A maximizing covariation approach to age-period-cohort analysis with application to smoking behavior

——Comparison of reliability with an intrinsic estimator——

Takashi Fujimoto

Abstract

In an additive age-period-cohort model, the behavior of the dependent variable comprises separable components for the three intrinsic effects and inseparable components capturing their linear dependency. This paper proposes a constraint only on the part of the parameter vector that generates the inseparable component to solve the identification problem. This parsimonious constraint will reduce the bias because the intrinsic effects are not biased. The three effects are identified while maximizing covariation between them; consequently, the inseparable component is divided equally between the birth factor (cohort effect) and the subsequent environmental factor (age＋period effect). The reliability of the identified estimator under this model was compared with that of the intrinsic estimator proposed by Fu (2000). The difference in reliability between the estimators in two models depends on the three effects.

Keywords: age-period-cohort analysis, inseparable component, maximizing covariation, reliability, smoking behavior, unbiasedness

1. Introduction

1.1. What is age-period-cohort analysis?

Individual's behavior changes over time. Changes in behavior can result from shifts in time period or simply from growing older. Behavior may also vary with birth cohort; i.e., a group of people born in a particular period may exhibit similar identifiable behaviors.

Behavior can be defined by effects peculiar to an individual's age, time period, and birth cohort. Changes caused by aging are the result of effects particular to each age. Changes caused by shifts in time period are the result of effects peculiar to each time period. Similarly, variations in behavior can be attributed to effects particular to each birth cohort. Age-Period-Cohort (APC) analysis attempts to measure these effects. Table 1 shows the smoking rate by age group and by time period. Such an age-period table, with the same time interval for both the age and period dimensions, is employed in the APC analysis.

1) APC: Age-period-cohort
MCE: Estimator under the maximizing covariation model
IE: Intrinsic estimator

Table 1 Smoking rate by age group and period in Japan.

(Gender)	Male					Female				
(Period)	1969	1979	1989	1999	2009	1969	1979	1989	1999	2009
(Age) 20-29	78.5	80.3	67.5	60.4	40.3	9.9	16.4	16.4	23.6	15.9
30-39	80.6	76.1	68.5	62.0	46.9	13.1	14.0	14.7	17.6	16.8
40-49	83.7	71.2	64.5	63.0	44.9	16.8	15.5	13.8	17.1	14.9
50-59	80.3	74.6	57.3	54.7	44.5	20.7	16.3	10.4	13.2	14.8
60-	71.1	62.0	49.5	38.6	27.8	19.8	15.4	8.6	6.8	6.2

[a] The data were obtained from the Japan Health Promotion & Fitness Foundation's website (http://www.health-net.or.jp) and retrieved from the survey of Japan Tobacco and Salt Public Corporation and Japan Tobacco Incorporated.

1.2. Identification problem

Decomposing behavior into the three dimensional effects is not simple. Regression analysis is commonly used to explain behavior according to the three listed effects. Blalock (1967) noted that an identification problem can occur in a regression model when one of the independent variables is taken as an exact function of other independent variables. Mason et al. (1973) noted that the identification problem should occur in APC analysis because the relationship Period − Age = Cohort indicates an exact linear dependency among the variables. The dependent variable is explained by two of the three effects because any one effect can be explained by the other two effects. In this case, a least squares solution cannot be identified because the regression design matrix is singular and so it has no unique inverse matrix (Rodgers, 1982). The identification problem pertains to the regression design matrix being one less than a full-rank matrix (Kupper et al., 1983).

Holford (1985) noted that the parameter in each effect consists of linear and curvature effects. The linear effect is the inestimable overall trend of the parameter, and the curvature effect is the estimable departure from the trend. Particular age, period, and cohort effects that depart from the overall trend can be found by considering the curvature effects (e.g., Tango and Kurashina, 1987). However, the linear effect is still not estimable.

1.3. Previous studies on parameter identification

It is necessary to impose constraints to solve the identification problem. Various constraints have been proposed.

The coefficients constraint approach has been used in many APC analyses. Mason et al. (1973) noted that if at least two age groups, periods, or cohorts have an identical effect (e.g., $A_1 = A_2$), then the three effects can be uniquely determined. However, different constraints lead to different estimates (e.g., Mason et al., 1973; Rodgers, 1982; Yang, Fu, and Land, 2004; Fu, 2008). Kupper et al. (1983) contended that the choice of constraint is a crucial determinant of the extent of bias. The effect of each variable can be identified by eliminating one of the three effects entirely (e.g., Firebaugh and Davis, 1988), which means constraining all age groups, all periods, or all cohorts to be equal. Rodgers (1982) noted that the imposition of multiple constraints, including the elimination of one of the three effects, brings us no closer to identifying the actual effects. Kupper et al. (1983) noted that the statistical methodology employed to demonstrate that considering one of the three effects to be unimportant can be seriously misleading. The Bayesian cohort model proposed by Nakamura (1986) is based on the assumption that the parameters change gradually, i.e., the successive parameters are not appreciably different. However, the parameters do not always change gradually.

The second method proposed to solve the identification problem is the proxy variable approach. Age-Period-Cohort characteristic models replace the cohort effect with cohort-related variables, such as the relative size of cohort (Kahn and Mason, 1987) or the percentage of births to unmarried mothres (O'Brien, Stockard, and Isaacson, 1999). The age or period effect can also be replaced. Farkas (1977) conceptualized the period effect as the unemployment rate. Heckman and Robb (1985) proposed a latent variable approach that replaces the three effects with proxy variables. Winship and Harding (2008) assumed that age, period, and cohort effects affect the outcome through several intermediary variables. However, proxy variables do not capture the actual effects accurately. The parameter identification is only an approximation.

The third approach assumes a nonlinear relationship between the parameters. James and Segal (1982) used a model that incorporates age-period interaction. Lee and Lin (1996) modeled cohort effects as an autoregressive process. Heuer (1997) and Holford (2006) used cubic splines for two or all three of the effects. O'Brien, Hudson, and Stockard (2008) proposed a mixed model that treats the age and period effects as fixed effects and the cohort effect as a random effect. Fu (2008) proposed a smoothing cohort model that estimates fixed age and period effects while smoothing the cohort effect with a nonparametric smoothing spline function. The smoothing cohort effect only introduces a small amount of binding for parameters and thus not much bias. However, assuming a special functional relationship for the three effects would make parameter estimates biased.

The forth approach is the intrinsic estimator model proposed by Fu (2000). The model identifies the parameter vector, which is orthogonal with the eigenvector in the null space of the design matrix. The eigenvector is independent of the dependent variable; therefore, the method imposes a constraint that involves almost no arbitrariness. Yang, Fu, and Land (2004) demonstrated that the intrinsic estimator model is statically more efficient than the coefficients constraint approach. Among the additive APC models, the intrinsic estimator model is thought to offer the most reliable constrained estimates.

1.4. Purpose of this paper

This paper introduces a new APC model. This model focuses on the fact that behavior comprises separable and inseparable components of the three dimensional intrinsic effects, and it imposes a constraint only on the part of the parameter vector that generates the inseparable component. This parsimonious constraint is a unique feature distinguishing this model from other models and will reduce the bias because the intrinsic effects are not biased. This model assumes that the inseparable component in each cell of the age-period table includes equally divided effects between the birth factor (cohort) and the subsequent environmental factor (age＋period).

The original idea of this model comes from Fujimoto (2011), which identified the three effects on food intake while maximizing the covariation between them. However, Fujimoto did not explain why a reliable solution can be obtained by maximizing covariation. In addition, Fujimoto only covered a special case in which the number of parameters corresponds to the number of equations that could be specified; therefore, no error term was included.

First, this paper generalizes Fujimoto's model to include an error term. Second, to show how this model performs, an analysis to determine the three effects of smoking behavior was performed and compared with the results of the intrinsic estimator model described by Yang, Fu, and Land (2004). Third, the unbiasedness and reliability of the identified estimator in comparison with the intrinsic estimator are demonstrated. Lastly, the reliability of the identified estimator under this model is explained.

2. Material and method

2.1. Model specification

The age-by-period data array analyzed in this paper is shown in Table 1. The effects of age, period, and cohort are estimated through eq. (1).

$$Q_{ij}=\mu+A_i+P_j+C_k+\varepsilon_{ij} \qquad (1)$$

Q_{ij} denotes the smoking rate for the ith age group for $i=1,\ldots,5$ at the jth time period for $j=1,\ldots,5$. μ denotes the average effect, which is the grand mean over all Q_{ij}. A_i denotes the ith row age effect or the parameter for the ith age group. P_j denotes the jth column period effect or the parameter for the jth time period. C_k denotes the kth diagonal cohort effect or the parameter for the kth cohort group for $k=1,\ldots,9$. ε_{ij} denotes the random error with expectation, $E(\varepsilon_{ij})=0$. Table 2 shows the correspondence between A_i, P_j, and C_k. The cohort effects, C_k, located on the diagonal lines have the same values. Eq. (1) is based on the assumption that the three dimensional effects are additive. This additive model is quite common in APC analyses because it does not require a specific assumption in the function.

Q_{ij} is assumed to be composed of the average, period, age, and cohort effects. The average effect is common to all cells, but the period, age, and cohort effects are particular to each cell. For example, in Table 1, the smoking rate for males, 78.5% for the group of people in their 20s in 1969, consists of the average effect, μ; the age effect, A_1; the period effect, P_1; and the cohort effect, C_5.

Then, zero-sum constraints are imposed on the parameters as in eq. (2), (3), and (4).

$$\sum_{i=1}^{5}A_i=0 \qquad (2)$$

$$\sum_{j=1}^{5}P_j=0 \qquad (3)$$

$$C_1+2C_2+3C_3+4C_4+5C_5+4C_6+3C_7+2C_8+C_9=0 \qquad (4)$$

In eq. (4), the weight assigned to each parameter is the number of times that the parameter appears on the age-period table. Then, eq. (5) can be obtained from eq. (2), (3), and (4).

$$\sum_{i=1}^{5}\sum_{j=1}^{5}A_{ij}=0 \quad \sum_{i=1}^{5}\sum_{j=1}^{5}P_{ij}=0 \quad \sum_{i=1}^{5}\sum_{j=1}^{5}C_{ij}=0 \qquad (5)$$

A_{ij}, P_{ij}, and C_{ij} denote A_i, P_j, and C_k, respectively, in cell (i, j) of the age-period table. Eq.

Table 2 Correspondence between A_i, P_j, and C_k.

	P_1	P_2	P_3	P_4	P_5
A_1	C_5	C_6	C_7	C_8	C_9
A_2	C_4	C_5	C_6	C_7	C_8
A_3	C_3	C_4	C_5	C_6	C_7
A_4	C_2	C_3	C_4	C_5	C_6
A_5	C_1	C_2	C_3	C_4	C_5

(5) indicates that A_{ij}, P_{ij}, and C_{ij} are assumed to have a zero sum or zero mean.

Eq. (1) can be transformed to eq. (6).

$$d_{ij}=A_i+P_j+C_k+\varepsilon_{ij} \tag{6}$$

d_{ij} denotes the difference between Q_{ij} and its constituents' average value, $\mu(d_{ij}=Q_{ij}-\mu)$, or deviation. A_i, P_j, or C_k are also deviations from their respective means of zero because zero-sum constraints, as shown in eq. (5), are imposed on them.

2.2. Least squares solution

Eq. (2), (3), (4), and (6) can be expressed as eq. (7) for the least squares regression model.

$$\boldsymbol{d}=\boldsymbol{Mx}+\boldsymbol{\varepsilon} \tag{7}$$

$\boldsymbol{d}$ denotes the column vector consisting of 25 d_{ij}'s. $\boldsymbol{M}$ denotes the regression design matrix (25×16) and $\boldsymbol{x}$ denotes the column vector consisting of 16 parameters. $\boldsymbol{\varepsilon}$ denotes the column vector consisting of 25 ε_{ij}'s. The exact form of eq. (7) is given by eq. (A.1) in Appendix A. It is worth noting that the parameter vector $\boldsymbol{x}$ cannot be identified using a least squares regression. The design matrix $\boldsymbol{M}$ is one less than full-rank because of the linear dependency. Therefore, the inverse matrix of $\boldsymbol{M}$ does not exist, and it is impossible to identify a unique solution.

The rank deficient of matrix $\boldsymbol{M}$ can be solved by generating C_9 exogenously. Then, eq. (7) is transformed to eq. (8).

$$\tilde{\boldsymbol{d}}=\tilde{\boldsymbol{M}}\tilde{\boldsymbol{x}}+\boldsymbol{\varepsilon} \tag{8}$$

$\tilde{\boldsymbol{d}}$ denotes the column vector consisting of 23 d_{ij}'s and 2 $(d_{ij}+C_9)$'s. $\tilde{\boldsymbol{M}}$ is the regression design matrix (25×15). $\tilde{\boldsymbol{x}}$ is the column vector consisting of 15 parameters. The exact form of eq. (8) is given by eq. (A.2) in Appendix A.

In eq. (8), a unique least squares solution can be obtained. The obtained parameters, which are functions of C_9, are shown in eq. (9). The parameters can also be obtained as functions of each individual parameter except A_3, P_3 and C_5.

$$\begin{array}{lll}
A_1=\alpha_1-0.50C_9 & P_1=\alpha_6+0.50C_9 & C_1=\alpha_{11}-C_9 \\
A_2=\alpha_2-0.25C_9 & P_2=\alpha_7+0.25C_9 & C_2=\alpha_{12}-0.75C_9 \\
A_3=\alpha_3 & P_3=\alpha_8 & C_3=\alpha_{13}-0.50C_9 \\
A_4=\alpha_4+0.25C_9 & P_4=\alpha_9-0.25C_9 & C_4=\alpha_{14}-0.25C_9 \\
A_5=\alpha_5+0.50C_9 & P_5=\alpha_{10}-0.50C_9 & C_5=\alpha_{15} \\
 & & C_6=\alpha_{16}+0.25C_9
\end{array} \tag{9}$$

$$C_7 = \alpha_{17} + 0.50C_9$$
$$C_8 = \alpha_{18} + 0.75C_9$$
$$C_9 = C_9$$

Values of α_x for $x=1,\ldots,18$ are functions of the d_{ij} values. Thus, the values of α_x are known constant values. α_1, for example, can be written as eq. (10).

$$a_1 = 0.02(4d_{11} - d_{12} - d_{13} - d_{15} + 4d_{21} - d_{22} - d_{23} - d_{24} - d_{25} + 4d_{31} - d_{32} - d_{33} - d_{34} - d_{35} + 4d_{41} - d_{42} - d_{43} - d_{44} - d_{45} + 24d_{51} - 6d_{52} - 6d_{53} - 6d_{54} - 6d_{55}) \quad (10)$$

The coefficients for C_9, referred to as β_x, are also known constant values.

2.3. Parameter identification

In eq. (6), d_{ij} is the deviation of Q_{ij} from the mean of its constituents, μ. A_i, P_j, and C_k are also deviations from their respective means of zero. When least squares solutions are obtained, if both sides of eq. (6) are squared and added across all cells on the age-period table, the variation function shown in eq. (11) can be obtained.

$$\sum_{i=1}^{5}\sum_{j=1}^{5} d_{ij}^2 = \sum_{i=1}^{5}\sum_{j=1}^{5} A_{ij}^2 + \sum_{i=1}^{5}\sum_{j=1}^{5} P_{ij}^2 + \sum_{i=1}^{5}\sum_{j=1}^{5} C_{ij}^2 + 2\sum_{i=1}^{5}\sum_{j=1}^{5} (A_{ij}+P_{ij})C_{ij} + \sum_{i=1}^{5}\sum_{j=1}^{5} \varepsilon_{ij}^2 \quad (11)$$

A_{ij}, P_{ij}, and C_{ij} denote A_i, P_j, and C_k, respectively, in cell (i, j) of the age-period table. The squared sum of the deviations denotes the variation in the data. Thus, the left side of eq. (11) is the total variation. On the right side, the first term denotes the variation of the age effect, the second term signifies the variation of the period effect, the third term is the variation of the cohort effect, the fourth term is the covariation between them, and the fifth term is the variation of error.

$\Sigma\Sigma A_{ij}P_{ij}$ has disappeared. The reason for this is shown in eq. (12). $\Sigma\Sigma A_{ij}P_{ij}$ can be transformed to $\Sigma A_i \Sigma P_j$ and corresponds to zero because of the zero-sum constraints ($\Sigma A_i = \Sigma P_j = 0$). $\Sigma\Sigma A_{ij}P_{ij} = 0$ implies that the vectors of $\boldsymbol{A_{ij}}$ and $\boldsymbol{P_{ij}}$ are orthogonal or that the correlation between A_{ij} and P_{ij} is zero on the age-period table.

$$\sum_{i=1}^{5}\sum_{j=1}^{5} A_{ij}P_{ij} = \sum_{i=1}^{5} A_i \sum_{j=1}^{5} P_j = 0 \quad (12)$$

The interaction term between the error and the parameter has also disappeared because the first-order condition to minimize $\Sigma\Sigma\varepsilon_{ij}^2$ is $\Sigma\Sigma(A_{ij}+P_{ij}+C_{ij})\varepsilon_{ij}=0$ in the least squares estimation. $\Sigma\Sigma\varepsilon_{ij}^2$ can be referred to as the least squared error.

The APC model introduced in this paper identifies the three effects while maximizing the covariation. The covariation, $2\Sigma\Sigma(A_{ij}+P_{ij})C_{ij}$, can be expressed as a function of C_9, which can be obtained by substituting eq. (9) for the covariation function. When $2\Sigma\Sigma(A_{ij}+P_{ij})C_{ij}$ is differentiated in C_9 and equated to zero, $\hat{C}_9$, which maximizes the covariation, can be identified. Any of the other parameters, where the covariation is maximized, can also be identified by substituting $\hat{C}_9$ for eq. (9).

To show how this model performs, an analysis was performed to determine the three effects of smoking behavior and compare them to the results of the intrinsic estimator model.

2.4. Parameter recovery experiment

The unbiasedness and reliability of the estimators under the maximizing covariation model and intrinsic estimator model were assessed by Monte Carlo simulation using repeated random samples.

The hypothetical parameter vector, $(A_1,\ldots,A_5, P_1,\ldots,P_5, C_1,\ldots,C_9)$, was generated artificially using random numbers while imposing two sets of constraints. The first are the zero-sum constraints. The second are the fluctuation ranges of parameters in each dimensional parameter vector. If the range is not constrained, unrealistic parameters are generated.

The hypothetical parameters were substituted for eq. (6), and hypothetical d_{ij} values were obtained. Then, the error, ε_{ij}, was assumed to be zero. The experiment was performed to determine whether the hypothetical parameters could be recovered from the hypothetical d_{ij} values. The hypothetical parameter that should be recovered will be referred to as the "true parameter," and the recovered parameter will be referred to as the "recovered parameter."

First, the unbiasedness of the estimator was demonstrated. It should be noted that $\sum\sum\varepsilon_{ij}^2$ is invariant in eq. (11), which corresponds to the fact that the goodness-of-fit measure, R^2, is invariant with respect to the choice of constraint being imposed to accomplish identification (Kupper et al., 1983). When the imposed constraint varies, the identified parameter vector varies, although $\sum\sum\varepsilon_{ij}^2$ does not vary. Even if the identified parameter vector varies, the original data will be recovered equally well because $\sum\sum\varepsilon_{ij}^2$ is invariant. Therefore, the difference between the identified and true parameter vectors represents bias, not error. The identified parameter vectors in all additive APC models are biased estimators in the sense that, even if the sample size is increased, the estimated parameter vector does not approach the true parameter vector. However, when repeated random samples are used, the expected value of bias, i.e., the difference between the identified and true parameters (e.g., $\hat{C}_9 - \mathit{true}\ C_9$), may converge to zero. If so, the identified estimator can be called an approximately unbiased estimator in the sense that overestimates and underestimates are equally likely in the repeated random samples.

Second, the reliability of the estimators under the maximizing covariation model and the intrinsic estimator model were demonstrated. The reliability implies consistency between the identified and true parameter vectors. The consistency was assessed in the repeated random samples.

2.5. Computation

The three effects under the maximizing covariation model were estimated using Wolfram Mathematica. The command lines are given in Appendix B.

The three effects under the intrinsic estimator model were estimated using Stata add-on program written by Schulhofer-Wohl and Yang (2006). The estimation was carried out through eq. (1) with imposing zero-sum constraints on the parameters as $\sum A_i=0$, $\sum P_j=0$, and $\sum C_k=0$. The constraint on the cohort effect parameters, $\sum C_k=0$, is different from that under the maximizing covariation model, $\sum\sum C_{ij}=0$.

3. Results

3.1. Three effects of smoking behavior

Let MCE denote the estimator under the maximizing covariation model. Let IE denote the

Table 3 Identified parameters in APC analysis for smoking behavior.

			Male		Female	
			MCE	IE	MCE	IE
Age effect	A_1	Age 20s	6.2	4.4	0.4	−0.2
	A_2	30s	6.3	5.4	0.6	0.3
	A_3	40s	3.6	3.6	2.3	2.3
	A_4	50s	−1.0	−0.1	1.2	1.4
	A_5	Over 60s	−15.2	−13.4	−4.4	−3.8
Period effect	P_1	Year 1969	13.9	15.6	0.3	0.9
	P_2	1979	9.6	10.5	1.6	1.9
	P_3	1989	−0.4	−0.4	−0.6	−0.6
	P_4	1999	−4.8	−5.6	1.0	0.7
	P_5	2009	−18.3	−20.1	−2.3	−2.9
Cohort effect	C_1	Born in 1900s	10.5	6.9	9.1	8.0
	C_2	1910s	5.5	2.9	3.9	3.1
	C_3	1920s	3.8	2.0	−1.0	−1.5
	C_4	1930s	−3.1	−3.9	−3.8	−4.1
	C_5	1940s	−1.6	−1.6	−3.3	−3.3
	C_6	1950s	1.8	2.7	0.0	0.3
	C_7	1960s	−1.4	0.4	1.1	1.7
	C_8	1970s	−3.0	−0.4	5.7	6.5
	C_9	1980s	−9.5	−6.0	3.1	4.2

intrinsic estimator described by Yang, Fu, and Land (2004). The estimated parameters for smoking behavior under both models are shown in Table 3. However, for a comparison of MCE and IE, the cohort effect parameters in IE are reparameterized to be $\sum\sum C_{ij}=0$ by deducting a constant value from each parameter.

The age effect can be understood as follows. The smoking rate is high for males in their 20s and 30s but decreases gradually as they grow older and significantly when they pass 60 years of age. For females, the age effect is different. Their smoking rate increases gradually with age, reaches a peak in their 40s, and decreases thereafter.

The period effect can be understood as follows. The smoking rate for males decreases remarkably with the shift in time period. The rate for females does not have a definite relationship with the time period but shows a decreasing trend in the past 10 years.

The cohort effect can be understood as follows. As an overall trend, younger male cohorts smoke less than older cohorts do. However, for the cohorts who were born before versus after World War II, that is, in the 1930s and the 1950s, respectively, younger cohorts smoke more than older cohorts. For females, the trend differs between cohorts born before and after World War II; the younger cohorts born in the prewar period smoke less than older cohorts born during the same period, whereas younger cohorts born in the postwar period smoke more than older cohorts born during the same period.

Regarding the MCE, the ratios of $\sum\sum A_{ij}^2+\sum\sum P_{ij}^2$ and $\sum\sum C_{ij}^2$ to their respective totals are

92% and 8% for males and 37% and 63% for females. On the one hand, this result indicates that the smoking behavior of males depends primarily on the age and period effects. On the other hand, for females, smoking behavior depends on the cohort effect rather than on the age and period effects. It is reasonable to conclude that the smoking behavior of males is easily influenced by the environmental factors surrounding smoking, such as age or time period, but that of females is not, instead depending mainly on the cohort effect.

Let us compare MCE with IE. The changes in parameters resulting from growing older, shifts in time period, or transitions of cohort are identical between MCE and IE. However, the parameter fluctuation ranges are not identical; that of the age effect and cohort effect are larger in MCE than in IE; that of the period effect is smaller in MCE than in IE.

3.2. Unbiasedness

The results of the parameter recovery experiment are as follows. Unbiasedness can be assessed by the distribution of $\hat{C}_9 - C_9^*$, where $\hat{C}_9$ or C_9^* denotes recovered or true C_9. The fluctuation range of hypothetical parameters was constrained as follows. The cohort effect parameters were allowed to fluctuate between −10.0 and 10.0. The age and period effect parameters were allowed to fluctuate between −5.0 and 5.0. The cohort effect parameters can fluctuate twice as much as the age and period effect parameters because the age-period table records the change of cohort effect for 80 years (1900s–1980s), whereas it records the changes in the age effect and period effect for 40 years (20s–60s and 1969–2009, respectively). It is assumed that the sizes of the three true effects, standardized by a time dimensional scale, are identical. The zero-sum constraints were imposed on the hypothetical parameters as $\sum A_i = 0$, $\sum P_j = 0$, $\sum C_k = 0$, and $\sum\sum C_{ij} = 0$.

The experiment was repeated 250 times. The results for the MCE and IE are shown in Fig. 1. The means of $\hat{C}_9 - C_9^*$ for the MCE and IE are −0.24 and −0.32, respectively (95% confi-

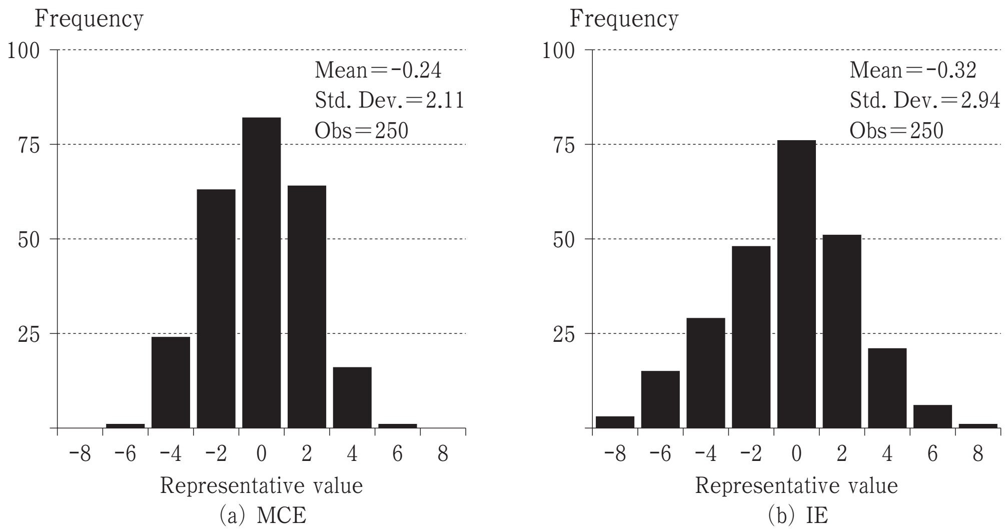

Fig. 1. Distribution of $\hat{C}_9 - C_9^*$.

dence intervals: -0.51−0.02 and -0.69−0.04, respectively). $\hat{C}_9 - C_9^*$ reflects bias, not error, because the error, ε_{ij}, is assumed to be zero. The distribution of biases, $\hat{C}_9 - C_9^*$, are symmetric around their zero means in both APC models. Thus, $\hat{C}_9$ is an approximately unbiased estimator of C_9 in the sense that overestimates and underestimates are equally likely in the repeated random samples.

If $\hat{C}_9$ is an approximately unbiased estimator of C_9, all parameters are identified as approximately unbiased estimators. Any of the parameters can be obtained by multiplying C_9 by a constant value β_x and adding a constant value α_x as shown in eq. (9).

3.3. Reliability

Reliability can be assessed by the consistency between the recovered and true parameter vectors. Three experiments were performed that employed different constraints for the fluctuation range of hypothetical parameters. The cohort effect parameters were allowed to fluctuate between −10.0 and 10.0 in all experiments. The age and period effect parameters were allowed to fluctuate between −5.0 and 5.0 in experiment 1, between −10.0 and 10.0 in experiment 2, and between −20.0 and 20.0 in experiment 3. The zero-sum constraints were imposed on the hypothetical parameters as $\sum A_i = 0$, $\sum P_j = 0$, $\sum C_k = 0$, and $\sum\sum C_{ij} = 0$. Each experiment was repeated 250 times.

First, the standard deviations of the $\hat{A}_1 - A_1^*$ values were compared between MCE and IE, where $\hat{A}_1$ or A_1^* denotes recovered or true A_1. The age effect is parameterized identically in both models as $\sum A_i = 0$, therefore, we do not need to mind the influence of the parameterization. The reliability decreases as the standard deviation increases, i.e., the bell curve shown in Fig. 1 becomes flatter. The results are shown in Table 4. In experiment 2, the p-value indicates no significant difference in the standard deviations between MCE and IE. However, the significant difference is observed in experiment 1 and 3. MCE is more reliable than IE in experiment 1, but less reliable in experiment 3.

Table 4 Standard deviation of $\hat{A}_1 - A_1^*$ values across 250 samples.

	Experiment	MCE	IE	p-value[a]
1	-5<A, P<5	1.06	1.43	0.00
2	-10<A, P<10	1.61	1.58	0.73
3	-20<A, P<20	2.71	1.99	0.00

[a] p-value for the two-sample variance-comparison test.

Second, to test reliability, ranking values based on size were given to each parameter in each dimensional parameter vector (e.g., if $A_4 > A_5 > A_2 > A_1 > A_3$, $A_1 = 4$, $A_2 = 3$, $A_3 = 5$, $A_4 = 1$, $A_5 = 2$), and the consistency of the recovered and true parameter vectors was compared. The consistency was demonstrated by Kendall's coefficient of concordance, referred to as W[2].

2) W is defined by the following formula.
$W = 12\sum_{l=1}^{m}(R_l - \bar{R})^2 / n^2(m^3 - m)$

Table 5 Mean of W values across 250 samples.

Experiment		Age			Period			Cohort		
		MCE	IE	p-value[a]	MCE	IE	p-value[a]	MCE	IE	p-value[a]
1	-5＜A, P＜5	0.94	0.92	0.00	0.95	0.93	0.01	0.97	0.95	0.00
2	-10＜A, P＜10	0.96	0.96	0.87	0.96	0.96	0.76	0.94	0.94	0.48
3	-20＜A, P＜20	0.97	0.98	0.21	0.97	0.98	0.00	0.90	0.92	0.00

[a] p-value for the mean-comparison test, paired data.

$W=1$ when the ranking values are perfectly matched between the recovered and true parameter vectors. When they are completely random, $W=0$. Table 5 shows the mean of the W values across 250 samples. In experiment 2, the p-value indicates no significant difference in the means between MCE and IE. However, the significant difference is observed in experiment 1 and in experiment 3 except age effect. MCE is more reliable than IE in experiment 1 because the ranking values can be recovered more accurately, whereas IE is more reliable in experiment 3.

It can be concluded that MCE is more reliable than IE when the sizes of the true age effect and period effect are smaller than that of the true cohort effect on the age-period table (in experiments 1), whereas MCE is less reliable than IE when the sizes of the true age effect and period effect are larger than that of the cohort effect (in experiment 3).

4. Discussion

In the following, we discuss why the identified estimator under the maximizing covariation model is reliable.

4.1. Illustration of variations

The variation function from eq. (11) can be rewritten as eq. (13).

$$\sum_{i=1}^{5}\sum_{j=1}^{5} d_{ij}^2-\sum_{i=1}^{5}\sum_{j=1}^{5}\varepsilon_{ij}^2=\sum_{i=1}^{5}\sum_{j=1}^{5}(A_{ij}+P_{ij})^2+\sum_{i=1}^{5}\sum_{j=1}^{5}C_{ij}^2+2\sum_{i=1}^{5}\sum_{j=1}^{5}(A_{ij}+P_{ij})C_{ij} \quad (13)$$

$\Sigma\Sigma A_{ij}^2+\Sigma\Sigma P_{ij}^2$ is transformed to $\Sigma\Sigma(A_{ij}+P_{ij})^2$ because $\Sigma\Sigma(A_{ij}+P_{ij})^2=\Sigma\Sigma A_{ij}^2+\Sigma\Sigma P_{ij}^2$ is established from eq. (12). The variation of error, $\Sigma\Sigma\varepsilon_{ij}^2$, is moved to the left side. The total variation, $\Sigma\Sigma d_{ij}^2$, is invariant. The total variation after deducting the error variation, $\Sigma\Sigma d_{ij}^2 - \Sigma\Sigma\varepsilon_{ij}^2$, is also invariant because the error variation is invariant. In the following sections, the total variation after deducting the error variation is simply referred to as the total variation.

$\Sigma\Sigma(A_{ij}+P_{ij})^2$, $\Sigma\Sigma C_{ij}^2$, and $2\Sigma\Sigma(A_{ij}+P_{ij})C_{ij}$ can be functions of C_9, which can be ob-

n denotes the number of components in the sample. The sample is composed of the recovered and true parameter vectors such that $n=2$. m denotes the number of parameters in each dimensionsl parameter vector: $m=5$ for age effect, $m=5$ for period effect, and $m=9$ for cohort effect. R_l for $l=1,\ldots,m$ denotes the total of ranking values of recovered and true lth parameters. For example, when the ranking valurs of recovered and true A_1's are 4 and 3, respecitively, $R_l=4+3$. $\bar{R}$ denotes the average of $R_l(\sum_{l=1}^{m} R_l/m)$).

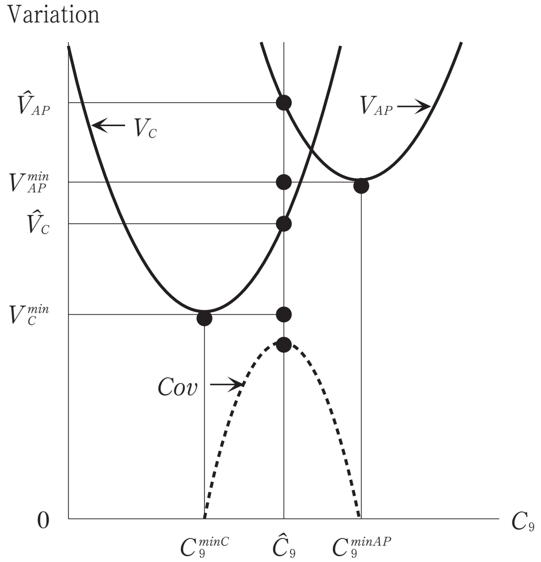

Fig. 2. Variation functions of C_9.

tained by substituting eq. (9) for them. Fig. 2 illustrates the relationship between C_9 and $\Sigma\Sigma(A_{ij}+P_{ij})^2$, $\Sigma\Sigma C_{ij}^2$, and $2\Sigma\Sigma(A_{ij}+P_{ij})C_{ij}$ using an age-period table. V_{AP} denotes $\Sigma\Sigma(A_{ij}+P_{ij})^2$, V_C denotes $\Sigma\Sigma C_{ij}^2$, and Cov denotes $2\Sigma\Sigma(A_{ij}+P_{ij})C_{ij}$.

When $C_9=\hat{C}_9$, the covariation is maximized. The variation of the age+period effect or the cohort effect is minimized at V_{AP}^{min} or V_C^{min} when $C_9=C_9^{minAP}$ or $C_9=C_9^{minC}$. C_9^{minAP} can be obtained by differentiating $\Sigma\Sigma(A_{ij}+P_{ij})^2$ in C_9 and equating it to zero. C_9^{minC} can also be obtained by differentiating $\Sigma\Sigma C_{ij}^2$ in C_9 and equating it to zero. It is then interesting to note that $\hat{C}_9=(C_9^{minAP}+C_9^{minC})/2$ is established, which indicates that $\hat{C}_9$ is located in the center of C_9^{minAP} and C_9^{minC}.

4.2. *How is variation decomposed?*

When the covariation is maximized, how is the total variation decomposed? V_{AP} can be considered to consist of V_{AP}^{min} and $V_{AP}-V_{AP}^{min}$. V_C can also be considered to consist of V_C^{min} and $V_C-V_C^{min}$. Then, the variation function in eq. (13) can be rewritten as eq. (14).

$$TV=V_{AP}^{min}+V_C^{min}+(V_{AP}-V_{AP}^{min})+(V_C-V_C^{min})+Cov \quad (14)$$

TV denotes the total variation. V_{AP}^{min} is explained only by the age+period effect because if not so V_{AP} can be less than V_{AP}^{min}. And V_{AP}^{min} is invariant because it can be a function of the invariants d_{ij}'s. Similarly, V_C^{min} is also explained only by the cohort effect and is invariant. Thus, V_{AP}^{min} and V_C^{min} are separable from the total variation as the intrinsic variations. However, $V_{AP}-V_{AP}^{min}$, $V_C-V_C^{min}$, and Cov cannot be separated without any constraints because their values depend on C_9. The total variation comprises the separable and inseparable components for the intrinsic variation of the age+period effect or cohort effect.

Then, it is interesting to note that $\hat{V}_{AP}-V_{AP}^{min}=\hat{V}_C-V_C^{min}$ is established, where $\hat{V}_{AP}$ and $\hat{V}_C$ are the variations of the age+period effect and the cohort effect, respectively, when the covariation is maximized. Therefore, the previously inseparable variation is separated into $V_{AP}-V_{AP}^{min}$, $V_C-V_C^{min}$ and Cov so that it can be assigned equally to $V_{AP}-V_{AP}^{min}$ and $V_C-V_C^{min}$.

4.3. How is behavior decomposed?

Eq. (13) shows that the total variation consists of the variation of the age+period effect, the variation of the cohort effect, and the covariation between them. The covariation merely represents the degree of linkage between the age+period effect and the cohort effect. Therefore, behavior in the APC model consists of the birth factor (cohort effect) and the subsequent environmental factor (age+period effect). The APC analysis has suffered from questions about how to decompose behavior into the two factors. When covariation is maximized, the age+period effect and the cohort effect are linked as closely as possible on the age-period table. Then, how is the behavior of the dependent variable decomposed?

Let A_{ij}^{min} or P_{ij}^{min} denote the age effect or period effect when $C_9=C_9^{minAP}$. Let C_{ij}^{min} denote the cohort effect when $C_9=C_9^{minC}$. A_{ij}, P_{ij}, or C_{ij} can be considered to consist of A_{ij}^{min} and $A_{ij}-A_{ij}^{min}$, P_{ij}^{min} and $P_{ij}-P_{ij}^{min}$, or C_{ij}^{min} and $C_{ij}-C_{ij}^{min}$. Then, eq. (6) can be rewritten as eq. (15).

$$d_{ij}=A_{ij}^{min}+P_{ij}^{min}+C_{ij}^{min}+(A_{ij}-A_{ij}^{min})+(P_{ij}-P_{ij}^{min})+(C_{ij}-C_{ij}^{min})+\varepsilon_{ij} \quad (15)$$

A_{ij}^{min}, P_{ij}^{min}, and C_{ij}^{min} are separable from the behavior of the dependent variable as the intrinsic effects, because V_{AP}^{min} and V_C^{min} are separable from the total variation as the intrinsic variations. However, $A_{ij}-A_{ij}^{min}$, $P_{ij}-P_{ij}^{min}$, and $C_{ij}-C_{ij}^{min}$ cannot be separated without any constraints because their values depend on C_9. The separable components in each cell of the age-period table, $A_{ij}^{min}+P_{ij}^{min}+C_{ij}^{min}$, obtained from the data in Table 1, are tabulated in Table 6. The inseparable components, $(A_{ij}-A_{ij}^{min})+(P_{ij}-P_{ij}^{min})+(C_{ij}-C_{ij}^{min})$, obtained from the data in Table 1, are also tabulated in Table 7. We find in Table 7 that the values located on the diagonal lines are identical.

Why are the identical values located on the diagonal lines? Eq. (15) can be rewritten as eq. (16).

$$d_{ij}-(A_{ij}^{min}+P_{ij}^{min}+C_{ij}^{min})-\varepsilon_{ij}=(A_{ij}-A_{ij}^{min})+(P_{ij}-P_{ij}^{min})+(C_{ij}-C_{ij}^{min}) \quad (16)$$

On the right side of eq. (16), d_{ij}, A_{ij}^{min}, P_{ij}^{min}, C_{ij}^{min}, and ε_{ij} are invariant. A least squares regression model can be formulated from eq. (16). Let $d_{ij}-(A_{ij}^{min}+P_{ij}^{min}+C_{ij}^{min})-\varepsilon_{ij}$ be the dependent variable; let $A_{ij}-A_{ij}^{min}$, $P_{ij}-P_{ij}^{min}$, and $C_{ij}-C_{ij}^{min}$ be independent variables. Then, $d_{ij}-(A_{ij}^{min}+P_{ij}^{min}+C_{ij}^{min})-\varepsilon_{ij}$ can be explained completely only by $C_{ij}-C_{ij}^{min}$, which in turn can be explained completely only by $(A_{ij}-A_{ij}^{min})+(P_{ij}-P_{ij}^{min})$. The exact linear dependency can be

Table 6 $A_{ij}^{min}+P_{ij}^{min}+C_{ij}^{min}$ of smoking behavior.

(Gender)		Male					Female				
(Period)		1969	1979	1989	1999	2009	1969	1979	1989	1999	2009
(Age)	20-29	18.4	20.5	10.2	7.0	-10.2	-2.6	1.7	0.4	6.2	0.0
	30-39	14.3	14.3	10.6	5.9	-6.4	-2.6	-1.2	-0.3	2.1	3.0
	40-49	15.5	7.3	1.6	3.5	-10.4	2.2	0.4	-1.6	2.9	0.5
	50-59	9.8	6.7	-7.3	-7.3	-14.6	6.3	2.4	-2.9	-1.2	-1.5
	60-	-2.3	-8.6	-17.5	-25.8	-35.1	6.2	2.0	-5.4	-6.9	-10.0

Table 7 $(A_{ij}-A_{ij}^{min})+(P_{ij}+P_{ij}^{min})+(C_{ij}-C_{ij}^{min})$ of smoking behavior.

(Gender)		Male					Female				
(Period)		1969	1979	1989	1999	2009	1969	1979	1989	1999	2009
(Age)	20-29	0.0	-2.9	-5.7	-8.6	-11.5	0.0	0.3	0.6	0.9	1.2
	30-39	2.9	0.0	-2.9	-5.7	-8.6	-0.3	0.0	0.3	0.6	0.9
	40-49	5.7	2.9	0.0	-2.9	-5.7	-0.6	-0.3	0.0	0.3	0.6
	50-59	8.6	5.7	2.9	0.0	-2.9	-0.9	-0.6	-0.3	0.0	0.3
	60-	11.5	8.6	5.7	2.9	0.0	-1.2	-0.9	-0.6	-0.3	0.0

found between $C_{ij}-C_{ij}^{min}$ and $(A_{ij}-A_{ij}^{min})+(P_{ij}-P_{ij}^{min})$, i.e., $C_{ij}-C_{ij}^{min}$ is taken as an exact function of $(A_{ij}-A_{ij}^{min})+(P_{ij}-P_{ij}^{min})$, as shown in eq. (17), which is the root of the identification problem.

$$(C_{ij}-C_{ij}^{min})=\delta\{(A_{ij}-A_{ij}^{min})+(P_{ij}-P_{ij}^{min})\} \quad (17)$$

δ denotes the coefficient that determines how to decompose the inseparable component.

Why must we impose the constraint on the parameter vector $(\boldsymbol{A_i}, \boldsymbol{P_j}, \boldsymbol{C_k})$ to achieve identification? We can solve the identification problem by imposing the constraint parsimoniously on the parameter vector $(\boldsymbol{A_i}-\boldsymbol{A_i^{min}}, \boldsymbol{P_j}-\boldsymbol{P_j^{min}}, \boldsymbol{C_k}-\boldsymbol{C_k^{min}})$, which will reduce the bias because the intrinsic effects, A_{ij}^{min}, P_{ij}^{min}, and C_{ij}^{min}, are not biased. For example, in the coefficients constraint approach in APC analysis, if the constraint is imposed on the parameter vector $(\boldsymbol{A_i}-\boldsymbol{A_i^{min}}, \boldsymbol{P_j}-\boldsymbol{P_j^{min}}, \boldsymbol{C_k}-\boldsymbol{C_k^{min}})$, e.g., $A_1-A_1^{min}=A_2-A_2^{min}$, the identified estimator may be more reliable than the estimator under the constraint imposed on the parameter vector $(\boldsymbol{A_i}, \boldsymbol{P_j}, \boldsymbol{C_k})$, e.g., $A_1=A_2$.

When the covariation is maximized, $\delta=1$ is established in eq. (17), which means that $(C_{ij}-C_{ij}^{min})=(A_{ij}-A_{ij}^{min})+(P_{ij}-P_{ij}^{min})$ is established in all cells of the age-period table. The result implies that the maximizing covariation model imposes a constraint only on the parameter vector that generates the inseparable component, and the inseparable component is equally divided between the age+period effect and cohort effect. The inseparable component can be explained only by either the cohort effect or the age+period effect. Accordingly, this APC model assumes that the inseparable component in each cell of the age-period table includes equally divided effects between the birth factor (cohort) and the subsequent environmental factor (age+period).

4.4. How is behavior decomposed under intrinsic estimator model?

When the parameter vector that generates the separable component, $(\boldsymbol{A_i^{min}}, \boldsymbol{P_j^{min}}, \boldsymbol{C_k^{min}})$, is deducted from the parameter vector $(\boldsymbol{A'_i}, \boldsymbol{P'_j}, \boldsymbol{C'_k})$ which is identified under the intrinsic estimator model, the parameter vector $(\boldsymbol{A'_i}-\boldsymbol{A_i^{min}}, \boldsymbol{P'_j}-\boldsymbol{P_j^{min}}, \boldsymbol{C'_k}-\boldsymbol{C_k^{min}})$ is obtained. Let us assume that the cohort effect parameter vector, $\boldsymbol{C'_k}$, is reparameterized to be $\sum\sum C'_{ij}=0$, and generate an age-period table from the parameter vector $(\boldsymbol{A'_i}-\boldsymbol{A_i^{min}}, \boldsymbol{P'_j}-\boldsymbol{P_j^{min}}, \boldsymbol{C'_k}-\boldsymbol{C_k^{min}})$ obtained from the data in Table 1. Then, Table 7 can be generated.

And then, $C'_{ij}-C_{ij}^{min}$ is taken as an exact function of $(A'_{ij}-A_{ij}^{min})+(P'_{ij}-P_{ij}^{min})$ as shown in eq. (17). When the data for males in Table 1 are used, $\delta=0.24$; when those for females are used, $\delta=-3.22$. We can find no consistency in the constraint (i.e., δ) to separate the inseparable component. It is impossible to determine δ statistically because the dependent variable can be explained completely only by either the cohort effect or the age+period effect. The value of δ must be given exogenously and is the crucial determinant of the extent of bias. We need a reasonable and consistent assumption to separate the inseparable component.

5. Conclusion

This paper revealed that, in an additive age-period-cohort model, the behavior of the dependent variable comprises separable components for the three dimensional intrinsic effects and inseparable components resulting from the linear dependency among them. This paper proposed to impose a constraint only on the part of the parameter vector that generates the inseparable component, then solve the identification problem. This parsimonious constraint will reduce the bias because the intrinsic effects are not biased.

This paper also revealed that the total variation of the dependent variable comprises the variation of the age+period effect, the variation of the cohort effect, and the covariation between them. The APC model introduced in this paper identifies the three effects while maximizing the covariation. The maximization enables the age+period effect and the cohort effect on behavior to be linked as far as possible on the age-period table. Consequently, the inseparable component is divided equally between the age+period effect and the cohort effect, which implies that only the part of the parameter vector that generates the inseparable component is constrained. The inseparable component can be explained only by either the cohort effect or the age+period effect. Accordingly, this APC model assumes that the inseparable component in each cell of the age-period table includes equally divided effects between the birth factor (cohort) and the subsequent environmental factor (age+period). We have no information on the factors determining the inseparable component; therefore, we have no option but to rely on this neutral assumption.

The identified estimator is biased in all additive APC models because the variation of error is invariant regardless of the choice of constraint being imposed to accomplish identification. Therefore, even if the sample size is increased, the estimated parameter vector will not approach the true parameter vector. However, when repeated random samples are used, the expected value of bias converges to zero. The identified estimator under this APC model can be considered as an approximately unbiased estimator in the sense that overestimates and underestimates are equally likely in the repeated random samples.

The reliability of the identified estimator under this APC model was compared with the intrinsic estimator. The results show that this APC model is more reliable when the sizes of the true age effect and period effect are smaller than that of the true cohort effect on the age-period table, but less reliable when the sizes of the true age effect and period effect are larger than that of the true cohort effect. It can be deduced that the reliability depends on how the inseparable component is divided into the birth factor (cohort) and the subsequent environmental factor (age+period), i.e., how the δ in eq. (17) is determined under a reasonable assumption.

Appendix A

$$
\begin{bmatrix} d_{11} \\ d_{12} \\ d_{13} \\ d_{14} \\ d_{15} \\ d_{21} \\ d_{22} \\ d_{23} \\ d_{24} \\ d_{25} \\ d_{31} \\ d_{32} \\ d_{33} \\ d_{34} \\ d_{35} \\ d_{41} \\ d_{42} \\ d_{43} \\ d_{44} \\ d_{45} \\ d_{51} \\ d_{52} \\ d_{53} \\ d_{54} \\ d_{55} \end{bmatrix}
=
\begin{bmatrix}
-1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
-1 & -1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
-1 & -1 & -1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
-1 & -1 & -1 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
-1 & -1 & -1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & -1 & -1 & -1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & -1 & -1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & -1 & -1 & -1 & -1 & -2 & -3 & -4 & -5 & -4 & -3 & -2 & -1 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix} A_2 \\ A_3 \\ A_4 \\ A_5 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \\ C_8 \\ C_9 \end{bmatrix}
+
\begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{14} \\ \varepsilon_{15} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{23} \\ \varepsilon_{24} \\ \varepsilon_{25} \\ \varepsilon_{31} \\ \varepsilon_{32} \\ \varepsilon_{33} \\ \varepsilon_{34} \\ \varepsilon_{35} \\ \varepsilon_{41} \\ \varepsilon_{42} \\ \varepsilon_{43} \\ \varepsilon_{44} \\ \varepsilon_{45} \\ \varepsilon_{51} \\ \varepsilon_{52} \\ \varepsilon_{53} \\ \varepsilon_{54} \\ \varepsilon_{55} \end{bmatrix}
\tag{A.1}
$$

$$
\begin{bmatrix} d_{11} \\ d_{12} \\ d_{13} \\ d_{14} \\ d_{15}+C_9 \\ d_{21} \\ d_{22} \\ d_{23} \\ d_{24} \\ d_{25} \\ d_{31} \\ d_{32} \\ d_{33} \\ d_{34} \\ d_{35} \\ d_{41} \\ d_{42} \\ d_{43} \\ d_{44} \\ d_{45} \\ d_{51}+C_9 \\ d_{52} \\ d_{53} \\ d_{54} \\ d_{55} \end{bmatrix}
=
\begin{bmatrix}
-1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
-1 & -1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
-1 & -1 & -1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
-1 & -1 & -1 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
-1 & -1 & -1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & -1 & -1 & -1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & -1 & -1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & -1 & -1 & -1 & -1 & -2 & -3 & -4 & -5 & -4 & -3 & -2 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix} A_2 \\ A_3 \\ A_4 \\ A_5 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \\ C_8 \\ C_9 \end{bmatrix}
+
\begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{14} \\ \varepsilon_{15} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{23} \\ \varepsilon_{24} \\ \varepsilon_{25} \\ \varepsilon_{31} \\ \varepsilon_{32} \\ \varepsilon_{33} \\ \varepsilon_{34} \\ \varepsilon_{35} \\ \varepsilon_{41} \\ \varepsilon_{42} \\ \varepsilon_{43} \\ \varepsilon_{44} \\ \varepsilon_{45} \\ \varepsilon_{51} \\ \varepsilon_{52} \\ \varepsilon_{53} \\ \varepsilon_{54} \\ \varepsilon_{55} \end{bmatrix}
\tag{A.2}
$$

Appendix B

```
(* Least Squares estimation *)
LeastSquares [{{-1, -1, -1, -1, -1, -1, -1, -1, 0, 0, 0, 1, 0, 0, 0}, {-1, -1, -1, -1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0}, {-1, -1, -1, -1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0}, {-1, -1, -1, -1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1}, {-1, -1, -1, -1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0}, {1, 0, 0, 0, -1, -1, -1, -1, 0, 0, 1, 0, 0, 0, 0}, {1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0}, {1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0}, {1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0}, {1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1}, {0, 1, 0, 0, -1, -1, -1, -1, 0, 1, 0, 0, 0, 0, 0}, {0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0}, {0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0}, {0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0}, {0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0}, {0, 0, 1, 0, -1, -1, -1, -1, 1, 0, 0, 0, 0, 0, 0}, {0, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0}, {0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0}, {0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0}, {0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0}, {0, 0, 0, 1, -1, -1, -1, -1, -2, -3, -4, -5, -4, -3, -2}, {0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0}, {0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0}, {0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0}}, {78.5, 80.3, 67.5, 60.4, 40.3-C9, 80.6, 76.1, 68.5, 62, 46.9, 83.7, 71.2, 64.5, 63, 44.9, 80.3, 74.6, 57.3, 54.7, 44.5, 71.1+C9, 62, 49.5, 38.6, 27.8}]

(* Substituting estimated parameters for covariation function *)
Eliminate [{cov==2*((A1+P1)*C5+(A1+P2)*C6+(A1+P3)*C7+(A1+P4)*C8+(A1+P5)*C9+(A2+P1)*C4+(A2+P2)*C5+(A2+P3)*C6+(A2+P4)*C7+(A2+P5)*C8+(A3+P1)*C3+(A3+P2)*C4+(A3+P3)*C5+(A3+P4)*C6+(A3+P5)*C7+(A4+P1)*C2+(A4+P2)*C3+(A4+P3)*C4+(A4+P4)*C5+(A4+P5)*C6+(A5+P1)*C1+(A5+P2)*C2+(A5+P3)*C3+(A5+P4)*C4+(A5+P5)*C5), A1+A2+A3+A4+A5==0, P1+P2+P3+P4+P5==0, C1+2C2+3C3+4C4+5C5+4C6+3C7+2C8+C9==0, A2==3.94-0.25C9, A3==3.60, A4==1.42+0.25C9, A5==-10.40+0.50C9, P2==11.98+0.25C9, P3==-0.40, P4==-7.14-0.25C9, P5==-23.09-0.50C9, C2==-1.63-0.75C9, C3==-0.97-0.50C9, C4==-5.44-0.25C9, C5==-1.63, C6==4.18+0.25C9, C7==3.40+0.50C9, C8==4.12+0.75C9}, {P1, P2, P3, P4, P5, A1, A2, A3, A4, A5, C1, C2, C3, C4, C5, C6, C7, C8}]

(* Differentiating covariation function in C9 and equating to zero*)
  D[(-3800057-119380C9-6250C9^2)/500, C9]
  Solve[(-119380-12500C9)500==0, C9]
```

References

Blalock, H. M. Jr., 1967. Status inconsistency, social mobility, status integration and structural effects. American Sociological Review 32, 790-801.

Farkas, G., 1977. Cohort, age, and period effects upon the employment of white females: evidence for 1957-1968. Demography 14, 33-42.

Fu, W. J., 2000. Ridge estimator in singular design with application to age-period-cohort analysis of disease rate. Communication in Statics Theory and Method 29, 263-78.

Fu, W. J., 2008. A Smoothing cohort model in age-period-cohort analysis with applications to homicide arrest rates and lung cancer mortality rates. Sociological Methods and Research 36, 327-61.

Fujimoto, T., 2011. Period-age-cohort model maximizing covariation applied to foods intake. Journal of Rural Economics 83(1), 1-14 (in Japanese).

Firebaugh, G. and K. E. Davis., 1988. Trends in anitiblack prejudice, 1972-1984. American Journal of Sociology 94, 251-72.

Heckman, J. and R. Robb., 1985. Using longitudinal data to estimate age, period, and cohort effects in earnings equations. In: Mason, W. M. and S. E .Fienberg (Ed.), Cohort Analysis in Social Research. Springer, New York, 137-50.

Heuer, C., 1997. Modeling of time trends and interactions in vital rates using restricted regression splines. Biometrics 53(1), 161-77.

Holford, T. R., 1985. An alternative approach to statiscal age-period-cohort analysis. Journal of Chronic Disease 38, 831-36.

Holford, T. R., 2006. Approaches to fitting age-period-cohort models with unequal intervals. Statistics in Medicine 25, 977-93.

James, I. R. and M. R. Segal., 1982. On a method of mortality analysis incorporating age-year interaction, with application to prostate cancer mortality. Biometrics 38, 433-43.

Kahn, J. R. and W. M. Mason., 1987. Political, alienation, cohort size, and the easterlin hypothesis. American Sociological Review 52, 155-69.

Kupper, L. L., J. M. Janis, I. A. Salama, C. N. Yoshizawa, and B. G. Greenberg., 1983. Age-period-cohort analysis: An illustration of the problems in assessing interaction in one observation per cell data. Communications in Statistics: Theory and Method, 12, 2779-2807.

Lee, W. C. and R. C. Lin., 1996. Autoregressive age-period-cohort models. Statistics in Medicine, 15: 27-81.

Mason, K. O., W. M. Mason, H. H. Winsborough, and K. Poole., 1973. Some methodological issues in cohort analysis of archival data. American Sociological Review 38, 242-58.

Nakamura, T., 1986. Bayesian cohort models for general cohort table analysis. Annals of the Institute Statistical Mathematics 38, 353-70.

O'Brien, R. M., J. Stockard, and L. Isaacson., 1999. The enduring effects of cohort size and percent of nonmarital births on age-specific homicide rates. 1960-1995. American Journal of Sociology 104, 1061-95.

O'Brien, R. M., K. Hudson, and J. Stockard., 2008. A mixed model estimation of age, period, and cohort effects. Sociological Methods & Research 36, 402-428.

Rodgers, W. L., 1982. Estimable functions of age, period, and cohort effects. American Sociological Review 47, 774-87.

Schulhofer-Wohl, S and Y. Yang., 2006. APC: Stata module for estimating age-period-cohort effects. Boston College Department of Economics Statistical Software Components. <http://ideas.repec.org/s/boc/bocode.html> (accessed 12.03.29).

Tango, T. and S, Kurashina., 1987. Age, period and cohort analysis of trends in mortality from major diseases in Japan, 1955 to 1979: Peculiarity of the cohort born in the early Showa Era. Statistics in Medicine 6, 709-726.

Winship, C. and D. J. Harding., 2008. A mechanism-based approach to the identification of age-period-cohort models. Sociological Methods & Research 36, 362-401.

Yang, Y., W. J. Fu, and K. C. Land., 2004. A Methodological Comparison of Age-Period-Cohort Models: Intrinsic Estimator and Conventional Generalized Linear Models. Sociological Methodology 34, 75-110.