

Forthcoming in *American Journal of Sociology*

Dec. 20, 2007

The Intrinsic Estimator for Age-Period-Cohort Analysis: What It Is and How To Use It

Yang Yang,^{*} Sam Schulhofer-Wohl,^{**} Wenjiang J. Fu,[#] and Kenneth C. Land^{***}

^{*}Department of Sociology, Population Research Center, and Center on Demography and Economics of Aging at NORC, The University of Chicago

^{**}Department of Economics and Woodrow Wilson School of Public and International Affairs, Princeton University

[#]Department of Epidemiology, Michigan State University

^{***}Department of Sociology and Center for Population Health and Aging, Duke University

Address all correspondence to: Yang Yang, Department of Sociology, University of Chicago, 1126 E. 59th St., Chicago, IL 60637; Phone: (773) 834-1113, Fax: (773) 702-4849, e-mail: yangy@uchicago.edu

The Intrinsic Estimator for Age-Period-Cohort Analysis:

What It Is and How To Use It

Abstract

The identification problem in Age-Period-Cohort (APC) analysis using multiple classification/accounting models has previously been widely addressed in demography, epidemiology, and statistics. A new approach to the statistical estimation of APC accounting models, called the Intrinsic Estimator (IE), recently has been developed. This paper (1) further describes the IE algebraically, geometrically, and verbally, (2) reviews properties of the IE as a statistical estimator, (3) provides model validation evidence for the IE both from an empirical example and from a simulation exercise, (4) relates the coefficients of the IE to those of conventional constrained APC models using formal definitions of statistical estimability, hypothesis testing and empirical applications that directly address a criticism that often has been lodged at general-purpose methods of APC analysis, and (5) introduces computer software for application of the IE that interested users can readily access. We conclude that the IE holds the potential for applications not only to APC analysis but also to similar problems of structural underidentification in sociology. This paper lays the foundations for such applications.

The Intrinsic Estimator for Age-Period-Cohort Analysis:

What It Is and How To Use It

INTRODUCTION

Age-period-cohort (APC) analysis has played a critical role in studying time-specific phenomena in sociology, demography, and epidemiology for the past 80 years (Mason and Wolfinger 2002). Broadly defined, APC analysis distinguishes three types of time-related variation in the phenomena of interest: *age effects*, or variation associated with different age groups, *period effects*, or variation over time periods that affect all age groups simultaneously, and *cohort effects*, or changes across groups of individuals who experience an initial event such as birth in the same year or years. These distinctions have important implications for measurement and analysis. The considerable regularity of age variations in many social outcomes across time and place reflects the developmental nature of true age changes. In contrast, period and cohort effects reflect the influences of social forces. Period variations often result from shifts in social, historical, and cultural environments. Cohort variations are conceived as the essence of social change and may reflect the effects of early life exposure to socioeconomic, behavioral, and environmental factors that act persistently over time to produce differences in life course outcomes for specific cohorts (Ryder 1965).

The Age-Period-Cohort (APC) accounting/multiple classification model was introduced to sociologists by Mason, Mason, Winsborough, and Poole (1973) and serves as a general methodology for cohort analysis when all three of age, period, and cohort are potentially of interest. This general methodology focuses on the APC analysis of data in the form of tables of

percentages or occurrence/exposure rates of events such as births, deaths, disease incidence, crimes, etc. In spite of its theoretical merits and conceptual relevance, APC analysis of tabulated data suffers from the “identification problem” induced by the exact linear dependency between age, period, and (birth) cohort: $\text{Period} = \text{Age} + \text{Cohort}$. This can be viewed as a special case of collinear regressors that produces, in this case, a singular matrix (of one less than full rank) used in the statistical estimation process. Since a singular matrix produces multiple estimators of the three effects, it is difficult to estimate the unique true separate effects.

A number of methodological contributions to the specification and estimation of age-period-cohort models have occurred in recent decades in a wide variety of disciplines, including social and demographic research (e.g., Fienberg and Mason 1978, 1985; Glenn 1976; Hobcraft et al. 1982; O’Brien 2000; Wilmoth 1990;) and biostatistics and epidemiology (e.g., Clayton and Schiffers 1987; Fu 2000; Holford 1992; Osmond and Gardner 1982; Robertson and Boyle 1998; Tarone and Chu 1992;). Various analytic approaches have produced ambiguous and inconsistent results. Researchers do not agree on methodological solutions to these problems and conclude that APC analysis is still in its infancy (Kupper et al. 1985; Mason and Wolfinger 2002).

Recent developments in APC methodology in biostatistics have emphasized the utility of estimable functions that are invariant to the selection of constraints on the parameters (Clayton and Schiffers 1987; Holford 1983, 1991, 1992; Kupper, Janis, Salama, Yoshizawa, and Greenberg 1983; Kupper, Janis, Karmous, and Greenberg 1985; Robertson et al. 1999). This is the approach applied by Fu (2000) in the derivation of a new APC estimator – termed the *intrinsic estimator* (IE) – based on estimable functions and the singular value decomposition of matrices.

Yang, Fu, and Land (2004) compared two approaches to the identification problem in APC accounting models, namely, the IE method and the *constrained generalized linear models* (CGLIM) estimator that has been conventional among demographers and other social scientists for more than two decades (Fienberg and Mason 1978, 1985; Mason and Smith 1985). Through data analyses of population mortality rates, Yang et al. illustrated some similarities and differences of these two methods in parameter estimates and model fit. They also discussed some key statistical properties of the intrinsic estimator compared with the constrained generalized linear models estimator.

Based on their results, Yang et al. (2004) concluded that the IE offers a useful alternative to conventional methods for the APC analysis of tables of rates. The conceptual foundations of the IE are sufficiently abstract and difficult to understand, however, that additional exposition and illustration are merited. Accordingly, the objectives of this paper are to (1) further describe the IE algebraically, geometrically, and verbally, (2) review properties of the IE as a statistical estimator, (3) report results of model validation assessments of the IE both from an empirical example and from a simulation exercise, (4) give some “how to use” advice and show how to relate the coefficients of the IE to those of conventional constrained APC models with applications to U.S. female mortality rates, 1960-1999, and (5) introduce computer software that interested users can readily access.

Since the IE is a general purpose method of APC analysis of potentially wide applicability in the social sciences, it merits recalling the criteria for acceptability of such a general purpose method recently articulated by Norval Glenn, a long-time critic of attempts to provide general solutions to the APC analysis problem. Glenn (2005: 20) stated that such a method “... may prove to be useful... if it yields approximately correct estimates ‘more often

than not,' if researchers carefully assess the credibility of the estimates by using theory and side information, and if they keep their conclusions about the effects tentative." These are strong criteria with which we agree. The purpose of this paper is to assess the extent to which the IE satisfies them.

THE ALGEBRA OF THE APC IDENTIFICATION PROBLEM

To fix ideas, we focus on the APC analysis of rectangular arrays of demographic rates arranged in conventional fashion with age intervals defining the rows and time periods defining the columns. Specifically, we analyze the same table of U.S. female mortality rates, 1960 through 1999, that was studied by Yang et al. (2004); see Table 1. The data from which this table was constructed were obtained from the Berkeley Human Mortality Database. As is conventional in demographic and epidemiological analyses of arrays of this type, both age and period are measured in five-year intervals, and the diagonal elements of the matrices correspond to birth cohorts.

[Table 1 About Here]

For example, the death rate for the age 35-39 group recorded in 1960-64 was 0.001835 or 183 per 100,000 females in the population. By comparison, the mortality rate for the same age group in 1995-99 was 119 per 100,000 females. The complication that arises from the APC analysis of such arrays of rates is that the former rate ($183/100000$) also corresponds to the mortality rate of the cohort of U.S. females born in the years 1925 to 1929 and the latter rate

(119/100000) also corresponds to the mortality rate for the birth cohort of the years 1960 to 1964.

The *Age-Period-Cohort accounting/multiple classification model* (Mason, et al. 1973) is a conventional tool for the analysis of age-by-time period arrays of demographic rates. Yang et al. (2004) described the algebraic structure of this model and, for the sake of completeness of the exposition, we repeat it here. For mortality rates, this model can be written in *linear regression form* as:

$$M_{ij} = D_{ij} / P_{ij} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ij} \quad (1)$$

where M_{ij} denotes the observed death rate for the i -th age group for $i = 1, \dots, a$ age groups at the j -th time period for $j = 1, \dots, p$ time periods of observed data; D_{ij} denotes the number of deaths in the ij -th group; P_{ij} denotes the size of the estimated population in the ij -th group, the population at risk of death; μ denotes the intercept or adjusted mean death rate; α_i denotes the i -th row age effect or the coefficient for the i -th age group; β_j denotes the j -th column period effect or the coefficient for the j -th time period; γ_k denotes the k -th diagonal cohort effect or the coefficient for the k -th cohort for $k = 1, \dots, (a+p-1)$ cohorts, with $k=a-i+j$; and ε_{ij} denotes a random error with expectation $E(\varepsilon_{ij}) = 0$.

Conventional APC models as represented in (1) fall into the class of generalized linear models (GLIM; see McCullagh and Nelder 1989 or McCulloch and Searle 2001 for expositions) that can take various alternative forms. First, model (1) can take a *log-linear regression form* via a *log link* as:

$$\log(E_{ij}) = \log(P_{ij}) + \mu + \alpha_i + \beta_j + \gamma_k \quad (2)$$

where E_{ij} denotes the expected number of deaths in cell (i,j) that is assumed to be distributed as a Poisson variate, and $\log(P_{ij})$ is the log of the exposure P_{ij} in (1) and is called the “offset” or

adjustment for the log-linear contingency table model. Models of this type are widely used in demography and epidemiology where the counts of demographic events such as deaths or the incidence of diseases generally follow Poisson distributions and the rates are estimated through log-linear models (Agresti 1996). A second alternative formulation of the model is to treat the underlying number of deaths as a binomial variate. The canonical link changes from a log link to a *logit link*, which yields a *logistic model*:

$$\theta_{ij} = \log\left(\frac{m_{ij}}{1-m_{ij}}\right) = \mu + \alpha_i + \beta_j + \gamma_k \quad (3)$$

where θ_{ij} is the log odds of death and m_{ij} is the probability of death in cell (i,j) . This model has been implemented more widely in demographic research (e.g., Mason and Smith 1985).

Regression models (1), (2), and (3) can be treated as *fixed-effects generalized linear models* after a reparametrization to center the parameters:

$$\sum_i \alpha_i = \sum_j \beta_j = \sum_k \gamma_k = 0 \quad (4)$$

After re-parameterization (4), model (1) can be written in the conventional matrix form of a least-squares regression:

$$Y = Xb + \varepsilon \quad (5)$$

where Y is a vector of mortality rates or log-transformed rates, X is the regression design matrix consisting of “dummy variable” column vectors for the vector (of dimension $m = 1 + (a-1) + (p-1) + (a+p-2)$) of model parameters b :

$$b = (\mu, \alpha_1, \dots, \alpha_{a-1}, \beta_1, \dots, \beta_{p-1}, \gamma_1, \dots, \gamma_{a+p-2})^T \quad (6)$$

with the T superscript denoting vector transposition, and where ε in (5) is a vector of random errors with mean 0. Note that the parameters $\alpha_a, \beta_p,$ and γ_{a+p-1} are not included in the parameter vector b because of the constraints (4) and can be uniquely determined by use of (4) in

conjunction with each estimator of b . Later herein, we illustrate by an empirical example that the use of reference categories is equivalent to the translation by a constant of the parameter estimates produced by the constraints (4) and thus of no substantive importance.

The ordinary least squares (OLS) estimator of the matrix regression model (5) is the solution b of the normal equations:

$$\hat{b} = (X^T X)^{-1} X^T Y \quad (7)$$

But this estimator *does not exist* (i.e., there is no uniquely defined vector of coefficient estimates). This is due to the fact that the design matrix X is singular with one less than full column rank (Kupper et al. 1985), due to the perfect linear relationship between the age, period and cohort effects:

$$\text{Period} - \text{Age} = \text{Cohort}.$$

Therefore, $(X^T X)^{-1}$ does not exist. This is the *model identification problem* of APC analysis. It implies that there are an infinite number of possible solutions of the matrix equation (7) (i.e., OLS estimators of model (5)), one for each possible linear combination of column vectors that result in a vector identical to one of the columns of X . Therefore, it is not possible to separately estimate the effects of cohort, age, and period without imposing at least one constraint on the coefficients in addition to the reparameterization (4).

Since the work of Fienberg and Mason (1978, 1985), the conventional approach to multiple classification APC models in demography has been a *coefficients-constraints approach*, which takes the form of placing (at least) one additional identifying constraint on the parameter vector (6), e.g., constraining the effect coefficients of the first two periods to be equal, $\beta_1 = \beta_2$. With this one additional constraint, the model (5) is just-identified, the matrix $(X^T X)$ becomes

nonsingular, and the least squares estimator (7) exists (as do related maximum likelihood estimators for models (1), (2), and (3)).¹

The main problems with this CGLIM approach have been discussed in a large methodological literature in demography, epidemiology, and statistics. First, the analyst needs to rely on external or side information to find constraints, but such information often does not exist or cannot easily be verified (Mason and Wolfinger 2002). Second, different choices of identifying constraints can produce widely different estimates of patterns of change across the A, P, and C categories. As demonstrated, for example, by Mason and Smith (1985) and Yang et al. (2004), estimates of model effect coefficients are sensitive to the choice of the equality coefficient constraint. Third, all just-identified models will produce the same levels of goodness-of-fit to the data, making model fit a useless criterion for selecting the best constrained model. But no other criteria or guidelines have been proposed. These problems often lead to the conclusion that it is impossible to obtain meaningful estimates of the distinct effects of age, period, and cohort in analysis of social change.

THE INTRINSIC ESTIMATOR – DESCRIPTION

So what is new about the Intrinsic Estimator? The consensus has been that the key problem for APC analysis is to identify an estimable function that uniquely determines the parameter estimates. But the controversy continues as to whether there exists such an estimable

¹ There are other solutions to APC identification problems. The *proxy variables approach* uses one or more proxy variables as surrogates for the age, period, or cohort coefficients (see, e.g., O'Brien [2000]). The *nonlinear parametric transformation approach* defines a nonlinear function of one of the age, period, or cohort variables so that its relationship to others is nonlinear (see, e.g., Fienberg and Mason [1985], Yang and Land [2006]). These approaches do not take the form of APC accounting model and are not directly comparable to the Intrinsic Estimator in statistical terms. As will be illustrated later, however, comparison of results obtained by different approaches provides a means of model validation.

function that solves the identification problem. The conventional wisdom is that only the nonlinear, but not the linear, components of APC models can be estimable (Holford 1985; Rodgers 1982). As noted in Fu (forthcoming 2008), however, there have been only numeric demonstrations, but no rigorous proofs, to support the idea that no estimable function exists. It should also be noted that Kupper et al. (1985) provided a condition for estimable functions and suggested that an estimable function satisfying this condition resolves the identification problem. Subsequent publications have shown in that the Intrinsic Estimator satisfies this condition and estimates the unique estimable function including both the linear and nonlinear components of the parameter vector of the multiple classification model (Fu 2000, 2007; Fu, Hall, and Rohan 2004; Yang et al. 2004).

Within the context of the foregoing description of the algebra of the APC identification problem in conventional linear regression models, we next describe the IE in three ways – algebraically, geometrically, and verbally.

Algebraic Definition

First, as concerns the algebraic definition of the IE, Yang et al. (2004) showed that, because the design matrix X is one less than full column rank, the parameter space of the unconstrained APC regression model (5) can be decomposed into the direct sum of two linear subspaces that are perpendicular to each other. One subspace corresponds to the unique zero eigenvalue of the matrix $X^T X$ of equation (7) and is of dimension one; it is termed the null subspace of the design matrix X . The other non-null subspace is the complement subspace orthogonal to the null space.

Due to this orthogonal decomposition of the parameter space, each of the infinite number of solutions of the unconstrained APC accounting model (5) can be written as

$$\hat{b} = B + sB_0, \quad (8)$$

where s is a scalar corresponding to specific solution and B_0 is a unique eigenvector of Euclidean norm or length one. The eigenvector B_0 does not depend on the observed rates Y , only on the design matrix X , and thus is completely determined by the numbers of age groups and period groups – regardless of the event rates. In other words, B_0 has a specific form that is a function of the design matrix. To give an explicit representation of B_0 , note that the exact linear dependency between age, period, and cohort variables in model (5) is mathematically equivalent to

$$XB_0 = 0. \quad (9)$$

This equation expresses the property that X is singular, i.e., there exists a linear combination of the columns of the design matrix X that equals a zero vector. Kupper et al. (1985) showed that B_0 has the algebraic form:

$$B_0 = \frac{\tilde{B}_0}{\|\tilde{B}_0\|}, \quad (10)$$

i.e., B_0 is the normalized vector of \tilde{B}_0 :

$$\tilde{B}_0 = (0, A, P, C)^T, \quad (11)$$

where $A = \left(1 - \frac{a+1}{2}, \dots, (a-1) - \frac{a+1}{2}\right)$, $P = \left(\frac{p+1}{2} - 1, \dots, \frac{p+1}{2} - (p-1)\right)$, and

$C = \left(1 - \frac{a+p}{2}, \dots, (a+p-2) - \frac{a+p}{2}\right)$ with a , p , and c denoting the number of age categories,

time periods, and cohorts in the age-by-time period array of rates. It is important to note that the vector B_0 is fixed or non-random, because it is a function solely of the dimension of the design

matrix X or the number of age groups (a) and periods (p). The fact that the fixed vector B_0 is independent of the response variable Y suggests that it should not play any role in the estimation of effect coefficients. But the conventional CGLIM approach violates this principle if the scalar s in equation (8) is nonzero.

The idea that B_0 should not affect the results is key point, as intuition suggests that the eigenvector corresponding to the zero eigenvalue should be an arbitrary vector. And, indeed, sB_0 is arbitrary. But B_0 is not arbitrary; it is fixed by the design matrix. Furthermore, by equation (8), any APC estimator, obtained by placing any identifying constraint(s) on the design matrix can be written as a linear combination $B + sB_0$, where B is the special estimator *termed the intrinsic estimator* (IE) that lies in the parameter subspace that is orthogonal to the null space and determined by the Moore-Penrose generalized inverse.²

One computational algorithm for the IE is a principal components regression method, whereby one (a) computes the eigenvalues and eigenvectors (principal components) of the matrix $X^T X$, (b) normalizes them to have unit length; (c) identifies the eigenvector B_0 corresponding to the unique eigenvalue 0; (d) estimates a regression model with response vector Y as in (5) and design matrix U whose column vectors are the principal components determined by the eigenvectors of non-zero eigenvalues, i.e., estimates a principal components regression model; and (e) then uses the orthonormal matrix of all eigenvectors to transform the coefficients of the principal components regression model to the regression coefficients of the intrinsic estimator B .

Geometric Representation

The parameter space of the unconstrained vector b can be decomposed into two parts that are orthogonal or independent to each other:

² See, e.g., Searle (1971: 16-19) for a definition of the Moore-Penrose generalized inverse and its properties.

$$b = b_0 + sB_0, \tag{12}$$

where $b_0 = P_{proj}b$ is a special parameter vector that is a linear function of b corresponding to the projection of the unconstrained parameter vector b to the non-null space of X . Specifically, the special parameter vector b_0 corresponding to $s = 0$ satisfies the geometric projection

$$b_0 = (I - B_0B_0^T)b \tag{13}$$

This projection is illustrated in Figure 1. Figure 1 shows the projection of two parameterizations, b_1 and b_2 onto the non-null parameter space (the vertical axis in Figure 1), which is independent of the real number s . The geometric representation in Figure 1 either can be thought of as a simple parameter space of dimension two or as multidimensional with the vertical axis representing a direction in a multidimensional non-null space. In either case, since the projection of any parameterization of b in Figure 1 yields the same parameter vector b_0 , the latter is estimable.³

[Figure 1 About Here]

Figure 1 also helps to illustrate geometrically that the IE may, in fact, also be viewed as a constrained estimator. But, in contrast to the equality constraints on two or more *coefficients* of the parameter vector b that are imposed in conventional approaches to the estimation of APC accounting models, the constraint imposed by the IE to identify model (5) is a constraint on the *geometric orientation of the parameter vector b in parameter space*. Specifically, the IE imposes the constraint that the direction in parameter space defined by the eigenvector B_0 in the null

³ Figure 1 provides only a geometric illustration. Full algebraic details of the proof that b_0 is estimable and the only estimable function that determines both the linear and nonlinear trends in the age, period, and cohort coefficients are given in Fu, Hall, and Rohan (2004).

space of the design matrix X have zero influence on the parameter vector b_0 (i.e., on the specific parameterization of the vector b that is estimated by the IE). Since B_0 is a fixed vector that is a function solely of the design matrix (e.g., the number of time periods of data in an analysis) and does not depend on the observed event rates or frequencies being analyzed, this would seem to be a reasonable constraint.

Corresponding to the projection of the parameter vector b onto b_0 , we have the following projection of the estimators of equation (13) onto the IE B

$$B = (I - B_0 B_0^T) \hat{b}. \quad (14)$$

This equation provides another algorithm for computing the IE, namely, compute an initial estimator \hat{b} of model (5), say by an equality constraint on two of the age, period, or cohort parameters, and then geometrically project \hat{b} to the IE B by removing the component in the B_0 direction.

Verbal Description

Statisticians have known since Kupper et al. (1985) that the dimension of the design matrix in the APC accounting model (i.e., the numbers of age groups and time periods) may affect the estimates obtained from CGLIM estimators. Put in the simplest possible terms, the basic idea of the IE is to remove the influence of the design matrix on coefficient estimates. As noted later herein, this approach produces an estimator that has desirable statistical properties.

The IE also can be viewed as a special form of principal components regression estimator (for a standard exposition of principal components regression, see, e.g., Sen and Srivastava 1990) that removes the influence of the null space of the design matrix X on the estimator. It specifically estimates a constrained parameter vector b_0 that is a linear function of the parameter

vector b of the unconstrained APC accounting model (5). This constrained parameter vector b_0 corresponds to the projection of the unconstrained parameter vector b onto the non-null subspace of the design matrix X .

Since the IE is a principal components estimator, one might well ask: Why not just calculate the eigenvectors of the matrix $X^T X$ by application of principal components, regress the observed rates on the sub-space spanned by these eigenvector, and leave it at that? The answer is that regression coefficient estimates in this subspace are not directly interpretable in terms of age, period, and cohort effects. Therefore, the IE uses the extra step of inverse orthonormal transformation of the coefficient estimates of the principal components regression back to the original space of age, period, and cohort coordinates. The inverse transformation is what makes the IE a special form of principal components estimator. It yields coefficients, as will be illustrated by numerical example below, that are directly interpretable as age, period, and cohort effects and that can be compared to corresponding effects estimated by the conventional imposition of equality constraints on parameters.

STATISTICAL PROPERTIES AND VALIDATION OF THE IE

Statistical Properties

Yang et al. (2004) stated and proved some properties of the IE as a statistical estimator that we briefly summarize here. For context, consider first the analysis of an APC dataset for a finite number of time periods, p . That is, suppose that an APC analysis is to be conducted for a

fixed matrix of observed rates or event counts. This implies that the corresponding design matrix X is fixed (i.e., X has a fixed number of age groups and time periods). The randomness in the error term ε of model (5) then corresponds to measurement errors in the rates or in the event counts and/or to intrinsic randomness in the rates or counts.

In this context of an age-by-time period table of population rates with a fixed number of time period, p , of data, Yang et al. (2004: 101) showed that the IE satisfies a condition for estimability of linear functions of the parameter vector b that was established by Kupper et al. (1985: Appendix B) and recently further elaborated by Fu (forthcoming 2008). *Estimable functions*⁴ are invariant with respect to whatever solution (see equation (7) above) is obtained to the normal equations; estimable functions are desirable as statistical estimators because they are linear functions of the unidentified parameter vector that can be estimated without bias, i.e., they have unbiased estimators.⁵ Specifically, the condition for estimability of a constraint on the parameter vector that was established algebraically by Kupper et al. (1985) is, in the notation defined above, that $l^T B_0 = 0$, where l^T is a constraint vector (of appropriate dimension) that defines a linear function $l^T b$ of b . Note that, since the IE imposes the constraint that $s = 0$, i.e., that the arbitrary vector B_0 has zero influence, $l^T = (I - B_0 B_0^T)$ for the IE. Since $B_0^T B_0 = 1$, it follows that $l^T B_0 = (I - B_0 B_0^T) B_0 = B_0 - B_0 B_0^T B_0 = B_0 - B_0 = 0$, i.e., the Kupper et al. condition holds for the IE. Note also that the Kupper et al. condition implies that any constrained estimator that is obtained by imposing an equality constraint on the parameter vector

⁴ See Searle (1971: 180-188) or McCulloch and Searle (2001: 120-121) for expositions of this concept.

⁵ In the history of discussions of the APC accounting model in sociology, Rodgers (1982a) was early to argue that analysts should seek estimable functions of the unidentified parameter vector (equation (6)); see also the comment by Smith, Mason, and Fienberg (1982) and the response by Rodgers (1982b). In some respects, the IE can be regarded as providing a practical, easily applicable method to produce estimates of estimable functions from data in the form of age by time period tables of rates, as called for by Rodgers over two decades ago. The estimability referred by Rodgers, however, essentially means identifiability that can be achieved by any linear constraints, which differs from statistical estimability defined for APC models by Kupper et al.

b and that contains any nonzero component due to the vector B_0 defined by the design matrix is not estimable, i.e., produces biased estimates of the A, P, and C effect coefficients.

Because the IE B satisfies the estimability condition for APC models, it follows that, for a fixed number of time periods of data, the IE B is an unbiased estimator of the special parameterization (or linear function) b_0 of b defined in equation (12).⁶ Thus, a first statistical property of the IE in the context of an APC analysis of demographic rates with a fixed number of time periods of data is that it produces unbiased estimates of the regression coefficients of the projected coefficient vector b_0 . Second, Yang et al. (2004:108) also showed that, for a fixed number of time periods of data, the IE is more statistically efficient (has a smaller variance) than any CGLIM estimator that is obtained from a nontrivial equality constraint on the unconstrained regression coefficient estimator b , i.e., any equality constraint that does not produce a projection of b onto b_0 .

In brief, the IE has nice finite-time-period properties. In addition to finite-time-period properties, the asymptotic properties of APC estimators as the number of time periods p of data increase have been studied by Fu, Hall, and Rohan (2004), Fu and Hall (2006), and Fu (forthcoming 2008). These properties derive largely from the fact that the eigenvector B_0 converges elementwise to zero with increasing numbers of time periods of data. The vector can converge elementwise to zero even though its length is fixed at one because the number of elements of the vector grows as we add time periods. Therefore, for any two estimators $\hat{b}_1 = B + s_1 B_0$ and $\hat{b}_2 = B + s_2 B_0$, where s_1 and s_2 are nonzero and correspond to different identifying constraints placed on model (5), as the number of time periods in an APC analysis increases, the difference between these two estimators \hat{b}_1 and \hat{b}_2 decreases towards zero, and, in

⁶ Yang et al. (2004: 107) also proved this property directly.

fact, that the estimators converge toward the IE B . Suffice it to say that the proof proceeds by demonstrating that the coordinates of B_0 are bounded by a quantity that is a function of the number of age groups and periods and this function converges to 0 as $p \rightarrow \infty$. Under suitable regularity conditions on the error term process and a fixed set of age categories with effect coefficients, this feature of B_0 yields a convergence of the IE asymptotically to these “true” effect coefficients. In addition, Fu and Hall (2006) argue that, as the number of time periods of data increases, there are definite bounds on the differences between the effect coefficients estimated by the IE and the effect coefficients of the true period and cohort processes. A corollary is that CGLIM estimators \hat{b} with nonzero s in Equation (8) of the parameter vector in model (5) are biased in finite-time-period APC analyses. Except under certain conditions specified by Fu, Hall, and Rohan (2004), this bias decreases as the number of time periods in the analysis increases, in which case the CGLIM estimators may converge to the IE estimator.

Remark 1: These statistical properties are not trivial and merit comment. Both the intrinsic estimator B and any other estimator $\hat{b} = B + sB_0$ with $s \neq 0$ obtained from an equality constraint are asymptotically consistent as the number of time periods of data increase without bound. Therefore, with a large number (e.g., 30 or 40) of periods of data, differences among estimators decline and it makes little difference which identifying constraint is employed. In most empirical APC analyses, however, there usually are a small number (e.g., 4 or 5) time periods of observations available for analysis. In these cases, the differences can be substantial.

Remark 2: As just noted, the IE, by its very definition and construction, satisfies the estimability condition. Other estimators using theoretically derived equality constraints on the parameter vector b may satisfy this condition either exactly or statistically (in a sense defined below). If other estimators do indeed satisfy the estimability condition, then they also produce

unbiased estimates of the A, P, and C effect coefficients. If not, then the estimates they produce are biased.

Remark 3: These properties provide a means for differentiating among estimators. That is, for tables of rates with a finite number of time periods of data, especially a small number (e.g., 4 or 5), an unbiased estimator should be preferred to a biased estimator, as the latter can be misleading with respect to the estimated trends across the age, time period, and cohort categories.

Remark 4: In contrast, as has been noted many times over the years in discussions of the APC accounting model (see, e.g., Pullum 1977, 1980; Rogers 1982, Smith et al., 1982), different just-identified models will generate the same data and yield exactly the same model fit. In particular, linear transformations of the estimated A, P, and C coefficients obtained by the IE (i.e., linear transformations of the elements of the B vector) will fit the observed data just as well as the IE.⁷ If, however, such a linear transformation of the A, P, and C coefficients (or any subset thereof) results in coefficients that depart sufficiently far from the coefficients in B that the transformed coefficients contain a significant component of the B_0 vector, i.e., contain a significantly nonzero s coefficient, then the resulting transformed vector will not be estimable, that is, it will not be unbiased.⁸ Thus, even though the transformed coefficients will reproduce the data just as well as those obtained by the IE, they will be biased and will give poor indications of the patterns of change across the age, period, and cohort categories used in the analysis. Therefore, goodness-of-fit to the data (as measured, e.g., by log-likelihood functions or

⁷ Geometrically, a linear transformation of the coefficient vector corresponds to a rotation of the vector in parameter space. Such a rotation will produce distorted and misleading indications of patterns of change across the age, period, and cohort categories used in the analysis.

⁸ This point is illustrated in our empirical analyses and is the basis of the statistical test for estimability derived below.

deviance statistics) *cannot* be used as a criterion for selecting among estimators. But estimability can be so used.

Remark 5: Because of its estimability and unbiasedness properties, the IE may provide a means of accumulating reliable estimates of the trends of coefficients across the age, period, and cohort categories of the APC accounting model. To provide intuition for this statement, recall the distinction between the *steady-state* and *general solutions* to the ordinary differential equation for, say, Hooke's Law for the motion of a displaced spring-mass system subject to an additional forcing motion in classical mechanics. This law has the algebraic form $F = -kx + a \cos(\omega t)$, where F denotes acceleration (second derivative of the motion with respect to time) of the mass, x denotes the distance of displacement, k is a constant unique to the particular spring under study and $a \cos(\omega t)$ is the forcing term (see, e.g., Marion and Thornton 1995, p. 125). If the mass is displaced by, say, a distance of two feet, it will oscillate back and forth with some influence of the length of the initial displacement, but it eventually will settle down to a characteristic pattern of oscillations that depends only on the driving force. By comparison, if the initial displacement of the mass is a distance of four feet, then the mass will display an initial set of larger oscillations back and forth that are different from the pattern observed for the two-foot displacement. But, after an initial series of oscillations, the driving force will cause the mass to settle into the same set of oscillations back and forth as those found after an initial two-foot displacement. Mathematically, the pattern of oscillations observed after the impacts of the initial lengths of displacements have worn off are termed the *steady-state solution* of the Hooke's Law differential equation, whereas the *general solution* of the equation consists of the steady-state solution plus a factor that takes into account the initial conditions or displacement of the spring. Because initial conditions can vary from application to application,

the general solution of the differential equation can be quite unique to the application. On the other hand, the steady-state part of the solution is invariant and generalizable to the motion of the system regardless of the initial conditions.

Analogously, the IE is essentially a steady-state solution to the APC accounting model estimation problem that factors out the initial conditions of the dimension of the matrix of observed data, namely, the number of age and time period categories that define the design matrix. Because the IE does not allow these “initial conditions” to influence the estimates it produces of the A, P, and C effect coefficients, they will be more invariant to changes in the design matrix, such as additional time period of data, than estimates produced by estimators that incorporate such influences. In this sense, the IE removes the part of the subjectivity in the estimator that is due to the shape of the data. We will illustrate this feature of the IE below in our model validation analyses.

Validation – Empirical Example

In brief, the IE possesses some valuable properties as a statistical estimator. But, given the long history of problems and pitfalls in proposed methods of APC analysis, it is reasonable to question whether this estimator gives numerical estimates of age, period, and cohort effect coefficients that are valid, i.e., reveal the true effects. This is a question of *model validation* – that is, does the identifying constraint imposed by the IE, the projection of the unconstrained APC accounting model vector b onto the non-null space of the design matrix X , as in equation (12), produce estimated coefficients that meaningfully capture the true age, time period, and cohort trends.

One approach to the question of validity is to compare results from an APC analysis of empirical data by application of the IE with results from an analysis of the same empirical data by application of another approach that uses a different family of models that do not use the identifying constraint of the IE or CGLIM estimators for the accounting model. As an instance of such an empirical comparison, we next describe analyses of verbal test score data from 15 cross-sections of the General Social Survey (GSS), 1974-2000.⁹ This is an extension of GSS verbal test data for 1974-1996 on which a series of articles published in the *American Sociological Review* in 1999 posed and disputed the existence of an intercohort decline in verbal vocabulary.

In the GSS surveys, a survey respondent's vocabulary knowledge is measured by a composite scale score named WORDSUM, which is constructed by adding the correct answers to ten verbal test questions and ranges from zero to 10. WORDSUM has a distribution that is approximately bell-shaped with a mean of about 6 and is reported in previous studies to have an internal reliability of .71 (Wilson and Gove 1999: 258).¹⁰ The data include 19,500 respondents who had measures on WORDSUM and other covariates across all survey years. Respondents' ages in the data pooled across all surveys vary from 18 to 89. The average completed education is 12.7 years. Fifty-seven percent of respondents are female, and 15% are black. There are 19 five-year birth cohorts. The oldest cohort member was born in 1890, and the youngest was born in 1980.

We apply the IE to data on verbal test scores grouped into five-year age groups and time periods shown in Table 2. In this age-by-period array, there are 12 five-year age categories from

⁹ Survey years including verbal ability are 1974, 1976, 1978, 1982, 1984, 1987, 1988, 1989, 1990, 1991, 1993, 1994, 1996, 1998, and 2000.

¹⁰ In an item analysis of individual words in WORDSUM, Alwin (1991: 628) found that some of the words have become more difficult over time. The general conclusion in this series of articles (Alwin 1999; Glenn 1994; Wilson and Gove 1999), however, is that word obsolescence does not account for observed changes in the test scores over time.

ages 20 to 75+,¹¹ five five-year period groups from 1976 to 2000, and $12 + 5 - 1 = 16$ cohorts born in 1901 to 1976.¹² This yields 60 degrees of freedom. The event/exposure rates (sample mean proportions) of correct answers to GSS vocabulary questions can be transformed by a log link and modeled by a log-linear regression for which the IE can be obtained. The events are the total number of correct answers for every age and period group and are calculated by multiplying the mean verbal scores by the number of individuals. These are nonnegative counts and can be considered to be distributed as Poisson variates. The population exposure is calculated as the product of number of people in each cell and the total number of possible correct answers (10). Vocabulary knowledge test scores are not available every year from 1974 to 2000. Therefore, for those missing years, we interpolated the mean verbal scores and numbers of individuals at risk based on the data of neighboring years.

[Table 2 and Figure 2 about here]

Figure 2 shows the results from application of the IE to the data of Table 2. Estimated coefficients and their 95% confidence intervals are plotted for successive categories within the age, period, and cohort classifications. Since they indicate changes in correct answer rates from one age group to the next, from one time period to the next, and from one cohort to the next, the estimated coefficients represent the temporal trends of vocabulary knowledge along each of these three dimensions, net of the effects of the other two. It can be seen that the age effect coefficients

¹¹ We excluded ages 18-19 to obtain age groups of equal interval length so that the diagonal elements of the age by period matrix refers to cohort members. The results are not influenced by omitting this cell of small sample size. Ages 75-89 are grouped into a last 75+ category in order to combine small population exposures for more stable estimates.

¹² The definition of the width of the time intervals used to define birth cohorts in APC analyses is somewhat arbitrary. For consistency with extant articles on trends in GSS vocabulary knowledge data, we use five-year birth cohorts.

depict a concave pattern. This corroborates the quadratic age effect found by Wilson and Gove (1999): low at youth, rising to a peak in the 40s, staying largely flat until the mid-50s and declining gradually into late life. Period effects are fairly flat for the first ten years from mid-1970s to mid-1980s, but followed by a jump into the mid-1990s that flattens out again until 2000. The slightly zigzag shape of the period curve shows some variation in vocabulary knowledge over time for the past 30 years. This is quite revealing given the absence of direct estimates of period effects in the previous studies. The cohort effects are more complicated and are characterized by a bimodal curve, with the peaks occurring for cohorts born in 1911-1916 and 1941-1946. The intercohort declines, as insisted by Glenn (1994, 1999) and Alwin (1991, 1999), are evident for cohorts in-between and afterwards, that is, among the post-World War I and post-World War II cohorts. But there are also increases in vocabulary knowledge for cohorts born before World War I and for the 1930-1950 cohorts. These graphical results give us an overview of the average APC trends that help to inform and specify models for additional APC analyses using individual-level data.

The GSS data on verbal test scores also were analyzed by Yang and Land (2006) and Yang (2006) using a very different approach to APC analysis – namely, a hierarchical age-period-cohort (HAPC) analysis in the form of a cross-classified random effects regression model. This approach proceeds by building a level-one fixed-effects regression model at the individual level of analysis and then a random-effects model for cohort and time period effects at level-two. The resulting analysis takes full advantage of the existence of individual-level data on all respondents from the GSS surveys, while, at the same time, nesting these within the time periods and cohorts to which they correspond. The mixed effects model does not require the assumption of additive and fixed age, period, and cohort effects used by the conventional linear regression

models that cause the identification problem. The resulting estimates of cohort and period effect coefficients estimated in this hierarchical regression models approach are average residual effects of the cohorts and period across all time period and cohorts, respectively, and are not constrained in any way to conform to the IE or any other identifying constraint required by the account model.

[Figure 3 about here]

The results of the HAPC analysis are shown in Figure 3, which is reproduced from Yang's (2006) Bayesian estimates of the HAPC model. While the coefficient metrics and time scales for the age and period effects in Figure 3 are in single years rather than the five-year groups of those from the IE analysis in Figure 2, the trends of estimated effect coefficients are quite similar across the graphs. That is, both Figure 2 and Figure 3 exhibit age effect curves that are quadratic and period effect curves that show a slight decline from the 1970s into the 1980s and then a slight rise into the 1990s. And the cohort effect curves are quite similar – showing peaks in the middle and early twentieth century followed by declines.

In brief, this comparison shows that the independent estimates from a hierarchical models analysis corroborate the estimated patterns of change across age, period, and cohort categories that are obtained by imposing the identifying assumption of IE analysis. Of course, this is only one example of comparative analysis, and additional empirical studies are needed before it can be concluded that the IE produces substantively meaningful and empirically valid results under various circumstances. And any statistical model has its limitations and surely will break down

under some circumstances. The identification of these limits for the IE is a topic that merits additional research.

Validation – Simulation Results

Following good statistical practice, we now consider another approach to model validation: simulation analyses. It is well known that given any age by period array of rates, it is impossible to know what true effects generated the data, because any number of just-identified constrained models fit the data equally well. It is, therefore, important that we apply an estimator such as the IE or CGLIM to artificial data wherein we know the true form of the underlying model that generated the data. We present here the results of some Monte Carlo simulation analyses to determine whether IE and CGLIM estimators recover the true parameters.

We first investigate whether the IE is unbiased in finite samples. The asymptotic properties of the IE apply as the number of periods in the dataset goes to infinity. Any dataset used in practice has only a finite number of periods. We explore whether the asymptotic results give good approximations to the behavior of the IE in finite samples by simulating datasets with five and 50 periods. The basic asymptotic result we expect is that, as the number of periods increases, estimated age effects should converge to the true age effects when using the IE but not necessarily when using other estimators.¹³

We fix the number of ages in all of our simulations at 9 without loss of generality. For a given number of periods P , we generate 1,000 datasets by Monte Carlo simulation in which the entries in the $P \times 9$ outcome matrix are distributed according to:

¹³ Regardless of the estimator used, estimated period and cohort effects cannot be expected to converge to their true values as the number of periods increases because adding a period to the dataset does not add information about the previous periods or about cohorts not present in the period just added.

$$y_{ij} \sim \text{Poisson}\{\exp[0.3 + 0.1(\text{age}_{ij} - 5)^2 + 0.1\sin(\text{period}_{ij}) + 0.1\cos(\text{cohort}_{ij}) + 0.1\sin(10 \cdot \text{cohort}_{ij})]\}$$

This equation for the data-generating process tells us what the true age, period and cohort effects are:

age effect at age a	$0.1(a - 5)^2$
period effect in period p	$0.1\sin(p)$
cohort effect in cohort c	$0.1\cos(c) + 0.1\sin(10c)$

We subtract constants from the above values so that the true effects are normalized to have mean zero in each category according to (4), where the constants can be calculated as the mean effects for each category (see footnote 15). We then estimate age, period and cohort effects in each simulated dataset using the IE and using three different CGLIM estimators: one with the first two age effects constrained to be equal (CGLIM_a), one with the first two period effects constrained to be equal (CGLIM_p), and one with the first two cohort effects constrained to be equal (CGLIM_c). To make the CGLIM results comparable to the IE results, we re-normalize, by using procedures described in the next section of the paper, each set of age, period and cohort effects to sum to zero for both the CGLIM estimators and the IE.

Table 3 reports our results on age effects. For each age effect in the model, we show the true value and, for each estimator, the mean, standard deviation and mean squared error (MSE) of the estimated effect across 1,000 simulations. By comparing the mean of the simulated estimates to the true values, we can see whether each estimator is unbiased. The standard deviation of the simulated estimates shows how much the estimated parameters vary from sample to sample. Mean squared error is the average squared difference between the estimated parameter and the truth; this measure of accuracy takes account of both bias and variance.

Three of the four estimators recover the qualitative U-shaped profile of the age effects. The CGLIM estimator that constrains the first two age effects to be equal fails to recover even the qualitative form of the age effects, because the constraint is far from correct: the first two true age effects are not nearly equal. However, even the CGLIM estimators that recover the qualitative shape of the age effects are far off the mark in quantitative terms; only for the IE is the mean of each estimated age effect close to the true value. The IE also exhibits substantially less sampling variation than the CGLIM estimators and has much lower mean squared error. The IE performs better than the CGLIM whether we use five or 50 periods of data.

[Table 3 about here]

Because the numbers of period and cohort effects in our samples are large, we report our results for these effects graphically in Figure 4 for the simulations with five time periods of data. The figure shows the mean of the simulated estimates and the mean squared error for each coefficient and each estimator. Again, the mean for the IE is close to the true value, while the CGLIM estimators are not close, and the mean squared error is much smaller for the IE than for CGLIM. It is noteworthy that the cohort effects are particularly poorly estimated by the CGLIM models.

We also investigated how well the estimators perform when there are no true period or cohort effects. Figure 5 shows results from simulations where the age and period effects were as described above but the cohort effects were all zero. Only CGLIM_p estimates are shown among all CGLIM estimates because they are the closest to the true effects. On average, the IE correctly estimates that there are no cohort effects in the data. The CGLIM estimators incorrectly

find cohort effects that are different from zero and change substantially across birth cohorts. Results were similar when we set all of the period effects to zero in the data: Only the IE detected that no period effects were truly present.

[Figures 4 and 5 about here]

In simulations not reported here, we found similar results for datasets with ten and twenty periods of data.¹⁴ The simulations thus show that, regardless of the sample size, the IE is more accurate than the CGLIM. The IE performs well even in datasets with just five periods, perhaps the smallest sample size that might be used in practice. By contrast, the CGLIM estimators give incorrect results even when there are as many as fifty periods, which would be an unusually large sample in many demographic applications.

THE INTRINSIC ESTIMATOR – IINTERPRETION AND USE

Empirical Applications, Interpretation of Model Parameters, and Hypothesis Testing

Given the desirable properties the IE possesses as a statistical estimator and its evident ability to produce valid estimates of age, period, and cohort effect coefficients, the question becomes one of how to interpret and use this estimator. The question of interpretation arises because the identifying constraint imposed by the IE on the unidentified APC accounting model parameter vector b – namely, projection onto the non-null (column) space of the design matrix X

¹⁴ Due to space constraints, only a limited set of simulation results are reported here. Additional results, including analyses of other models in the GLM family and formal tests for unbiasedness will be reported in subsequent work.

– appears to be a constraint of purely algebraic convenience, devoid of substantive meaning. By contrast, conventional equality-constrained estimators of APC accounting models often are motivated by substantive hypotheses derived from theory or prior studies that indicate that certain coefficients are, say, of the same magnitude and hence can be constrained to be equal. As noted earlier herein, however, if the equality constraints are not, in fact, valid, the constraints result in a non-estimable function of b and produce CGLIM estimates of APC effect coefficients that may be wildly off the mark.

To focus the discussion, consider the effect coefficient estimates reported in Table 4 for the mortality data for U.S. females, 1960-1999, given in Table 1. The first two columns of Table 4 report the coefficient estimates and standard errors (S.E.s) for the IE applied to these data. The next two columns report the corresponding estimates for a CGLIM model, which (for reasons that will be made clear below) we label CGLIM 3, wherein the coefficients for the respective first categories of the age, period, and cohort groups are taken as the reference categories and have effects set to zero and the identifying constraint is that the second birth cohort (C2), 1870-74, is constrained to have the same effect coefficient (zero) as the first cohort (C1), 1865-69. The bottom rows of Table 4 give the overall model fit statistics for these two models. Earlier we noted (in Remark 4) that all just-identified models that incorporate effect coefficients for the full array of age, period, and cohort categories will fit the data equally well. This is evident in the model fit statistics in Table 4. Thus, to reiterate the point made earlier, one cannot use fit statistics to discriminate among just-identified models. Rather, some other criterion must be employed. The criterion applied in this paper is that the constrained vector must be estimable.

[Table 4 about here]

The numerical estimates in Table 4 show that the constraint imposed in the CGLIM 3 model (the effect of the 1865 cohort constrained to equal that of the 1870 cohort) is statistically valid (i.e., within sampling variability) for these data – the estimated effect for the 1865-69 cohort in the IE column is .939 and that for the 1870-74 cohort is .937 with the S.E. of the former being .060 and latter being .031. Thus, the difference of these two effect coefficient estimates, .002, is well within sampling error and therefore effectively zero. In other words, this equality constraint produces an estimable function in a statistical sense (to be made precise below). The consequence is that the coefficient vectors for these two models are statistically identical up to a centering or normalizing transformation. This equivalence is demonstrated numerically in the last column of Table 4, which gives the numerical values of the corresponding re-centered/re-normalized CGLIM 3 effect coefficients, that is, the CGLIM 3 age, period, and cohort coefficients transformed by subtracting their respective group means so that the transformed coefficients sum to zero.¹⁵ The resulting re-centered CGLIM 3 and IE effect coefficients generally agree up to two or three digits. Again, this equivalence is due to the fact that the equality constraint imposed in the CGLIM 3 model is statistically valid for this dataset. The equivalence is illustrated in Figure 6, which shows graphs of the values of the age, period, and cohort effect coefficients for the CGLIM 3 and IE models. The patterns of the respective groups of effects are virtually identical.

¹⁵ CGLIM coefficients can be transformed as follows to the normalization that age, period and cohort effects each sum to zero. Under the original CGLIM normalization, let a_i be the estimated age effect for age i , let p_j be the estimated period effect for period j , and let c_k be the estimated cohort effect for cohort k . Let d be the estimated intercept. Transforming these estimates to a different normalization means that we want to find new coefficients a_i' , p_j' , c_k' and d' such that (1) the predicted values for each data point do not change, and (2) the age, period and cohort effects each sum to zero. The solution is to subtract the mean of the original age effects from each a_i to obtain a new age effect a_i' – this guarantees that the new age effects sum to zero – but then add the mean of the original age effects to the intercept, which guarantees that the predicted values do not change. The same process can be used to transform the period and cohort effects.

[Figure 6 about here]

Other possible identifying assumptions do not fare so well. To illustrate this, Table 4 reports comparable effect coefficient estimates for two alternative CGLIM models: CGLIM 1 which identifies the model by constraining the effect coefficient for the ages 25-29 group (A2) to be the same as that of the ages 20-24 group (A1), and CGLIM 2, which achieves identification by constraining the effect coefficient for the 1965-69 time period (P2) to equal that for the 1960-64 period (P1). Figure 7 shows graphs of the respective sets of effect coefficients for these two CGLIM models, and, for comparison, shows the corresponding effect coefficients of CGLIM 3, which imposes the constraint $C1 = C2$, as reported above. For all three sets of effect coefficients, it can be seen that those produced by the CGLIM 1 and CGLIM 2 models bracket those produced by the CGLIM 3 model – that is, the two alternative CGLIM models yield effect coefficients that diverge substantially from those given by the CGLIM 3 model. The divergences are substantial for the age and cohort effects and dramatic for the period effects. The reason for this behavior is that the equality constraints used to produce the CGLIM 1 and CGLIM 2 models do not produce statistically estimable functions and corresponding coefficient estimates.

To proceed more systematically, a procedure is needed for assessing whether two estimated coefficient vectors are within sampling error of being equal. There are several ways to accomplish this. For instance, centered CGLIM effect coefficients could be compared *element by element* with the corresponding estimated IE effect coefficients in the first column to determine whether the former are within, say, two standard errors of the latter. Using the standard errors of the IE coefficients, it can be seen, that all of the centered effect coefficients for the CGLIM 3 model are within two standard errors of the IE coefficients. By comparison, this is not the case for the centered effect coefficients of the CGLIM 1 and CGLIM 2 models. This confirms the

inference above from visual inspection of Figures 6 and 7, namely, that the CGLIM 3 model produces estimates of age, period, and cohort effects that are quite consistent with those estimated by the IE and thus is a statistically estimable function, whereas models CGLIM 1 and CGLIM 2 are not.

[Figure 7 about here]

An alternative procedure is to define a test statistic based on the *entire vector of coefficients*. To do so, we state the null hypothesis:

$$H_0 : E(\hat{b}^T B_0) = (B + sB_0)^T B_0 = s = 0 . \quad (15)$$

In words, this null hypothesis is that the expected value of the product of the estimated and re-normalized CGLIM vector \hat{b} and the eigenvector B_0 that is fixed by the design matrix is zero.

Due to the orthogonality of the vectors B and B_0 , this, as equation (15) indicates, is equivalent to the hypothesis that the expected value of the scalar s is equal to zero. Using the geometric projection illustrated in Figure 1, this test is equivalent to testing whether the estimated parameter vector under a given set of constraints (b_1 or b_2 or b_3) lies significantly far away from the estimable function, b_0 , so that one can infer that its horizontal projection results from a nonzero s .

To specify a test for this null hypothesis, we build upon a well-known asymptotic distribution property of the maximum likelihood estimator (MLE) used to estimate \hat{b} : under broad regularity conditions, as sample size or the number of time periods of data increases, the MLE of \hat{b} is consistent and asymptotically normally distributed (McCulloch and Searle 2001:306). This property facilitates the definition of an asymptotic t-test for the null hypothesis (15) as:

$$t = \frac{s - 0}{se(s)} = \frac{s}{se(s)}, \quad (16)$$

where $se(s)$ denotes the estimated (asymptotic) standard error of the scalar s .¹⁶ To obtain this test statistic, note that the numerator s can be computed, as indicated in equation (15), by calculating the product of the vectors \hat{b} and B_0 . Then the denominator can be computed by transforming the asymptotic variance-covariance matrix $\hat{\Sigma}^{-1}$ (i.e., the inverse of the Fisher information matrix) that is obtained in the process of estimating b by maximum likelihood to obtain \hat{b} . Because the vector B_0 is orthogonal to the design matrix (see equation (9) above), the variance of the scalar s can be computed by imposition of the usual quadratic form transformation applied to obtain the variance of the restricted maximum likelihood estimator from the MLE, namely $B_0^T \hat{\Sigma}^{-1} B_0$.¹⁷ Taking the square root of this transformation yields an estimate of the standard deviation of $s = \hat{b}^T B_0$, which, when divided by the degrees of freedom of the model, $df = ap - (1 + (a-1) + (p-1) + (a+p-2))$, produces an estimate of the standard error of s in the denominator of equation (16).

Applied to the three alternative CGLIM models for which the re-normalized coefficients are given in Table 4, we obtain the following t-ratios: 0.742 ($p = 0.458$) for CGLIM 1, -4.822 (p

¹⁶ This asymptotic t-test is based on asymptotic normality properties of the maximum likelihood estimator, which assume large numbers of degrees of freedom, i.e., large APC matrices. A systematic study of the statistical power of the test for small- to moderate-sized matrices is beyond the scope of the present paper. Even without this, the test can provide useful information for assessing the plausibility of constraints based on theory or side information from other studies and for comparing alternative constraints, as illustrated here.

¹⁷ Formally, the asymptotic normality property of the MLE yields $\hat{b}^{asy} \sim N(b, \Sigma^{-1})$, i.e., the sampling distribution of the estimated coefficient vector of a model that is identified by imposition of a theoretically motivated equality constraint on two or more coefficients is asymptotically multivariate normal with a mean (expected) parameter vector b and a variance-covariance matrix Σ^{-1} , where Σ^{-1} is estimated by $\hat{\Sigma}^{-1}$, the inverse of the Fisher information matrix. By properties of the MLE (see, e.g., McCulloch and Searle, 2001: 309), a linear transformation of \hat{b} using B_0 then yields $\hat{b}^T B_0 \stackrel{asy}{\sim} N(0, B_0^T \Sigma^{-1} B_0)$, where $\hat{b}^T B_0 = s$. From this, the asymptotic standard error of s can be computed as the square root of $B_0^T \hat{\Sigma}^{-1} B_0$ divided by degrees of freedom as indicated in the text.

< 0.001) for CGLIM 2, and -0.076 ($p = 0.939$) for CGLIM 3. The results show that the CGLIM 1 and CGLIM 3 models estimates are not statistically different from those of the IE, but the CGLIM 2 model estimates are. Because the degrees of freedom of the APC models are small, it may be appropriate to use a p value larger than the conventional value of .05 for assessing the t -ratios. In this sense, the CGLIM 1 model with a p value of 0.5 deviates from the IE model more than the CGLIM 3 model with a p value of 0.9, although both models clearly are acceptable by the conventional .05 criterion. From a substantive point of view, the equality constraint used by the CGLIM 1 model – the effects are equal for ages 20-24 and 25-29 – is consistent with the demographic evidence from life table studies that the mortality risk is generally low and roughly the same in the 20s. Reiterating the point made in Remark 4 above, the results of applying the asymptotic t -test to the CGLIM1, CGLIM2, and CGLIM3 models also show that estimators that incorporate significant components of the B_0 vector (and thus the design matrix) may give misleading results concerning the patterns of change of the estimated effect coefficients across the age, period, and cohort categories. This is illustrated graphically in Figure 7. It shows that the patterns of changes of the coefficients for the CGLIM1 and CGLIM3 models, for which the foregoing asymptotic t -test are not significant, are quite close to those of the IE and thus lead to similar conclusions. On the other hand, those of the CGLIM2 model, which has an asymptotic t -ratio that indicates a statistically significant departure from the IE coefficients, show patterns of change that depart substantially from those of the IE.

To show that not all constraints based on age effect coefficients produce acceptable t -ratios, consider an alternative constraint on the coefficients of the two oldest age groups, a constraint that might be motivated by recent studies of the deceleration of mortality at the oldest-old ages (see, e.g., Vaupel [1997]). To see how this different constraint compares to the current

CGLIM 1 model, we estimated CGLIM 1B model constraining the last two age groups to be equal in coefficients ($A_{15} = A_{16}$). The estimates result in a large t-ratio of 23.9 ($p < .001$), suggesting that this equality constraint does not produce an estimable function.

In brief, the foregoing test for statistical estimability leads to new avenues for using and interpreting the IE as applied to a particular dataset. On the one hand, an analyst can apply the IE in an *exploratory data analysis* manner, in which the objective is to ascertain good estimates of the patterns of age, period, and cohort effects in a table or set of tables of demographic rates.¹⁸ In such an exercise, the analyst does not approach the data with strong prior notions about particular patterns of effects that should be evident in the data, but rather seeks to let the patterns emerge from application of the IE, taking advantage of the fact that the IE is an estimable function and thus has desirable statistical properties.

On the other hand, using the definition of statistical estimability and the test described above, an analyst can use the IE in a more *confirmatory data analysis* manner, in which a vector of effect coefficients estimated from application of the IE is used as a benchmark to assess whether a corresponding vector of coefficients estimated from the imposition of one or more theoretically or substantively motivated constraints to achieve model identification is acceptable. In this type of exercise, an analyst approaches a table or tables of demographic rates with a definite hypothesis or set of hypotheses about the underlying age, period, and cohort effects that generated the data. The analyst can use the vector of effect coefficients estimated by the IE to assess the empirical plausibility of the hypotheses. In this way, the definition of statistical estimability and the test described above directly address a criticism that often has been lodged against general-purpose methods of APC analysis, namely, that they provide no avenue for

¹⁸ The distinction between “exploratory” and “confirmatory” data analysis dates back in statistics at least to the classic work of Mosteller and Tukey (1977).

testing specific substantive hypotheses and thus are mere devices of algebraic convenience that may be misleading.

Computational Tools

Other interested users may desire computer programs for conveniently applying the IE to their data. To date, programs for estimating the IE have been written as add-on files to two commercially available software packages, S-Plus (Venables and Ripley 2000) and Stata (Rabe-Hesketh and Everitt 2004). These two programs can be accessed as follows.

The S-Plus program can be obtained by writing to Wenjiang Fu (fuw@epi.msu.edu) and requesting a copy.

An ado-file to calculate the IE in Stata may be obtained by typing “`ssc install apc`” on the Stata command line on any computer connected to the Internet, or downloaded from the Statistical Software Components archive at <http://ideas.repec.org/s/boc/bocode.html>. The program uses much the same syntax as Stata’s `glm` command for generalized linear models. For example, a user whose dataset contains a dependent variable “y”, an exposure variable “x”, an age variable “a” and a period variable “t” can fit a Poisson model with age, period and cohort effects by typing:

```
apc_ie y, exposure(x) family(poisson) link(log) age(a) period(t)
```

Stata will then display coefficient estimates, standard errors, confidence intervals, the log likelihood value, and a variety of other statistics. The program is documented more fully in a help file that can be read by typing “`help apc_ie`” in Stata after installation. The package also includes a command “`apc_cglm`” for calculating CGLIM estimators that impose equality constraints on pairs of coefficients; for details, type “`help apc_cglm`” in Stata after installation.

CONCLUSIONS

The problem of obtaining reliable estimates of the patterns of changes across the age groups, time periods, and cohorts has long provided an intriguing challenge in many contexts in the social sciences, demography and epidemiology. This paper has presented a number of properties and studied the performance of the Intrinsic Estimator for the APC accounting model in the context of age-by-time period tables of rates to address this challenge. As noted at the beginning of this paper, Glenn (2005: 20) has stated several strong criteria for judging the acceptability and utility of a general-purpose method of APC analysis. The intrinsic estimator appears to satisfy these criteria. That is, as shown herein, the IE has passed both empirical and simulation tests of validity and can be used to test theoretically motivated hypotheses and to incorporate and test side information from other studies. The IE therefore may provide a useful tool for the accumulation of scientific knowledge about the distinct effects of age, period, and cohort effects in social research.

Indeed, since the APC underidentification problem is an instance of a larger family of structural underidentification problems, the potential range of application of the IE may be even larger. *Structural underidentification* problems occur when a conceptualization of effects of structural arrangements leads to an exact linear dependency among the effects. An example is the classical problem in mobility analysis of distinguishing the effects of socioeconomic mobility or distance moved on an outcome variable from the effects of origin and destination statuses (see, e.g, Duncan, 1966). Another example pertains to the estimation of the effects of years of labor force experience separately from current age and age at labor force entry. These and similar problems of structural underidentification occur frequently in sociology and related disciplines.

The intrinsic estimator and/or algebraic approaches similar to the IE may prove useful in applications to such problems.

Is the intrinsic estimator then a “complete solution” to the structural identification problem in APC and similar models? No. Structural identification problems are just that — points of underidentification of parameters due to the very nature of the underlying models. There is not now, and can never be, a complete resolution of such problems. But there can be variations among approaches to structural identification problems with respect to statistical properties. Because of its desirable properties as a statistical estimator, including its ability to produce good estimates of the underlying patterns of change in the age, period, and cohort coefficients in the accounting model with a small number of time periods of data, as demonstrated above, the IE adds a potentially useful method to the toolkit available for these analyses. Does this mean that researchers should naively apply this method to tables of rates and expect to obtain meaningful results? No. Every statistical model has its limits and will break down under some conditions. APC analysis is well known to be treacherous for reasons articulated by Glenn (2005) and should, in all cases, be approached with great caution and awareness of its many pitfalls. Replications, in particular, are called for when additional data become available.

Last but not the least, the limitations of the IE in theoretical and empirical studies are related to the limitations of the APC accounting model (Smith 2004). This model is based on the assumption of additivity of age, period, and cohort effects that not only incurs the identification problem but also may be a poor approximation of how social change occurs. More new models and methods are needed for testing other theories of social change. The recent development of hierarchical APC models is useful for studying the *contextual effects* of cohort membership and

period events on a wide range of social processes (Yang 2006; Yang and Land 2006; Yang forthcoming). Other conceptually appealing models, such as the *cohort inversion* model and the *continuously-accumulating/evolving cohort effects* model (Hobcraft et al. 1982), should also be explored and more fully developed in future research.

References

- Agresti, A. 1996. *An Introduction to Categorical Data Analysis*. Wiley Series in Probability and Statistics. New York: John Wiley & Sons, Inc.
- Alwin, D.F. 1991. "Family of Origin and Cohort Differences in Verbal Ability." *American Sociological Review* 56:625-38.
- _____. and R.J. McCammon. 1999. "Aging Versus Cohort Interpretations of Intercohort Differences in GSS Vocabulary Scores." *American Sociological Review* 64:272-86.
- Clayton, D. and E. Schifflers. 1987. "Models for Temporal Variation in Cancer Rates. II: Age-period-cohort Models". *Statistics in Medicine* 6: 469-81.
- Duncan, O.D. 1966. "Methodological Issues in the Analysis of Social Mobility." in *Social Structure and Mobility in Economic Development*, edited by N.J. Smelser and S. M. Lipset. Chicago: Aldine, pp.51-97.
- Fienberg, S.E. and W.M. Mason. 1978. "Identification and Estimation of Age-period-cohort Models in the Analysis of Discrete Archival Data." in *Sociological Methodology 1979* edited by K.F. Schuessler. San Francisco: Jossey-Bass, pp. 1-67.
- _____. 1985. "Specification and Implementation of Age, Period, and Cohort Models." Pp. 45-88 in *Cohort Analysis in Social Research*, edited by W.M. Mason and S.E. Fienberg. New York: Springer-Verlag.
- Fu, W. J. 2000. "Ridge Estimator in Singular Design with Application to Age-Period-Cohort Analysis of Disease Rates." *Communications in Statistics--Theory and Method* 29:263-278.

- Fu, W.J. 2007. "A Smoothing Cohort Model in Age-Period-Cohort Analysis with Applications to Homicide Arrest Rates and Lung Cancer Mortality Rates." *Sociological Methods and Research*: in press.
- Fu, W. J. and P. Hall. 2006. "Asymptotic properties of estimators in age-period-cohort analysis", *Statistics and Probability Letters* 76: 1925-1929.
- Fu, W.J. Hall, P. and Rohan, T. 2004. "Age-period-cohort analysis: structure of estimators, estimability, sensitivity and asymptotics", Technical report, Department of Epidemiology, Michigan State University.
- Glenn, N.D. 1976. "Cohort Analysts' Futile Quest: Statistical Attempts to Separate Age, Period, and Cohort Effects." *American Sociological Review* 41:900-904.
- _____. 1994. "Television Watching, Newspaper Reading, and Cohort Differences in Verbal Ability." *Sociology of Education* 67:216-30.
- _____. 1999. "Further Discussion of the Evidence for An Intercohort Decline in Education-Adjusted Vocabulary." *American Sociological Review* 64:267-71.
- _____. 2005. *Cohort Analysis*. Thousand Oaks, CA: Sage.
- Goldstein, H. 1950. *Classical Mechanics*. Reading, MA: Addison-Wesley.
- Hobcraft, J., J. Menken, and S. Preston. 1982. "Age, Period, And Cohort Effects in Demography: A Review." *Population Index* 48:4-43.
- Holford, T.R. 1983. "The Estimation of Age, Period and Cohort Effects for Vital Rates." *Biometrics* 39:311-324.
- _____. 1991. "Understanding the Effects of Age, Period, and Cohort on Incidence and Mortality Rates." *Annual Review of Public Health* 12:425-57.

- _____. 1992. "Analysing the Temporal Effects of Age, Period, and Cohort." *Statistical Methods in Medical Research* 1:317-37.
- 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.
- Kupper, L.L., J.M. Janis, A. Karmous, and B.G. Greenberg. 1985. "Statistical Age-Period-Cohort Analysis: A Review and Critique." *Journal of Chronic Disease* 38:811-830.
- Marion, J.B., and S.T. Thornton. 1995. *Classical Dynamics of Particles and Systems*, fourth edition. Fort Worth, TX: Harcourt Brace College Publishers.
- Mason, K.O., W.H. Mason, H.H. Winsborough, and K. Poole. 1973. "Some Methodological Issues in Cohort Analysis of Archival Data." *American Sociological Review* 38: 242-58.
- Mason, W.H. and H.L. Smith. 1985. "Age-Period-Cohort Analysis and the Study of Deaths from Pulmonary Tuberculosis." Pp. 151-228 in *Cohort Analysis in Social Research*, edited by W.M. Mason and S.E. Fienberg. New York: Springer-Verlag.
- Mason, W.H. and N.H. Wolfinger. 2002. "Cohort Analysis." Pp. 2189-2194 in *International Encyclopedia of the Social & Behavioral Sciences*. New York: Elsevier.
- McCullagh P. and J.A. Nelder. 1989. *Generalized Linear Models*, 2nd edition. New York: Chapman and Hall.
- McCulloch, C. E. and S. R. Searle. 2001. *Generalized, Linear, and Mixed Models*. New York: Wiley-Interscience.
- Mosteller, F. and J. W. Tukey. 1977. *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.

- O'Brien, R.M. 2000. "Age Period Cohort Characteristic Models." *Social Science Research* 29:123-139.
- Osmond, C. and M.J. Gardner. 1982. "Age, Period, and Cohort Models Applied to Cancer Mortality Rates." *Statistics in Medicine* 1:245-59.
- Pullum, T. W. 1977. "Parametrizing Age, Period, and Cohort Effects, With Applications to U.S. Delinquency Rates, 1964-1973." *Sociological Methodology*:116-140.
- _____. 1980. "Separating Age, Period, and Cohort Effects in White U.S. Fertility, 1920-1970." *Social Sciences Research* 9:25-244.
- Rabe-Hesketh, S. and B. Everitt. 2004. *A Handbook of Statistical Analysis using Stata*. Third Edition. New York: Chapman & Hall.
- Robertson, C. and P. Boyle. 1998. "Age-period-cohort Analysis of Chronic Disease Rates I: Modeling Approach." *Statistics in Medicine* 17:1305-23.
- Robertson, C., S. Gandini, and P. Boyle. 1999. "Age-Period-Cohort Models: A Comparative Study of Available Methodologies." *Journal of Clinical Epidemiology* 52:569-583.
- Rodger, W. L. 1982a. "Estimable Functions of Age, Period, and Cohort Effects." *American Sociological Review* 47:774-787.
- _____. 1982b. "Reply to Comments by Smith, Mason, and Fienberg." *American Sociological Review* 47:793-796.
- Searle, S. R. 1971. *Linear Models*. New York: Wiley.
- Sen, A. and M. Srivastava. 1990. *Regression Analysis: Theory, Methods, and Applications*. New York: Springer-Verlag.
- Smith, H. L. 2004. "Response: Cohort Analysis Redux." Pp. 111-119 in *Sociological Methodology*. Boston, MA: Blackwell

- Smith, H. L., W. M. Mason, and S. E. Fienberg. 1982. "More Chimeras of the Age-Period-Cohort Accounting Framework: Comment on Rodgers." *American Sociological Review* 47:787-793.
- Strehler, B. L. and A. S. Mildvan. 1960. "General Theory of Mortality and Aging." *Science* 132:14-21.
- Tarone, R. and K.C. Chu. 1992. "Implications of Birth Cohort Patterns in Interpreting Trends in Breast Cancer Rates." *Journal of National Cancer Institute* 1992:1402-10.
- _____. 2000. "Age-period-cohort Analysis of Breast-, Ovarian-, Endometrial- and Cervical-cancer Mortality Rates for Caucasian Women in the USA." *Journal of Epidemiology and Biostatistics* 5:221-31.
- Vaupel, J.W. 1997. "Trajectories of Mortality in Advanced Ages." in *Between Zeus and the Salmon*, edited by K.W. Wachter and C. E. Finch. Washington, DC: National Academy Press, pp. 17-37.
- Venables, W.N. and B.D. Ripley. 2000. *Modern Applied Statistics with S-PLUS*. 3rd Edition. New York: Springer.
- Wilmoth, J. R. 1990. "Variation in Vital Rates by Age, Period, and Cohort." *Sociological Methodology* 20: 295-335.
- Wilson, J.A. and W.R. Gove. 1999. "The Intercohort Decline in Verbal Ability: Does it Exist?" and reply to Glenn and Alwin & McCammon. *American Sociological Review*.
- Yang, Y. Forthcoming. "Social Inequalities in Happiness in the U.S. 1972-2004: An Age-Period-Cohort Analysis." *American Sociological Review* (April 2008).
- Yang, Y. 2006. "Bayesian Inference for Hierarchical Age-Period-Cohort Models of Repeated Cross-Section Data." *Sociological Methodology* 36: 39-74.

Yang, Y. and K.C. Land. 2006. "A Mixed Models Approach to the Age-Period-Cohort Analysis Of Repeated Cross-Section Surveys, With an Application to Data on Trends in Verbal Test Scores." *Sociological Methodology* 36: 75-97.

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.

Table 1. Deaths per 100,000: United States Females 1960-1999

Age/Year	1960-64	1965-69	1970-74	1975-79	1980-84	1985-89	1990-94	1995-99
0	518	444	371	309	259	227	192	160
5	39	36	34	27	23	20	18	16
10	31	30	29	25	22	20	19	18
15	53	58	60	54	48	48	46	43
20	70	73	72	64	57	54	51	48
25	88	86	82	70	64	64	64	59
30	123	123	113	90	80	83	86	82
35	183	185	172	135	115	113	119	119
40	275	282	267	221	186	170	166	170
45	419	422	405	344	304	277	261	252
50	631	625	591	524	483	454	419	395
55	906	905	868	770	741	711	671	631
60	1,414	1,337	1,282	1,156	1,126	1,111	1,053	1,014
65	2,228	2,120	1,967	1,740	1,686	1,656	1,590	1,543
70	3,608	3,400	3,112	2,698	2,606	2,563	2,445	2,428
75	5,886	5,593	5,143	4,380	4,087	3,996	3,796	3,778
80	9,678	9,162	8,395	7,337	6,946	6,641	6,302	6,351
85	15,665	14,993	13,757	12,016	11,487	11,223	10,444	10,934
90	23,635	22,839	21,261	18,991	18,253	18,361	17,395	17,666
95+	33,478	32,491	31,170	28,986	28,613	29,536	28,973	29,796

Table 2. Verbal Test Correct Rates (Exposure): GSS 1976 – 2000

Age/Year	1976-80	1981-85	1986-90	1991-95	1996-2000
20-24	0.554 (8203)	0.531 (8489)	0.526 (5553)	0.548 (4950)	0.555 (5045)
25-29	0.596 (9671)	0.554 (10653)	0.577 (6946)	0.589 (6480)	0.575 (6780)
30-34	0.630 (8643)	0.634 (8718)	0.588 (7096)	0.613 (7745)	0.596 (7740)
35-39	0.625 (6776)	0.643 (7946)	0.632 (6799)	0.634 (7960)	0.622 (8400)
40-44	0.622 (5619)	0.633 (5996)	0.642 (5548)	0.652 (6940)	0.629 (8355)
45-49	0.609 (4595)	0.603 (5112)	0.617 (4734)	0.675 (5915)	0.653 (6775)
50-54	0.617 (5454)	0.597 (4560)	0.594 (3314)	0.617 (4770)	0.661 (5885)
55-59	0.630 (5439)	0.594 (5335)	0.589 (3207)	0.623 (3590)	0.676 (4225)
60-64	0.625 (4915)	0.596 (4830)	0.561 (3319)	0.610 (3430)	0.604 (3700)
65-69	0.605 (4010)	0.576 (498)	0.575 (3633)	0.648 (3000)	0.596 (2985)
70-74	0.523 (3420)	0.571 (3333)	0.610 (2789)	0.587 (3305)	0.607 (3540)
75+	0.549 (4394)	0.532 (4264)	0.538 (3584)	0.573 (4890)	0.563 (4970)

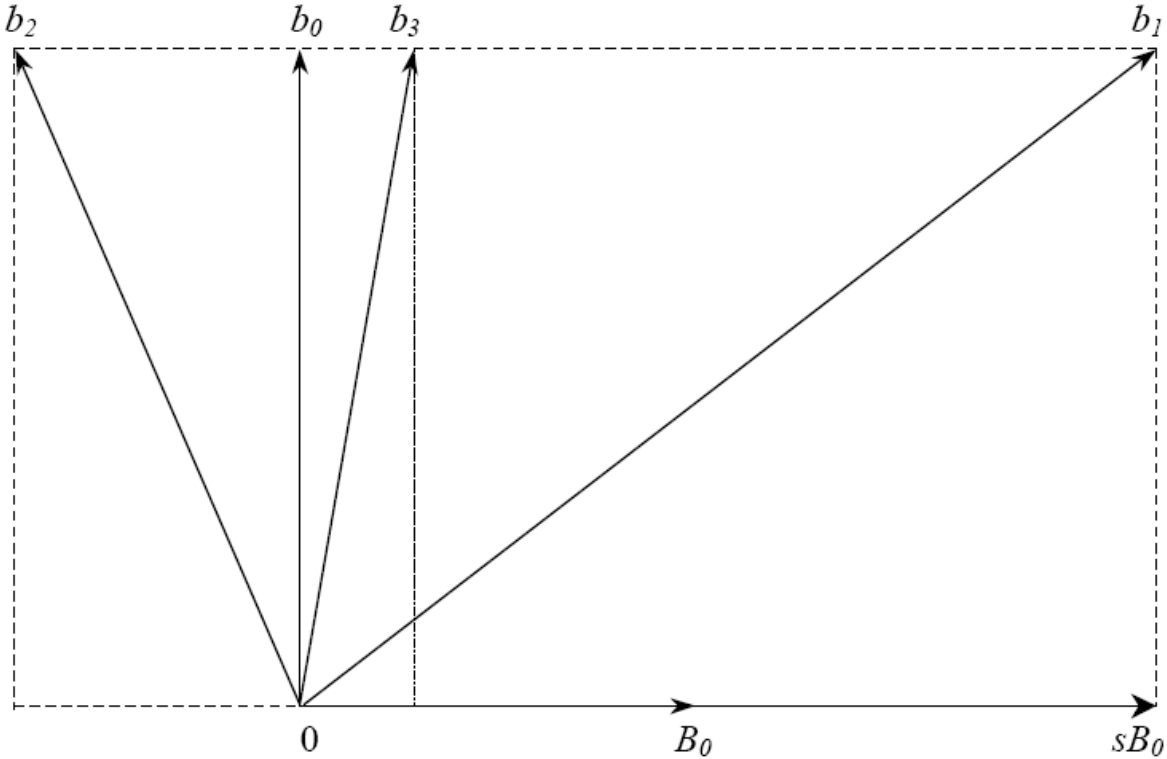
Table 3 Simulation Results of the IE and CGLIM Estimators: Age Effects (n = 1,000)

Age	True	IE		CGLIM (a1=a2)		CGLIM (p1=p2)		CGLIM (c1=c2)	
		p = 5	p = 50	p = 5	p = 50	p = 5	p = 50	p = 5	p = 50
1 mean	0.933	0.924	0.942	-1.903	-1.876	1.214	1.078	1.714	1.211
sd		0.276	0.060	1.265	0.348	1.366	1.308	5.294	2.965
MSE		0.076	0.004	9.646	8.011	1.942	1.730	28.609	8.862
2 mean	0.233	0.217	0.238	-1.903	-1.876	0.435	0.340	0.810	0.440
sd		0.280	0.077	1.265	0.348	1.041	0.983	3.967	2.223
MSE		0.079	0.006	6.164	4.569	1.123	0.977	16.053	4.980
3 mean	-0.267	-0.303	-0.267	-1.717	-1.676	-0.158	-0.199	0.092	-0.132
sd		0.343	0.094	0.798	0.209	0.742	0.653	2.665	1.486
MSE		0.119	0.009	2.740	2.028	0.562	0.431	7.221	2.223
4 mean	-0.567	-0.629	-0.574	-1.335	-1.279	-0.556	-0.540	-0.431	-0.507
sd		0.401	0.112	0.536	0.146	0.526	0.351	1.384	0.751
MSE		0.165	0.013	0.878	0.528	0.276	0.124	1.932	0.567
5 mean	-0.667	-0.721	-0.674	-0.721	-0.674	-0.721	-0.674	-0.721	-0.674
sd		0.428	0.121	0.428	0.121	0.428	0.121	0.428	0.121
MSE		0.186	0.015	0.186	0.015	0.186	0.015	0.186	0.015
6 mean	-0.567	-0.599	-0.576	0.108	0.128	-0.671	-0.610	-0.796	-0.644
sd		0.397	0.106	0.518	0.141	0.514	0.343	1.387	0.749
MSE		0.158	0.011	0.723	0.503	0.275	0.119	1.974	0.566
7 mean	-0.267	-0.240	-0.266	1.173	1.143	-0.386	-0.334	-0.636	-0.401
sd		0.337	0.097	0.813	0.218	0.758	0.661	2.660	1.484
MSE		0.114	0.009	2.733	2.034	0.588	0.440	7.203	2.217
8 mean	0.233	0.304	0.237	2.424	2.350	0.086	0.135	-0.289	0.035
sd		0.288	0.077	1.114	0.295	1.076	0.981	3.997	2.226
MSE		0.088	0.006	6.039	4.569	1.177	0.971	16.233	4.989
9 mean	0.933	1.048	0.940	3.875	3.758	0.757	0.804	0.257	0.671
sd		0.268	0.060	1.437	0.390	1.366	1.304	5.272	2.962
MSE		0.085	0.004	10.716	8.132	1.894	1.716	28.221	8.835

Table 4. IE and CGLIM Estimates, U.S. Female Adult Mortality, 1960-1999

	Intrinsic Estimator		CGLIM 3 (C1=C2)		Centered CGLIM 1 (A1=A2)	Centered CGLIM 1B (A15=A16)	Centered CGLIM 2 (P1=P2)	Centered CGLIM 3 (C1=C2)
	Effect	S.E.	Effect	S.E.	Effect	Effect	Effect	Effect
<i>Intercept</i>	-4.571	0.007	-5.803	1.058	-4.571	-4.571	-4.571	-4.571
<i>Age</i>								
20-24	-2.032	0.032	0.000	0.000	-1.749	0.807	-2.410	-2.044
25-29	-1.995	0.029	0.039	0.087	-1.749	0.466	-2.322	-2.005
30-34	-1.819	0.026	0.217	0.156	-1.611	0.263	-2.096	-1.828
35-39	-1.555	0.023	0.482	0.228	-1.385	0.148	-1.782	-1.563
40-44	-1.245	0.020	0.794	0.301	-1.112	0.080	-1.421	-1.250
45-49	-0.915	0.017	1.125	0.375	-0.821	0.031	-1.041	-0.919
50-54	-0.596	0.015	1.446	0.449	-0.540	-0.028	-0.672	-0.599
55-59	-0.295	0.013	1.749	0.524	-0.276	-0.105	-0.320	-0.296
60-64	0.026	0.011	2.071	0.598	0.007	-0.163	0.051	0.027
65-69	0.349	0.010	2.395	0.672	0.292	-0.219	0.424	0.351
70-74	0.701	0.009	2.750	0.747	0.607	-0.245	0.827	0.705
75-79	1.072	0.008	3.122	0.821	0.940	-0.253	1.249	1.078
80-84	1.484	0.008	3.535	0.896	1.313	-0.220	1.710	1.491
85-89	1.896	0.009	3.949	0.970	1.688	-0.187	2.173	1.904
90-94	2.273	0.011	4.328	1.045	2.027	-0.188	2.601	2.284
95+	2.652	0.014	4.708	1.118	2.368	-0.188	3.030	2.664
<i>Period</i>								
1960-64	-0.115	0.009	0.000	0.000	-0.247	-1.440	0.062	-0.109
1965-69	-0.064	0.008	0.049	0.075	-0.159	-1.011	0.062	-0.060
1970-74	-0.042	0.007	0.069	0.149	-0.099	-0.610	0.033	-0.040
1975-79	-0.081	0.007	0.029	0.224	-0.100	-0.270	-0.056	-0.080
1980-84	-0.034	0.007	0.074	0.298	-0.015	0.156	-0.059	-0.034
1985-89	0.044	0.007	0.150	0.373	0.101	0.612	-0.032	0.041
1990-94	0.094	0.007	0.199	0.447	0.189	1.041	-0.032	0.090
1995-99	0.197	0.008	0.301	0.522	0.330	1.522	0.021	0.192
<i>Cohort</i>								
1865	0.939	0.060	0.000	0.000	1.355	5.104	0.385	0.921
1870	0.937	0.031	0.000	0.000	1.315	4.723	0.433	0.921
1875	0.900	0.020	-0.035	0.094	1.241	4.308	0.447	0.886
1880	0.834	0.016	-0.100	0.165	1.137	3.863	0.431	0.821
1885	0.748	0.014	-0.184	0.239	1.013	3.398	0.395	0.737
1890	0.653	0.012	-0.277	0.313	0.880	2.925	0.351	0.644
1895	0.556	0.011	-0.373	0.387	0.746	2.450	0.305	0.548
1900	0.449	0.011	-0.478	0.461	0.601	1.964	0.248	0.443
1905	0.331	0.011	-0.595	0.536	0.445	1.467	0.180	0.327
1910	0.243	0.011	-0.682	0.610	0.318	1.000	0.142	0.240
1915	0.148	0.012	-0.775	0.685	0.186	0.527	0.098	0.146
1920	0.060	0.013	-0.861	0.759	0.060	0.060	0.060	0.060
1925	-0.018	0.015	-0.938	0.834	-0.056	-0.397	0.032	-0.016
1930	-0.124	0.017	-1.042	0.908	-0.200	-0.882	-0.024	-0.121
1935	-0.249	0.019	-1.165	0.983	-0.362	-1.385	-0.098	-0.244
1940	-0.415	0.022	-1.330	1.057	-0.567	-1.930	-0.214	-0.409
1945	-0.567	0.024	-1.480	1.132	-0.756	-2.460	-0.315	-0.559
1950	-0.678	0.026	-1.589	1.207	-0.905	-2.950	-0.376	-0.668
1955	-0.732	0.030	-1.642	1.281	-0.997	-3.383	-0.380	-0.721
1960	-0.805	0.035	-1.713	1.356	-1.108	-3.834	-0.402	-0.792
1965	-0.918	0.044	-1.825	1.431	-1.258	-4.325	-0.464	-0.903
1970	-1.063	0.060	-1.969	1.506	-1.442	-4.850	-0.560	-1.048
1975	-1.230	0.103	-2.134	1.583	-1.646	-5.395	-0.676	-1.213
<i>t ratio</i>					0.742	23.946	-4.822	-0.076
<i>p-value</i>					0.458	0.000	0.000	0.939
<i>Deviance</i>		18903.2		18903.2	18903.2	18903.2	18903.2	18903.2
<i>DF</i>		84		84	84	84	84	84
<i>Overdispersion</i>		224.3		224.3	224.3	224.3	224.3	224.3

Figure 1. Geometric Projection of Parameters Vectors



Note: b_1 , b_2 , and b_3 are three sets of parameter vectors that result from the imposition of various constraints on elements of the parameter vector b of the APC regression model; sB_0 represents the arbitrary term in these parameter vectors; projection of b_1 , b_2 , and b_3 to the vertical axis yields the same estimable function b_0 . The vertical vector b_0 is orthogonal to and therefore independent of the component sB_0 , which represents effects of the design matrix. While b_0 is estimable, none of the others, $b_0 + sB_0$ with $s \neq 0$ such as b_1 , b_2 , or b_3 are. As shown later in the text, however, a statistical test can be defined and applied to ascertain whether or not the vector of estimated coefficients corresponding to constraints such as those used to define, say, b_3 are sufficiently close to those corresponding to b_0 , as estimated by the IE, that it can be concluded that the scalar $s = 0$, and, correspondingly, that the equality constraint that produced b_3 also produces an estimable function in a statistical sense.

Figure 2. Intrinsic Estimator Coefficient Estimates and 95% Confidence Intervals of the Age, Period, and Cohort Effects on Vocabulary Ability Using GSS Data, 1976 –2000

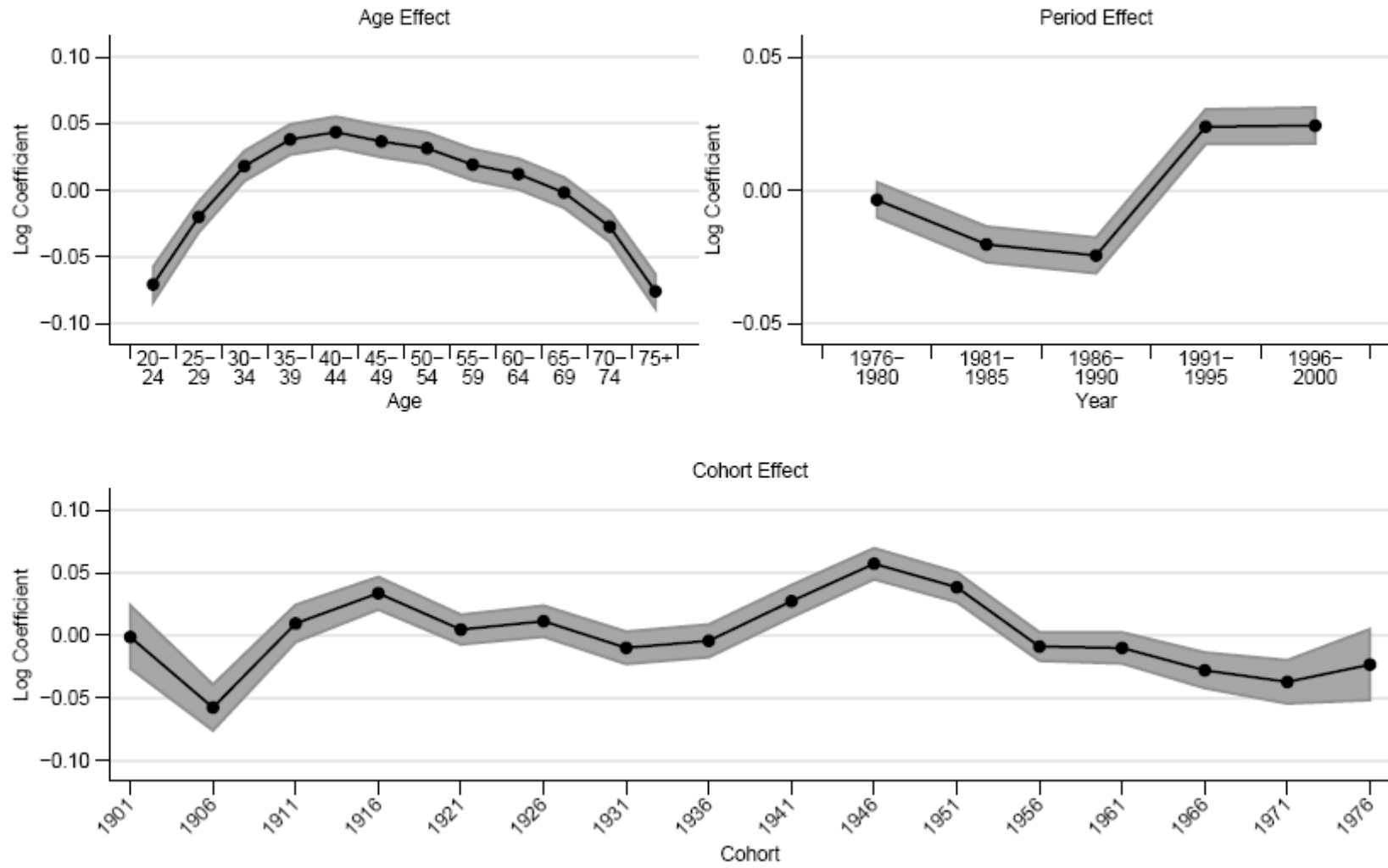
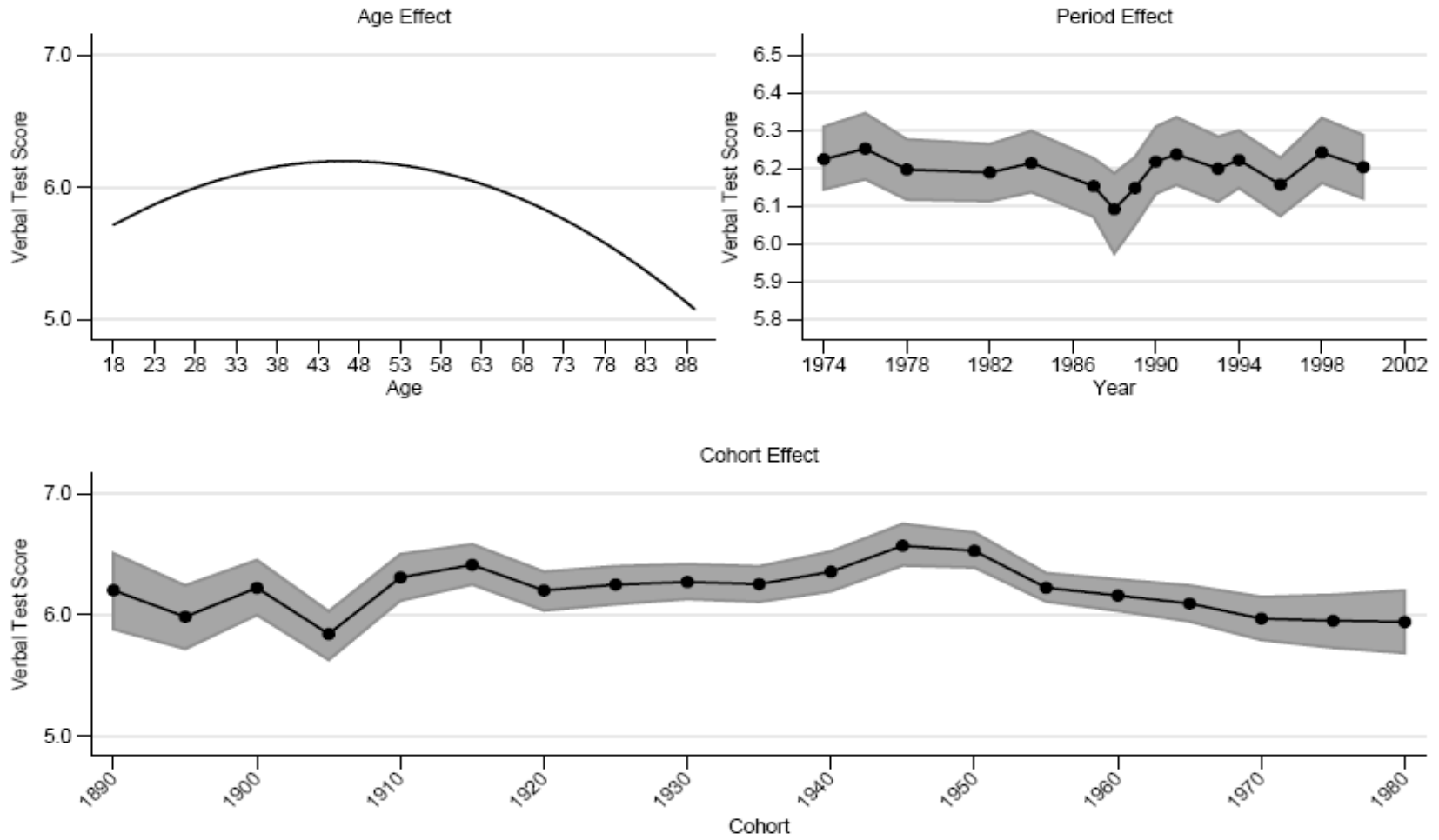


Figure 3. Bayes Estimates of Age, Period, and Cohort Effects on Verbal Scores



Source: Adapted from Yang (2006)

Figure 4. Simulation Results of the IE and CGLIM Estimators: Period and Cohort Effects (n = 1,000)

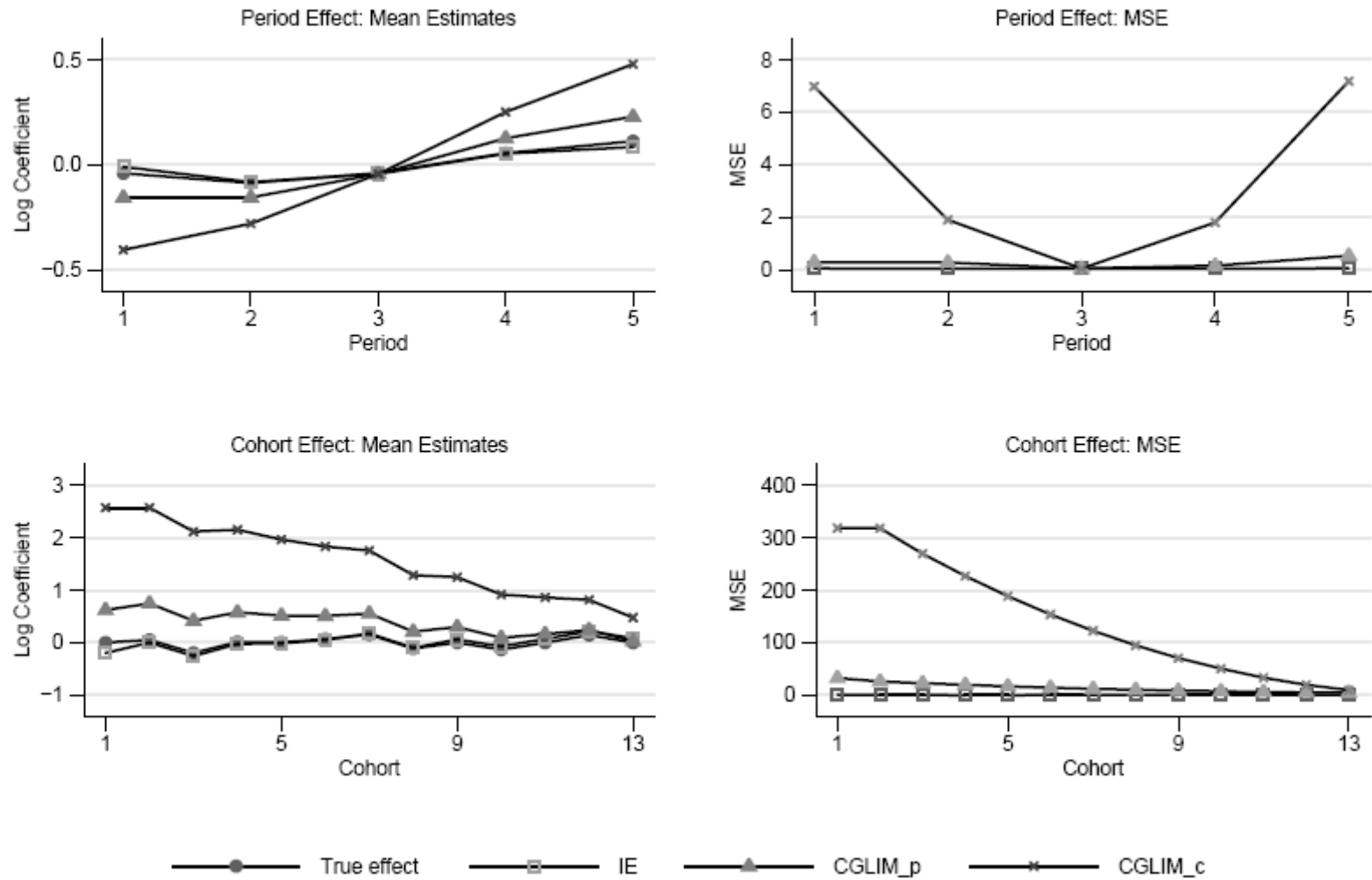
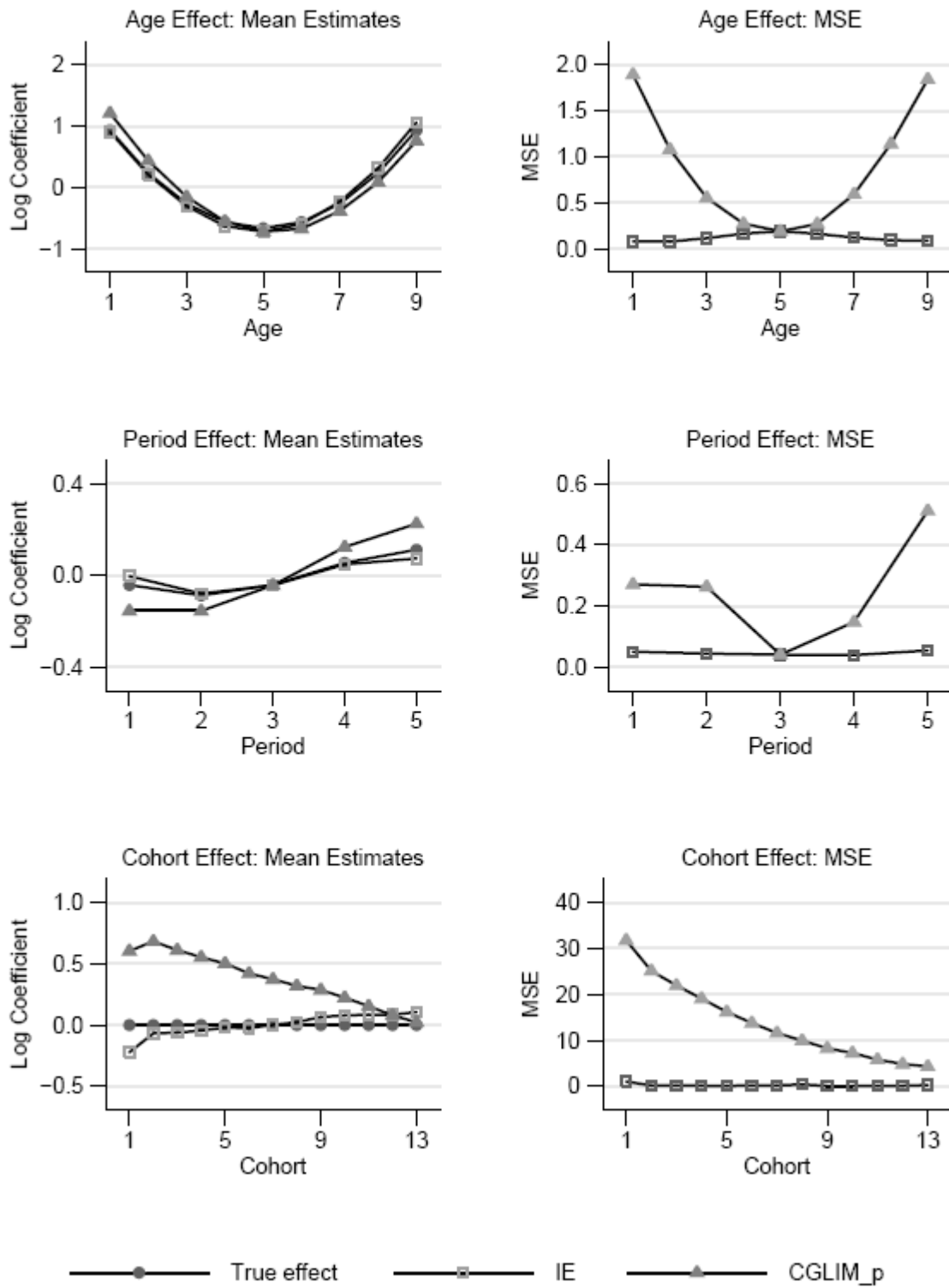


Figure 5. Simulation Results of the IE and CGLIM Estimators: True Cohort Effects = 0



**Figure 6. Comparison of Graphs of Age, Period, and Cohort Effect Coefficients.
CGLIM 3 and IE Models**

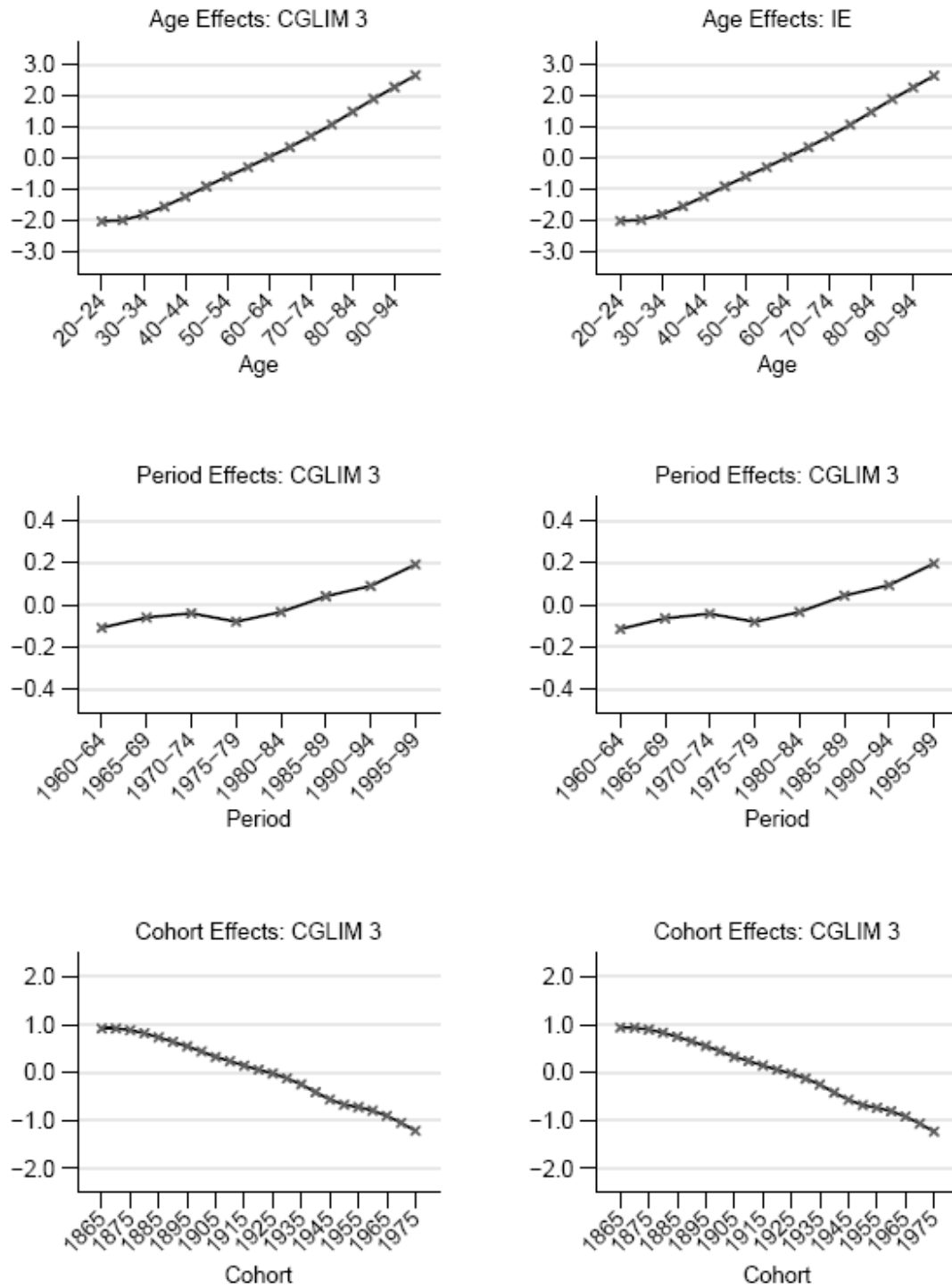


Figure 7. Age, Period, and Cohort Effects Estimated for Three CGLIM Models

