Assessing Validity and Scope of the Intrinsic Estimator Approach to the Age-Period-Cohort Problem

Liying Luo[1,2]

1. Department of Sociology, University of Minnesota

2. Division of Biostatistics, School of Public Health, University of Minnesota

DRAFT: Do not cite without permission

August 3, 2011

Direct correspondences to Liying Luo, 909 Social Sciences, 267 19th Avenue S, Minneapolis, MN 55455; luoxx054@umn.edu.

Introduction

For well over a century, social scientists who strive to understand the temporal variation in the long run have attempted to separate age effects from period and cohort effects on a variety of social phenomena such as mortality, disease rate, family structure, social mobility, and inequality (Manson, Manson, Winsborough, and Poole 1973, Holford 1983, O'Brien 2000, Winship and Harding 2008, Fu 2000). Whereas *age effects* represent the variation associated with growing old and is considered to reflect the developmental nature over one's life course, cohort and period represent the social aspects that shape the temporal patterns of these phenomena. Specifically, *period effects* refer to the effects that are due to social and historical shifts such as economic recessions and prevalent unemployment, and are applied to all age groups simultaneously. *Cohort effects* are defined as the formative effects of social events on individuals at a specific period during their life course (Ryder 1965). Age-Period-Cohort (APC) models where the three variables are simultaneously considered in a statistical equation have been the conventional framework for quantifying age, period, and cohort effects. Unfortunately, statistical APC models suffer from *identification problem*: once the knowledge about any two of the three variables is available, the value for the third one is determined, that is, Cohort = Period – Age. Because of the exact linear dependency between age, period, and cohort, APC models are inestimable, so there are no unique estimates of the "distinct" effects of the three variables.

Many researchers consider the interrelationship between age, period, and cohort an intriguing methodological challenge and have developed various methods to deal with the identification problem inherent in APC models. A large body of literature going back to the 1970s has examined the estimation problem inherent in the APC model. Manson, Manson, Winsborough, and Poole (1973) introduced the APC multiple classification model and suggested the Constrained Generalized Linear Model (CGLM) as a means of estimating the linear effects of age, period, and cohort. While this method has been used as a general solution to the APC problem, it has been criticized as being sensitive to the choice of constraints

(Glenn 1976, Rodgers 1982a, 1982b). Recently, Fu (2000) and Yang, Fu and Land (2004) proposed a new method for distinguishing the age, period, and cohort effects. They applied the Moore-Penrose generalized inverse to statistical modeling of APC effects, termed the resulting method the *Intrinsic Estimator* (IE), and suggested it as an all-purpose solution to the APC conundrum with potentially wide application in the social sciences. On the one hand, although the authors did not claim that their method is a general solution to the identification problem in APC analysis, this method has been hailed (Schwadel 2011, Huang, Yue, and Yang 2008, Smith 2004) because it appears to have fewer assumptions and to produce less biased estimates than the conventional CGLM. On the other hand, Winship and Harding (2008) and Harding (2009) noted that this highly technical solution has not been rigorously examined with statistical and sociological theories, so it is still undetermined whether the IE approach provides a general solution to the APC problem. Based on Kupper, Janis, Karmous, and Greenberg's work, O'Brien (2011a) provided insights into the constraint that IE assumes and showed that the IE is not an unbiased estimator of the underlying, true age, period, and cohort effects. Yet scholars disagree about what exactly the IE constraint is and how this constraint affects estimation in real-world research (Fu, Land, and Yang 2011, O'Brien 2011b). Moreover, important statistical properties such as unbiasedness and consistency of the IE method have not been examined with simulation experiments that demonstrate how it performed in various circumstances, nor have the implications of the constraint that IE imposes and its application scopes for empirical research been fully understood.

The objective of this paper is to better understand the statistical properties of the IE technique and its implications for empirical research by making explicit the assumption that IE imposes implicitly. In this paper, I will identify the constraint that IE imposes on parameter estimation with mathematical derivations, reveal the non-trivial implications of this constraint for empirical APC analysis, and examine the validity of IE using simulation methods. Specifically, I will discuss the following statistical features and properties of the IE estimator:

(1) Constraint: What is the constraint that IE imposes given that "the IE may in fact also be viewed as a constrained estimator" (Yang et al. 2008, p 1706)? In other words, what are the specific forms of the IE constraint for data sets with varying numbers of age, period, and cohort groups?

(2) The implication of this constraint with respect to:

   a. Unbiasedness: Is the expectation of IE the "true" age, period, and cohort effects being estimated? In statistical terms, does the equation $E(b_{IE}) = b_{true}$ hold? If not, does the biased IE estimates correctly reflected the trend of age, period, and cohort?

   b. Consistency: Does the IE estimator converge to the "true" effects when age, period, and cohort are all of interest? In other words, as the sample size increases, does IE get closer and closer to the "true" value?

   c. Application scope: Fu and associates (2011) suggested that "the important statistical issue about APC modeling is how to identify the trend that helps to resolve the real-world problem for a given APC data set" (p 3). Compared to CGLM, does IE yield better, if not unbiased, estimates to recover the age, period, and cohort patterns that may be observed in empirical research?

Beyond the assessment of the IE technique, I also attempt to extend the discussion to the APC problem in general. Indeed, the APC problem belongs to a class of analyses where the predictors are exactly linearly dependent. Another example is the analysis of social mobility that attempts to distinguish the effects on an outcome of interest of origin socioeconomic status (SES), destination SES, and the discrepancy between them. Another sociological example is estimation of the effects on marriage duration of husband's education, wife's education, and the educational difference between them. Examples in other disciplines include analysis of the effects on economic growth of market supplies, consumers' needs, and the gap in between, and that of the effects on earnings of current age, age at labor force entry, and years in labor force. In each of these cases, the value of any one of three variables is completely determined by the other two. These studies are common and share similar theoretical problems. By pointing out the theoretical and logical problems of separating the effects of age, period, and cohort, this paper will make theoretical and methodological contributions to the literatures of related disciplines.

The Identification Problem

Although researchers are generally aware of the identification problem from a statistical perspective, the conceptual aspect that underlies this challenge and its implication for real-world research have not been emphasized. Thus to understand the nature of statistical methods for APC models including IE, I will first explicate the identification problem inherent in APC analysis that the IE estimator is intended to address.

Table 1 presents a set of tabulated data that are frequently used in APC analysis (Mason et al. 1973, Rogers 1982a, Glenn 2005, Yang et al. 2008). This table is a rectangular array of U.S. male mortality rates from 1970 through 2004 as measures of the outcome. The rows are defined by age intervals, columns defined by time periods, and the diagonals represent birth cohorts.

[Table 1 about here][1]

Age, period, and cohort effects are confounded in the tabulated data. For example, the mortality rate in 1995-99 for the 40-44 age group was 329 per 100,000 males in the U.S. population, a death rate that also corresponds to the cohort who were born in 1950-54 and became 40-44 years old in 1995-99. Similarly, the mortality rate in 1970-74 for the same age group was 461 per 100,000 in the America's male population, which also corresponds to the birth cohort of 1930-34.

[1] Source: The Berkeley Human Mortality Database (http://www.mortality.org).

Researchers have conventionally used the Analysis of Variance (ANOVA) model below where the three time-related variables are included as independent entities to separate age, period, and cohort effects:

$$g(E(Y_{ij})) = \mu + \alpha_i + \beta_j + \gamma_k, \tag{1}$$

for age groups $i = 1,2, \dots, a$, periods $j = 1,2, \dots, p$, and cohorts $k = 1,2, \dots, (a + p - 1)$, where $\sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{p} \beta_j = \sum_{k=1}^{a+p-1} \gamma_k = 0$. $E(Y_{ij})$ denotes the expected value of the outcome of interest *Y* for the *i*th age group in the *j*th period of time; *g* is the "link function" in the language of generalized linear models; α_i denotes the mean difference from the global mean μ associated with the *i*th age category; β_j denotes the mean difference from μ associated with the *j*th period; γ_k denotes the mean difference due to the membership of the *k*th cohort. The usual ANOVA constraint applies where the sum of coefficients for each effect is set to zero.

For a normally distributed outcome Y_{ij}, the ANOVA model above can also be written in a generic regression fashion:

$$Y = Xb + \varepsilon, \tag{2}$$

where *Y* is a vector of outcomes; *X* is the design matrix consisting of an intercept and categorical variables indexing respondents' age, period, and cohort group; *b* denotes a parameter vector whose elements correspond to the effects of a given age, cohort, or period group; and ε denotes random errors with a certain distribution. Then the estimated age, period, and cohort effects can be obtained using the ordinary least squares (OLS) method:

$$\hat{b} = (X^T X)^{-1} X^T Y \tag{3}$$

This approach has been strenuously challenged because age, period, and cohort are exactly linearly related, that is, Cohort = Period – Age. Specifically, in the APC model above, when age, period, and cohort are all of potential interest, the inverse of the matrix $(\mathrm{X}^{\mathrm{T}}\mathrm{X})^{-1}$ does not exist because of the age-period-cohort linear dependency, and thus the parameter vector *b* is inestimable. This is the so-called identification problem in APC analysis. One serious consequence is that no unique set of coefficients can be obtained because there are an infinite number of solutions, and the fits to the data from these solutions are all *identical*.

This identification problem can be shown more explicitly. For simplicity, suppose the data we have are perfect, without random error or measurement error; then the problem is mathematical rather than statistical and the regression model is:

$$Y = Xb. \tag{4}$$

Due to the linear dependency between age, period, and cohort, there exists a nonzero vector b_0, a linear function of the design matrix *X*, such that the product of the design matrix and the vector equals zero:

$$Xb_0 = 0. \tag{5}$$

In other words, b_0 represents the null space of the design matrix *X*, which has dimension equal to one. It follows that the parameter *b* can be decomposed into components in two orthogonal subspaces:

$$b = b_1 + s \cdot b_0, \tag{6}$$

where *s* is an arbitrary real number corresponding to a specific solution to equation (3), and b_1 is a linear function of the parameter vector *b*, corresponding to the projection of *b* on the non-null space of the design matrix *X,* orthogonal to the null space.

Given equations (4) and (6), the following equation must hold:

$$Y = Xb = X(b_1 + s \cdot b_0) = Xb_1 + s \cdot Xb_0. \tag{7}$$

But $Xb_0 = 0$ and thus $s \cdot Xb_0 = 0$, so equation (7) is true for all values of *s*. That is, *s* can be any arbitrary real number, and each value of *s* corresponds to a specific solution to equation (4). Therefore, an infinite number of possible solutions for *b* exist, and it is impossible to find a unique solution without additional constraints on *b.* To illustrate, I use an example with three age groups, three periods, and five cohorts and suppose that error is zero for ease of presentation (and without loss of generality). Table 2 presents three different parameter vectors $b^T = \left(u, \alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_{2,} \gamma_3, \gamma_4, \gamma_5\right)$ arising from three different values of s, namely 0, 2, and 10. In Figure 1a the observed value in each cell is represented in terms of the unknown parameters α_i, β_j, and γ_k from model (1) for $i = 1, 2, 3;\ j = 1, 2, 3;\ k = 1, 2, \ldots, 5$. Figure 1b presents the fitted values based on Table 2's three different b^Ts in the same tabular form as Figure 1a. Note that the three sets of fitted values are *identical* although the parameter vectors are not the same. Taken collectively, Table 2 and Figure 1 demonstrate that for a single data set, an infinite number of possible solutions for the age, period, and cohort effects exist, and each solution corresponds to a specific value of s. Therefore, any possible solution, or alternatively, none of these solutions, can be viewed as reflecting the "true" effects of age, period, and cohort even though different values of s gives radically different age and period effects.

[Table 2 about here]

[Figure 1 about here]

In social science research, real data inevitably contain random as well as measurement errors so researchers will not have the same perfect fit as in the idealized data above, but the identification problem remains: no unique set of coefficient estimates can be obtained when the predictor variables are perfectly linearly dependent, nor can any set of these coefficients be interpreted as reflecting the "true" effects. I argue that the identification problem is not methodological or statistical in nature; rather, it reflects a fundamental problem in theoretical thinking and logical reasoning. I will address this argument further in the discussion section. Now I will consider the IE and other solutions to the identification problem based on imposing a constraint on b.

The Constraint Approach

The CGLM and the IE

A large body of literature in sociology, demography, and epidemiology dating back to the 1970s has been devoted to statistically distinguishing the effects of age, period, and cohort. Mason et al. (1973)

explicated the “identification problem” in cohort analysis and set forth the Constrained Generalized Linear Model (CGLM), a coefficient-constrained method that has been used as a conventional approach to the identification problem in APC analysis. This method places (at least) one identifying restriction on the parameter vector *b* in equation (2). Usually the effects of the first two age groups, periods, or cohorts are constrained to be equal. With this one additional constraint, the APC model becomes just-identified, and the least squares and maximum likelihood estimators exist. The theoretical foundation of CGLM is to use external, extra information to make APC models estimable.

The CGLM approach has been criticized for relying on external or side information to find constraints when such information often does not exist or cannot easily be verified. In addition, different choices of identifying constraints can produce widely different estimates of patterns of change across the age, period, and cohort categories. That is, estimates of the model effect coefficients are quite sensitive to the choice of the equality coefficient constraint (Rodgers 1982a, Glenn 2005).

Recent publications on cohort analysis authored by Fu (2000, 2006) and Yang, Fu and Land (2004) derived a new APC estimator, called the *Intrinsic Estimator* (IE), based on methodological developments in biostatistics. It turns out that the IE implicitly imposes a constraint on *b*, like CGLM. This assumption is easy to overlook at least in part because on the surface the assumption necessary to apply the technique seem technical and remote. Although the assumption is, to some extent, recognized by the authors (see Yang et al. 2008, p 1706), the meaning and implications of this assumption have not been properly understood. As I will illustrate, this obscure assumption can have major effects on the coefficient estimates of the APC model. Specifically, I will explain the meaning of the IE assumption and its non-trivial implications, and then examine how IE performs in various circumstances with mathematical derivatives and simulation experiments.

The Constraint Imposed by the IE Method

IE is an application of the Moore-Penrose Generalized Inverse[2] to the APC problem. It can also be viewed as an extension of Principal Component (PC) Analysis, a multi-purpose technique that can be used to deal with identification problems when explanatory variables are highly correlated. By transforming correlated explanatory variables to a set of orthogonal linear combinations of these variables, called principal components, PC analysis is a useful tool to reduce data redundancy and to develop predictive models.

In contrast, the goal of IE is neither data reduction nor prediction, but estimation of the effects of, and capturing the general trend of, age, period, and cohort. The computational algorithm of IE includes five steps: (1) transform the design matrix X to the PC space using its eigenvector matrix; (2) in the PC space, identify the "null eigenvector" – the special eigenvector that corresponds to an eigenvalue of zero – and the corresponding null subspace (with one dimension) and non-null subspace (with *m-1* dimensions, where *m* denotes the number of coefficients to be estimated); (3) in the non-null subspace of *m-1* dimensions, regress the outcome of interest using OLS or maximum likelihood (ML) on the *m-1* PCs to obtain *m-1* coefficient estimates; (4) extend the *m-1* coefficient vector to the whole PC space of dimension *m* by adding an element corresponding to the null eigenvector direction and arbitrarily setting its coefficient to zero; and (5) use the eigenvector matrix to transform the extended coefficient vector estimated in the PC space, including the added zero element, back to the original age-period-cohort space to obtain estimates for age, period, and cohort effects (see Yang et al. 2004, Yang 2008).

The fourth step, "extend the *m-1* coefficient vector to the whole PC space of dimension *m* by adding an element corresponding to the null eigenvector direction and arbitrarily setting its coefficient to zero," carries the key assumption of the IE approach to APC analysis. This assumption has implicit yet

[2] See, e.g., Searle (1971, p 16-19) for definition and properties.

enormous implications for the validity and application of the IE approach. Specifically, setting the "coefficient of the null eigenvector", *s*, to zero is equivalent to assuming

$$b \cdot b_0 = 0, \tag{8}$$

i.e., the projection of *b* on b_0 is zero, where *b* and b_0 were defined in equation (6). Kupper et al. (1985) provided a closed-form representation for the eigenvector b_o. When the responses are in the order by row, then[3]

$$b_0 = (0, A, P, C)^T \ , \tag{9}$$

where

$$A = (1 - \tfrac{1+a}{2}, \dots, (a-1) - \tfrac{1+a}{2})$$

$$P = (\tfrac{1+p}{2} - 1, \dots, \tfrac{1+p}{2} - (p-1))$$

$$C = (1 - \tfrac{a+p}{2}, \dots, (a+p-2) - \tfrac{a+p}{2}).$$

For example, when a=3 and p=3, that is, for three age categories and three time periods, b_o is:

$$b_0 = (0, -1,0,1,0, -2, -1,0,1)^T, \tag{10}$$

where $A = (-1,0)$, $P = (1,0)$, and $C = (-2, -1, 0,1)$. Suppose that age, period, and cohort each has effects on the outcome of interest that show a linear trend. Denote these trends as k_a, k_p, k_c, respectively, the intercepts for the three variables as i_a, i_p, and i_c, and the overall mean as *u*. Thus the effects associated with the three age categories are i_a, $i_a + k_a$, and $i_a + 2 \cdot k_a$, respectively. Similarly, the effects related to the three time periods are i_p, $i_p + k_p$, and $i_p + 2 \cdot k_p$, respectively. For the five cohort groups, the effects are i_c, $i_c + k_c$, $i_c + 2 \cdot k_c$, $i_c + 3 \cdot k_c$, and $i_c + 4 \cdot k_c$, respectively. Then the parameter vector, *b*, can be written as:

$$b = (u, i_a, i_a + k_a, i_p, i_p + k_p, i_c, i_c + k_c, i_c + 2 \cdot k_c, i_c + 3 \cdot k_c)^T, \tag{11}$$

[3] Yang et al. (2004, 2008) use $b_0^* = \frac{b_0}{\|b_0\|}$, where $\|b_0\|$ is the length of b_0 so b_0^* has a length of 1. b_0 is used in this paper because it doesn't differ from b_0^* in nature and is simpler for computation.

where the last category of each variable is omitted as the reference group. According to the constraint for age effects in model (1), we know that

$$\sum_{i=1}^{a} \alpha_i = i_a + (i_a + k_a) + (i_a + 2 \cdot k_a) = 3 \cdot i_a + 3 \cdot k_a = 0, \tag{12}$$

which is equivalent to

$$i_a = -k_a. \tag{13}$$

Similarly, it can be shown using the constraint for period and cohort effects in model (1) that

$$i_p = -k_p, \tag{14}$$

and

$$i_c = -2 \cdot k_c. \tag{15}$$

Using on equations (13), (14), and (15), equation (11) can be simplified as:

$$b = (u, -k_a, 0, -k_p, 0, -2 \cdot k_c, -k_c, 0, k_c)^T. \tag{16}$$

Since the constraint that IE implicitly imposes is $b \cdot b_0 = 0$, according to equations (10) and (16), the specific form of its assumption for APC data with three age categories, three time periods, and five cohort groups are

$$b \cdot b_0 = u \cdot 0 + (-k_a) \cdot (-1) + 0 \cdot 0 + \left(-k_p\right) \cdot 1 + 0 \cdot 0 + (-2 \cdot k_c) \cdot (-2) + (-k_c) \cdot (-1) + 0 \cdot 0 + k_c \cdot 1 = k_a + 0 - k_p + 0 + 4 \cdot k_c + k_c + k_c = k_a - k_p + 6 \cdot k_c. \tag{17}$$

To illustrate the implications of this implicit constraint, I simulate normally distributed data sets as follows. For those at age *i* in period *j*, the mean response variable is $10 + k_a \cdot age_i + k_p \cdot period_j + k_c \cdot cohort_{ij}$ and the standard deviation equals 0.1. The number of age and period groups fixed at three each. I consider three sets of true k_a, k_p, k_c, and for each selection of true k_a, k_p, k_c, I simulated 1,000 such data sets by drawing random errors. As shown in Table 3, for the first set of true k_a, k_p, k_c, the age effects for the three age categories are -1, 0, and 1, respectively, so k_a, the linear trend in age effects, equals 1. The period effects in the same dataset are -7, 0, and 7 respectively, so k_p is 7. Similarly, since the cohort effects for -2, -1, 0, 1, and 2, respectively, k_c is 1. Note that for this data set,

$$k_a - k_p + 6 \cdot k_c = 1 - 7 + 6 \cdot 1 = 0\,, \tag{18}$$

i.e., the relationship between the linear trend in the age, period, and cohort effects satisfies equation (17), the constraint implicit in IE. For the other two sets of true k_a, k_p, k_c in Table 3, however, equation (17) does not hold. Specifically, for the second set, k_a=1, k_p=7, and k_c=10, so

$$k_a - k_p + 6 \cdot k_c = 1 - 7 + 6 \cdot 10 = 54 \neq 0; \tag{19}$$

And for the third set, k_a=3, k_p=1, and k_c=4, so:

$$k_a - k_p + 6 \cdot k_c = 3 - 1 + 6 \cdot 4 = 26 \neq 0. \tag{20}$$

[Table 3 about here]

Figure 2 presents the IE estimates for the three sets of true k_a, k_p, k_c, averaged over the 1,000 simulated data sets, along with the "true" effects graphically. The bias of IE is estimated by the difference between the truth and the IE estimates, averaged over the 1,000 data sets. In Figure 2a, IE yields good estimates for the first true k_a, k_p, k_c because equation (17) holds. In contrast, as shown in Figure 2b and 2c, IE returns poor estimates that are vastly different from the "true" effects for the second and third sets of true k_a, k_p, k_c, because theses true effects do not satisfy IE's constraint.

[Figure 2 about here]

In fact, for any coefficient-constraint approach such as CGLM and IE, "the choice of constraint is the crucial determinant of the accuracy in the estimated age, period, and cohort effects" (Kupper et al. 1985, p 822). Since the constraint assumption largely affects estimation results, no matter what constraint a statistical method assumes, it produces accurate estimates only when its assumptions approximate the true structure of the data at hand. It follows that when there are three age groups, three periods, and five cohorts, only when the linear effects of age, period, and cohort satisfy equation (17) can IE yields accurate estimates. However, in APC analysis researchers usually have no *a priori* knowledge about structure of age, period, cohort effects. IE, therefore, unfortunately has extremely limited application and researchers will not know what datasets it can be applied to. Thus it is no better than CGLM in this respect.

More importantly, the statistical exposition above indicates that the constraint that IE assumes also depends on the design matrix *X*, i.e., on the number of age, period, and cohort groups. For example, if we add one age group to our example, so we now have four age groups, three periods, and six cohorts, and the IE constraint is

$$b \cdot b_0 = 11 \cdot k_a - 4 \cdot k_p + 45 \cdot k_c = 0. \quad (21)$$

Compared to equation (17) for the case of three age groups, three periods, and five cohorts, equation (21) shows that adding an age group dramatically changes the constraint. Readers can verify that adding to or reducing the number of periods or cohorts will also greatly alter the constraint that the IE implicitly assumes.

These examples demonstrate that not only does IE rely on a constraint like the CGLM method does, but also unlike CGLM, where the constraint (e.g., equal age effects for the first two age groups) is explicit and based on theoretical account or side information, the constraint of IE is implicit and varies depending on the number of age, period, and cohort groups. Although the constraint that the IE technique assumes has been described as minimum (Schwadel 2011, Yang et al. 2008), in fact the assumption is so exquisitely precise that a slight inconsistency between the assumption and the reality or a small increase or decrease in the number of age, period, and cohort categories can have a tremendous impact on parameter estimation. This instability of IE's implicit constraint further limits the application of IE for analyzing data in combination with lack of *a priori* or side information about the age, period, and cohort effects.

In statistical theory, the limitation of the IE approach results from a misinterpretation of the constraint that IE imposes on parameter estimation. It is true that b_0, the null eigenvector, is determined by the design matrix, but it is incorrect to assume that therefore b_0 "is independent of the response variable *Y*" or "it should not play any role in the estimation of effect coefficient" (Yang et al. 2008, p 1705). Rather, both the null eigenvector and non-null eigenvectors (whose eigenvalues are nonzero) are determined by the design matrix, that is, by the number of age categories, time periods, and birth cohorts. To this extent, it is no less likely that the data contains a significant component in the b_0 direction than in the directions of the non-null eigenvectors. The fact that *s*, the coefficient for b_0, can be an arbitrary real number simply tells us that variation in the direction of b_0 is not estimable. If the data have variation in this direction, the IE method will mistakenly attribute that variation to (linear functions of) other columns in the design matrix, resulting in significant errors in estimation.

The Implications of the IE Constraint

Biasedness

By definition, an estimator δ is an unbiased estimator of a parameter θ if the expectation of δ over the distribution that depends on θ is equal to θ, or $E_\theta(\delta) = \theta$. If IE is an unbiased estimator, the expected value of IE must be the true age, period, and cohort effects. The following mathematical computation shows, however, that the expectation of the IE estimator is not the true effects.

As noted in the section above, the key computation of IE is to extend the coefficient vector in the PC space

$$(b')^T = (b'_0, b'_1, b'_2, \dots, b'_{m-1}) \tag{22}$$

by adding a zero element such that

$$(b'_{new})^T = (b'_0, b'_1, b'_2, \dots, b'_{m-1}, 0). \tag{23}$$

In the context of the foregoing discussion of the identification problem, the b'_{new} corresponds to the projection of the coefficient vector *b* in the non-null space, that is, b_1 as in equation (6). The IE method then uses the eigenvector matrix arising from the design matrix *X* to transform the extended coefficient vector b'_{new} including the added zero element, back to the original age-period-cohort space to obtain coefficient estimates for age, period, and cohort corresponding to the OLS and ML estimates in the PC space.

Given that OLS and ML estimators have been proven unbiased in simpler— identifiable— problems with normally distributed errors, as in equation (2), and since IE uses these methods to obtain estimates for b_1, whose projection in the PC space corresponds to extended coefficient vector b'_{new}, IE yields unbiased estimates for b_1. In other words,

$$E(b_{IE}) = b_1. \tag{24}$$

Now based on the preceding discussion on the identification problem of APC models, the "true" parameter space b can be decomposed into two orthogonal subspaces corresponding to b_1 and b_0 in equation (6), which is equivalent to

$$b_1 = b - s \cdot b_0. \tag{25}$$

Substituting equations (25) in (24) results in

$$E(b_{IE}) = b_1 = b - s \cdot b_0. \tag{26}$$

Equation (26) means that the expectation of the IE estimator will be different from the true effects unless $s \cdot b_0 = 0$ or $s = 0$. IE assumes s=0; Thus, IE is a biased estimator when the true value of s is anything but 0. The larger the absolute value of s, the more biased the IE estimate becomes.

This theoretical demonstration of the bias of IE has important implications for empirical APC analysis. For researchers who are interested in separating age, period, and cohort effects, there exists little theoretical or empirical knowledge about what b_0, the "null eigenvector," may imply about the outcome variable. In specific applications, then, the IE estimator must be assumed to be biased, resulting in misleading conclusions about the "true" age, period, and cohort effects unless proven otherwise.

Consistency

In statistics, for a consistent estimator δ, δ must converge in probability to the unknown parameter as the sample size grows. If unbiased, consistency follows immediately. A biased estimator can be consistent if its bias decreases as the sample size increases. However, the bias of IE, $s \cdot b_0$, does not necessarily shrink as the sample size grows. Thus, the IE method is not a consistent estimator of the coefficient vector *b*.

This theoretical argument on consistency can be illustrated with simulation results. Simulation is considered a helpful technique to illustrate the statistical properties of an estimator such as unbiasedness, consistency, or efficiency when these properties cannot be computed in closed form. When applying simulation methods to study the probabilistic features of consistency for a biased estimator, if the estimator is consistent, one should expect to see that as the sample size increases, the bias should be reduced, and the estimates must be closer and closer to the value specified in the simulation function. It follows that provided IE is a consistent estimator, its bias, $s \cdot b_0$, should become smaller and smaller, and the estimates should become closer and closer to the "true" effects of age, period, and cohort.

To assess IE against the consistency criterion, I generate a data set of 1,000 cross tables in R 2.13.0 using the simulation function below:

$$y_{ij} \sim Normal\{u = 10 + 3 \cdot age_i + 4 \cdot period_j + 1 \cdot cohort_{ij}, \sigma = 0.1\}, \quad (27)$$

where the response variable Y_{ij} for people at age *i* in period *j* has a Normal distribution with mean $(10 + 3 \cdot age_i + 4 \cdot period_j + 1 \cdot cohort_{ij})$ and standard variation 0.1. For this specification, k_a= 3, k_p=4, and k_c=1. I begin with three age categories, three time periods, and five cohort groups, and then increase the number of periods to 10 and 100 while the number of age categories is fixed at three. If IE is a consistent estimator, as the number of periods increases, the resulted estimates should get closer and closer to the simulated values, the true effects that we know based on the simulation function.

Figure 3 presents the estimated effects along with the true values for the three data where the number of time periods is set at three, 10, and 100, respectively. These figures show that as the number of time period increases from three to 100, the estimated effects are not converging to the "true" effects. Specifically, the IE estimates for the age effects are grossly incorrect, suggesting an opposite linear trend. For the period and cohort effects, although the IE method correctly captures the direction of the trend, there is no evidence that these estimates are converging to the "true" period and cohort effects.

[Figure 3 about here]

On the one hand, Yang et al. (2008) correctly note the estimation of the period and cohort effects will not improve with more time periods because "adding a period to the data set does not add information about the previous periods or about cohorts not present in the period just added" (p 1718). On the other hand, the estimated coefficients for age effects using the IE did appear to become closer and closer to the "true" value as the number of period increases when they simulated data using the following function:

$$y_{ij} \sim Poisson\{\exp\,[0.3 + 0.1(age_i - 5)^2 + 0.1\sin(period_j) + 0.1\cos(cohort_{ij}) + 0.1\sin\,(10 \cdot cohort_{ij})]\}, \tag{28}$$

where the response variable Y_{ij} has a Poisson distribution with mean parameter $\left[0.3 + 0.1(age_i - 5)^2 + 0.1\sin(period_j) + 0.1\cos(cohort_{ij}) + 0.1\sin(10 \cdot cohort_{ij})\right]$.

The reason why IE estimates of the age effects converge to the "true" effects as the number of period increases appears to rest on the fact that, as illustrated in Figure 4, the linear trend in the true age effects is flat, and the true period and cohort effects in their simulations becomes flatter as the number of period increases from five to 50. This is because Yang and colleagues use a sine function for the period effects and a combination of sine and cosine functions for cohort effects; as the number of period increases, these sine and functions become closer to a flat line that satisfies the IE constraint. As I point out above, the implicit constraint that IE imposes is

$$b \cdot b_0 = l \cdot k_a + m \cdot k_p + n \cdot k_c = 0 \tag{29}$$

where k_a, k_p, k_c, are defined as in equation (11) and can be obtained by fitting a linear regression line for the "true" age, period, and cohort effects; *l, m, n* are real numbers that are determined by the number of age, period, and cohort groups. As shown in Figure 4, for the data set generated by function (28) with only five time periods, k_a, k_p, and k_c (the slopes of the dotted line in the graphs) are 0.0000, -0.053, -0.004, respectively, a situation where equation (29) does not hold ($b \cdot b_0 = -0.339$). When the number of period increases to 50, with k_a remaining unchanged, k_p, and k_c are all almost zero (-0.00047 and -0.00029, respectively). As a result, equation (29) is approximately achieved ($b \cdot b_0 = -0.036$) regardless of the values of *l, m*, and *n* because k_p and k_c are close to zero. In other words, the implicit constraint of IE is approximately satisfied for simulated data where the linear trends in age, period, and cohort effect are approaching zero when the number of time period increases. For social data where neither the linear trends in the three variables, nor the relationship between these trends and number of periods are known to researchers, adding more periods or cohorts promises nothing about the accuracy about coefficient estimation for either age effects or period or cohort effects. Even with an unrealistically large number of periods (e.g. 100 periods), as I have shown in Figure 3c and will show in Figure 5, the IE produces poor estimates that are vastly different from the "true" effects.

The simulation results help to illustrate the theoretical argument that IE is not a consistent or unbiased estimator of "true" age, period, and cohort effects when the three variables are all of potential interest. For a biased and inconsistent estimator such as IE, no matter how large the sample size becomes, one cannot obtain estimates that are close to the "true" effects. Therefore, even with a sufficiently large sample, researchers using IE to estimate age, period, and cohort effects are not guaranteed to have desirable results that are close to the "true" value.

Application Scope: CGLM vs. IE

While recognizing that “the IE may in fact also be viewed as a constrained estimator” (Yang et al. 2008, p 1706), the IE’s inventors claim that IE has clear advantages over the conventional CGLM approach based on simulations in which IE estimates are closer to the “true” effects of age, period, and cohort than the CGLM results. However, the age, period, and cohort effects in the simulated data in Yang et al’s (2008) paper approximately meet the constraint that IE imposes. For other kinds of age, period, and cohort effects not satisfying IE’s implicit constraint, IE will not perform better than CGLM and may perform much worse. Thus, IE is not necessarily better than the CGLM method, because the restriction that IE imposes is essentially no different from the constraints assumed in the CGLM method.

To illustrate, I show simulations, as Yang and colleagues did, to compare the CGLM and IE estimates but the data-generating mechanisms will satisfy the constraint assumed by the CGLM but not the implicit constraint assumed by IE. Moreover, I will simulate from four models that accommodate social theories and conform to empirical reality. The first data set is simulated to represent the common observation that overall health status for adults deteriorates as they grow older, and that recent development in health knowledge and technology have improved health conditions and life expectancy for the entire population, so people born in more recent years are likely to be healthier than older cohorts. On the other hand, the sociological, epidemiological, and demographic literature has also suggested that age, period, or cohort effects may not all exist (Winship and Harding 2008, Fabio, Leober, Balasubramani, Roth, Fu, and Farrington 2006, Preston and Wang 2006). The other three simulations are used accordingly to generate data that approximate the likely empirical situation where one of the three variables has little impact on the outcome of interest.

Specifically, I fix the number of age groups at nine and periods at 50 in all of my simulations with little loss of generality. I then generate 1,000 data sets from each of the following four models for the 9 by 50 outcome matrix for nine age groups and 50 periods.

$$y_{ij} \sim Normal\{10 + 0.4 \cdot age_i - 0.1 \cdot age_i^2 + 1 \cdot period_j + 1.5 \cdot cohort_{ij}, \sigma = 0.1\} \quad (30)$$

$$y_{ij} \sim Normal\{10 + 1 \cdot period_j + 1.5 \cdot cohort_{ij}, \sigma = 0.1\} \quad (31)$$

$$y_{ij} \sim Normal\{10 + 0.4 \cdot age_i - 0.1 \cdot age_i^2 + 1.5 \cdot cohort_{ij}, \sigma = 0.1\} \quad (32)$$

$$y_{ij} \sim Normal\{10 + 0.4 \cdot age_i - 0.1 \cdot age_i^2 + 1 \cdot period_j, \sigma = 0.1\} \quad (33)$$

For instance, in equation (30), the outcomes for people with age *i* in period *j* are normally distributed with mean $10 + 0.4 \cdot age_i - 0.1 \cdot age_i^2 + 1 \cdot period_j$ and standard deviation $\sigma = 0.1$. In equation (31), (32), and (33), one of the age, period, and cohort effects are not present while the effects for the other two variables are the same as in equation (30). Note that none of these models satisfies the constraint that IE implicitly assumes; specifically, for the first model, $b \cdot b_0 = 83.80$; for the second model, $b \cdot b_0 = 83.94$; for the third model, $b \cdot b_0 = 145.49$; and for the last model, $b \cdot b_0 = -61.83$.

Figure 5a through Figure 5d present and compare the IE estimates and CGLM estimates using two different constraints to the true effects for the four models. The coefficient estimates obtained using the IE method are largely away from true effects for all the models because for all four models, the implicit constraint that IE assumes is not satisfied. In Figure 5c, when there is no period effect in the data generating mechanism (32), the IE estimates suggest a substantially positive period effect on top of the inaccurate estimates for age and cohorts effects. In contrast, the CGLM assuming equal age effects for the first and third age group produces close estimates for all four models. It is equally important to note that the performance of the CGLM estimator also depends on whether its assumption approximates the true effects. For instance, in Figure 5c, whereas the CGLM that assumes equal age effects for the first and third group yields good estimates, the same method with a different constraint, that is, the age effects are the same for the first and second group, results in estimates as poor as the IE estimates.

[Figure 5 about here]

Conclusion and Discussion

This paper focuses on the Intrinsic Estimator (IE), a statistical method intended to separate the effects of age, period, and cohort on sociological, demographic, and epidemiological outcomes. I have discussed the nature and application scope of IE theoretically and illustrated it with simulation data. This paper has shown that IE implicitly assumes a very specific constraint on the age, period, and cohort effects. This assumption not only depends on the number of age, period, and cohort groups, but also is extremely difficult, if not impossible, to verify in empirical research. This feature of IE is no different from the constraints assumed in the CGLM method except that the CGLM constraint does not change automatically as the number of age, period, and cohort groups changes. The conclusion is that IE's strategy of circumventing the identification problem inherent in the APC model (1), by using an obscure constraint, can yield grossly incorrect estimates and thus is potentially misleading.

More than a quarter century ago, Glenn (1976) argued that attempts to statistically quantify effects of age, period, and cohort are "a futile quest." Yet this quest has not been abandoned and, ironically, numerous authors have cited Glenn's work as the authority for encouraging statistical APC models. The intellectual origin of this continued quest lies in the belief that despite the exact equality, Cohort = Period - Age, there must exist simultaneously separate effects of age, period, and cohort underlying the social phenomena that we observe. In other words, the identification problem is considered as merely a statistical barrier that makes estimation difficult while the theoretical merits of APC analysis are still relevant and scientists are legitimately motivated to search for a technical solution. To (re)navigate the future development of cohort analysis, this paper aims beyond the evaluation of specific statistical techniques such the IE and the CGLM approaches to emphasize the theoretical and methodological problems of separating age, period, and cohort effect using the APC multiple classification model (1).

The fundamental goal for researchers who are interested in APC analysis is to investigate the "true, simultaneously independent effects" of age, period, and cohort effects. What does "*effect*" mean? In general, for two random variables *W* and *Z*, we say *W* has an effect on *Z*, if (a) had *W* had taken a different value than it actually did, and (b) everything else temporally prior to or simultaneous with *W* remained unchanged, then *Z* would have taken a different value. In the APC problem specifically, we may say that period, for example, has a distinct effect on the outcome if (a) had period had taken a different value than it actually did, and (b) age and cohort as well as other variables temporally prior to or simultaneous with period remained unchanged, then the outcome of interest would have taken a different value. The two sufficient and necessary conditions to establish a (causal) effect between two variables, however, cannot be satisfied in the case of APC analysis because when two of the three variables remain fixed (condition b), it is impossible for the third variable to take a different value (condition a). For example, once we have a person's birth year and the time of measurement, we know his or her age using the exact equality Cohort = Period – Age. It is not possible to manipulate the value of one of the APC variables while holding the other two constant. Therefore, even in theory and logic, there cannot be "simultaneously independent" effects of age, period, and cohort. Note that I am not arguing that age, period, or cohort effects do not exist; rather, my argument is that with the presence of any two of the three variables, the third one does not make additional contribution to explaining the variation in the phenomenon of scientific interest and thus it does not have "conditional" effects.

The current APC literature emphasizes the "unusual" statistical challenge to quantify the linear effects of age, period, and cohorts (Holford 1983, Kupper et al. 1985) but fails to recognize theoretical and logical problems inherent in APC models. When there is no logical basis for the existence of simultaneously independent effects of age, period, and cohort, the statistical methods are doomed to fail. On the one hand, the identification problem, i.e., that no unique set of solutions exists, is the immediate consequence of including all three variables (age, period, and cohort) in a statistical model. It is important to understand that the identification problem inherent in APC models is not a statistical challenge in nature; rather, it is an adverse consequence of the fundamental problem in theoretical framing and logical reasoning about causal effects of age, period, and cohort. To this extent, the identification problem "is a blessing for social science" (Heckman and Robb, 1985) because it warns scientists that they want something – a general statistical decomposition of data – for nothing. Therefore, the search for a definitive separation of age, period, and cohort effects is not only a futile but also a false quest.

On the other hand, even if the identification problem could be circumvented by using constraints, research assuming that age, period, and cohort groups are "independently" related to the outcome of interest will lead to conclusions that are logically paradoxical. A possible conclusion, for example, may be: "with age and cohort [two variables temporally prior to or simultaneous with time period] remaining unchanged, the outcome of scientific interest varies (or does not vary) across different time periods." This conditional interpretation of the statistical coefficients, however, makes no sense because, for example, the variable of period cannot vary after accounting for the person's age and his or her birth year in the APC model. In other words, the differences among period groups have been completely captured by the variation in age and period. The same problem occurs for understanding the age and cohort effect coefficients, as structured in model (1).

Furthermore, according to Karl Popper, to be scientific in nature any argument and conclusion must be disprovable. Consider the data in Figure 1 where three very different processes yield an identical data set. This can also be interpreted as meaning that for the same data, there exist an infinite number of equally good estimates. None of these estimates can be said to be more accurate or to have a better chance of revealing the "true" underlying mechanism than others, because there is no way to define true, simultaneously independent effects of age, period, and cohort. Nor can we ever replicate or falsify the conclusions using empirical data with different statistical techniques because the estimates depend completely on the constraints that a specific method imposes. In other words, APC research that aims to recover "simultaneously independent" age, period, and cohort effects is empirically unfalsifiable and theoretically meaningless.

References

Alwin, Duane F. 1991. "Family of Origin and Cohort Differences in Verbal Ability." *American Sociological Review* 56 (5):pp. 625-638.

Fabio, Anthony, Rolf Leober, G. K. Balasubramani, Jeffrey Roth, Wenjiang Fu, and David P. Farrington. 2006. "Why some Generations are More Violent than Others: Assessment of Age, Period, and Cohort Effects." *American Journal of Epidemiology* 164 (2):151-160.

Fu, Wenjiang. 2000. ‘‘Ridge Estimator in Singular Design with Applications to Age-Period-Cohort Analysis of Disease Rates.’’ *Communications in Statistics Theory and Method* 29:263-78.

Fu, Wenjiang J., and Peter Hall. 2006. "Asymptotic Properties of Estimators in Age-Period-Cohort Analysis." *Statistics & Probability Letters* 76 (17):1925-1929.

Fu, Wenjiang J., Kenneth C. Land, and Yang Yang. 2011. "On the Intrinsic Estimator and Constrained Estimators in Age-Period-Cohort Models." *Sociological Methods & Research* 40 (3):453-466.

Glenn, Norval D. 1974. "Aging and Conservatism." *Annals of the American Academy of Political and Social Science* 415 (, Political Consequences of Aging):pp. 176-186.

———. 1976. "Cohort Analysts' Futile Quest: Statistical Attempts to Separate Age, Period and Cohort Effects." *American Sociological Review* 41 (5):900-904.

———. 2005. *Cohort Analysis* . Thousand Oaks, Calif.: Sage Publications.

Harding, David J. 2009. "Recent Advances in Age-Period-Cohort Analysis. A Commentary on Dregan and Armstrong, and on Reither, Hauser and Yang." *Social Science & Medicine* 69 (10):1449-1451.

Heckman, James and Richard Robb. 1985. "Using Longitudinal Data to Estimate Age, Period, and Cohort Effects in Earnings Equations.’ " Pp. 137-50 in *Cohort Analysis in Social Research*, edited by W. M. Mason and S. E. Fienberg. New York: Springer-Verlag.

Holford, Theodore R. 1983. "The Estimation of Age, Period and Cohort Effects for Vital Rates." *Biometrics* 39 (2):pp. 311-324.

Huang, Hong-Chien, Jack C. Yue, Sharon S. Yang. 2008. “An Empirical Study of Mortality Models in Taiwan.” *Asian-Pacific Journal of Risk and Insurance* 3(1): Article 8.

Kupper, Lawrence L., Joseph Janis, Azza Karmous, and Bernard G. Greenberg. 1985. "Statistical Age-Period-Cohort Analysis: A Review and Critique." *Journal of Chronic Diseases* 38 (10):811-830.

Mason, Karen Oppenheim, William M. Manson, H.H. Winsborough, W. Kenneth Poole. 1973. "Some Methodological Issues in Cohort Analysis of Archival Data." *American Sociological Review* 38 (2):242-258.

O'Brien, Robert. M. 2000. "Age Period Cohort Characteristic Models." *Social Science Research* 29 (1):123.

O'Brien, Robert M. 2011a. "Constrained Estimators and Age-Period-Cohort Models." *Sociological Methods & Research* 40 (3):419-452.

O'Brien, Robert M. 2011b. "Intrinsic Estimators as Constrained Estimators in Age-Period-Cohort Accounting Models." *Sociological Methods & Research* 40 (3):467-470.

O'Brien, Robert. M. 2000. "Age Period Cohort Characteristic Models." *Social Science Research* 29 (1):123.

Popper, Karl Raimund. 1968. *The Logic of Scientific Discovery*. London: Hutchinson.

Preston, Samuel H., and Haidong Wang. 2006. "Sex Mortality Differences in the United States: The Role of Cohort Smoking Patterns." *Demography* 43 (4):631-646.

Rodgers, Willard L. 1982a. "Estimable Functions of Age, Period, and Cohort Effects." *American Sociological Review* 47 (6):pp. 774-787.

———. 1982b. "Reply to Comment by Smith, Mason, and Fienberg." *American Sociological Review* 47 (6):pp. 793-796.

Ryder, Norman B. 1965. "The Cohort as a Concept in the Study of Social Change." *American Sociological Review* 30 (6):pp. 843-861.

Schwadel, Philip. 2011. "Age, Period, and Cohort Effects on Religious Activities and Beliefs." *Social Science Research* 40 (1):181-192.

Searle, Shayle R. 1971. *Linear Models*. New York: John Wiley & Sons.

Smith, Herbert L. 2004. "Response: Cohort Analysis Redux." *Sociological Methodology* 34 (1):111-119.

Winship, Christopher, and David J. Harding. 2008. "A Mechanism-Based Approach to the Identification of Age--Period--Cohort Models." *Sociological Methods & Research* 36 (3):362-401.

Yang, Yang, Wenjiang J. Fu, and Kenneth C. Land. 2004. "A Methodological Comparison of Age-Period-Cohort Models: The Intrinsic Estimator and Conventional Generalized Linear Models." *Sociological Methodology* 34 (1):75-110.

Yang, Yang, Sam Schulhofer-Wohl, Wenjiang J. Fu, Kenneth C. Land. 2008. "The Intrinsic Estimator for Age-Period-Cohort Analysis: What it is and how to use it." *American Journal of Sociology* 113 (6):1697-1736.

Tables and Figures

Table 1. Deaths per 100,000: U.S. males, 1970-2004

Age	1970-74	1975-79	1980-84	1985-89	1990-94	1995-99	2000-04
0-	493	384	311	279	237	198	181
5-	48	39	33	28	24	20	17
10-	50	43	36	35	31	27	23
15-	157	143	125	120	125	104	93
20-	218	199	178	166	163	144	139
25-	200	189	178	178	178	142	134
30-	221	193	188	213	225	180	148
35-	303	257	230	260	284	236	205
40-	461	392	339	335	359	329	302
45-	738	632	547	503	486	461	453
50-	1,128	1,005	894	807	730	660	650
55-	1,783	1,523	1,410	1,296	1,160	1,036	952
60-	2,693	2,395	2,144	2,025	1,834	1,654	1,473
65-	3,951	3,484	3,251	3,013	2,779	2,528	2,254
70-	5,854	5,321	4,915	4,659	4,158	3,902	3,498
75-	8,487	7,590	7,243	6,941	6,310	5,897	5,462
80-	12,358	11,444	11,089	10,809	10,124	9,615	8,650
85-	18,182	16,752	16,427	16,274	15,527	15,453	14,442
90+	28,050	26,450	26,095	26,479	26,009	26,653	26,389

Table 2. Different values of *s* and the corresponding coefficents for age, period, and cohort.

s	**Age Effects**			**Period Effects**			**Cohort Effects**				
	1	2	3	1	2	3	1	2	3	4	5
0	2	0	-2	-1	0	1	-1	-0.5	0	0.5	1
2	0	0	0	1	0	-1	-5	-2.5	0	2.5	5
10	-8	0	8	9	0	-9	-21	-10.5	0	10.5	21

Note: 1. *s* is an arbitrary real number correspoding to a specific solution to equation (3).
2. Numbers in each row correspond to a set of age, period, and cohort coefficients corresponding to a specific value of *s,*and thereof, a specific set of solution.

Table 4. Linear trends in age, period, and cohort effects.

Data Set	Age Effects				Period Effects				Cohort Effect		
	1	2	3	k_a	1	2	3	k_p	1	2	3
1	-1	0	1	1	-7	0	7	7	-2	-1	0
2	-1	0	1	1	-7	0	7	7	-20	-10	0
3	-3	0	3	3	-1	0	1	1	-8	-4	0

Note: k_a, k_p, and k_c represent the linear trend in age, period, and cohort effects, respectively.

		Period		
		1	2	3
	1	$\mu + \alpha_1 + \beta_1 + \gamma_3$	$\mu + \alpha_1 + \beta_2 + \gamma_4$	$\mu + \alpha_1 + \beta_3 + \gamma_5$
Age	2	$\mu + \alpha_2 + \beta_1 + \gamma_2$	$\mu + \alpha_2 + \beta_2 + \gamma_3$	$\mu + \alpha_2 + \beta_3 + \gamma_4$
	3	$\mu + \alpha_3 + \beta_1 + \gamma_1$	$\mu + \alpha_3 + \beta_2 + \gamma_2$	$\mu + \alpha_3 + \beta_3 + \gamma_3$

Figure 1a. Tabular data: observed cells and the unobserved parameters α_i, β_j, and γ_k from model (1).

		Period		
		1	2	3
	1	11	12.5	14
Age	2	8.5	10	11.5
	3	6	7.5	9

Figure 1b. Tabular data: identical fitted values produced by the three different paramters vectors in Table 2.

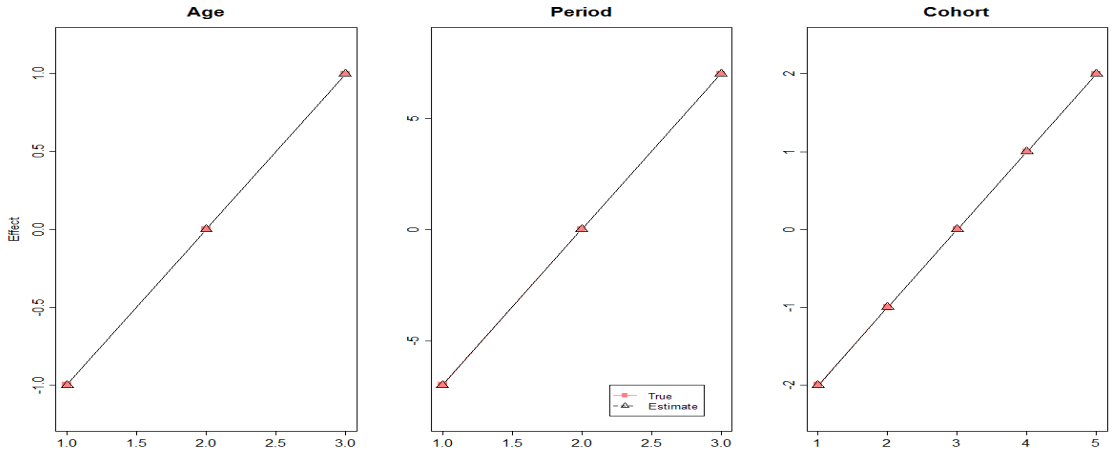

Figure 2a. Simulation results: IE estimates for dataset 1 in Table 3.

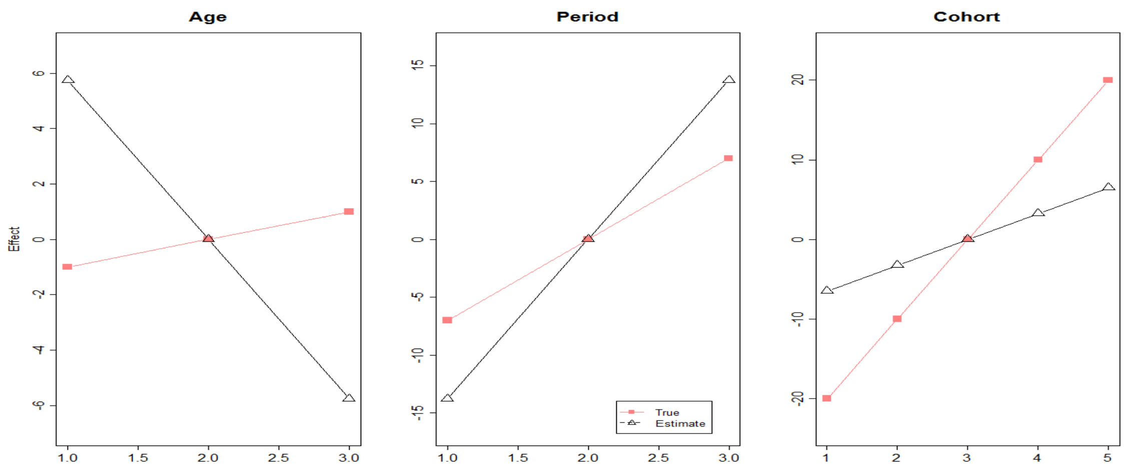

Figure 2b. Simulation results: IE estimates for dataset 2 in Table 3.

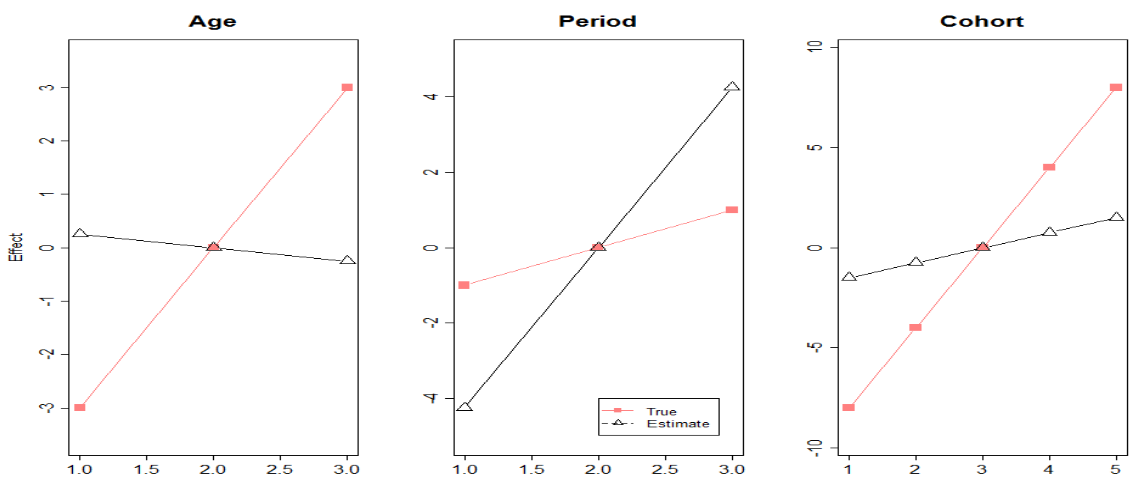

Figure 2c. Simulation results: IE estimates for dataset 3 in Table 3.

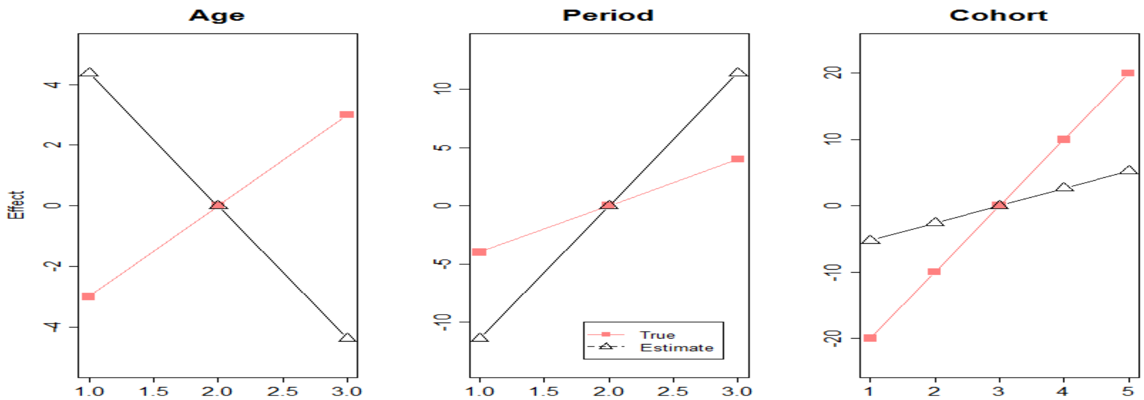

Figure 3a. Simulation results: Inconsistency of the IE estimator; number of periods = 3; simulation function: $y_{ij} \sim Normal\{u = 10 + 3 \cdot age_{ij} + 4 \cdot period_{ij} + 10 \cdot cohort_{ij}, \sigma = 0.1\}$

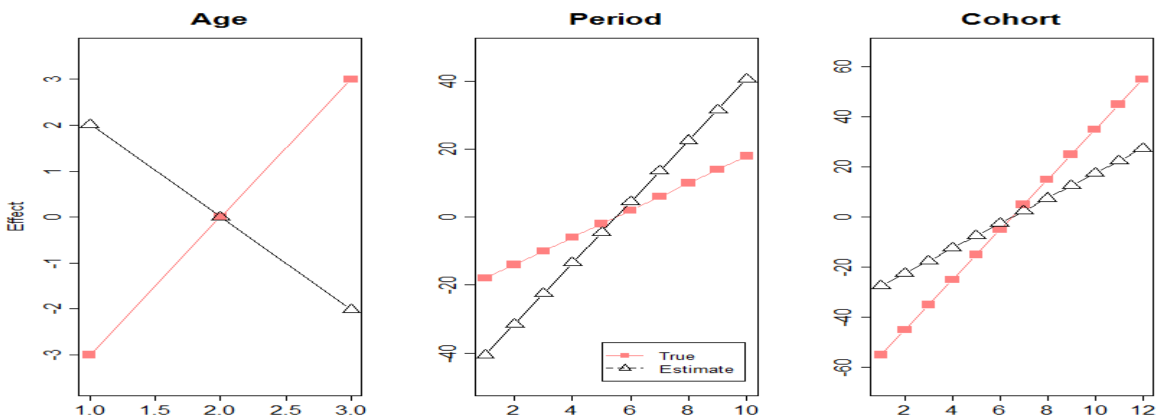

Figure 3b. Simulation results: Inconsistency of the IE estimator; number of periods = 10; simulation function: $y_{ij} \sim Normal\{u = 10 + 3 \cdot age_{ij} + 4 \cdot period_{ij} + 10 \cdot cohort_{ij}, \sigma = 0.1\}$

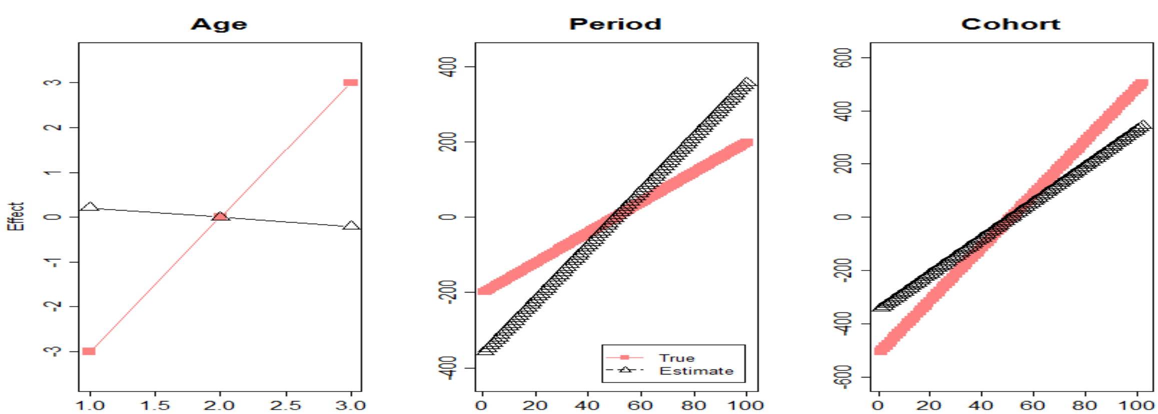

Figure 3c. Simulation results: Inconsistency of the IE estimator; number of periods = 100; simulation function: $y_{ij} \sim Normal\{u = 10 + 3 \cdot age_{ij} + 4 \cdot period_{ij} + 10 \cdot cohort_{ij}, \sigma = 0.1\}$

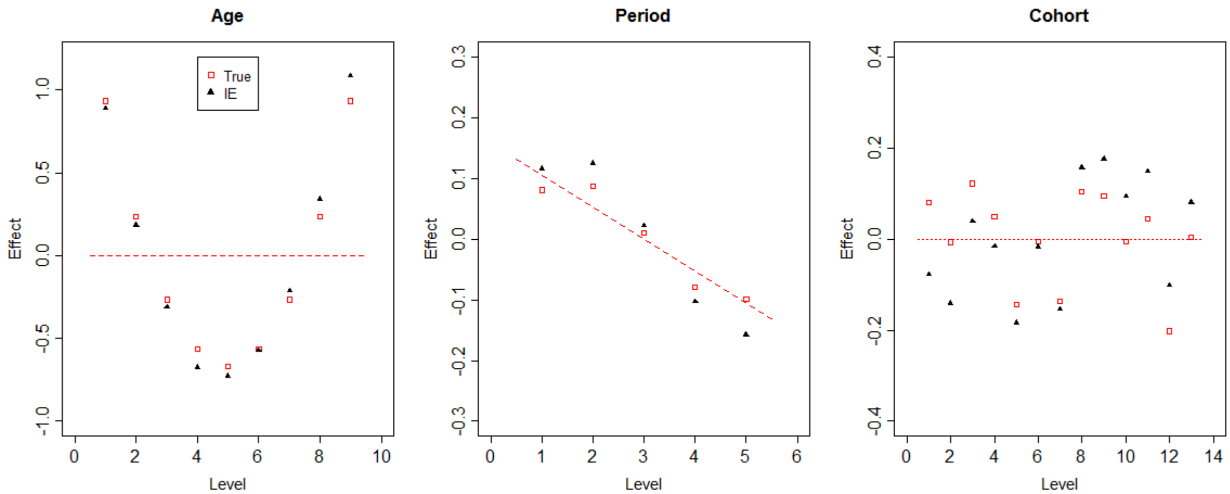

Figure 4a. Linear trends in age, period, and cohort effects in Yang et al.'s (2008) simulated data when the number of period = 5

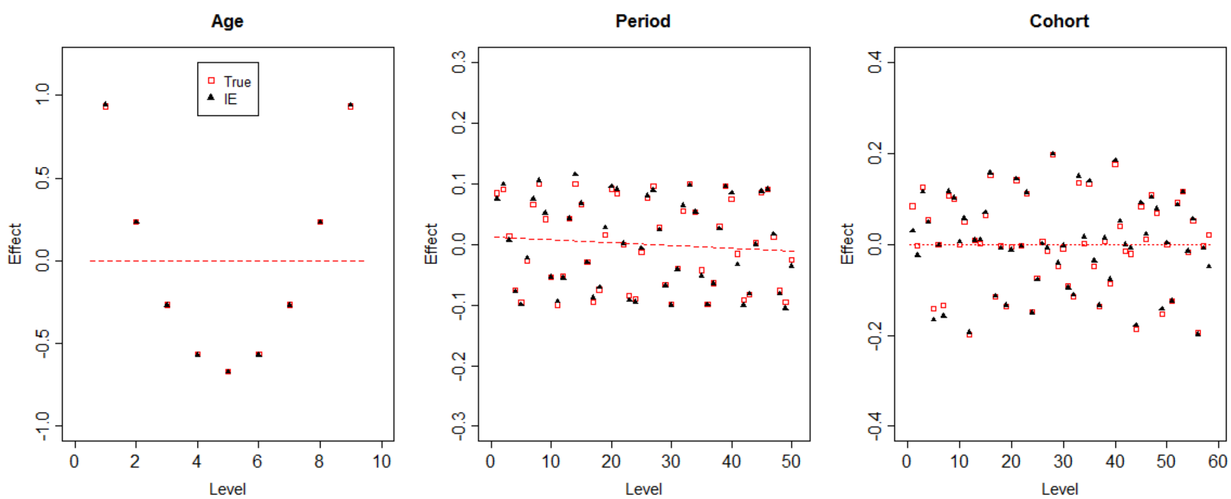

Figure 4b. Linear trends in age, period, and cohort effects in Yang et al.'s (2008) simulated data when the number of period = 50

Note: Dotted lines are fitted linear regression slopes in true effects of age, period, and cohorts, respectively.

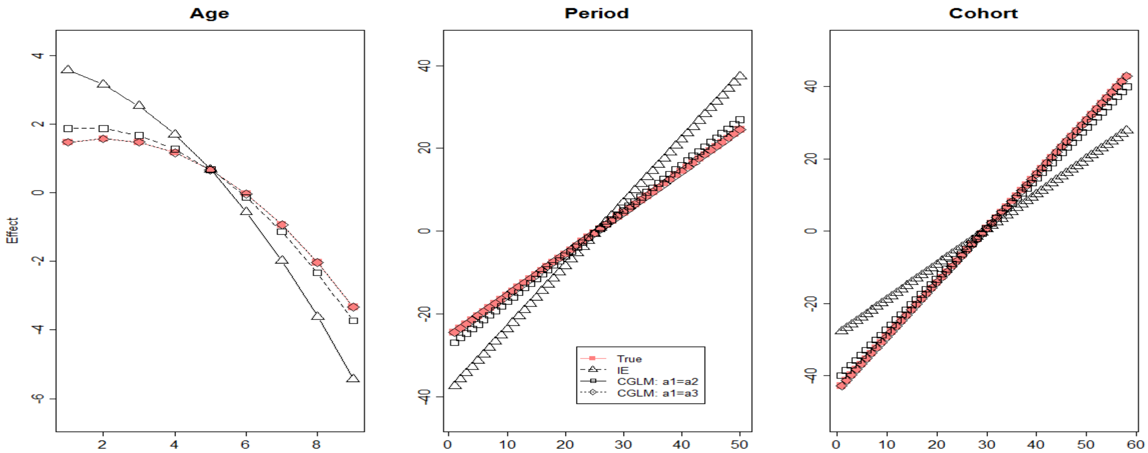

Figure 5a. Simulation results: IE vs. CGLM; number of periods = 50; simulation function: $y_{ij} \sim Normal\{u = 10 + 0.4 \cdot age_{ij} - 0.1 \cdot age_{ij}^2 + 1 \cdot period_{ij} + 1.5 \cdot cohort_{ij}, \sigma = 0.1\}$

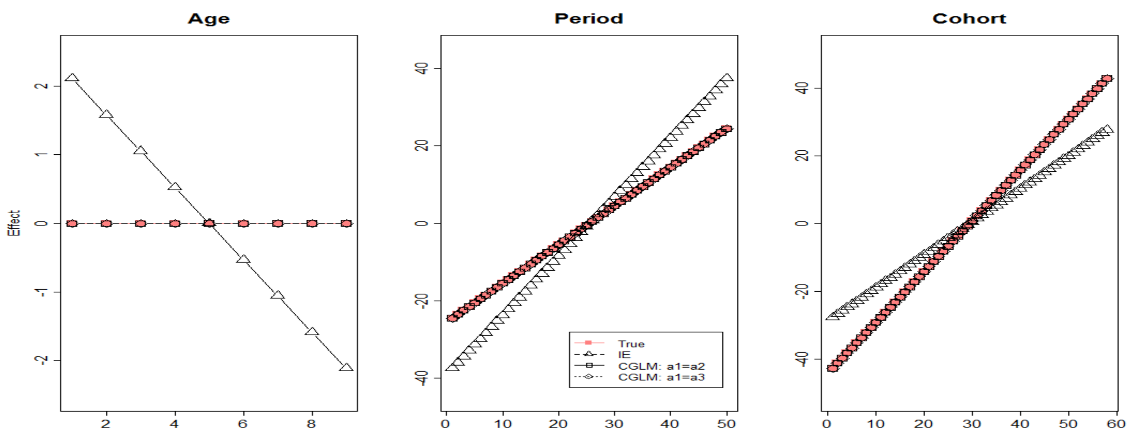

Figure 5b. Simulation results: IE vs. CGLM; number of periods = 50; simulation function: $y_{ij} \sim Normal\{u = 10 + 1 \cdot period_{ij} + 1.5 \cdot cohort_{ij}, \sigma = 0.1\}$

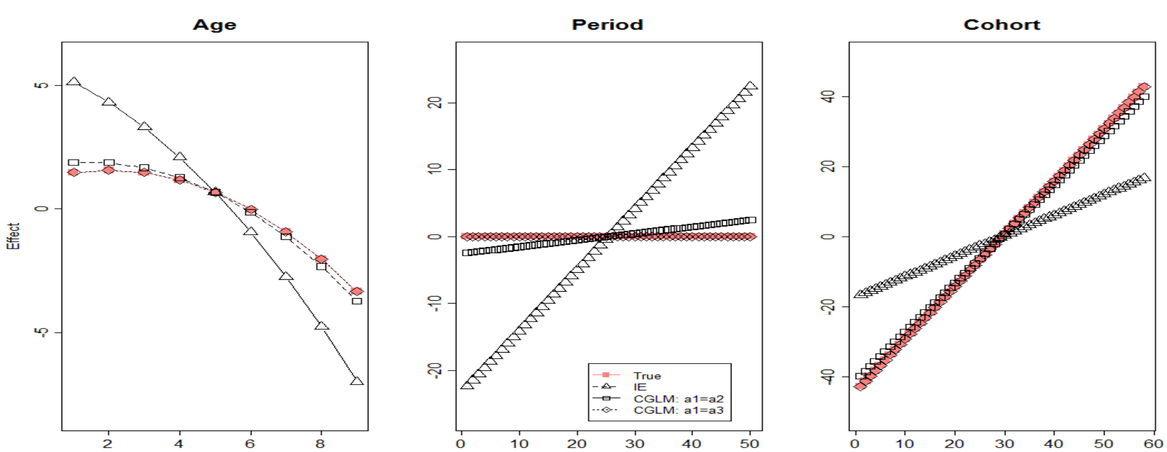

Figure 5c. Simulation results: IE vs. CGLM; number of periods = 50; simulation function: $y_{ij} \sim Normal\{u = 10 + 0.4 \cdot age_{ij} - 0.1 \cdot age_{ij}^2 + 1.5 \cdot cohort_{ij}, \sigma = 0.1\}$