

STRUCTURAL PARAMETER ESTIMATION USING GENERALIZED ESTIMATING EQUATIONS FOR REGRESSION CREDIBILITY MODELS

BY

CHI HO LO, WING KAM FUNG AND ZHONG YI ZHU

ABSTRACT

A generalized estimating equations (GEE) approach is developed to estimate structural parameters of a regression credibility model with independent or moving average errors. A comprehensive account is given to illustrate how GEE estimators are worked out within an extended Hachemeister (1975) framework. Evidenced by results of simulation studies, the proposed GEE estimators appear to outperform those given by Hachemeister, and have led to a remarkable improvement in accuracy of the credibility estimators so constructed.

KEYWORDS

Credibility theory, Hachemeister model, generalized estimating equations (GEE), regression model, random effects models.

1. INTRODUCTION

Theoretical development has led the subject of credibility to be closely connected with random effects models. A regression credibility model (Hachemeister, 1975; Norberg, 1986), characterized by a within-panel correlation structure through the introduction of random regression coefficients, is essentially a generalized linear model (GLM) (McCullagh and Nelder, 1989) with solely random effects. Whereas independence is assumed in a classical GLM, generalized estimating equations (GEE) were developed to serve as an extension of the classical framework to handle random effects and correlated errors (Liang and Zeger, 1986). GEE were firstly introduced to estimate regression coefficients in a GLM provided that the covariance structure is either known or estimated *a priori*. Later developments have seen the use of GEE to estimate parameters which specify the covariance structure (Prentice and Zhao, 1991; Liang et al., 1992).

As a regression credibility model can be treated as a GLM with random effects (Nelder and Verall, 1997), it may also be handled by the GEE methodology.

The Bühlman (1967) and Bühlmann-Straub (1970) models are two important credibility models which can be formulated under the regression credibility framework. Under these models, statistical dependence between the observed data of the same risk entity is brought by the random effects shared in common. Recently, Lo, Fung and Zhu (2006) suggest to use the GEE method to estimate the structural parameters of these two popular models. The proposed GEE estimators are found to perform well in simulation studies. Meanwhile, Norberg (1982) applied the estimated generalized least squares method to a special case where, as in our proposed GEE approach, variances and covariances can be expressed as linear functions of structural parameters. The regression credibility model was first introduced by Hachemeister (1975). Since then, regression models have been the subject of interest in a number of credibility literature (see, for example, Bühlmann and Gisler (1997) and Bühlmann and Bühlmann (1999)). In this paper, we aim at using GEE to estimate the structural parameters of the Hachemeister (1975) regression model.

The remaining of the paper is organized as follows. Section 2 provides the model specification and the credibility estimator. The GEE methodology is introduced in Section 3. Section 4 describes the structural parameter estimation by means of the Hachemeister and the proposed GEE methods. Section 5 comes up with empirical studies by simulations and a real data analysis which reveal the advantages of GEE estimation of the variance and covariance parameters over Hachemeister's. In addition to the improved accuracy using GEE we do not observe the problems of reaching negative variances and correlations beyond 1 which may be delivered by Hachemeister's method (Dannenburg et al., 1996; Bühlmann and Gisler, 2005). Some concluding remarks are given in the final section.

2. THE REGRESSION CREDIBILITY MODEL

We consider the following model:

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, 2, \dots, n. \quad (1)$$

Each element y_{ij} in the $n_i \times 1$ vector \mathbf{y}_i is the observed value of some measurable characteristic for entity i in the j th observation period. \mathbf{X}_i is a $n_i \times m$ design matrix of known constants. The dimension of the vector of regression coefficients $\boldsymbol{\beta}_i$, labelled m in our paper, is essentially the number of covariates present in our model, including the constant term. Unlike classical regression models, $\boldsymbol{\beta}_i$'s are of random nature, i.i.d. with common mean $\boldsymbol{\beta}$ and covariance matrix \mathbf{F} for all i , i.e., $E(\boldsymbol{\beta}_i) = \boldsymbol{\beta}$ and $V(\boldsymbol{\beta}_i) = \mathbf{F}$. The error vectors $\boldsymbol{\varepsilon}_i$'s are assumed to be independently distributed with mean $\mathbf{0}$ and covariance matrix $\sigma^2 \mathbf{V}_i = \sigma^2 \mathbf{W}_i^{-1/2} \boldsymbol{\Gamma}_i \mathbf{W}_i^{-1/2}$, where \mathbf{W}_i is a diagonal matrix of known weights and $\boldsymbol{\Gamma}_i$ is a correlation matrix. When the errors are assumed to be independent, i.e.,

$\Gamma_i = \mathbf{I}$ for all i , model (1) is reduced to Hachemeister's (1975) regression credibility model. The following about \mathbf{y}_i can be obtained:

- (a) \mathbf{y}_i and \mathbf{y}_j are statistically independent for $i \neq j$;
- (b) $E(\mathbf{y}_i | \boldsymbol{\beta}_i) = \mathbf{X}_i \boldsymbol{\beta}_i$ and $\boldsymbol{\mu}_i = E(\mathbf{y}_i) = \mathbf{X}_i \boldsymbol{\beta}_i$;
- (c) $V(\mathbf{y}_i) = \mathbf{X}_i \mathbf{F} \mathbf{X}_i' + \sigma^2 \mathbf{W}_i^{-1/2} \Gamma_i \mathbf{W}_i^{-1/2}$.

Recently, Cossette and Luong (2003) obtained generalized least squares estimators for regression credibility models. Nonetheless, the model we propose here is quite different from the one considered in Cossette and Luong. First of all, Cossette and Luong's model has fixed all n_i 's at the same value, which carries an implication that an equal number of observations have to be made for each entity i . Our proposed model (1) does not restrict the values of n_i 's, the same practice as that in the Bühlmann-Straub and Hachemeister frameworks. Another difference, a crucial one, lies in the way the design matrix \mathbf{X}_i is treated. In Cossette and Luong's model, the design matrix is common to all i . In other words, no difference is admitted regarding the values of the covariates for different entities. In insurance setting, quite often the covariates would have values depending upon a particular entity; obvious examples include the age, gender, marital status and smoking behavior of an insured. In Sundt's (1987) credibility regression analysis, he attempts to use power and weight-to-price ratio to account for the variations in claim amounts for various car models. Cossette and Luong's model is not apt to cases involving the aforementioned factors as covariates. Meanwhile, model (1), formulated in line with Hachemeister's, steps in as a good candidate for use. Further, to implement Cossette and Luong's estimation methods, the weight matrices \mathbf{W}_i 's have to be the same for all entities. This constraint turns out to be very demanding. Contrariwise, we do not impose this constraint as we develop the estimators using GEE.

The credibility estimator (Hachemeister, 1975) for the regression coefficients $\boldsymbol{\beta}_i$ in model (1) is indeed a linear Bayes estimator in statistical terminology (Norberg, 1980). It has been shown by Rao (1975) that this Bayes estimator minimizes the quadratic loss function. The credibility estimator for $\boldsymbol{\beta}_i$ is a convex combination between individual experience and the collective mean:

$$\hat{\boldsymbol{\beta}}_i^{(B)} = \mathbf{Z}_i \hat{\boldsymbol{\beta}}_i^{(GLS)} + (\mathbf{I} - \mathbf{Z}_i) \boldsymbol{\beta}, \quad (3)$$

where

$$\mathbf{Z}_i = \mathbf{F} [\mathbf{F} + \sigma^2 (\mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i)^{-1}]^{-1} \quad (4)$$

is the credibility matrix, and

$$\hat{\boldsymbol{\beta}}_i^{(GLS)} = (\mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{X}_i)^{-1} \mathbf{X}_i' \mathbf{V}_i^{-1} \mathbf{y}_i \quad (5)$$

is the generalized least squares estimator for $\boldsymbol{\beta}_i$.

3. THE GEE METHODOLOGY

3.1. The GLM in brief

Model (1) can be viewed as an extension to the generalized linear model (GLM) with an identity link between the mean response and the linear predictor, i.e., $\mu = E(y) = \mathbf{x}'\boldsymbol{\beta}$ (McCullagh and Nelder, 1989). The probability distribution of the response y belongs to the exponential family which takes the following density function:

$$f_Y(y; \theta, \lambda) = \exp \left\{ \frac{y\theta - b(\theta)}{a(\lambda)} + c(y, \lambda) \right\}, \quad (6)$$

where θ is called the canonical parameter which determines the mean, and λ is called the dispersion parameter which controls the dispersion of the distribution. The mean and variance are, respectively, $\mu = b'(\theta)$ and $V(y) = b''(\theta)a(\lambda)$. Many popular distributions including the normal, Poisson, binomial, gamma, inverse gaussian and geometric distributions belong to the exponential family.

A common approach to estimate the regression parameters $\boldsymbol{\beta}$ is by means of the maximum likelihood method. For a random sample with observations y_1, \dots, y_n , the log-likelihood function is $\mathcal{L} = \sum_{i=1}^n \{ [y_i \theta_i - b(\theta_i)]/a(\lambda) + c(y_i, \lambda) \}$. The maximum likelihood estimate for $\boldsymbol{\beta}$ is then obtained by solving the score equations

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\beta}} = \sum_{i=1}^n \frac{\partial \mu_i}{\partial \boldsymbol{\beta}} V(y_i)^{-1} (y_i - \mu_i) = \mathbf{0}. \quad (7)$$

Some previous attempts to link credibility theory to GLMs include Nelder and Verrall (1997) and Frees et al. (1999). Discussions on the applications of GLMs can be found in McCullagh and Nelder (1989), Hardin and Hilbe (2001) and He, Fung and Zhu (2005) among others.

3.2. The use of GEE

While the GLM extends the classical linear model by having a more general family of distributions, the assumption of data independence is retained. This assumption fails when correlation between data exists. In the credibility context, correlation arises from the random effects as well as the dependency of errors. The generalized estimating equations (GEE) methodology (Liang and Zeger, 1986) has been developed to extend the application of GLMs to handle correlated data.

In order to introduce the GEE approach, we now shift our attention back to correlated or clustered data, the type of data analyzed in the regression credibility model. A second subscript j is added where appropriate to indicate the j th observation for an entity. In the classical GLM framework, observations

are taken as independent, thus giving the following score equations which are merely a sum of contributions from each observation y_{ij} :

$$\sum_{i=1}^n \sum_{j=1}^{n_i} \frac{\partial \mu_{ij}}{\partial \boldsymbol{\beta}} V(y_{ij})^{-1} (y_{ij} - \mu_{ij}) = \mathbf{0}. \quad (8)$$

In matrix form, (8) may be rewritten as

$$\sum_{i=1}^n \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}} \right)' \text{Diag}[V(\mathbf{y}_i)]^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i) = \mathbf{0}, \quad (9)$$

where $\mathbf{y}_i = (y_{i1}, \dots, y_{in_i})'$, $E(\mathbf{y}_i) = \boldsymbol{\mu}_i = (\mu_{i1}, \dots, \mu_{in_i})'$ and $\text{Diag}[V(\mathbf{y}_i)]$ denotes the diagonal matrix formed by the diagonal elements in $V(\mathbf{y}_i)$. Note that the matrix $\text{Diag}[V(\mathbf{y}_i)]$ is restricted to have zero off-diagonal elements because y_{ij} 's are assumed to be independent in the classical GLM. To take into account the dependency due to random effects and/or a correlated error structure, equation (9) can be turned into a system of GEE with $\text{Diag}[V(\mathbf{y}_i)]$ replaced by $V_w(\mathbf{y}_i)$ which needs not be diagonal:

$$\sum_{i=1}^n \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}} \right)' V_w(\mathbf{y}_i)^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i) = \mathbf{0}. \quad (10)$$

The matrix $V_w(\mathbf{y}_i)$ is called the working covariance matrix in the GEE terminology. A nice property of the GEE methodology is that the GEE estimators can still be consistent even if $V_w(\mathbf{y}_i)$ is not correctly specified. For a justification of using equation (10) to handle correlation and other properties of the GEE estimators, readers may see Liang and Zeger (1986) and Zeger and Liang (1986).

In credibility theory, our main interest is to estimate the variance and covariance parameters. To accomplish this by means of GEE, we first define $s_{ijk} = (y_{ij} - \mu_{ij})(y_{ik} - \mu_{ik})$ and $\boldsymbol{\alpha}$ as the vector of all variance and covariance parameters to be estimated. Then we build the following system of estimating equations analogous to (10) to solve for $\boldsymbol{\alpha}$:

$$\sum_{i=1}^n \mathbf{D}_i' \mathbf{H}_i^{-1} (\mathbf{s}_i - \boldsymbol{\xi}_i) = \mathbf{0}, \quad (11)$$

where $\mathbf{s}_i = \text{vech}(\mathbf{y}_i - \boldsymbol{\mu}_i)(\mathbf{y}_i - \boldsymbol{\mu}_i)' = (s_{i11}, s_{i12}, \dots, s_{i1n_i}, s_{i22}, s_{i23}, \dots)'$, $\boldsymbol{\xi}_i = E(\mathbf{s}_i)$, $\mathbf{D}_i = \partial \boldsymbol{\xi}_i / \partial \boldsymbol{\alpha}$ and $\mathbf{H}_i = V_w(\mathbf{s}_i)$. The vech operation constructs a vector by concatenating the lower-triangular elements of a matrix (see Graham (1981) for more details about the vech operation and also the vec operation, which will appear in Section 4.2). The equations may be solved numerically, for example, using Newton-Raphson. For more details on using GEE to estimate the variance and covariance parameters, readers may refer to Hardin and Hilbe (2003) and Prentice and Zhao (1991).

The vectors \mathbf{s}_i and $\boldsymbol{\xi}_i$ in (11) above indeed denote the empirical and the corresponding true covariance matrices of \mathbf{y}_i in vector form. When it happens that, as in our case, each element of $\boldsymbol{\xi}_i$ can be expressed as a linear combination of the variance and covariance parameters in $\boldsymbol{\alpha}$, the derivative matrix \mathbf{D}_i will simply be made up of constants independent of any unknown parameters. \mathbf{H}_i , the working covariance matrix, may contain information about fourth-order moments of \mathbf{y}_i . An example of \mathbf{H}_i is $V(\mathbf{s}_i)$ with components specified by:

$$\begin{aligned} \text{Cov}(s_{ijk}, s_{ilm}) &= E[(y_{ij} - \mu_{ij})(y_{ik} - \mu_{ik})(y_{il} - \mu_{il})(y_{im} - \mu_{im})] \\ &\quad - \text{Cov}(y_{ij}, y_{ik}) \text{Cov}(y_{il}, y_{im}). \end{aligned} \tag{12}$$

To solve the system of GEE (11), knowledge about $\boldsymbol{\mu}_i$, which depends on $\boldsymbol{\beta}$, is assumed. Hence, $\boldsymbol{\beta}$ has to be estimated via external means *a priori*. While $\boldsymbol{\beta}$ may simply be taken as a simple or weighted average of $\hat{\boldsymbol{\beta}}_i^{(GLS)}$'s defined in Section 2, we, in accordance with the empirical evidence of the relative performance, suggest to apply another system of GEE to provide an estimate of $\boldsymbol{\beta}$:

$$\sum_{i=1}^n \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}} \right)' \mathbf{U}_i^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i) = \mathbf{0}. \tag{13}$$

On setting \mathbf{U}_i as $V(\mathbf{y}_i)$ in (13), the working covariance matrix then includes parameters which appear in $\boldsymbol{\alpha}$. Hence, in order to estimate $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ simultaneously, we have to solve (11) and (13) alternately by updating the estimates of $\boldsymbol{\alpha}$ and then applying the updated values to re-estimate $\boldsymbol{\beta}$ in each iteration.

4. STRUCTURAL PARAMETER ESTIMATION

4.1. The Hachemeister and modified Hachemeister estimators

Hachemeister (1975) extended the Bühlmann-Straub model by introducing covariates to help account for heterogeneity in the expected risk level, $E(y_{ij} | \Theta = \theta_i) = \mathbf{x}'_{ij} \mathbf{b}(\theta_i)$, where θ_i denotes the realized yet unobservable risk characteristic borne by a risk entity i , and $\mathbf{b}(\theta_i) = \boldsymbol{\beta}_i$ in our notation. As in the Bühlmann-Straub model, the conditional variance of y_{ij} is $V(y_{ij} | \Theta = \theta_i) = s^2(\theta_i)/w_{ij}$. The conditional independence assumption remains to apply:

$$\text{Cov}(y_{ij}, y_{ik} | \Theta = \theta_i) = 0 \quad \text{for } j \neq k. \tag{14}$$

Hence, the covariance structure for the errors is assumed to be known up to a scalar multiplier, i.e., $V(\mathbf{y}_i | \Theta = \theta_i) = s^2(\theta_i) \mathbf{W}_i^{-1}$. The structural parameters then include $\boldsymbol{\beta} = E[\mathbf{b}(\Theta)]$, $\mathbf{F} = V[\mathbf{b}(\Theta)]$ and $\sigma^2 = E[s^2(\Theta)]$, which are important for constructing the credibility matrix and ultimately the credibility estimator. The form of the credibility matrix stated in Hachemeister (1975),

$$\mathbf{Z}_i = (\mathbf{F}\mathbf{X}'_i \mathbf{W}_i \mathbf{X}_i + \sigma^2 \mathbf{I})^{-1} \mathbf{F}\mathbf{X}'_i \mathbf{W}_i \mathbf{X}_i, \tag{15}$$

can be proved to be tantamount to (4) when the errors are assumed to be independent.

Unbiased estimators have been proposed for the structural parameters. First of all, the design matrices and the vectors of observations are respectively stacked up to form two large single units:

$$\mathbf{X} = [\mathbf{X}'_1, \mathbf{X}'_2, \dots, \mathbf{X}'_n]' \quad \text{and} \quad \mathbf{y} = [\mathbf{y}'_1, \mathbf{y}'_2, \dots, \mathbf{y}'_n]'. \quad (16)$$

Next, a super matrix is constructed with individual exposure matrices as building blocks along the principal diagonal:

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_1 & & & \mathbf{0} \\ & \mathbf{W}_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{W}_n \end{bmatrix}. \quad (17)$$

A weighted least square estimate of $\boldsymbol{\beta}$ is obtained straight-forwardly (Hachmeister, 1975), making use of (16) and (17):

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \mathbf{X}'\mathbf{W}\mathbf{y}. \quad (18)$$

An unbiased estimator of σ^2 is taken as the simple average of n unbiased estimators $\hat{\sigma}_i^2$ each of which makes use of a particular risk entity's observations:

$$\hat{\sigma}^2 = n^{-1} \sum_{i=1}^n \hat{\sigma}_i^2 = n^{-1} \sum_{i=1}^n (n_i - m)^{-1} (\mathbf{y}_i - \mathbf{X}_i \hat{\boldsymbol{\beta}}_i)' \mathbf{W}_i (\mathbf{y}_i - \mathbf{X}_i \hat{\boldsymbol{\beta}}_i), \quad (19)$$

where $\hat{\boldsymbol{\beta}}_i$ is the weighted least squares estimator. The estimator for the covariance matrix of \mathbf{F} involves some cumbersome expressions. The following quantities are introduced in advance:

$$\mathbf{G} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \sum_{i=1}^n (\mathbf{X}'_i \mathbf{W}_i \mathbf{X}_i) (\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}}) (\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}})',$$

$$\mathbf{\Pi} = \mathbf{I} - \sum_{i=1}^n (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} (\mathbf{X}'_i \mathbf{W}_i \mathbf{X}_i) (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} (\mathbf{X}'_i \mathbf{W}_i \mathbf{X}_i)$$

and

$$\mathbf{C} = \mathbf{\Pi}^{-1} \left[\mathbf{G} - (n-1) (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \hat{\sigma}^2 \right].$$

An estimator for \mathbf{F} is given by

$$\hat{\mathbf{F}} = (\mathbf{C} + \mathbf{C}') / 2. \quad (20)$$

Note that \mathbf{C} itself is also an unbiased estimator for \mathbf{F} (Hachemeister, 1975). It is averaged with its transpose to be the final estimator since \mathbf{F} has to be symmetric while \mathbf{C} , however, is not in general. Plugging the *Hachemeister* estimators of $\boldsymbol{\beta}$, σ^2 and \mathbf{F} given by (18)-(20) into (3) and (4), and having $\mathbf{V}_i^{-1} = \mathbf{W}_i$ due to the conditional independence assumption (14), we obtain the *Hachemeister* estimators for $\boldsymbol{\beta}_i$ and \mathbf{Z}_i respectively.

Recalling that m is the dimension of the vector $\boldsymbol{\beta}$, the unstructured covariance matrix \mathbf{F} which is symmetric will have $m(m + 1) / 2$ unknown parameters to be estimated. Take the case of $m = 2$ for illustration. Define $\phi_{jj} = V(\beta_{ij})$ and $\phi_{jk} = \text{Cov}(\beta_{ij}, \beta_{ik})$ for $j \neq k$ respectively. The distinct elements needed to specify \mathbf{F} are, namely, ϕ_{11}, ϕ_{12} and ϕ_{22} . It is known that the *Hachemeister* method may give negative variance estimates ($\hat{\phi}_{jj}$) and correlation estimates ($\hat{\phi}_{jk} / (\hat{\phi}_j \hat{\phi}_k)^{1/2}$) with magnitudes greater than 1 (Dannenburg et al., 1996; Böhmann and Gisler, 2005). To avoid these problems, we suggest to replace (20) with

$$\hat{\phi}_{jj} = \sum_{i=1}^n w_i (\hat{\beta}_{ij} - \bar{\beta}_{.j})^2 / \left(\sum_{i=1}^n w_i - 1 \right) \tag{21}$$

and

$$\hat{\phi}_{jk} = \sum_{i=1}^n w_i (\hat{\beta}_{ij} - \bar{\beta}_{.j}) (\hat{\beta}_{ik} - \bar{\beta}_{.k}) / \left(\sum_{i=1}^n w_i - 1 \right), \tag{22}$$

where $w_i = \sum_k w_{ik}$ and $\bar{\beta}_{.j}$ is the weighted mean for the j th coefficient. These modified forms are termed *Hachemeister-M* estimators.

4.2. GEE-I and GEE-II estimators

We now introduce the use of GEE to estimate the unknown parameters in model (1) such that a credibility estimator of $\boldsymbol{\beta}_i$ can be reached. We first consider model (1) under the original *Hachemeister* framework in which the errors ε_i 's are independent, i.e., $\boldsymbol{\Gamma}_i = \mathbf{I}$ for all i . As a result, \mathbf{V}_i is known and equal to the inverse of the exposure matrix, i.e., \mathbf{W}_i^{-1} . Define $\boldsymbol{\alpha} = (\boldsymbol{\phi}', \sigma^2)'$ as the vector storing all variance and covariance parameters to be estimated, where $\boldsymbol{\phi}$ is a vector specifying the unknown parameters in \mathbf{F} ; for example in the $m = 2$ case, with \mathbf{F} being a 2×2 symmetric matrix, the unstructured covariance matrix \mathbf{F} is specified by $\boldsymbol{\phi} = (\phi_{11}, \phi_{12}, \phi_{22})'$.

The system of GEE (11) is applied to solve for $\boldsymbol{\alpha}$. The explicit forms of the components which appear in (11) are shown as follows. As stated in Section 3.2, \mathbf{s}_i is explicitly equal to

$$\text{vech}(\mathbf{y}_i - \boldsymbol{\mu}_i)(\mathbf{y}_i - \boldsymbol{\mu}_i)' = \begin{bmatrix} (y_{i1} - \mathbf{x}'_{i1}\boldsymbol{\beta})^2 \\ (y_{i1} - \mathbf{x}'_{i1}\boldsymbol{\beta})(y_{i2} - \mathbf{x}'_{i2}\boldsymbol{\beta}) \\ \vdots \\ (y_{i1} - \mathbf{x}'_{i1}\boldsymbol{\beta})(y_{in_i} - \mathbf{x}'_{in_i}\boldsymbol{\beta}) \\ (y_{i2} - \mathbf{x}'_{i2}\boldsymbol{\beta})^2 \\ (y_{i2} - \mathbf{x}'_{i2}\boldsymbol{\beta})(y_{i3} - \mathbf{x}'_{i3}\boldsymbol{\beta}) \\ \vdots \\ (y_{in_i} - \mathbf{x}'_{in_i}\boldsymbol{\beta})^2 \end{bmatrix}. \tag{23}$$

ξ_i , the expectation of \mathbf{s}_i , has entries being the linear combinations of the elements in $\boldsymbol{\alpha}$. The entries of ξ_i are specified in the following for the case of $m = 2$:

$$\begin{aligned} & E[(y_{ij} - \mathbf{x}'_{ij}\boldsymbol{\beta})(y_{ik} - \mathbf{x}'_{ik}\boldsymbol{\beta})] \\ &= \begin{cases} x_{ij1}^2\phi_{11} + 2x_{ij1}x_{ij2}\phi_{12} + x_{ij2}^2\phi_{22} + \sigma^2w_{ij}^{-1} & \text{for } j = k, \\ x_{ij1}x_{ik1}\phi_{11} + (x_{ij1}x_{ik2} + x_{ij2}x_{ik1})\phi_{12} + x_{ij2}x_{ik2}\phi_{22} & \text{for } j \neq k. \end{cases} \end{aligned} \tag{24}$$

It is more convenient to compute ξ_i using (2) on noting that $\xi_i = \text{vech } V(\mathbf{y}_i)$, but a componentwise specification like (24) is useful for the derivation of the derivative matrix $\mathbf{D}_i = \partial\xi_i/\partial\boldsymbol{\alpha}$. Since, as shown from (24), the entire ξ_i is linear in $\boldsymbol{\alpha}$, \mathbf{D}_i is entirely made of known constants.

The middle component in the system of GEE (11), denoted as \mathbf{H}_i , is the working covariance matrix which conveys information about the fourth-order moments of the vector of observations \mathbf{y}_i . A convenient choice based on normality is provided by Fuller (1987):

$$\mathbf{H}_i = 2\boldsymbol{\psi}_{n_i}(V(\mathbf{y}_i) \otimes V(\mathbf{y}_i))\boldsymbol{\psi}'_{n_i}, \tag{25}$$

where $\boldsymbol{\psi}_{n_i} = (\boldsymbol{\Phi}'_{n_i}\boldsymbol{\Phi}_{n_i})^{-1}\boldsymbol{\Phi}'_{n_i}$ and $\boldsymbol{\Phi}_p$ is a matrix of dimensions $p^2 \times p(p+1)/2$. Defining the vec operation as the stacking up of the columns of a matrix, $\boldsymbol{\Phi}_p$ facilitates the vec-vech conversion for any $p \times p$ symmetric matrix \mathbf{A} such that $\text{vec } \mathbf{A} = \boldsymbol{\Phi}_p \text{vech } \mathbf{A}$. From practical experience, this choice of \mathbf{H}_i works reasonably well for distributions not substantially different from normal.

As stated in Section 3.2, another system of GEE (13) is used to work out the estimate for $\boldsymbol{\beta}$. The working covariance matrix \mathbf{U}_i which is set as $V(\mathbf{y}_i)$ may be computed using (2) by setting $\boldsymbol{\Gamma}_i = \mathbf{I}$, while the derivative matrix $\partial\boldsymbol{\mu}_i/\partial\boldsymbol{\beta}$ is merely the design matrix \mathbf{X}_i of model (1). As $V(\mathbf{y}_i)$ in (13) is in return specified using $\boldsymbol{\alpha}$, the estimates for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are obtained by solving (11) and (13) recursively until convergence. We term these estimates the *GEE-I* estimates.

The *GEE-II* estimators described below are obtained when a moving average error structure for ε_{ij} 's is considered. With an *MA*(1) error structure having correlation ρ , the vector of unknown variance and covariance parameters takes the form $\alpha = (\phi', \sigma^2, \sigma^2 \rho)'$. For $m = 2$, the specification of the entries of ξ_i in (24) is replaced by

$$E[(y_{ij} - \mathbf{x}'_{ij}\boldsymbol{\beta})(y_{ik} - \mathbf{x}'_{ik}\boldsymbol{\beta})] = \begin{cases} x_{ij1}^2\phi_{11} + 2x_{ij1}x_{ij2}\phi_{12} + x_{ij2}^2\phi_{22} + \sigma^2w_{ij}^{-1}, & j = k, \\ x_{ij1}x_{ik1}\phi_{11} + (x_{ij1}x_{ik2} + x_{ij2}x_{ik1})\phi_{12} \\ \quad + x_{ij2}x_{ik2}\phi_{22} + \rho\sigma^2w_{ij}^{-1/2}w_{ik}^{-1/2}, & j = k - 1, j = k + 1, \\ x_{ij1}x_{ik1}\phi_{11} + (x_{ij1}x_{ik2} + x_{ij2}x_{ik1})\phi_{12} + x_{ij2}x_{ik2}\phi_{22}, & \text{otherwise.} \end{cases} \quad (26)$$

(26) reveals that, as in the independent error case, ξ_i is still linear in α when moving average dependencies are taken for the errors ε_{ij} 's. This, again, results in the derivative matrix \mathbf{D}_i being entirely composed of known, constant elements. Evaluating the GEE components under the *MA*(1) error structure, the *GEE-II* estimates for α and $\boldsymbol{\beta}$ can be obtained through solving (11) and (13) recursively.

5. EMPIRICAL STUDIES

Three sets of simulation studies are included. The first is a linear time trend model related to the well-known study given in Hachemeister (1975). Each study contains 500 replicates of simulated data. The independent error and a correlated *MA*(1) error structure with $\rho = 0.4$ are taken. For each study, the estimation methods, *Hachemeister*, *Hachemeister-M*, *GEE-I* and *GEE-II* as illustrated in Section 4, are considered. The constraints of Cossette and Luong's (2003) estimation method set forth in Section 2 inhibit us from incorporating their approach as well in our investigation.

5.1. Study 1: the linear time trend model

Hachemeister (1975) adopted a linear time trend model to accommodate the private passenger bodily injury's claim data observed during the years 1970-1973 in the United States:

$$y_{ij} = \beta_{i1} + \beta_{i2}j + \varepsilon_{ij}. \quad (27)$$

It may be implemented by model (1) through setting $n = 5$, $n_i = 12$, $m = 2$ and $\mathbf{x}_{ij} = (1, j)'$. We have generated 500 replicates of data which resemble the actual claim data, taking the following parameter values chosen by Dannenburg et al. (1996) in their simulation study of the Hachemeister model:

$$\boldsymbol{\beta} = (1400, 150)', \quad \sigma^2 = 300^2, \quad \phi_{11} = 100^2, \quad \phi_{12} = 0, \quad \phi_{22} = 20^2.$$

The weights w_{ij} 's range from 250 to 9500 in Hachemeister's data.

Tables 1 and 2 present the estimation results associated with an independent error structure and a correlated error structure respectively. A clear advantage is observed for the GEE approach over Hachemeister's estimation regarding both the credibility estimator of the regression coefficients as well as the variance and covariance parameters in both cases. The modified Hachemeister method does not suffer from negative variance estimates which the Hachemeister does, but is still outperformed by the two GEE methods I and II.

Whilst a significant advantage has been recorded for the GEE approach over Hachemeister's in estimating the parameters which specify $\mathbf{F} = V(\boldsymbol{\beta}_i)$ (MSE is reduced by 30%-70% each in estimating ϕ_{11} , ϕ_{12} and ϕ_{22}), the improvement in estimating $\boldsymbol{\beta}_i^{(B)}$ is modest (MSE reduction is less than 20%). Besides, albeit a considerably larger MSE of the estimators for specifying \mathbf{F} , Hachemeister is quite trustworthy in estimating σ^2 , notably including the case of $\rho = 0.4$ in which the independent error assumption is violated. Note that, on the other hand, when $\rho = 0.4$, the performance of the GEE-I method, which assumes independence of the errors, is still comparable to or even better than that of GEE-II. This serves as an empirical justification of the use of GEE-I when knowledge about the correlated error structure is absent.

5.2. Study 2: a model with 2 covariates ($m = 2$)

In this study, in contrast to Study 1, we have set a different design matrix \mathbf{X}_i for each i , which is more adequate for practical use. A different scenario is also reflected on the relative size between the number of entities n and the number of observations n_i made per entity. While in Study 1 $n < n_i$, here we choose $n = 25$ and $n_i = 5$ respectively. Parameter settings are as follows:

$$\boldsymbol{\beta} = (20, 10)', \quad \sigma^2 = 4^2, \quad \phi_{11} = 3^2, \quad \phi_{12} = 4, \quad \phi_{22} = 3^2.$$

Each weighting element w_{ij} is generated with a Poisson mean λ_i which follows a uniform distribution defined on the interval (5, 100) and differs for each i . Likewise, each quantity x_{ij2} corresponding to the second covariate (the first covariate being the constant term) is simulated with a normal distribution around a 'mean-level' value which is uniformly selected from the interval (-5, 5) and specific to each i . Estimation results for the cases of independent and correlated error structures are shown in Tables 3 and 4 respectively.

Unlike results of the last study, enormous discrepancies as regards the performance of the credibility estimator for $\boldsymbol{\beta}_i$ between the GEE approach and Hachemeister's estimation method occur (relative efficiency beyond 20). As a reasonable conjecture, the level of precision in estimating ϕ_{11} , ϕ_{12} and ϕ_{22} applying Hachemeister's formulae is not enough to produce satisfactory estimates of \mathbf{Z}_i 's and $\boldsymbol{\beta}_i$'s — 15% of the estimates of the covariance matrix \mathbf{F} are found to be not positive definite. In contrast, all parameter estimates of ϕ_{11} ,

TABLE 1
ESTIMATION RESULTS OF STUDY I ASSOCIATED WITH AN INDEPENDENT ERROR STRUCTURE

Parameter		Method			
		GEE-I	GEE-II	Hachemeister- <i>M</i>	Hachemeister
β_{i1}	Bias	-0.1149	-0.0473	-0.0076	-0.0075
	MSE	2.85×10^1 (1.22 [▲])	3.30×10^1 (1.05)	3.41×10^1 (1.01)	3.46×10^1
β_{i2}	Bias	0.0256	0.0672	-0.0063	-0.0061
	MSE	6.16×10^{-1} (1.18)	6.39×10^{-1} (1.14)	6.96×10^{-1} (1.04)	7.25×10^{-1}
z_{i11}	Bias	-0.0135	-0.0098	0.0143	0.0065
	MSE	9.23×10^{-4} (2.76)	9.40×10^{-4} (2.71)	1.43×10^{-3} (1.78)	2.55×10^{-3}
z_{i12}	Bias	0.0459	-0.0153	-0.0068	-0.1294
	MSE	1.46×10^{-2} (3.34)	2.85×10^{-2} (1.71)	3.86×10^{-2} (1.26)	4.87×10^{-2}
z_{i21}	Bias	-0.0283	0.0008	-0.0067	-0.0038
	MSE	1.79×10^{-5} (1.21)	1.42×10^{-5} (1.53)	1.64×10^{-5} (1.32)	2.17×10^{-5}
z_{i22}	Bias	-0.0086	0.0040	0.0005	0.0005
	MSE	3.24×10^{-4} (4.07)	1.03×10^{-3} (1.28)	1.12×10^{-3} (1.17)	1.32×10^{-3}
ϕ_{11}	Bias	-1257	-1194	751	688
	MSE	3.77×10^7 (2.29)	3.94×10^7 (2.19)	4.02×10^7 (2.15)	8.63×10^7
ϕ_{12}	Bias	-8.14	4.20	22.4	34.3
	MSE	7.25×10^5 (1.81)	6.75×10^5 (1.95)	8.71×10^5 (1.51)	1.31×10^6
ϕ_{22}	Bias	78.5	102	38.1	30.8
	MSE	6.24×10^4 (2.33)	5.83×10^4 (2.50)	6.95×10^4 (2.10)	1.46×10^5
σ^2	Bias	-623	-208	927	927
	MSE	2.67×10^8 (1.14)	2.94×10^8 (1.04)	3.06×10^8 (1.00)	3.06×10^8

▲ Relative efficiency of the estimator. Hachemeister's estimator serves as the baseline.

TABLE 2
ESTIMATION RESULTS OF STUDY 1 ASSOCIATED WITH A CORRELATED ERROR STRUCTURE
($MA(1)$ WITH $\rho = 0.4$)

Parameter		Method			
		GEE-I	GEE-II	Hachemeister- M	Hachemeister
β_{i1}	Bias	0.0943	0.0626	0.7049	0.5413
	MSE	5.21×10^1 (1.08)	5.57×10^1 (1.01)	5.61×10^1 (1.01)	5.64×10^1
β_{i2}	Bias	-0.0008	0.0141	0.0063	0.0074
	MSE	9.83×10^{-1} (0.93)	9.63×10^{-1} (0.95)	9.11×10^{-1} (1.00)	9.15×10^{-1}
z_{i11}	Bias	0.0029	-0.0093	0.0043	-0.0164
	MSE	8.04×10^{-4} (1.66)	2.15×10^{-3} (0.62)	9.82×10^{-4} (1.36)	1.34×10^{-3}
z_{i12}	Bias	0.0455	-0.0241	-0.1108	-0.0533
	MSE	2.68×10^{-2} (1.40)	4.46×10^{-2} (0.84)	3.25×10^{-2} (1.15)	3.75×10^{-2}
z_{i21}	Bias	0.0015	-0.0038	-0.0041	-0.0008
	MSE	1.13×10^{-5} (1.69)	2.75×10^{-5} (0.70)	1.67×10^{-5} (1.14)	1.91×10^{-5}
z_{i22}	Bias	-0.0014	-0.0095	0.0095	0.0111
	MSE	4.06×10^{-4} (1.70)	8.19×10^{-4} (0.84)	6.12×10^{-4} (1.13)	6.89×10^{-4}
ϕ_{11}	Bias	-1068	1254	828	721
	MSE	3.54×10^7 (3.52)	4.83×10^7 (2.58)	5.94×10^7 (2.10)	1.25×10^8
ϕ_{12}	Bias	-28.8	30.5	28.7	44.6
	MSE	7.47×10^5 (1.44)	7.71×10^5 (1.39)	8.02×10^5 (1.34)	1.07×10^6
ϕ_{22}	Bias	65.4	73.6	41.1	18.8
	MSE	5.83×10^4 (2.54)	5.91×10^4 (2.51)	9.16×10^4 (1.62)	1.48×10^5
σ^2	Bias	-12574	-1081	11268	11268
	MSE	3.81×10^8 (1.09)	3.94×10^8 (1.06)	4.16×10^8 (1.00)	4.16×10^8

TABLE 3
ESTIMATION RESULTS OF STUDY 2 ASSOCIATED WITH AN INDEPENDENT ERROR STRUCTURE

Parameter		Method			
		GEE-I	GEE-II	Hachemeister- <i>M</i>	Hachemeister
β_{i1}	Bias	0.0072	-0.0085	0.3421	0.3714
	MSE	3.80×10^{-1} (>1000)	3.87×10^{-1} (>1000)	4.93×10^2 (1.06)	5.23×10^2
β_{i2}	Bias	-0.0014	-0.0015	0.0715	0.0853
	MSE	2.79×10^{-2} (>500)	2.86×10^{-2} (>500)	1.30×10^1 (1.24)	1.61×10^1
z_{i11}	Bias	0.0121	0.0093	-0.0922	0.1845
	MSE	7.27×10^{-4} (>100000)	7.67×10^{-4} (>100000)	9.73×10^1 (1.70)	1.65×10^2
z_{i12}	Bias	-0.0065	-0.0048	0.2588	0.1753
	MSE	3.82×10^{-4} (>100000)	3.98×10^{-4} (>100000)	6.07×10^1 (1.32)	8.02×10^1
z_{i21}	Bias	0.0002	0.0000	-0.0542	-0.0427
	MSE	3.27×10^{-5} (>100000)	3.50×10^{-5} (>100000)	2.96 (1.41)	4.17
z_{i22}	Bias	0.0007	0.0005	0.0157	-0.0198
	MSE	2.08×10^{-5} (>100000)	2.14×10^{-5} (>100000)	1.88 (1.52)	2.85
ϕ_{11}	Bias	0.0568	0.0532	0.6208	0.3382
	MSE	7.15 (5.84)	7.22 (5.78)	2.65×10^1 (1.58)	4.17×10^1
ϕ_{12}	Bias	-0.4282	-0.4563	-0.0843	-0.1754
	MSE	4.00 (4.83)	3.89 (4.97)	1.16×10^1 (1.67)	1.93×10^1
ϕ_{22}	Bias	0.3859	0.4011	0.2931	-0.1692
	MSE	7.05 (1.78)	7.18 (1.75)	8.32 (1.51)	1.25×10^1
σ^2	Bias	-0.3585	0.0907	-0.2459	-0.2459
	MSE	6.74 (1.00)	7.93 (0.85)	6.72 (1.00)	6.72

TABLE 4
ESTIMATION RESULTS OF STUDY 2 ASSOCIATED WITH A CORRELATED ERROR STRUCTURE
($MA(1)$ WITH $\rho = 0.4$)

Parameter		Method			
		GEE-I	GEE-II	Hachemeister- M	Hachemeister
β_{i1}	Bias	0.0060	-0.0083	0.0469	0.0317
	MSE	3.84×10^{-1} (25.8)	3.22×10^{-1} (30.8)	9.51 (1.04)	9.89
β_{i2}	Bias	0.0006	0.0083	0.0245	0.0125
	MSE	2.62×10^{-2} (23.6)	1.94×10^{-2} (31.8)	5.91×10^{-1} (1.05)	6.18×10^{-1}
z_{i11}	Bias	-0.0018	-0.0054	-0.0452	-0.0199
	MSE	7.16×10^{-4} (>10000)	5.54×10^{-4} (>10000)	6.13 (1.45)	8.90
z_{i12}	Bias	0.0008	0.0010	-0.0472	-0.0215
	MSE	3.63×10^{-4} (>1000)	2.72×10^{-4} (>1000)	8.26×10^{-1} (1.10)	9.11×10^{-1}
z_{i21}	Bias	-0.0024	-0.0001	0.0074	0.0093
	MSE	4.69×10^{-5} (>10000)	2.31×10^{-5} (>10000)	5.90×10^{-1} (1.10)	6.49×10^{-1}
z_{i22}	Bias	0.0016	-0.0002	0.0059	-0.0026
	MSE	2.27×10^{-5} (>1000)	1.30×10^{-5} (>1000)	4.31×10^{-2} (1.20)	5.15×10^{-2}
ϕ_{11}	Bias	0.7305	-0.6488	0.3218	0.2146
	MSE	6.42 (6.45)	6.44 (6.43)	2.08×10^1 (1.99)	4.14×10^1
ϕ_{12}	Bias	-0.5343	-0.5481	-0.2594	-0.0955
	MSE	4.16 (4.14)	4.19 (4.11)	8.31 (2.07)	1.72×10^1
ϕ_{22}	Bias	-0.6766	-0.5135	0.2715	-0.084
	MSE	6.94 (1.66)	7.02 (1.64)	8.14 (1.41)	1.15×10^1
σ^2	Bias	-2.85	-0.3467	-2.724	-2.724
	MSE	1.43×10^1 (1.01)	8.48 (1.71)	1.45×10^1 (1.00)	1.45×10^1

TABLE 5

ESTIMATION RESULTS OF STUDY 3 ASSOCIATED WITH AN INDEPENDENT ERROR STRUCTURE

Parameter	Method				
	GEE-I	GEE-II	Hachemeister- <i>M</i>	Hachemeister	
β_{i1}	Bias	0.0088	-0.0072	0.1844	0.2454
	MSE	1.58 (104)	1.71 (96.6)	1.37×10^2 (1.20)	1.65×10^2
β_{i2}	Bias	-0.0001	-0.0001	0.0105	0.0083
	MSE	5.75×10^{-2} (112)	6.04×10^{-2} (107)	6.16 (1.05)	6.46
β_{i3}	Bias	-0.0005	-0.0007	0.0418	0.0573
	MSE	5.30×10^{-2} (183)	5.54×10^{-2} (175)	9.36 (1.03)	9.68
z_{i11}	Bias	0.0084	0.0066	0.0072	-0.0050
	MSE	2.85×10^{-3} (>10000)	4.61×10^{-3} (>5000)	1.97×10^1 (1.71)	3.36×10^1
z_{i12}	Bias	0.0019	0.0021	-0.0053	-0.0039
	MSE	2.87×10^{-4} (180)	3.75×10^{-4} (137)	3.95×10^{-2} (1.30)	5.15×10^{-2}
z_{i13}	Bias	-0.0009	0.0000	-0.0036	-0.0065
	MSE	1.96×10^{-4} (409)	2.12×10^{-4} (378)	6.20×10^{-2} (1.29)	8.02×10^{-2}
z_{i21}	Bias	0.0013	-0.0019	0.0084	-0.0093
	MSE	5.53×10^{-5} (>1000)	1.10×10^{-4} (>10000)	0.852 (1.71)	1.46
z_{i22}	Bias	0.0009	-0.0011	0.0007	-0.0003
	MSE	2.50×10^{-5} (80.9)	3.84×10^{-5} (52.6)	1.04×10^{-3} (1.94)	2.02×10^{-3}
z_{i23}	Bias	-0.0001	-0.0001	-0.0008	0.0016
	MSE	5.27×10^{-6} (>500)	6.80×10^{-6} (493)	1.84×10^{-3} (1.82)	3.35×10^{-3}
z_{i31}	Bias	0.0008	0.0000	0.0049	0.0052
	MSE	8.78×10^{-5} (>5000)	9.81×10^{-5} (>5000)	3.41×10^{-1} (1.81)	6.16×10^{-1}
z_{i32}	Bias	0.0004	0.0003	0.0012	-0.0013
	MSE	1.31×10^{-5} (70.1)	1.32×10^{-5} (69.1)	6.64×10^{-4} (1.38)	9.14×10^{-4}
z_{i33}	Bias	-0.0005	-0.0002	-0.0024	0.0009
	MSE	1.80×10^{-5} (460)	1.81×10^{-5} (456)	6.47×10^{-3} (1.28)	8.25×10^{-3}
ϕ_{11}	Bias	0.3182	0.3566	-0.3483	0.3626
	MSE	10.1 (8.86)	10.2 (8.74)	4.2×10^1 (2.12)	8.92×10^1
ϕ_{22}	Bias	-0.7843	-0.7651	0.0251	0.0083
	MSE	6.96 (1.66)	7.06 (1.64)	7.24 (1.60)	1.16×10^1
ϕ_{33}	Bias	0.3527	0.3489	-0.2544	0.1849
	MSE	6.19 (2.97)	6.25 (2.95)	9.81 (1.88)	1.84×10^1
σ^2	Bias	-0.1422	0.5181	-0.1435	-0.1435
	MSE	9.69 (1.01)	2.71×10^1 (0.36)	9.78 (1.00)	9.78

TABLE 6

ESTIMATION RESULTS OF STUDY 3 ASSOCIATED WITH A CORRELATED ERROR STRUCTURE
(MA(1) WITH $\rho = 0.4$)

Parameter		Method			
		GEE-I	GEE-II	Hachemeister- <i>M</i>	Hachemeister
β_{i1}	Bias	-0.0083	0.0054	0.1657	0.1992
	MSE	1.40 (>500)	1.32 (>500)	8.95×10^2 (1.04)	9.27×10^2
β_{i2}	Bias	0.0041	-0.0016	0.0755	0.0623
	MSE	6.50×10^{-2} (364)	6.27×10^{-2} (377)	2.06×10^1 (1.15)	2.37×10^1
β_{i3}	Bias	0.0000	0.0005	0.0414	-0.0457
	MSE	3.82×10^{-2} (>500)	3.30×10^{-2} (>500)	1.51×10^1 (1.29)	1.95×10^1
z_{i11}	Bias	-0.0019	-0.0040	-0.0783	0.0092
	MSE	5.05×10^{-3} (>5000)	3.61×10^{-3} (>10000)	2.16×10^1 (1.93)	4.16×10^1
z_{i12}	Bias	-0.0052	-0.0024	-0.0140	0.0084
	MSE	2.51×10^{-4} (>500)	2.10×10^{-4} (>500)	1.06×10^{-1} (1.74)	1.85×10^{-1}
z_{i13}	Bias	0.0024	-0.0002	0.0083	-0.0136
	MSE	1.61×10^{-4} (>1000)	8.39×10^{-5} (>5000)	3.15×10^{-1} (1.66)	5.25×10^{-1}
z_{i21}	Bias	0.0036	0.0009	0.0119	-0.0128
	MSE	9.89×10^{-5} (>5000)	7.96×10^{-5} (>5000)	3.34×10^{-1} (1.54)	5.13×10^{-1}
z_{i22}	Bias	-0.0009	-0.0008	-0.0016	0.0038
	MSE	1.75×10^{-5} (141)	2.51×10^{-5} (98.0)	2.00×10^{-3} (1.23)	2.46×10^{-3}
z_{i23}	Bias	0.0001	0.0001	-0.0029	-0.0099
	MSE	5.61×10^{-6} (>1000)	2.71×10^{-6} (>1000)	3.75×10^{-3} (2.09)	7.84×10^{-3}
z_{i31}	Bias	-0.0064	-0.0000	0.0087	0.0155
	MSE	1.31×10^{-4} (>1000)	5.68×10^{-5} (>5000)	2.87×10^{-1} (1.37)	3.92×10^{-1}
z_{i32}	Bias	-0.0003	-0.0001	0.0009	-0.0019
	MSE	1.09×10^{-5} (215)	6.15×10^{-6} (381)	1.13×10^{-3} (2.07)	2.35×10^{-3}
z_{i33}	Bias	0.0012	0.0001	0.0001	0.0003
	MSE	1.47×10^{-5} (467)	8.30×10^{-6} (>500)	5.06×10^{-3} (1.36)	6.87×10^{-3}
ϕ_{11}	Bias	-0.0709	-0.3129	0.1235	0.2477
	MSE	9.33 (10.4)	9.00 (10.8)	4.36×10^1 (2.24)	9.74×10^1
ϕ_{22}	Bias	-0.5284	-0.5495	-0.000	0.0000
	MSE	6.93 (2.00)	7.13 (1.94)	8.11 (1.71)	1.39×10^1
ϕ_{33}	Bias	-0.3143	-0.3082	0.2045	0.1782
	MSE	6.15 (3.13)	6.25 (3.08)	9.44 (2.04)	1.92×10^1
σ^2	Bias	5.2142	-0.8094	-5.532	-5.532
	MSE	3.40×10^1 (1.00)	1.92×10^1 (1.77)	3.40×10^1 (1.00)	3.40×10^1

ϕ_{12} and ϕ_{22} using either GEE-I or GEE-II fall in an admissible range to construct a valid estimate of the covariance matrix \mathbf{F} .

In another study holding equivalent settings as described above with the only exception that the ‘mean-level’ values used for simulating values of the second covariate are uniformly selected from (0, 10) instead, an immense rise of MSE by more than 35 times in estimating β_i results using Hachemeister’s estimation formulae. Such apparent inconstancy in performance of estimation when there is translation to a covariate is not observed using either GEE-I or GEE-II — the performance of estimators for all parameters of interest basically remains at the same level as the original.

5.3. Study 3: a model with 3 covariates ($m = 3$)

The previous two studies have assessed the performance of the proposed GEE approach in two quite different scenarios, yet they remain to be cases of only two covariates present in the model. Hence, a third study is included to evaluate its application in the case of a multiple regression model.

The model used in this study consists of a linear time trend component and another covariate of which the corresponding quantities are simulated with a normal distribution around some entity-specific ‘mean-level’ values uniformly selected from the interval (0, 10). There are 25 entities and 5 observations recorded in consecutive periods for each entity. The exposures w_{ij} ’s are generated from exactly the same mechanism as in Study 2. Other parameter settings take place as follows:

$$\boldsymbol{\beta} = (20, 10, 10)', \quad \sigma^2 = 4^2, \quad \phi_{11} = 3^2, \quad \phi_{22} = 3^2, \quad \phi_{33} = 3^2.$$

While in the first two studies it is in line with the normal conjecture that better performance is held by the GEE method with a correct correlated error structure assumed, the manifestation is more clear-cut here when a higher dimension is attributed to the vector of regression coefficients $\boldsymbol{\beta}_t$. Hachemeister’s estimation lags behind the two GEE methods considerably. About 17% of all replicates have generated negative estimates of variance parameters in \mathbf{F} . The modified Hachemeister method can avoid from these shortcomings, but its performance is still far behind the GEE estimators. Nevertheless, consistent with those of the previous examples, the simulation results reaffirm that Hachemeister has proposed a good estimator of σ^2 .

5.4. Injury claim data

To illustrate the the Hachemeister and GEE methods, we use the Hachemeister data in which y_{ij} represents the average amount of w_{ij} claims for private passenger body injuries observed for the twelve periods during 1970-1973 in the USA. Five states of observations were provided; see Hachemeister (1975) for the full data. A linear time trend model was suggested.

TABLE 7
PARAMETER ESTIMATES FOR THE REGRESSION MODEL OF INJURY CLAIM DATA
OBTAINED BY HACHEMEISTER, GEE-I AND GEE-II METHODS

	Hachemeister	GEE-I	GEE-II
$\hat{\beta}^{(1)}$	1579	1483	1481
$\hat{\beta}^{(2)}$	44.2	34.8	35.3
$\hat{\sigma}^2$	7062 ²	6563 ²	6043 ²
$\hat{\phi}_{11}$	109 ²	103 ²	100 ²
$\hat{\phi}_{12}$	4191	1135	1149
$\hat{\phi}_{22}$	25.8 ²	22.3 ²	22.0 ²

These data can be modelled using (1), taking $n = 5$, $n_i = 12$, $m = 2$ and $\mathbf{x}_{ij} = (1, j)'$. The two common mean parameters in $\boldsymbol{\beta} = (\beta^{(1)}, \beta^{(2)})'$ correspond to the expected levels of the intercept and slope respectively. These data have also been studied by Dannenburg et al. (1996) using the Hachemeister method.

Table 7 lists the parameter estimates obtained by the Hachemeister and GEE methods. The two GEE methods give similar estimates which are, however, rather different from those obtained by the Hachemeister method. Similar to the finding observed in Dannenburg et al. (1996), the estimated correlation for $\beta_i^{(1)}$ and $\beta_i^{(2)}$, i.e., $\hat{\phi}_{12} / (\hat{\phi}_{11} \cdot \hat{\phi}_{22})^{1/2}$, is 1.49 which is greater than 1 when Hachemeister's estimation is employed. However, this abnormal phenomenon is not observed in the GEE estimation.

6. CONCLUDING REMARKS

In this paper, we have introduced the use of GEE to estimate structural parameters of a regression credibility model. Such estimators are found to perform well for the Hachemeister model, including an extended case that a moving average error structure is present. Improvement in the estimation of the structural parameters leads to an ameliorated accuracy in estimating the credibility matrices, and, finally, yields the credibility estimators for the regression coefficients with less variability.

Our proposed GEE method may adapt to encapsulating simple conditions stated *a priori* governing the covariance structure of $\boldsymbol{\beta}_i$. Flexibility may increase at the expense of a linear dependence of the form of the covariance matrix \mathbf{F} on unknown parameters. Moreover, though not accentuated in the previous sections, no additional complications are involved in exercising the proposed GEE method to accommodate panel data in the general case that $n_i \neq n_j$ for different i and j . This is in line with the model assumption of the Hachemeister framework.

ACKNOWLEDGEMENTS

The authors thank two referees for valuable comments which largely improved the presentation of the paper. This work was partially supported by a grant from the Hong Kong Research Grants Council HKU 7408/06H. The authors thank Danwei Huang for computing assistance.

REFERENCES

- BÜHLMANN, H. (1967) Experience rating and credibility. *ASTIN Bulletin* **4**, 199-207.
- BÜHLMANN, H. and STRAUB, E. (1970) Glaubwürdigkeit für Schadensätze (Credibility for loss ratios). *Bulletin of the Swiss Association of Actuaries* **70**(1), 111-133.
- BÜHLMANN, H. and GISLER, A. (1997) Credibility in the regression case revisited. *ASTIN Bulletin* **27**, 83-98.
- BÜHLMANN, H. and GISLER, A. (2005) *A Course in Credibility Theory and its Applications*. Springer: Heidelberg.
- BÜHLMANN, P. and BÜHLMANN, H. (1999) Selection of credibility regression models. *ASTIN Bulletin* **29**, 245-270.
- COSSETTE, H. and LUONG, A. (2003) Generalized least squares estimators for covariance parameters for credibility regression models with moving average errors. *Insurance: Mathematics and Economics* **32**, 281-293.
- DANNENBURG, D.R., KAAS, R. and GOOVAERTS, M.J. (1996) *Practical Actuarial Credibility Models*. Institute of Actuarial Science and Econometrics, University of Amsterdam, Amsterdam, the Netherlands.
- FREES, E.W., YOUNG, V.R. and LUO, Y. (1999) A longitudinal data analysis interpretation of credibility models. *Insurance: Mathematics and Economics* **24**, 229-247.
- FULLER, W. (1987) *Measurement Error Models*. Wiley, New York.
- GRAHAM, A. (1981) *Kronecker Products and Matrix Calculus: with Applications*. Horwood, Chichester.
- HACHEMEISTER, C.A. (1975) Credibility for regression models with application to trend. In: Kahn, P.M. (Ed.), *Credibility: Theory and Applications*. Academic Press, New York, pp. 129-163.
- HARDIN, J.W. and HILBE, J.M. (2001) *Generalized Linear Models and Extensions*. Stata Press, Texas.
- HARDIN, J.W. and HILBE, J.M. (2003) *Generalized Estimating Equations*. Chapman and Hall, Boca Raton.
- HE, X., FUNG, W.K. and ZHU, Z.Y. (2005) Robust estimation in generalized partial linear models for clustered data. *Journal of the American Statistical Association* **100**, 1176-1184.
- LIANG, K.Y. and ZEGER, S.L. (1986) Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13-22.
- LIANG, K.Y., ZEGER, S.L. and QAQISH, B. (1992) Multivariate regression analyses for categorical data. *Journal of the Royal Statistical Society Series B*, **54**, 3-40.
- LO, C.H., FUNG, W.K. and ZHU, Z.Y. (2006) Generalized estimating equations for variance and covariance parameters in credibility models. *Insurance: Mathematics and Economics*, **39**, 99-113.
- MCCULLAGH, P. and NELDER, J. (1989) *Generalized Linear Models*. Chapman and Hall, London.
- NELDER, J.A. and VERRALL, R.J. (1997) Credibility theory and generalized linear models. *ASTIN Bulletin* **27**, 71-82.
- NORBERG, R. (1980) Empirical Bayes credibility. *Scandinavian Actuarial Journal*, 177-194.
- NORBERG, R. (1982) On optimal parameter estimation in credibility. *Insurance: Mathematics and Economics* **1**, 73-89.
- NORBERG, R. (1986) Hierarchical credibility: analysis of a random effect linear model with nested classification. *Scandinavian Actuarial Journal*, 204-222.
- PRENTICE, R.L. and ZHAO, L.P. (1991) Estimating equations for parameters in means and covariances of multivariate discrete and continuous responses. *Biometrics* **47**, 825-839.

- RAO, C.R. (1975) Simultaneous estimation of parameters in different linear models and applications to biometric problems. *Biometrics* **31**, 545-554.
- SUNDT, B. (1987) Two credibility regression approaches for the classification of passenger cars in a multiplicative tariff. *ASTIN Bulletin* **17**, 41-70.
- ZEGER, S.L. and LIANG, K.Y. (1986) Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* **42**, 121-130.

WING KAM FUNG

Department of Statistics and Actuarial Science,

The University of Hong Kong

Pokfulam Road, Hong Kong, China

Tel.: +852 2859 1988

Fax: +852 2858 9041

E-mail: wingfung@hku.hk