possibilità che studiando periodi storici differenti si potrebbero ottenere risultati dei test out-of-sample diversi in funzione della capacità di stimare i valori futuri delle grandezze socioeconomiche con maggiore o minore affidabilità.

## Capitolo 6

# Conclusione

Nell'ultimo trentennio sono stati proposti diversi modelli per la stima dei tassi di mortalità. Questo aumento dell'interesse per la stima dei tassi futuri è legato al crescente interessamento da parte degli Holders del longevity risk.

Il miglioramento della qualità della vita e la possibilità per la maggior parte della popolazione mondiale di reperire i beni di prima necessità ha aumentato in maniera notevole l'aspettativa di vita, creando molti problemi però in ambito previdenziale, sanitario e nella sussistenza stessa degli individui successivamente al loro pensionamento.

Nel 1992 è stato introdotto un modello estrapolativo vastamente accettato dagli accademici, il modello di Lee-Carter. Negli anni a seguire sono state proposte varie modifiche di questo modello che prevedevano sostanzialmente l'introduzione di ulteriori variabili che ne rinforzassero la capacità di previsione.

Nonostante i vari tentativi non è stato identificato un modello universale che possa essere considerato ottimale per tutti i Paesi del mondo. L'incapacità di raggiungere questo scopo è dovuta all'eterogeneità dell'andamento dei tassi di mortalità e delle grandezze socioeconomiche nei vari Paesi.

G.Niu e B.Melenberg hanno proposto nel 2014 un modello che prevedeva l'introduzione del PIL pro capite nel modello di Lee-Carter.

Applicando questo modello ai recenti dati italiani bisogna però prendere in considerazione che la correlazione fra il PIL pro capite ed il fattore latente del Lee-Carter è calata molto in epoca recente per effetto dei danni causati all'economia dalla crisi

---

mondiale del 2007. L'Italia, assieme al Giappone, è uno dei Paesi con più alta aspettativa di vita. Questo dato è il risultato di un fattore culturale e di una educazione alimentare fra le migliori al mondo. Questi fattori, essendo già intrinseci nella cultura del nostro paese, non sono legati in alcun modo con fattori macroeconomici. Essendo la mortalità legata a più fattori macroeconomici, socioeconomici e culturali abbiamo provato ad introdurre dei modelli affini a quello proposto da G.Niu e B.Melenberg ma sostituendo il PIL con altre grandezze che si ipotizza possano avere effetti impattanti sui tassi di mortalità.

I dati utilizzati nella tesi fanno riferimento ai tassi di mortalità ed ai decessi italiani per uomini, donne ed unisex per il periodo dal 1990 al 2016 per tutte le età dai 40 ai 90 anni. Sono state scartate le morti in età più elevate e quelle in età giovanile per eliminare gli effetti distorsivi che possono essere legati alle mortalità infantile, accidentale e la mortalità nelle età più estreme.

Gli indicatori socioeconomici studiati sono PIL pro capite, tasso di disoccupazione, debito pubblico sul PIL, indice di sviluppo umano depurato dal legame con l'aspettativa di vita, spesa sanitaria pubblica sul PIL, inflazione e stipendi.

Prima di tutto è stato applicato il modello di Lee-Carter ai nostri dati e successivamente è stata osservata la correlazione fra il fattore latente ed i tassi di mortalità per ogni età con gli indicatori socioeconomici proposti.

Sulla base dei risultati ottenuti vengono eliminati il tasso di disoccupazione ed il debito pubblico sul PIL. Il PIL non mostra una fortissima correlazione, se paragonata agli altri indici, ma decidiamo di non scartarlo per osservare l'applicazione del modello Niu-Melenberg nella forma originale.

Sui rimanenti cinque indici ed i fattori latenti del Lee-Carter per uomini, donne ed unisex viene testata la stazionarietà e l'ordine di integrazione. Tutte le serie rifiutano l'ipotesi nulla di non stazionarietà alla prima differenziazione. Tentiamo dunque di osservarne la relazione a lungo termine attraverso il Johansen test per la cointegrazione. Dai risultati non appare evidente una cointegrazione, ma contemporaneamente non rifiutano una possibile relazione di lungo termine fra i tassi di mortalità e le grandezze socioeconomiche.

Successivamente, per osservare la capacità di modellizzare la mortalità con gli indici

studiati, introduciamo dei modelli OLS dove gli indicatori socioeconomici sono le variabili esplicative. In questi modelli si osserva dunque il coefficiente di determinazione, che ci fornisce informazioni sulla varianza spiegata dai modelli. Tutti i modelli OLS, ad eccezione di quello con PIL come variabile esplicativa, presentano un  $R^2$  molto elevato. Dai risultati ottenuti sembrerebbe che questi indici possano fornire maggiori informazioni del PIL.

Introduciamo dunque dei modelli per la mortalità affini a quello di Niu-Melenberg, dove invece che il PIL pro capite vengono utilizzati gli indicatori socioeconomici proposti.

Su questi modelli si osserva poi la capacità di fitting attraverso il BIC, il MAPE ed il MAD. I risultati ottenuti sono misti, nessun modello sembra prevalere per tutti i sessi e tutte le misure di fitting, in generale si osserva che per gli uomini il modello con l'indice di sviluppo umano depurato sembra essere quello che garantisce il miglior fitting mentre per le donne ed unisex i risultati risultano essere molto più eterogenei e non si riesce a definire quale sia fra i modelli proposti il migliore.

Abbiamo utilizzato poi i modelli osservati per stimare i tassi di mortalità futuri per individui di 40 e 80 anni. La previsione, effettuata per ogni sesso, mostra risultati molto simili per tutti i modelli con andamenti dei tassi di mortalità decrescenti. Introducendo però degli intervalli di confidenza al 95% si osserva come alcuni modelli, come quello con l'inflazione e la spesa sanitaria, presentino intervalli di confidenza molto più ampi per la stima dei tassi di mortalità futuri. Da questo risultato si osserva come i modelli introdotti stimino un valore atteso molto simile per i tassi di mortalità ma con degli intervalli di confidenza molto differenti. Dai grafici sembrerebbe che i modelli con indice di sviluppo e con gli stipendi siano in grado di stimare con un maggior grado di affidabilità i tassi di mortalità futuri.

Per testare la capacità di previsione dei modelli introduciamo dei test out-of-sample. Vengono divisi i dati in due gruppi, in-sample dal 1990 al 2007 e out-of-sample dal 2008 al 2016. Utilizzando i dati in-sample abbiamo stimato nuovamente i parametri. Successivamente abbiamo stimato i tassi di mortalità per il periodo tra il 2008 ed il 2016 ed attraverso delle misure di errore MAPE e MAD si è confrontata la capacità di previsione dei vari modelli. Anche in questo caso per gli uomini il modello migliore

sembra essere il modello che utilizza l'indice di sviluppo umano depurato come fattore osservato mentre per unisex si ha che il modello originale di Niu-Melenberg stima con minore errore i tassi di mortalità futuri. Per le donne si ottiene per il MAPE un errore minore con il modello con gli stipendi mentre per il MAD con il modello originale di Niu-Melenberg.

In conclusione dunque, ci associamo ai vari studi già condotti finora, non esiste tuttora un modello universale che si possa considerare migliore degli altri e con maggiore capacità predittiva. L'introduzione degli indicatori socioeconomici ha portato ad un miglioramento della capacità predittiva del modello di Lee-Carter, però allo stesso tempo presenta il problema di una relativa instabilità, come è successo per il PIL con la crisi mondiale. Sulla base dei dati osservati dunque possiamo affermare con certezza che gli indicatori socioeconomici siano in grado di catturare parte della stocasticità legata agli andamenti futuri dei tassi di mortalità, ma, non possiamo dire a priori quale indicatore socioeconomico risulti universalmente migliore, ogni Stato e periodo storico potrebbero prediligere un indicatore differente.

L'effetto delle grandezze socioeconomiche impatta in maniera differente sulle differenti cause, ad esempio, come osservato da J.A.Tapia Granados(2009), il numero di morti per incidenti sul lavoro ed automobilistici aumenta nei periodi di espansione economica, mentre, un aumento della spesa sanitaria può ridurre il numero di morti per malattie cardiache o del fegato ma non impatta in alcun modo sulle morti per incidenti.

Sulla base dei risultati ottenuti e della letteratura studiata, come per esempio K.Hanewald(2011), possiamo affermare che lo studio dei tassi di mortalità futuri dovrebbe incentrarsi su una modellizzazione che sia in grado di prendere in considerazione singolarmente gli andamenti delle maggiori cause di morte.

La strada da perseguire probabilmente sarà quella di introdurre modelli sempre più complessi che prendano in considerazione differenti grandezze socioeconomiche e che siano in grado di modellizzare l'effetto delle variazioni di queste grandezze sulle differenti cause di morte.

# Bibliografia

- [1] Bacaër, N. (2010). *A Short History of Mathematical Population Dynamics* 1. Springer-Verlag London.
- [2] Boonen, J.T., & Li, H. (2017). Modeling and forecasting mortality with economic growth: a multi-population approach. *Demography* 54: 1921-1946.
- [3] Brenner, M. H. (2005). Commentary: Economic growth is the basis of mortality rate decline in the 20th century: Experience of the United States 1901–2000. *International Journal of Epidemiology* 34, 1214–1221.
- [4] Brouhns, N., Denuit, M., & Vermunt, J. (2002). A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics* 31: 373-393.
- [5] Cairns, A. J. G., Blake, D., & Dowd, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance*, 73, 687–718.
- [6] Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., Ong, A. & Balevich, I. (2009). A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal* 13(1): 1-35.
- [7] Cairns, A.J.G., Kallestrup-Lamb, M., Rosenskjold, C.P.T., Blake, D. & Dowd, K.(2016). *Modelling Socio-Economic Differences in the Mortality of Danish Males Using a New Auence Index*. Heriot-Watt University.

- 
- [8] Ciecka, J.E. (2008). Edmond Halley's Life Table and Its Uses. *Journal of Legal Economics* 15(1)9: 65-74.
- [9] D'Amato, V., Haberman, S., Piscopo, G. & Russolillo, M. (2014). Computational framework for longevity risk management. *Computer Management Science* 11: 111-137.
- [10] De Moivre, A. (1725). *Annuities upon Lives*.
- [11] Deng Y., Brockett, P.L. & MacMinn, R.D. 2012. Longevity/mortality risk modelling and securities pricing. *The Journal of Risk and Insurance* 79(30): 697-721.
- [12] Ettner, S. L. (1996). New evidence on the relationship between income and health. *Journal of Health Economics* 15: 67-85.
- [13] French, D., & O'Hare, C. (2014). Forecasting death rates using exogenous determinants. *Journal of Forecasting* 33(8): 640-650.
- [14] Gaille, S., & Sherris, M. 2010. Age patterns and trends in mortality by cause of death and implications for modelling longevity risk. The University of New South Wales, Australian School of Business Research Paper: 2010ACTL12.
- [15] Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London* 115: 513-585.
- [16] Guirrieri, S.S. (2008). *Analisi econometrica delle serie storiche con R*. Università degli Studi di Palermo: 39-50.
- [17] Halley, E. (1693). An Estimate of the Degrees of Mortality of Mankind, Drawn from the Curious Tables of the Births and Funerals at the City of Breslaw, with an Attempt to Ascertain the Price of Annuities upon Lives. *Philosophical Transactions* 17: 596-610
- [18] Hanewald, K. (2011). Explaining mortality dynamics: The role of macroeconomics fluctuations and cause of death trends. *North American Actuarial Journal* 15: 290-314.

- 
- [19] Heligman, L., Pollard, J.H. (1980). The age pattern of mortality. *Journal of the Institute of Actuaries* 107: 49-80.
- [20] Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. *Journal of Econometrics* 54: 159-178.
- [21] Lee, R.D., & Carter, L.R. (1992). Modelling and forecasting U. S. mortality. *Journal of the American Statistical Association* 87 (419): 659-671.
- [22] Li, N., & Lee, R. D. (2005). Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography* 42, 575-594.
- [23] Lucchetti, R.(2015). *Appunti di analisi delle serie storiche*: 107-128.
- [24] Makeham, W. M. (1860). On the Law of Mortality and the Construction of Annuity Tables". *J. Inst. Actuaries and Assur. Mag.* 8: 301-310.
- [25] Neumayer, E. (2005). Commentary: the economic business cycle and mortality. *International journal of epidemiology*, 34 (6). 1221-1222.
- [26] Niu, G., & Melenberg, B. (2014). Trends in mortality decrease and economic growth. *Demography* 51: 1755-1773.
- [27] O'Hare, C., & Li, Y. (2011). Explaining young mortality. *Insurance: Mathematics and Economics* 50: 12-25.
- [28] Phillips, P. C. B., & Perron, P. (1988). Testing for a Unit Root in Time Series Regression. *Biometrika* 75 (2): 335-346.
- [29] Plat, R. (2009). On stochastic mortality modelling. *Insurance: Mathematics and Economics* 45(3): 393-404.
- [30] Preston, S. H (1975). The Changing Relation between Mortality and Level of Economic Development. *Population Studies*. 29 (2): 231-248.

- 
- [31] Renshaw, A.E., & Haberman, S. (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics* 38: 556-570.
- [32] Roy, A. 2012. Innovative approaches to managing longevity risk in Asia: Lessons from the west. Asian Development Bank Institute Working Paper Series: 353.
- [33] Seklecka, M., Md Lazam, N., Pantelous, A., & O'Hare C. (2017), Mortality Effects of Economic fluctuations in the selected Eurozone countries. SSRN Electronic Journal.
- [34] Swift, R. (2011). The relationship between health and GDP in OECD countries in the very long run. *Health Economics* 20: 306–322.
- [35] Tapia Granados, JA. (2005). Increasing mortality during the expansions of the US economy, 1990–1996. *Int J Epidemiol* 34: 1194–1202.
- [36] Tapia Granados, J.A. (2008). Macroeconomic Fluctuations and Mortality in Postwar Japan. *Demography* 45: 323-343.
- [37] Tapia Granados, J.A., Ionides, E.L. (2008b). The reversal of the relation between economic growth and health progress: Sweden in the 19th and 20th centuries. *Journal of Health Economics* 27: 544-563.
- [38] Tapia Granados, J. A., & Roux, A. V. D. (2009). Life and death during the Great Depression. *Proceedings of the National Academy of Sciences* 106: 17290–17295.
- [39] Tapia Granados, J.A. (2012). Economic growth and health progress in England and Wales: 160 years of a changing relation. *Social Science and Medicine* 74: 688-695.
- [40] Yang, S.S., Yue, J.C., Huang, H.C. (2010). Modelling longevity risks using a principal component approach: A comparison with existing stochastic mortality models. *Insurance: Mathematics and Economics* 46: 254-270.

## Sitografia

- [1] Coughnal, G., Epstein, D., Ong, A., Sinha, A., Balevich, I., Hevia-Portocarrero, J., Gingrich, E., Khalaf Allah, M. & Joseph, P. (2007). LifeMetrics: A Toolkit for Measuring and Managing Longevity and Mortality Risks. London: JP Morgan, in <http://www.macs.hw.ac.uk/~andrewc/lifemetrics/> (visitato il 12 Novembre 2018)
- [2] European Southern Observatory, in <http://www.eso.org> (visitato il 5 Novembre 2018)
- [3] Human Development Reports, in <http://hdr.undp.org/en/data> (visitato il 10 Novembre 2018)
- [4] International Monetary Fund, in <https://www.imf.org/external/index.htm> (visitato il 7 Novembre 2018)
- [5] Levantesi S. & Menzietti M. (2011, 2 Dicembre). I modelli di proiezione della mortalità. [http://www.ordineattuari.it/media/5704/Incontro%20Pensioni%202011%20-%20Lev-Menz\\_0AB.pdf](http://www.ordineattuari.it/media/5704/Incontro%20Pensioni%202011%20-%20Lev-Menz_0AB.pdf) (visitato il 3 Dicembre 2018)
- [6] Organisation for Economic Co-operation and Development, in <https://stats.oecd.org/Index.aspx> (visitato il 6 Novembre 2018)
- [7] Ortiz-Ospina, E. & Beltekian, D. (2018, 14 Agosto). Why do women live longer than men?, in <https://ourworldindata.org/why-do-women-live-longer-than-men> (visitato il 22 Gennaio 2019)

- 
- [8] Peracchi, F. (2012). Cointegrazione. in [http://www.treccani.it/enciclopedia/cointegrazione\\_%28Dizionario-di-Economia-e-Finanza%29/](http://www.treccani.it/enciclopedia/cointegrazione_%28Dizionario-di-Economia-e-Finanza%29/) (visitato il 15 Dicembre 2018)
- [9] Smargiassi, G. (2015). Inflazione e deflazione, in [http://www.treccani.it/enciclopedia/inflazione-e-deflazione\\_%28Enciclopedia-Italiana%29/](http://www.treccani.it/enciclopedia/inflazione-e-deflazione_%28Enciclopedia-Italiana%29/) (visitato il 27 Dicembre 2018)
- [10] Statistiche Istat, <http://dati.istat.it/> (visitato il 5 Novembre 2018)
- [11] The world Bank, in [https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?end=2007&locations=US-CA-NL-JP-GB&name\\_desc=false&start=1960](https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?end=2007&locations=US-CA-NL-JP-GB&name_desc=false&start=1960) (visitato il 5 Gennaio 2019)

# Ringraziamenti

Prima di tutto vorrei ringraziare la Prof.ssa Susanna Levantesi che mi ha aiutato nelle stesura della tesi. Poi vorrei ringraziare la mia famiglia che mi ha sempre sostenuto, ascoltato, aiutato e spronato in questi anni. Un grazie speciale va anche a tutti i miei amici, ai compagni di una vita del "vittoria", al Team Pincay, ai miei amici di Madrid, di Dortmund, di Lisbona ed a tutti i miei colleghi universitari con cui ho condiviso ansie e gioie. In particolar modo vorrei poi ringraziare i miei amici più cari che mi aiutano e sopportano da una vita: Peolo, Mimmo, Federico, Matteo, Mirco, Andrea e Frank.