The Joint Meeting of the 2011 Taipei International Statistical Symposium and IASC-CAR (Joint2011)

A Discussion on Environmental Risks for Breast and Liver Cancer by Analyzing (Age, period)-tabulated Data

Nobutane Hanayama

Department of Information Technology, Shobi University

Shimomatsubara 655, Kawagoe, Saitama 3501153, Japan

E-mail: nob-hanayama@jcom.home.ne.jp

Outline

(1) Background

(2) Aim

(3) Procedure

(4) Data

(5) Lexis diagram

(6) Age period cohort model

(7) Age period environment model

(8) Identifiability problem

(9) Results

(10) Conclusion

Background

Dietary Soy Isoflavones Inhibit Estrogen Effects in the Postmenopausal Breast

Cancer Res 2006;66:1241-1249.

Charles E. Wood1, Thomas C. Register1, Adrian A. Franke2, Mary S. Anthony1, and J. Mark Cline1

In the presence of estrogen higher doses of dietary soy isoflavones may alter estrogen receptors signaling and induce selective antagonistic effects in the breast.

Background

High soybean consumption increases risk of liver cancer in women: survey

http://www.hepcaustralia.com.au/symptoms-news/high-soybean-consumption-increases-risk-of-liver-cancer-in-women-survey

Women who eat a lot of soybean products are three to four times more likely to develop liver cancer than those who eat only a small amount.

Estrogen heightens the risk of breast cancer, but protects against liver cancer. It is believed that excessive consumption of isoflavones may block this property.

Aim

Our aim in this study is

to examine whether there exist environmental factors which decrease the breast cancer risk and increase the liver cancer risk for postmenopausal females aged 55 and over based on the analysis of (age, period)-tabulated data.

Procedure

Our procedure is

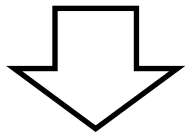
(1) estimating the environmental risks of breast cancer and liver cancer for female based on (age, period)-tabulated data,

(2) testing for the significance of the correlation coefficient between environmental risks for breast cancer and liver cancer.

Procedure

However

Hepatitis B virus infection is the main causes of liver cancer. Liver cancer correlates with the frequency of chronic hepatitis B virus infection.



It is needed to eliminate the effect of infection to hepatitis B virus from environment effects.

Procedure

Difference between sexes

Isoflavone consumption and subsequent risk of hepatocellular carcinoma in a population-based prospective cohort of Japanese men and women

Int J Cancer. 2009 Apr 1;124(7):1644-9

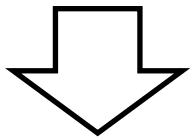
Kurahashi N, Inoue M, Iwasaki M, Tanaka Y, Mizokami M, Tsugane S; JPHC Study Group

In women, genistein and daidzein were dose-dependently associated with an increased risk of hepatocellular carcinoma (HCC, liver cancer). No association between isoflavones and HCC was observed in men.

Procedure

Difference between sexes

In women, genistein and daidzein were dose-dependently associated with an increased risk of hepatocellular carcinoma (HCC, liver cancer). No association between isoflavones and HCC was observed in men.



The difference of environmental effect on liver cancer between men and women is effect of isoflavone on breast cancer for women.

Procedure

Our modified procedure is

(1) estimating the environmental risks of breast cancer for female and liver cancer for both sexes,

(2) calculating difference of environmental risks between men and women,

(3) testing for the significance of the correlation coefficient between environmental effects on breast cancer and the difference of environmental effect on liver cancer between men and women.

(Age, period)-tabulated data

Table 1. Deaths of liver cancer in 100 thousand Japanese males aged 55 and over

age	Period							
	1970-74	1975-79	1980-84	1985-89	1990-94	1995-99	2000-04	2005-09
55～59	14.5	13.7	12.7	13.7	12.5	10.5	8.2	6.3
60～64	22.8	21.1	21.8	23.0	25.0	26.4	18.4	12.7
65～69	35.8	32.9	32.3	35.4	38.4	47.2	40.4	27.0
70～74	50.5	46.9	46.6	48.6	51.9	62.0	64.4	54.5
75～79	66.4	62.1	63.9	62.0	65.6	75.6	78.9	78.1

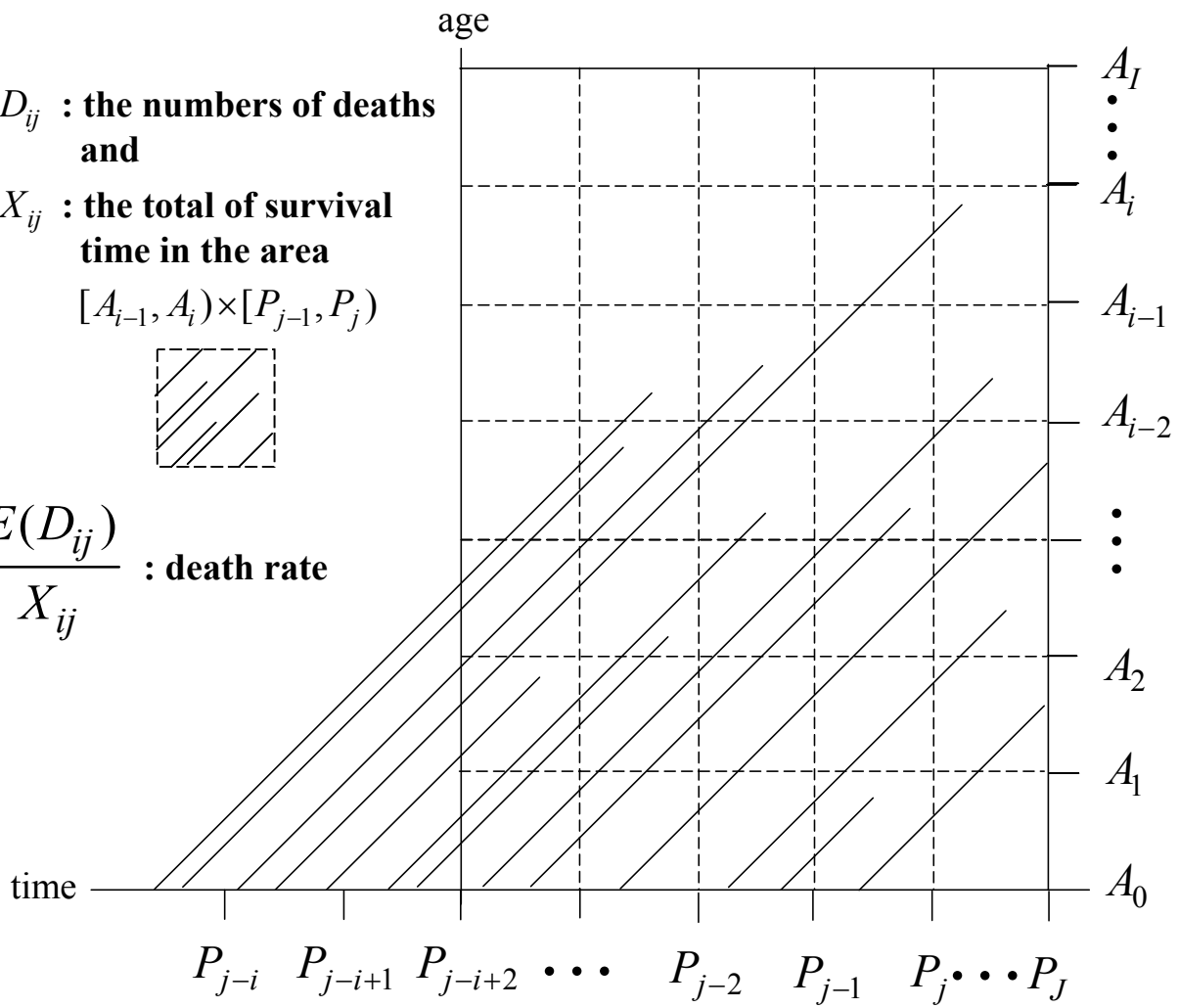
Deaths of liver and breast cancer in 100 thousand Japanese females aged 55 and over are given in the same formula.

Lexis diagram

D_{ij} : **the numbers of deaths and**

X_{ij} : **the total of survival time in the area** $[A_{i-1}, A_i) \times [P_{j-1}, P_j)$

$$\lambda_{ij} = \frac{E(D_{ij})}{X_{ij}}$$: **death rate**



Age period cohort model

$$\log \lambda_{ij} = \mu + \alpha_i + \beta_j + \gamma_{j-i}$$

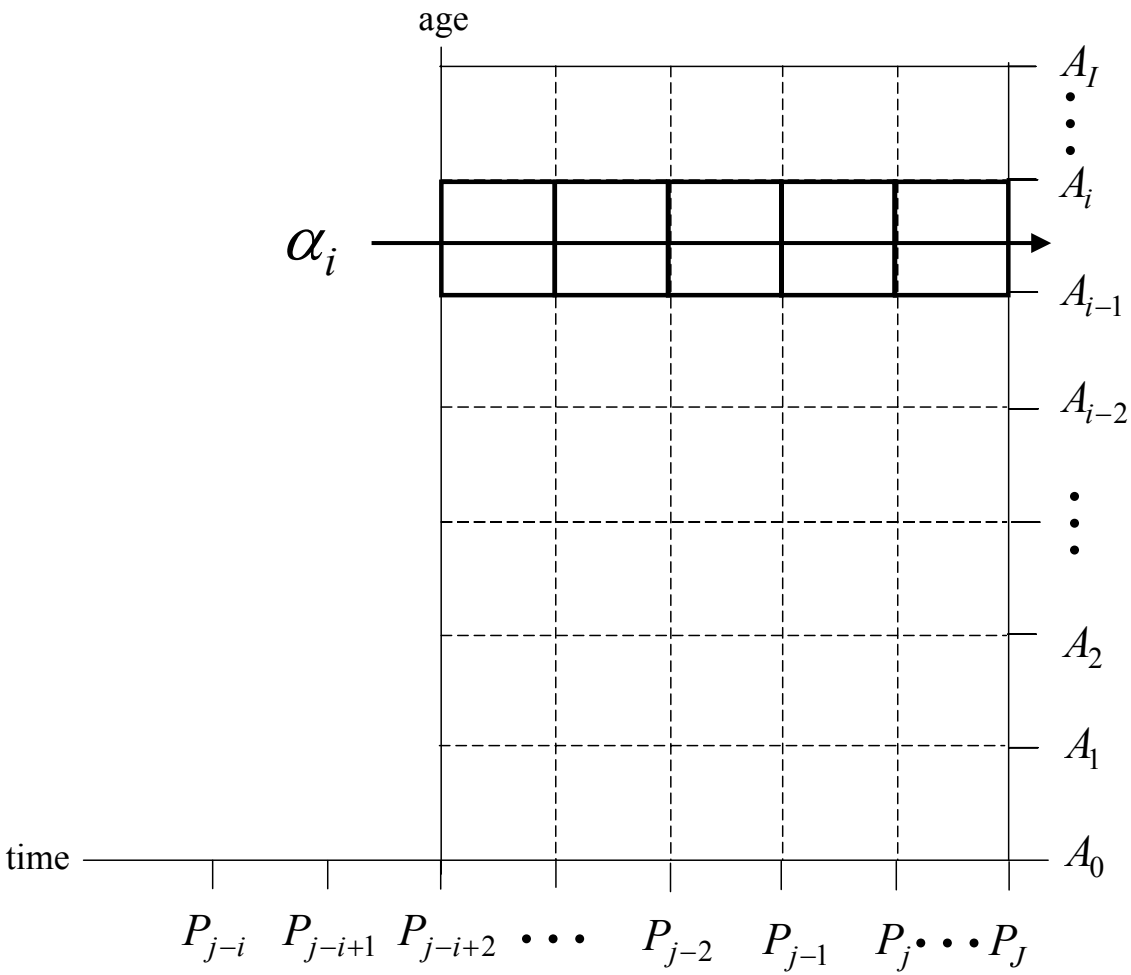
α_i : **age effect**

β_j : **period effect**

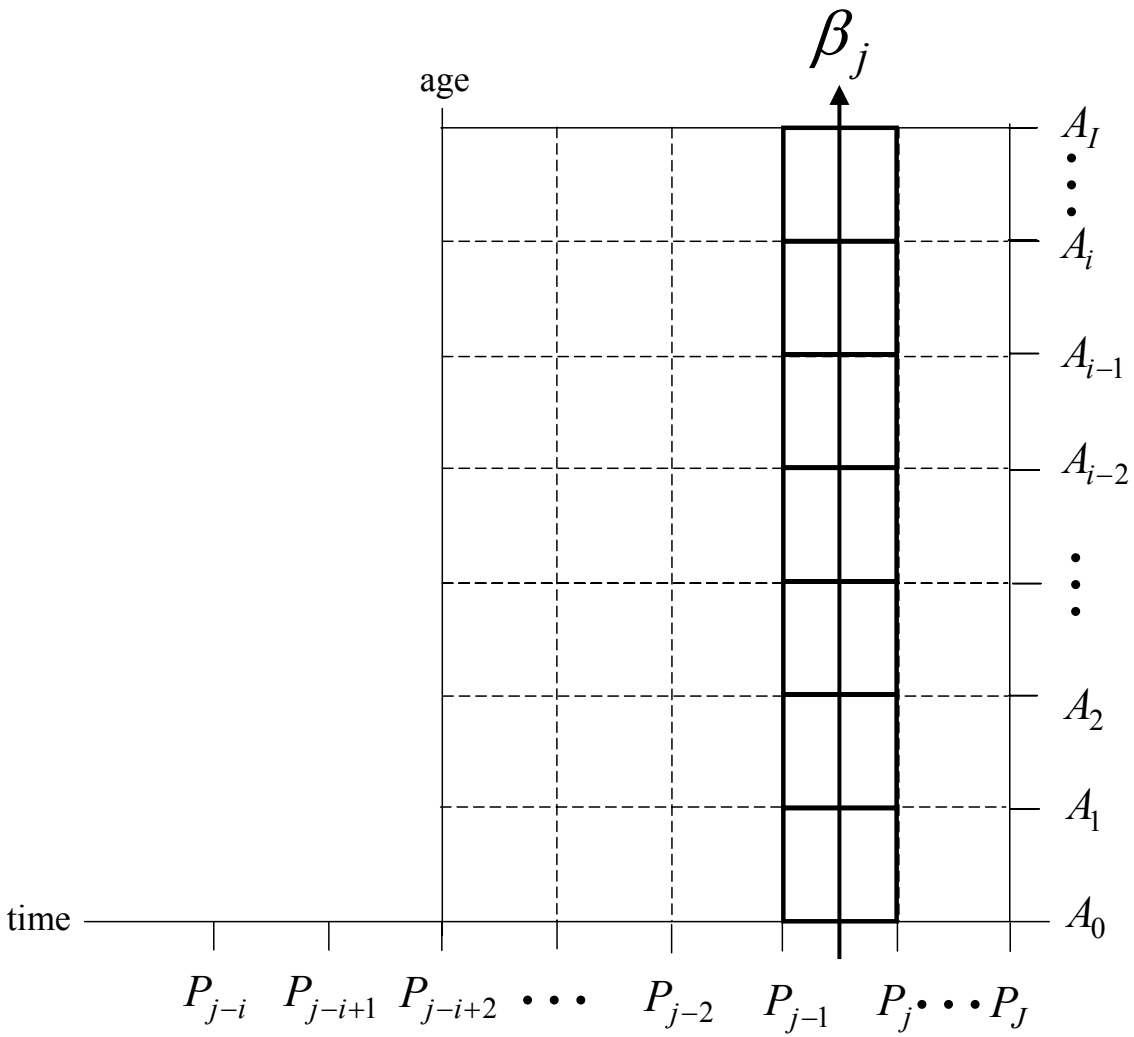
γ_{j-i} : **cohort effet**

Glenn (1977), Fienberg and Mason (1978), Osmond and Gardener (1982), Holford (1983), Kupper and Janis (1985), Clayton and Schifflers (1987), Tango and Kurashina (1987)

Age effect


age
A_I
A_i
α_i
A_{i-1}
A_{i-2}
A_2
A_1
A_0
time
P_{j-i} P_{j-i+1} P_{j-i+2} • • • P_{j-2} P_{j-1} P_j • • • P_J

Period effect

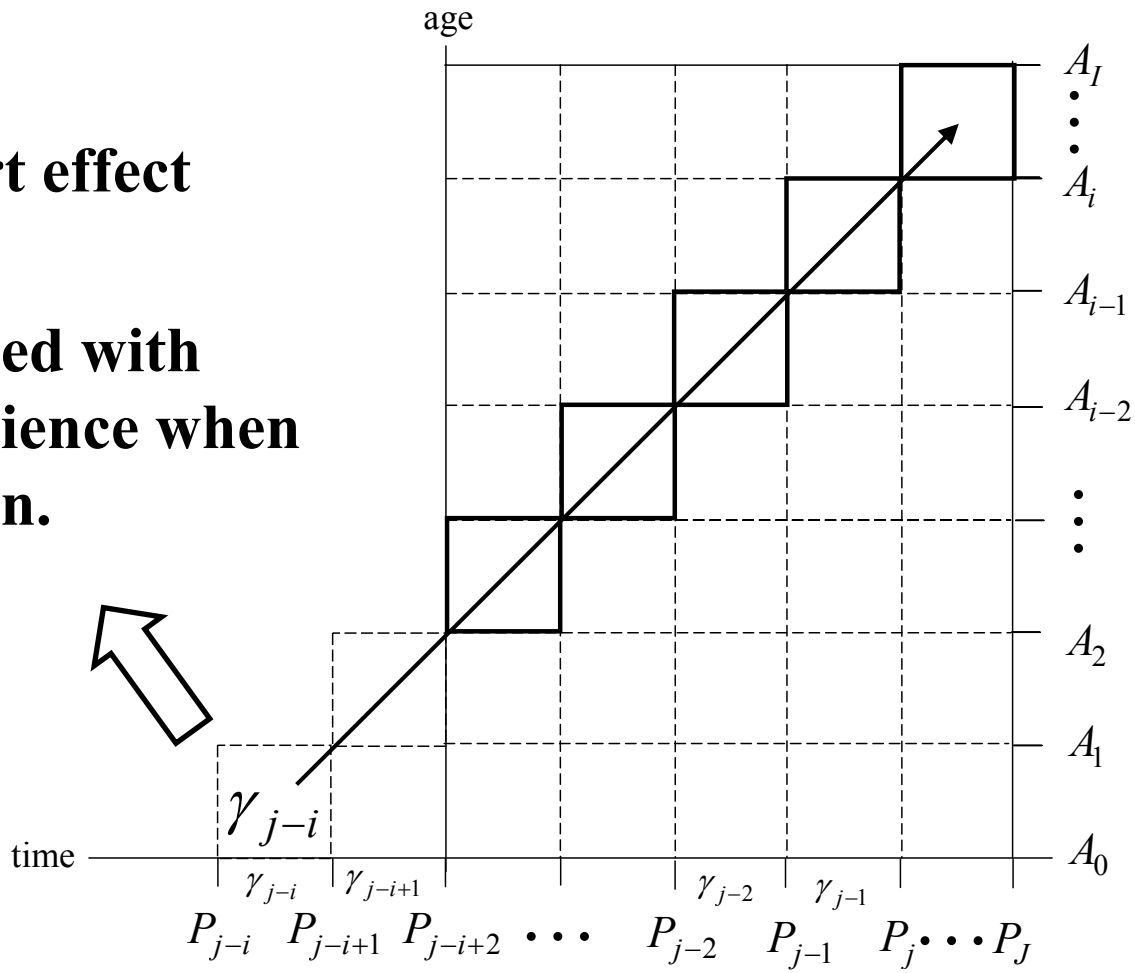


Cohort effect

Birth cohort effect

⇧

Effect associated with people's experience when they were born.



Questions about birth cohort effect

(i) Human experiences in their babyhood may be the most significant for determining their features.

Are experiences in other stages of life also significant?

(ii) It is supposed that the effect of experience in their babyhood appear immediately after they experience them.

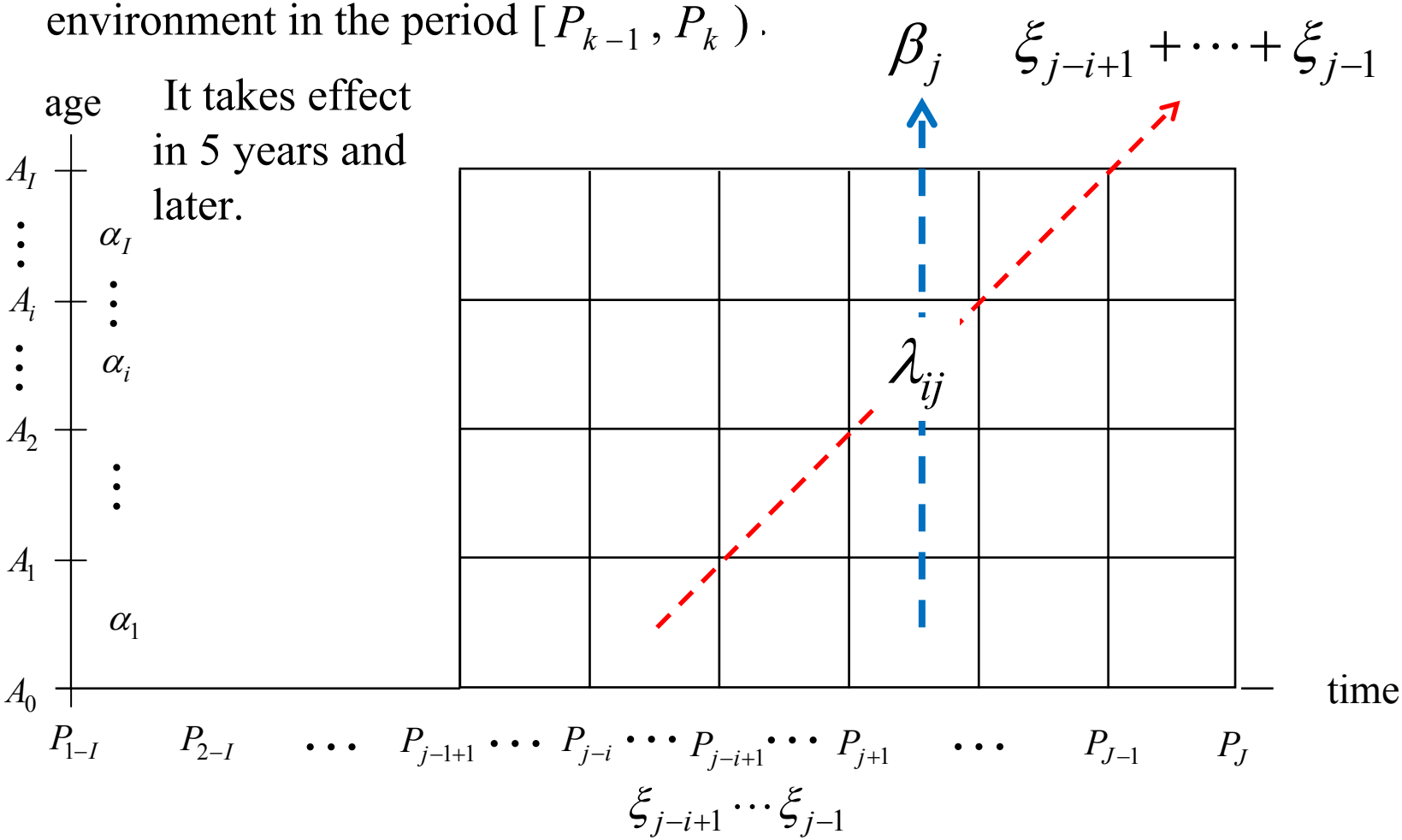
Is such a effect equal to the period effect?

⇩

Are period and cohort effect overlapped in their definitions?

Environment effect

ξ_k: effect associated with exposure to the environment in the period $[P_{k-1}, P_k)$.



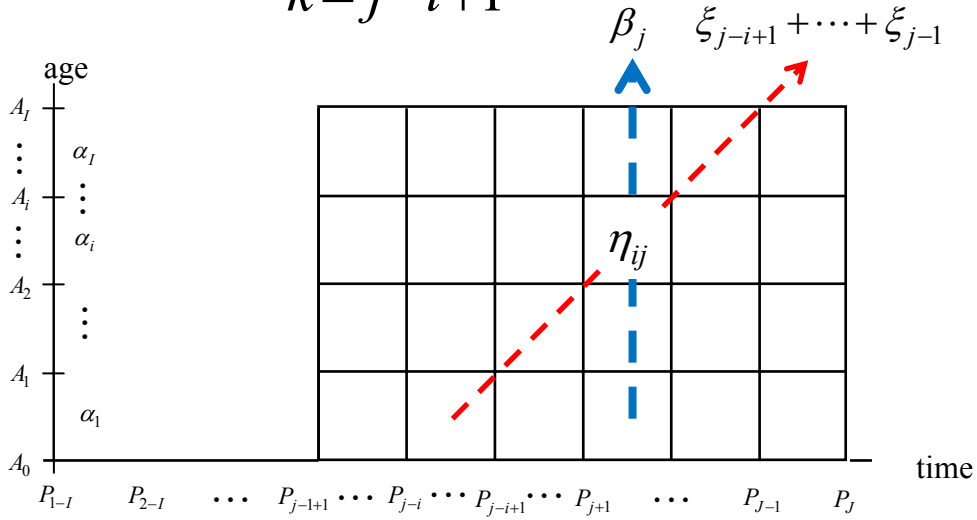
Age period environment model

$$\log \lambda_{ij} = \mu + \alpha_i + \beta_j + \sum_{k=j-i+1}^{j-1} \xi_k$$

α_i : age effect

β_j : period effect

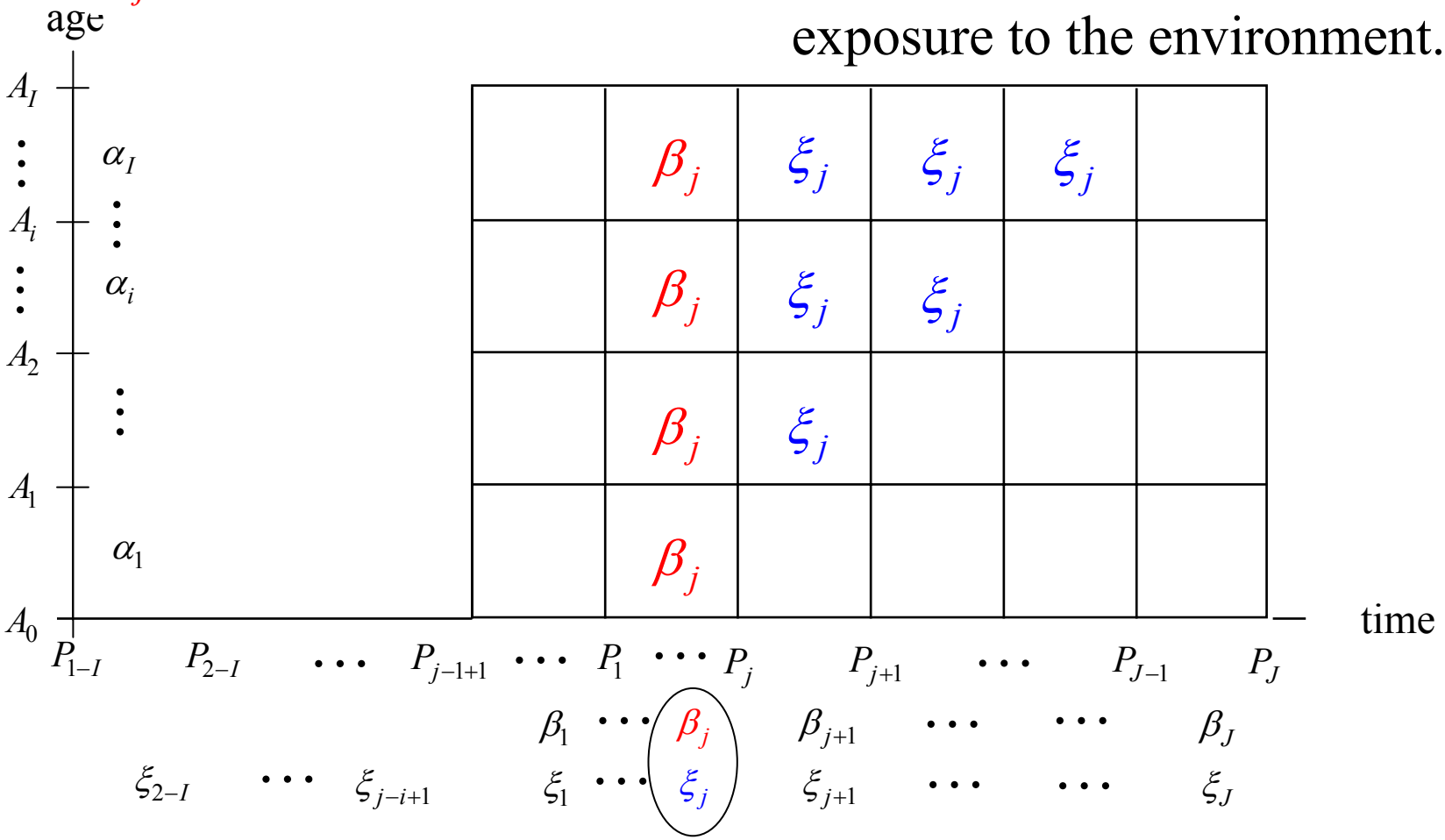
ξ_k : environment effect



Environment effects from people's birth to 5 years before present period are accumulated.

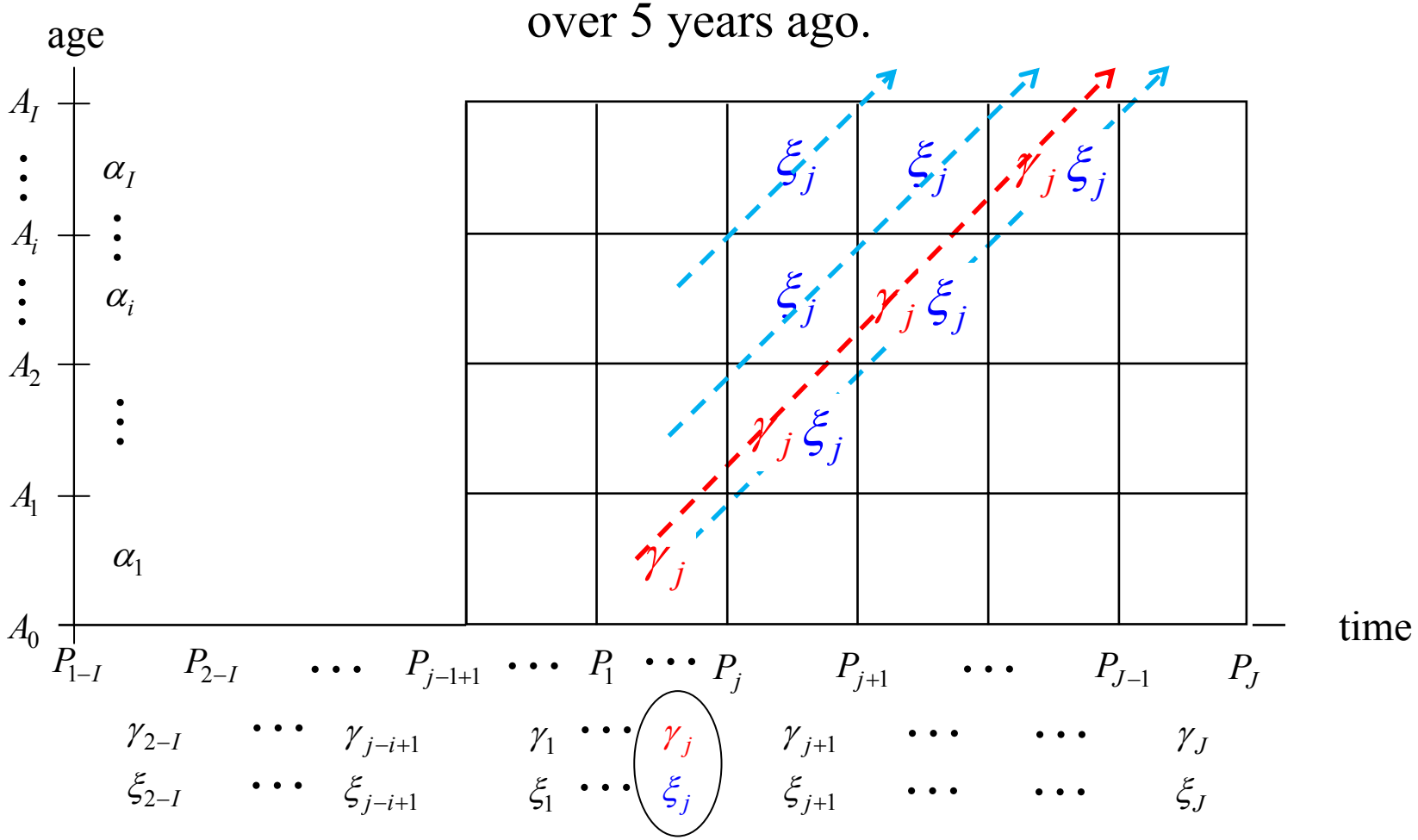
Period and environment effect

β_j appear immediately ξ_j appear 5 years later since exposure to the environment.



Period and environment effect

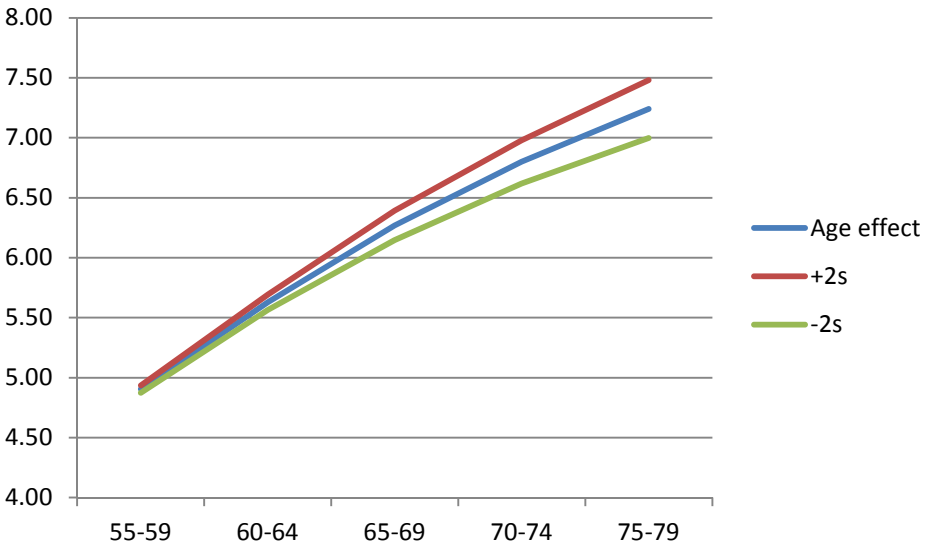
γ_j : affects one cohort ξ_j affects all cohort exposed to the environment over 5 years ago.



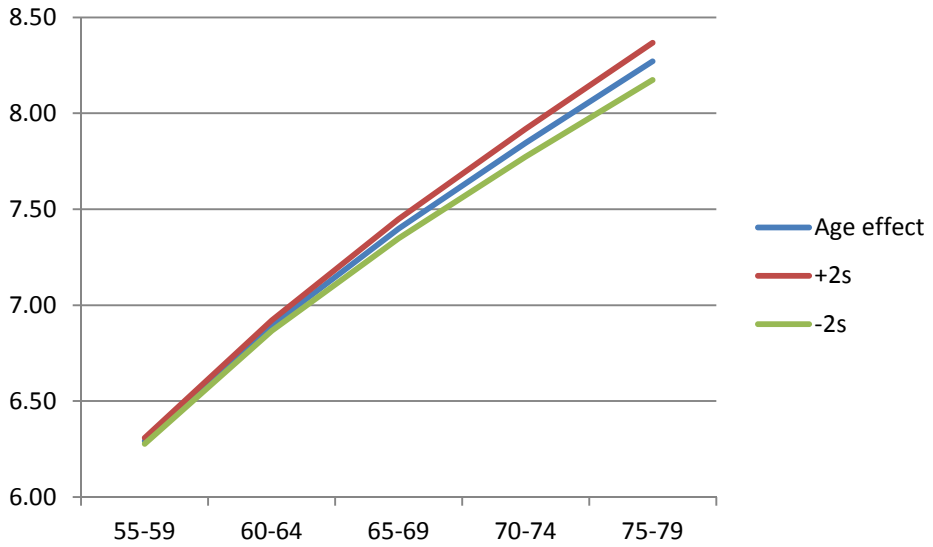
Results

Age effect on liver cancer

Female



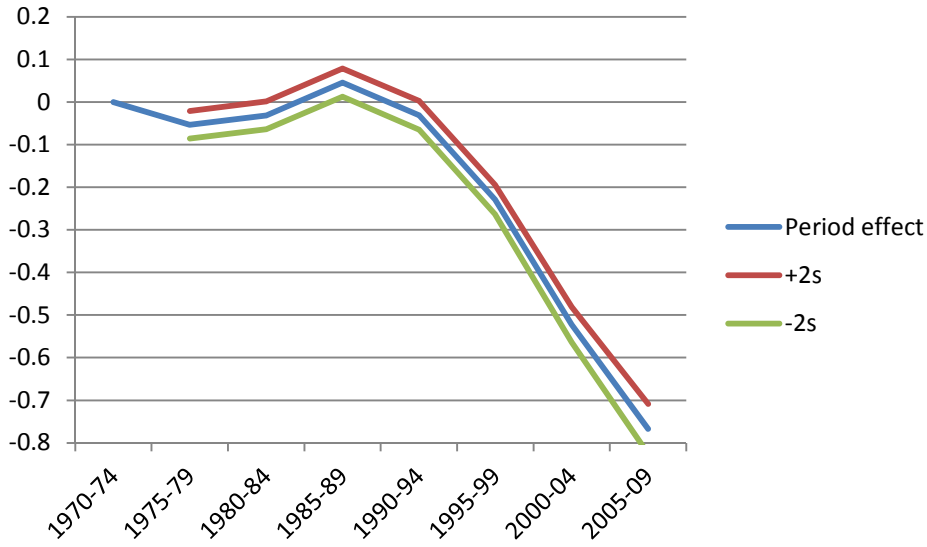
Male



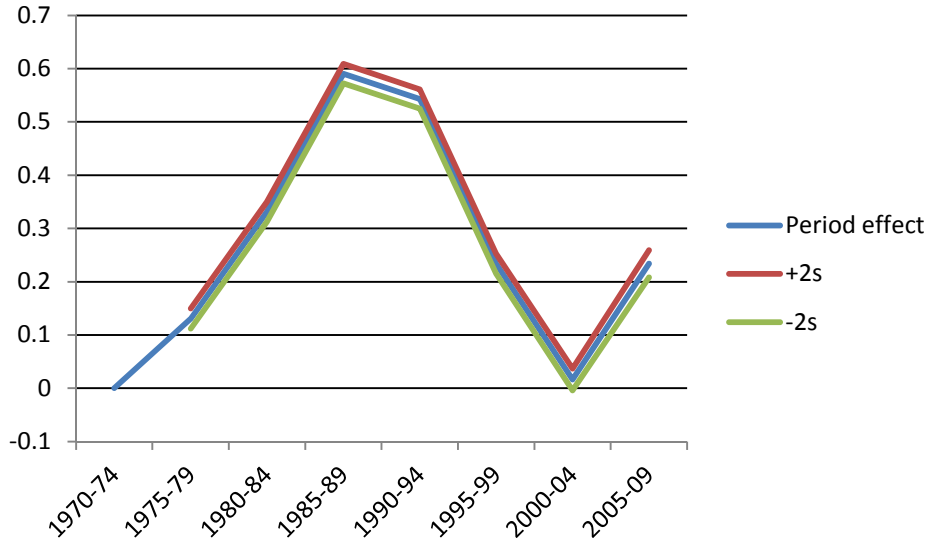
Results

Period effect on liver cancer

Female



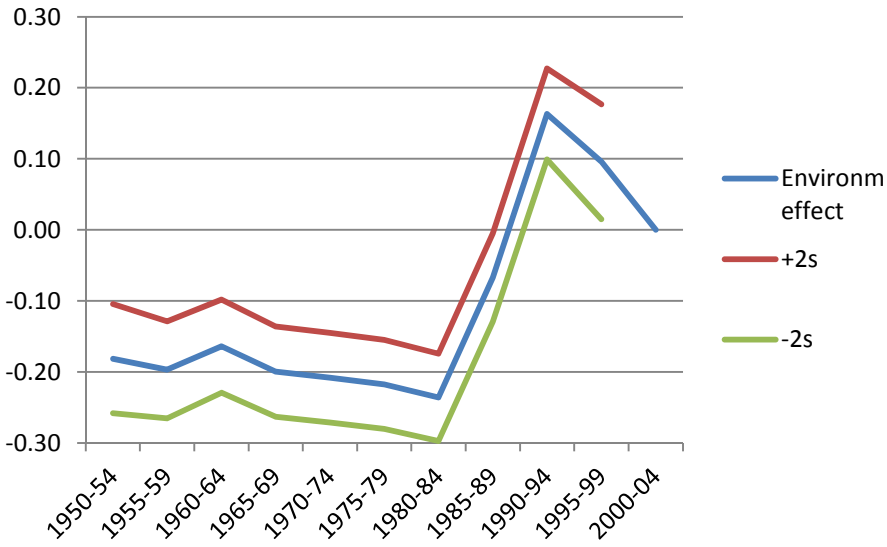
Male



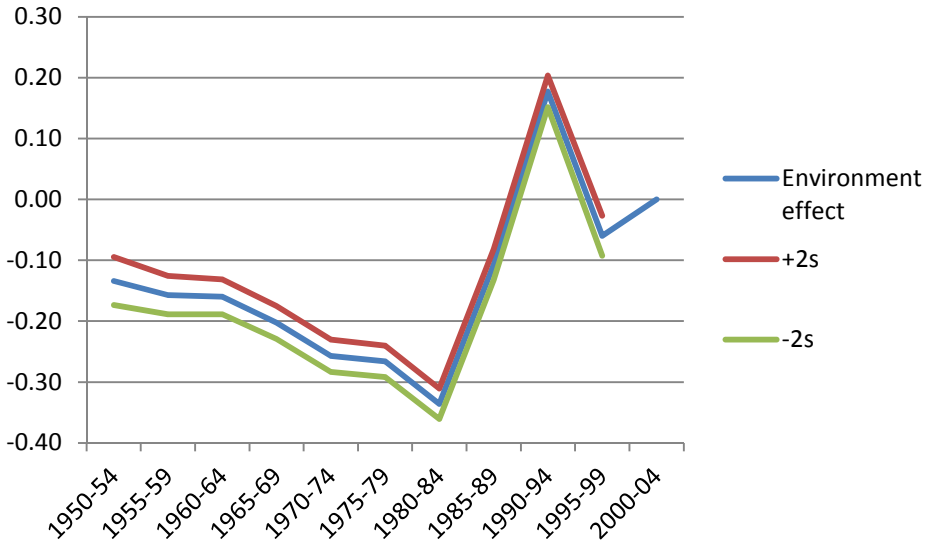
Results

Environment effect on liver cancer

Female

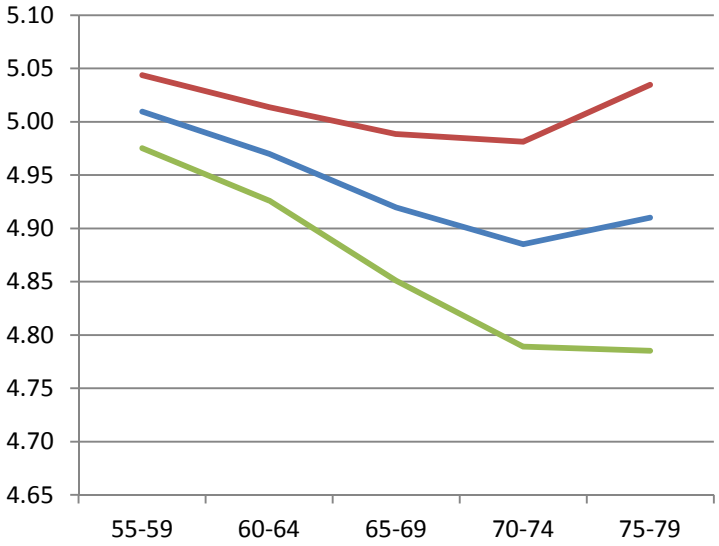


Male

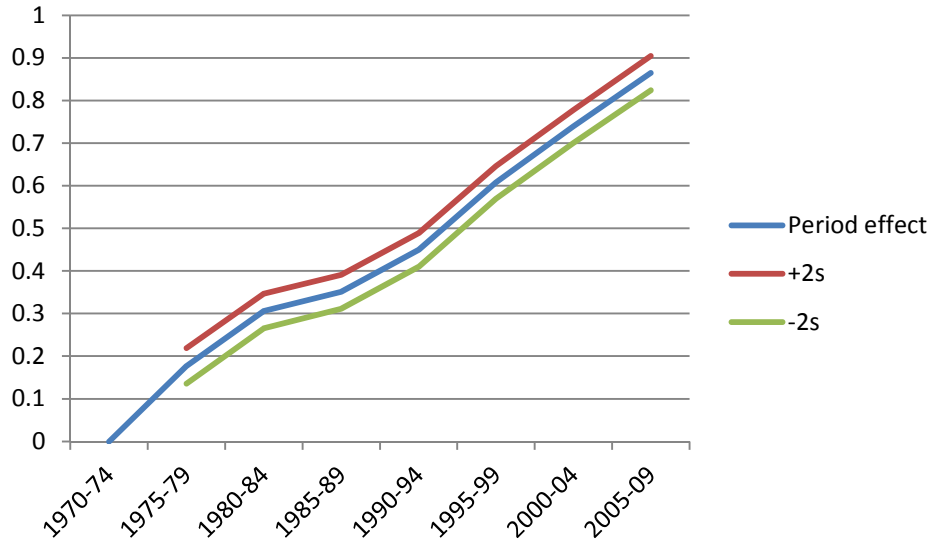


Results

Age effect on breast cancer

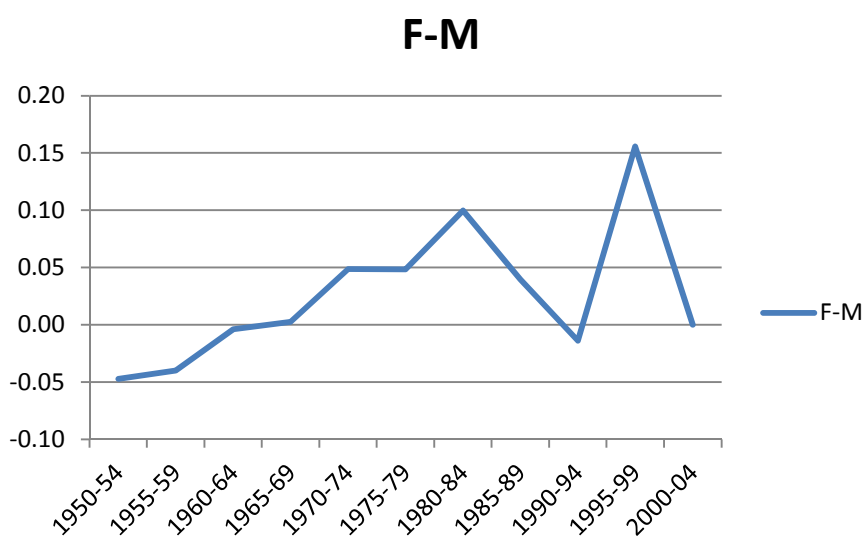


Period effect on breast cancer

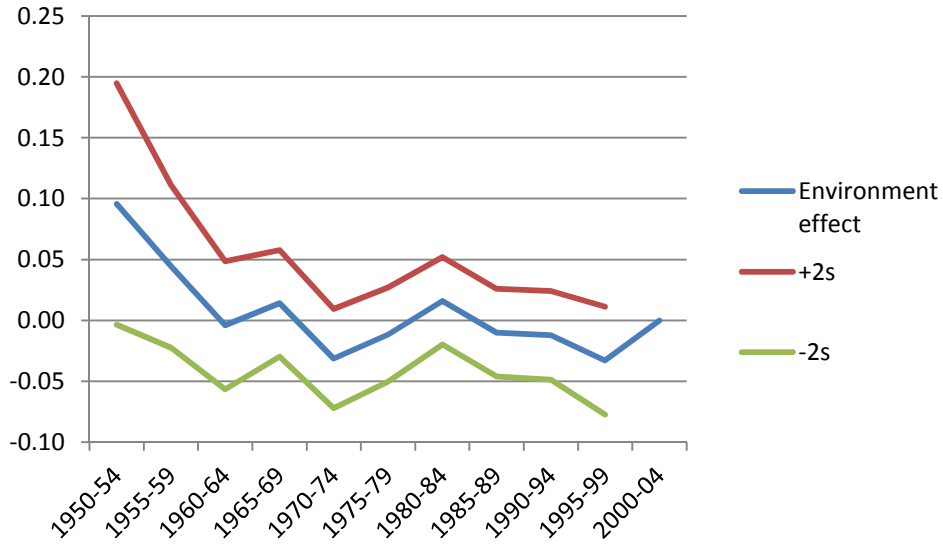


Results

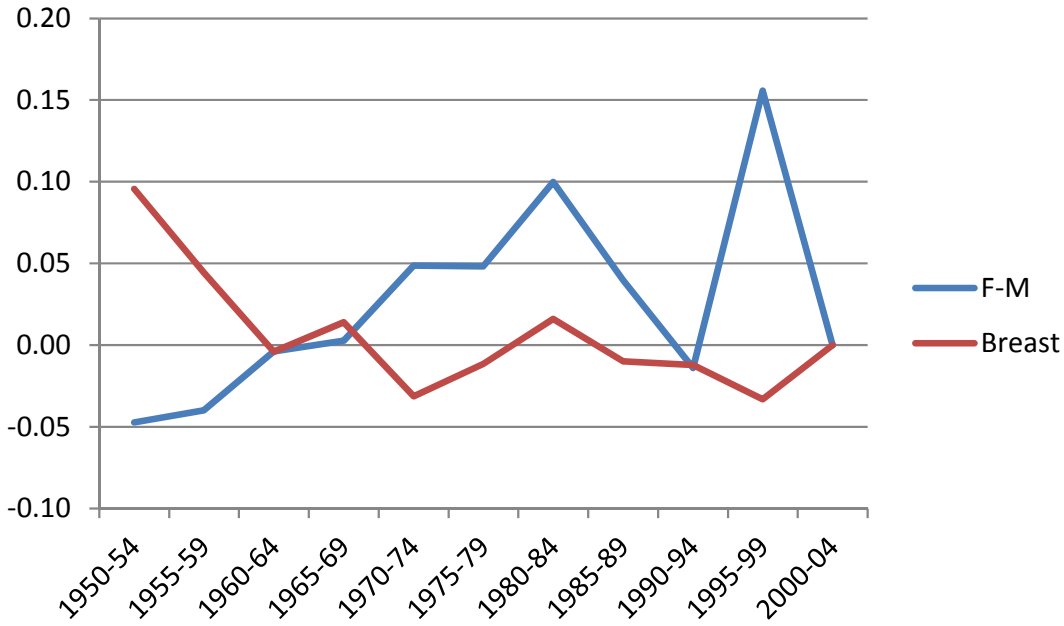
Difference of environment effect on liver cancer between sexes



Environment effect on breast cancer



Results



Those two factors have opposite trends.

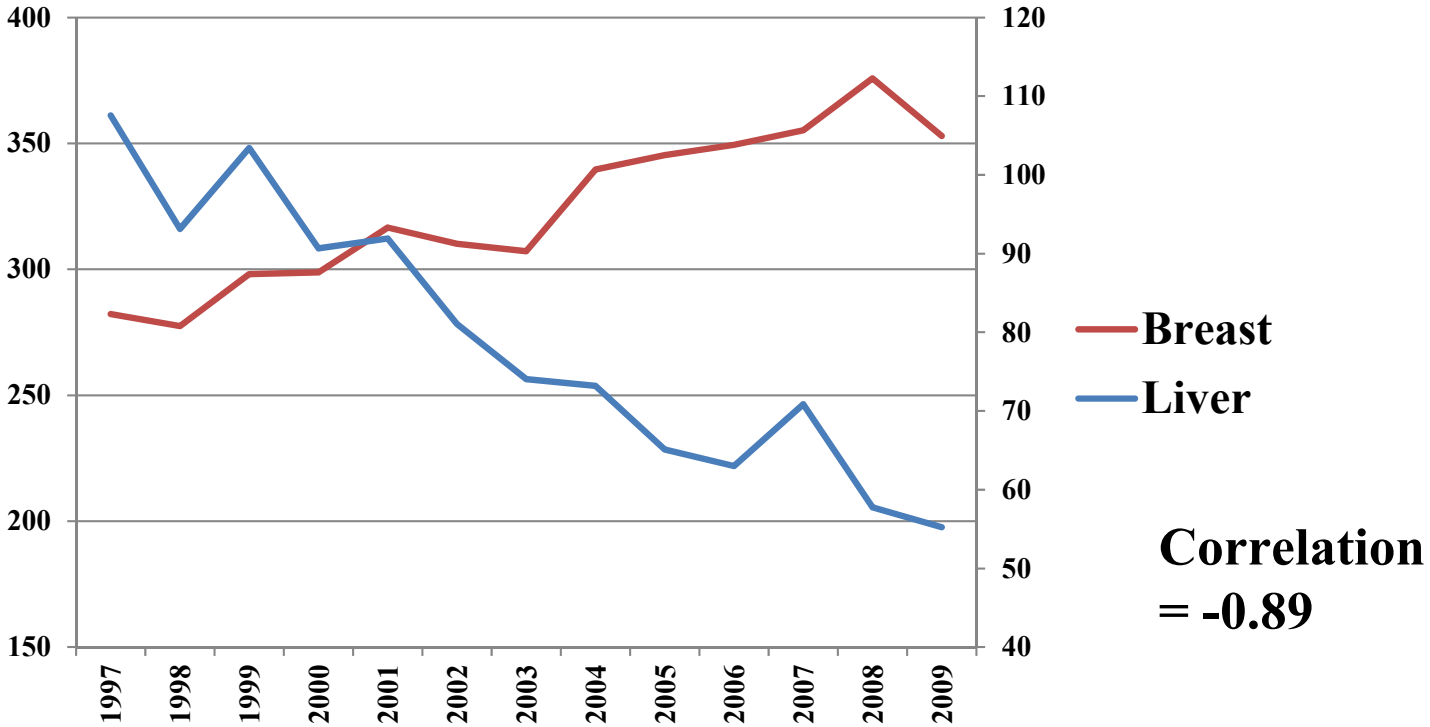
Correlation Coefficient	t Stat	P-value
-0.62	-2.37	0.04

Conclusion

It has been shown that there exist environmental factors which decrease the breast cancer risk and increase the liver cancer risk for postmenopausal females aged 55 and over.

Background

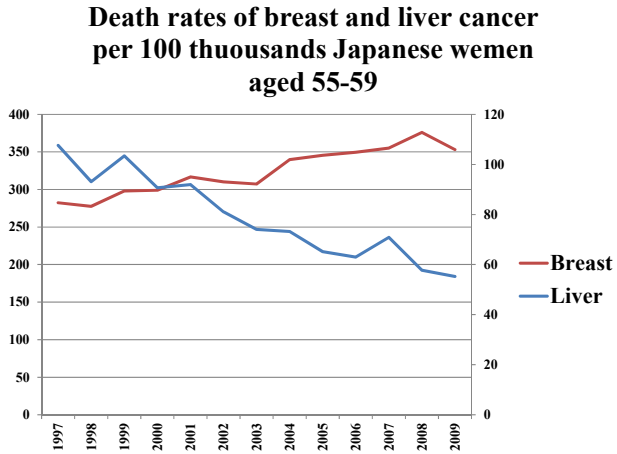
Death rates of breast and liver cancer per 100 thuousands Japanese wemen aged 55-59



Correlation = -0.89

Background

There exist environmental factors which decrease the breast cancer risk and increase the liver cancer risk for postmenopausal females.



There is some doubt that it is just a quasi-correlation?

Age period cohort model

$$\log \lambda_{ij} = \mu + \alpha_i + \beta_j + \gamma_{j-1+1}$$

$$\begin{pmatrix} \eta_{11} \\ \eta_{12} \\ \eta_{13} \\ \eta_{21} \\ \eta_{22} \\ \eta_{23} \\ \eta_{31} \\ \eta_{32} \\ \eta_{33} \\ \eta_{41} \\ \eta_{42} \\ \eta_{43} \\ \eta_{51} \\ \eta_{52} \\ \eta_{53} \end{pmatrix} = \left(\begin{array}{c|ccccc|ccc|ccccccc}
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \hline
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \hline
1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hline
1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0
\end{array}\right) \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \xi_{-3} \\ \xi_{-2} \\ \xi_{-1} \\ \xi_0 \\ \xi_1 \\ \xi_2 \end{pmatrix}.$$

Age period environment model

$$
\begin{pmatrix} \eta_{11} \\ \eta_{12} \\ \eta_{13} \\ \eta_{21} \\ \eta_{22} \\ \eta_{23} \\ \eta_{31} \\ \eta_{32} \\ \eta_{33} \\ \eta_{41} \\ \eta_{42} \\ \eta_{43} \\ \eta_{51} \\ \eta_{52} \\ \eta_{53} \end{pmatrix} = \left(\begin{array}{cccccc|ccc|cccccc} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right) \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \xi_{-3} \\ \xi_{-2} \\ \xi_{-1} \\ \xi_0 \\ \xi_1 \\ \xi_2 \end{pmatrix}.
$$

The age-period-environment moedl is equivalent to APC model in terms of space spanned by column vectors in their design matrix.

Equivalency of the two models

Design matrix

APE **APC**

$$\begin{pmatrix}
1&1&0&0&0&0&1&0&0&0&0&0&0&0&0\\
1&1&0&0&0&0&0&1&0&0&0&0&0&0&0\\
1&1&0&0&0&0&0&0&1&0&0&0&0&0&0\\
1&0&1&0&0&0&1&0&0&0&0&0&1&0&0\\
1&0&1&0&0&0&0&1&0&0&0&0&0&1&0\\
1&0&1&0&0&0&0&0&1&0&0&0&0&0&1\\
1&0&0&1&0&0&1&0&0&0&0&1&1&0&0\\
1&0&0&1&0&0&0&1&0&0&0&0&1&1&0\\
1&0&0&1&0&0&0&0&1&0&0&0&0&1&1\\
1&0&0&0&1&0&1&0&0&0&1&1&1&0&0\\
1&0&0&0&1&0&0&1&0&0&0&1&1&1&0\\
1&0&0&0&1&0&0&0&1&0&0&0&1&1&1\\
1&0&0&0&0&1&1&0&0&1&1&1&1&0&0\\
1&0&0&0&0&1&0&1&0&0&1&1&1&1&0\\
1&0&0&0&0&1&0&0&1&0&0&1&1&1&1
\end{pmatrix} =
\begin{pmatrix}
1&1&0&0&0&0&1&0&0&0&0&0&0&1&0&0\\
1&1&0&0&0&0&0&1&0&0&0&0&0&0&1&0\\
1&1&0&0&0&0&0&0&1&0&0&0&0&0&0&1\\
1&0&1&0&0&0&1&0&0&0&0&0&1&0&0&0\\
1&0&1&0&0&0&0&1&0&0&0&0&0&1&0&0\\
1&0&1&0&0&0&0&0&1&0&0&0&0&0&1&0\\
1&0&0&1&0&0&1&0&0&0&0&1&0&0&0&0\\
1&0&0&1&0&0&0&1&0&0&0&0&1&0&0&0\\
1&0&0&1&0&0&0&0&1&0&0&0&0&1&0&0\\
1&0&0&0&1&0&1&0&0&0&1&0&0&0&0&0\\
1&0&0&0&1&0&0&1&0&0&0&1&0&0&0&0\\
1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0\\
1&0&0&0&0&1&1&0&0&1&0&0&0&0&0&0\\
1&0&0&0&0&1&0&1&0&0&1&0&0&0&0&0\\
1&0&0&0&0&1&0&0&1&0&0&1&0&0&0&0
\end{pmatrix}
\begin{pmatrix}
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&0&0&0&0&-1&-1\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&-1\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&1&1&1&1&1&1\\
0&0&0&0&0&0&0&0&0&0&1&1&1&1&1\\
0&0&0&0&0&0&0&0&0&0&0&1&1&1&1\\
0&0&0&0&0&0&0&0&0&0&0&0&1&1&1\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&1\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&1
\end{pmatrix}$$

Equivalency of the two models

Design matrix

APC **APE**

$$\begin{pmatrix}
1&1&0&0&0&0&1&0&0&0&0&0&0&1&0&0\\
1&1&0&0&0&0&0&1&0&0&0&0&0&0&1&0\\
1&1&0&0&0&0&0&0&1&0&0&0&0&0&0&1\\
1&0&1&0&0&0&1&0&0&0&0&0&1&0&0&0\\
1&0&1&0&0&0&0&1&0&0&0&0&0&1&0&0\\
1&0&1&0&0&0&0&0&1&0&0&0&0&0&1&0\\
1&0&0&1&0&0&1&0&0&0&0&1&0&0&0&0\\
1&0&0&1&0&0&0&1&0&0&0&0&1&0&0&0\\
1&0&0&1&0&0&0&0&1&0&0&0&0&1&0&0\\
1&0&0&0&1&0&1&0&0&0&1&0&0&0&0&0\\
1&0&0&0&1&0&0&1&0&0&0&1&0&0&0&0\\
1&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0\\
1&0&0&0&0&1&1&0&0&1&0&0&0&0&0&0\\
1&0&0&0&0&1&0&1&0&0&1&0&0&0&0&0\\
1&0&0&0&0&1&0&0&1&0&0&1&0&0&0&0
\end{pmatrix}
=
\begin{pmatrix}
1&1&0&0&0&0&1&0&0&0&0&0&0&0&0&0\\
1&1&0&0&0&0&0&1&0&0&0&0&0&0&0&0\\
1&1&0&0&0&0&0&0&1&0&0&0&0&0&0&0\\
1&0&1&0&0&0&1&0&0&0&0&0&1&0&0&0\\
1&0&1&0&0&0&0&1&0&0&0&0&0&1&0&0\\
1&0&1&0&0&0&0&0&1&0&0&0&0&0&1&0\\
1&0&0&1&0&0&1&0&0&0&0&1&1&0&0&0\\
1&0&0&1&0&0&0&1&0&0&0&0&1&1&0&0\\
1&0&0&1&0&0&0&0&1&0&0&0&0&1&1&0\\
1&0&0&0&1&0&1&0&0&0&1&1&1&0&0&0\\
1&0&0&0&1&0&0&1&0&0&0&1&1&1&0&0\\
1&0&0&0&1&0&0&0&1&0&0&0&1&1&1&0\\
1&0&0&0&0&1&1&0&0&1&1&1&1&0&0&0\\
1&0&0&0&0&1&0&1&0&0&1&1&1&1&0&0\\
1&0&0&0&0&1&0&0&1&0&0&1&1&1&1&0\\
1&0&0&0&0&1&0&0&1&0&0&1&1&1&1&1
\end{pmatrix}
\begin{pmatrix}
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&1&0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&1&0\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&1\\
0&0&0&0&0&0&0&0&0&0&1&-1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&-1&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&-1&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&-1&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&-1
\end{pmatrix}$$

Equivalency of the two models

Correspondence between parameters

APC←→APE

$$\beta_j^{APC} = \beta_j^{APE} - \sum_{k=j}^{J-1} \xi_k \;;\; j = 1, \cdots, J-1$$

$$\beta_J^{APC} = \beta_J^{APE}$$

$$\gamma_k = \xi_k + \cdots + \xi_{J-1} \qquad ; k = 2-I, \cdots, J-1$$

APE←→APC

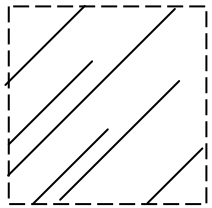
$$\beta_j^{APE} = \beta_j^{APC} + \gamma_j \qquad ; j = 1, \cdots, J$$

$$\xi_k = \gamma_k - \gamma_{k+1} \qquad ; k = 2-I, \cdots, J-1$$

Lexis diagram

The likelihood function for models assumed on (age period)-tabulated data is proportional to

$$L \equiv \prod_{i=1}^{I} \prod_{j=1}^{J} \lambda_{ij}^{\,y_{ij}} \exp(-\lambda_{ij} \cdot N_{ij})$$

X_{ij}  is approximated by person-years N_{ij}.

$[A_{i-1}, A_i) \times [P_{j-1}, P_j)$

The likelihood estimation results in that for Poisson regression models.

Identifiability problem

The likelihood function is a function only through λ_{ij} .

Therefore, it is needed to examine that for every set of

$$\{\lambda_{11}, \cdots, \lambda_{IJ}\} \in \Lambda$$

the values of

$$\{\mu, \alpha_1, \cdots, \alpha_I, \beta_1, \cdots, \beta_J, \gamma_{1-I}, \cdots, \gamma_{J-1}, \delta_2, \cdots, \delta_I\} \in \Theta$$

are uniquely determined in the model.

If not, plausible constraints on the parameters are needed.

Identifiability problem

For any given $\mu,\ \alpha_i,\ \beta_j$ *and* γ_{j-i} $(\alpha_1 = \beta_1 = \xi_{2-I} = 0)$

$$\mu + \alpha_i + \beta_j + \gamma_{j-i} = \mu + \alpha_i^*(d) + \beta_j^*(d) + \gamma_{j-i}^*(d)$$

holds for any d **, where**

$$\alpha_i^*(d) = \alpha_i + (i-1)d \quad \beta_i^*(d) = \beta_j - (j-1)d$$

$$\gamma_{j-i}^*(d) = \gamma_{j-i} - (j-i)d$$

⇩

Linear trends in age, period and cohort effects are not estimable.

Identifiability problem

In the proposed model under the constraint

$$\alpha_1 = \beta_1 = \xi_{2-I} = 0$$

the parameters are uniquely determined for every set of λ_{ij} **.**

⇩

The age, period and environment model are free from the non-identifiability problem from which APC model suffers.