Common Factor Decomposition of Cause-specific Mortality Rates Using the Cointegration Analysis

Viktoriya Glushko
Séverine Arnold (-Gaille)
Faculty of Business and Economics (HEC Lausanne)
University of Lausanne

Longevity 14
Amsterdam
2018

Table of contents

Table of contents

Motivation for the work

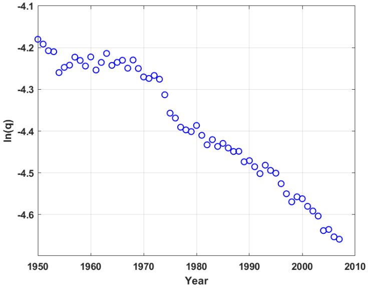
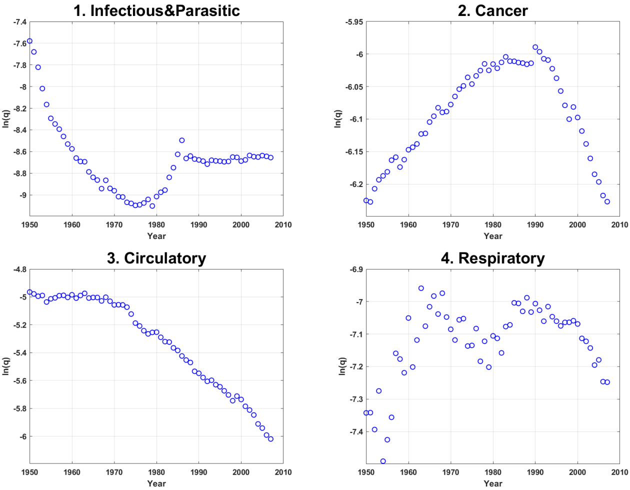


Figure: Total mortality rates after age standardization, US males, WHO mortality database

- No news for everyone, especially actuaries:
  $\rightarrow$ Mortality rates are decreasing over time
- Important amount of research trying to undestand and describe the dynamics of this process

Example of cause-specific mortality rates



Example of cause-specific mortality rates

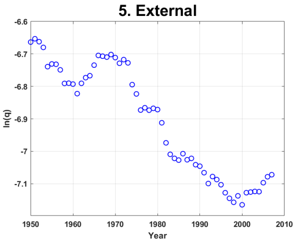


Figure: Cause-specific mortality rates after age standardization, US males, WHO mortality database

- Numerous effects going in different directions
  $\rightarrow$ Significant loss of information if only total death rates are analysed

Objective

- Better understand and model the mortality rates
- By taking into account the cause-specific mortality rates
- Which obviously are NOT independent (competing risks)

Data

Countries:

- USA (1950 - 2007)
- Japan (1950 - 2009/2013)
- France (1952 - 2008/2011)
- England & Wales (1950 - 2009/2013)
- Australia (1950 - 2004)

Causes of death:

- 1. Infectious & parasitic diseases
- 2. Cancer
- 3. Diseases of the circulatory system
- 4. Diseases of the respiratory system
- 5. External cau ses (mainly accidents)
- (6. Other causes)

Table of contents

The model

[Hamilton(1994), Lütkepohl(2005)]:

Vector autoregressive model VAR(p)

$$\mathbf{y}_t = \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{\Phi}_2 \mathbf{y}_{t-2} + \cdots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \epsilon_t$$

where

$\mathbf{y}_t$ is a $(n \times 1)$ vector of observations

$\mathbf{\Phi}_t$ is a $(n \times T)$ parameter matrix

$$E(\epsilon_t) = \mathbf{0}$$

$$E(\epsilon_t \epsilon_l) = \begin{cases} \mathbf{\Omega} & \text{for } t = l \\ \mathbf{0} & \text{for } t \neq l \end{cases}$$

The model

- Process $\mathbf{y}_t$ must be stationary
- By eye inspection and formal unit roots tests: most cause-specific mortality rates are not stationary

  $\rightarrow$ VAR(p) is not directly applicable

If non-stationary variables

- Apply first differences, i.e. detrend the series :

$$\Delta \mathbf{y}_t = \mathbf{y}_t - \mathbf{y}_{t-1}$$

  → Some information is inevitably lost

- Alternative way: use cointegrating relations

  → What if some linear combination of non-stationary variables is stationary? = cointegration relation

  → Then there exist long-term equilibrium relationships among the variables

  → Which can be included into the autogressive model

New model

Vector Error Correction Model of the cointegrated system

$$\Delta \mathbf{y}_t = \xi_1 \Delta \mathbf{y}_{t-1} + \xi_2 \Delta \mathbf{y}_{t-2} + \cdots + \xi_{p-1} \Delta \mathbf{y}_{t-p+1} + \mathbf{\Pi} \mathbf{y}_{t-1} + \epsilon_t$$

→ Equation valid only if $\mathbf{\Pi y}_{t-1}$ is stationary

→ Cointegrated term preserves the information on the long-term equilibrium

Johansen procedure

[Johansen (1994)] :

- Allows to find the matrix $\mathbf{\Pi}$ as

$$
\begin{aligned}
\mathbf{\Pi} &= \alpha\beta', \text{ matrix of rank } r \text{ where} \\
& \quad r \text{ is the number of cointegration relations} \\
\beta &= \text{a } (n \times r) \text{ matrix containing } r \text{ vectors} \\
& \quad \text{each representing a cointegration relation} \\
\alpha &= \text{a } (n \times r) \text{ loading matrix}
\end{aligned}
$$

- And to test if any deterministic element should be included

Table of contents

Algorithm

- Test for the number of lags in VAR(p)
- Apply formal tests for unit roots
- Test for the number of cointegration relations r and the form of the deterministic elements
- Calculate the matrix $\mathbf{\Pi}$ and the rest of the VECM coefficients
- Check the model fit, i.e. test the residuals for autocorrelation and normality

Preliminary results

Arnold[2016], 5 causes :

- For all countries and sexes the Johansen procedure showed at least 1 cointegration relation
- Sometimes with, sometimes without the time trend

→ Lont-term equilibrium relations between the causes exist

Worse model fit if the cause "Other" is added :(

Testing significance

$$\Delta \mathbf{y}_t = \mathbf{\Pi y}_{t-1} + \xi_1 \Delta \mathbf{y}_{t-1} + \epsilon_t$$

$$\text{For } r = 1 : \mathbf{\Pi y}_{t-1} = \alpha\beta' \mathbf{y_{t-1}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 \end{bmatrix} \mathbf{y}_{t-1} =$$

$$= \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{bmatrix} (\beta_1 y_{1t-1} + \beta_2 y_{2t-1} + \beta_3 y_{3t-1} + \beta_4 y_{4t-1} + \beta_5 y_{5t-1})$$

- Step 1 : test if a particular cause enters the long-term cointegration relation, i.e. if β_i is significant
  $\rightarrow$ Also possible to test a combination of causes

Testing significance - β_i

Arnold[2016], 5 causes :

- In 7 out of 10 datasets the causes Infectious & Parasitic and External were together not significant
  $\rightarrow$ This corresponds to the intuitive guess that exo- and endogenous causes should behave differently over the long-term
  $\rightarrow$ Then, the long-term equilibrium relations can be seen as somehow corresponding to the aging process in the humain body

Testing significance - α_i

- Test if the cointegration relation has an impact on a particular cause, i.e. if α_i is significant
- p - values of α_i for males :

	d(IP)t	d(Canc)t	d(Circ)t	d(Resp)t	d(Ext)t
US males	0.000 ***	0.054 .	0.343	0.004 **	0.126
JP males	0.041 *	0.518	0.000 ***	0.000 ***	0.947
FR males	0.048 *	0.610	0.000 ***	0.000 ***	0.049 *
EW males	0.040 *	0.002 **	0.078 .	0.000 ***	0.008 **
AU males	0.021 *	0.005 **	0.002 **	0.000 ***	0.611

Testing significance - α_i

- p - values of α_i for females :

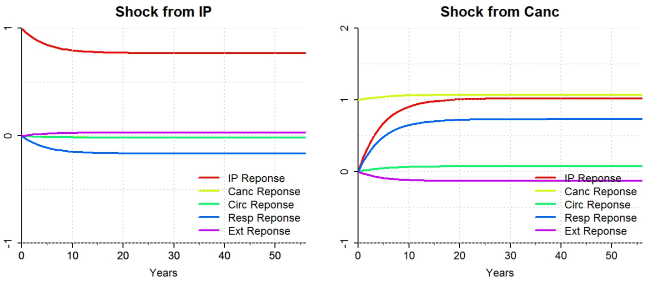
	d(IP)t	d(Canc)t	d(Circ)t	d(Resp)t	d(Ext)t
US females	0.000 ***	0.000 ***	0.505	0.041 *	0.569
JP females	0.047 *	0.021 *	0.000 ***	0.000 ***	0.622
FR females	0.174	0.054 .	0.001 ***	0.000 ***	0.015 *
EW females	0.000 ***	0.015 *	0.089 .	0.028 *	0.030 *
AU females	0.000 ***	0.801	0.003 **	0.987	0.011 *

Testing significance - α_i

- Infectious & Parasitic and Respiratory causes seem to be the most impacted by the cointegration relation : α_i is significant in 9 out of 10 datasets
- External cause seems to be the least impacted : α_i is significant in 5 out of 10 datasets
- Results for Cancer and Circulatory causes are more difficult to interpret : α_i is significant in 6 out of 10 datasets

Impulse-response analysis

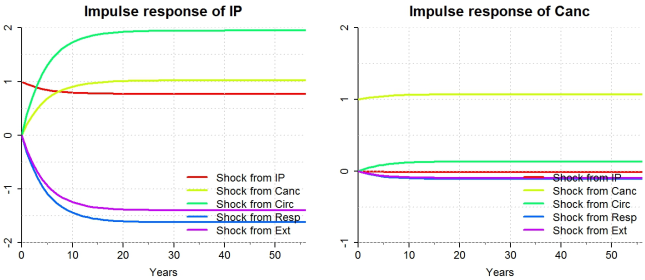
- How did a cause-specific mortality rate respond to a random shock from another cause (one at a time)?
- Example: US males, random shock from the IP, i.e. vector of starting values for the VECM is $y_0 = (1, 0, 0, 0, 0)$



→ Weak reaction to shock from IP, IP and Resp sensitive to shock from Cancer

Impulse-response analysis

- Alternative way: how sensitive is a particular cause-specific mortality rate to a random shock from every other cause (one at a time)?



→ While the IP mortality rate is rather reactive, mortality rate due to cancer is highly stable

Table of contents

Concluding remarks

- Cointegration analysis proves the dependence between the cause-specific mortality rates and the existence of long-term equilibrium relations
- The common trend between the endogenous causes of death can be seen as a mathematical reflection of the biological process of aging
- We can see the extent to which a particular cause enters the long-term cointegration relation, but also how much it in its turn is influenced by the cointegration relation(s)
- Next steps: study further the decomposition between the long and short-term elements of the VECM, including common factors

Bibliography

- S. Arnold and M. Sherris. International cause-specific mortality rates: new insights from a cointegration analysis. ASTIN Bulletin, January 2016, pp 1- 30.
- S. Johansen : The role of the constant and linear terms in cointegration analysis of nonstationary variables, Econometric Reviews, 1994, 13:2, 205-229
- J. D. Hamilton. Time Series Analysis. Princeton University Press, 1994.
- H. Lütkepohl. New Introduction to Multiple Time Series Analysis. Crown Publishing Group, 2005.
- W. Kuo, C. Tsai, W.-K. Chen. An empirical study on the lapse rate: The cointegration approach. The Journal of Risk and Insurance, 2003, 70(3), 489-508

Thank you for your attention!

Viktoriya.Glushko@unil.ch