A Scaled Linear Mixed Model for Multiple Outcomes

Xihong Lin,[1,*] Louise Ryan,[2] Mary Sammel,[3] Daowen Zhang,[4]
Chantana Padungtod,[5] and Xiping Xu[5]

[1]Department of Biostatistics, University of Michigan, Ann Arbor, Michigan 48109, U.S.A.
[2]Department of Biostatistical Science, Dana-Farber Cancer Institute, and
Department of Biostatistics, Harvard University, Boston, Massachusetts 02115, U.S.A.
[3]Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine,
Philadelphia, Pennsylvania 19104, U.S.A.
[4]Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695, U.S.A.
[5]Department of Environmental Health, Harvard University, Boston, Massachusetts 02115, U.S.A.
*email: xlin@sph.umich.edu

SUMMARY. We propose a scaled linear mixed model to assess the effects of exposure and other covariates on multiple continuous outcomes. The most general form of the model allows a different exposure effect for each outcome. An important special case is a model that represents the exposure effects using a common global measure that can be characterized in terms of effect sizes. Correlations among different outcomes within the same subject are accommodated using random effects. We develop two approaches to model fitting, including the maximum likelihood method and the working parameter method. A key feature of both methods is that they can be easily implemented by repeatedly calling software for fitting standard linear mixed models, e.g., SAS PROC MIXED. Compared to the maximum likelihood method, the working parameter method is easier to implement and yields fully efficient estimators of the parameters of interest. We illustrate the proposed methods by analyzing data from a study of the effects of occupational pesticide exposure on semen quality in a cohort of Chinese men.

KEY WORDS: Asymptotic relative efficiency; Effect sizes; Estimating equations; Global effect; Maximum likelihood.

1. Introduction

The problem of analyzing multiple outcomes arises frequently in many fields of biomedical research. For example, in phase II clinical trials, multiple endpoints for treatment efficacy are often obtained (Pocock, Geller, and Tsiatis, 1987), and in teratology, multiple birth defects are often associated with prenatal exposure to some agent (Sammel and Ryan, 1996). In such settings, one is often interested in studying whether various outcomes are affected by exposure to the same degree. This will often be the case, e.g., when different outcomes are essentially measuring the same underlying event (e.g., treatment efficacy or severity of birth defect) from different perspectives. If there is evidence that outcomes are similarly affected by exposure, then it will be of interest to test whether or not the effect differs from zero and to estimate the effect.

Several authors have discussed the problem of constructing a global test for common dose effects on multiple outcomes (O'Brien, 1984; Pocock et al., 1987). In general, one would expect such tests to have good power in settings where exposure tends to have a generalized effect on all the outcomes (Legler, Lefkopoulou, and Ryan, 1995) and to be more powerful than separate tests using individual outcomes. This is because the combined evidence of several outcomes for the exposure effect need not be as extreme as for a single outcome (Pocock et al., 1987). While theory for global testing is well studied, estimation of global effects has received less attention. In the binary data setting, general estimating equation (GEE) methods have been used to estimate common dose effects (Lefkopoulou, Moore, and Ryan, 1989).

Estimation of a common dose effect for continuous data is often complicated by the fact that outcomes are measured on different scales and a global measure of the exposure effects on the original scales of the outcomes could be misleading. One approach is to use the two-stage factor analysis method. Specifically, one first applies factor analysis to multiple outcome variables and calculates factor scores, then regresses them on covariates (Bartholomew, 1987). Sammel and Ryan (1996) considered a latent variable model, where multiple outcomes are modeled as functions of a latent variable and a simple regression model is used to relate the latent variable and the exposure variable. A drawback of this approach is its nonrobustness since the mean parameters depend heavily on the covariance parameters. To address this problem, Sammel, Lin, and Ryan (1999) proposed a multivariate linear mixed model, which maintains some of the features of the latent

variable model while disentangling the mean and covariance parameters.

The example that motivates this paper involves a study of male reproductive health in China (Padungtod et al., 1999). In developing countries, farmers and pesticide factory workers are often exposed to organophosphate pesticides. Although the doses are not typically high enough to be life threatening, many believe that long-term occupational exposure might cause adverse health effects, particularly related to reproductive function. The objective of the Padungtod study was to investigate the effect of occupational organophosphate pesticide exposure on semen quality among Chinese workers.

The study consisted of 43 Chinese male workers, among whom 20 were exposed and 23 were not exposed to pesticides. The 20 exposed workers were randomly chosen from a pesticide factory in Anqing City, China. The 23 unexposed workers had similar work practices to the exposed subjects and were randomly chosen from a nearby textile factory. The two groups of subjects were also comparable in terms of duration of employment and socioeconomic status. Semen samples were collected for the study subjects and analyzed for sperm concentration, percentage of sperm with normal motility movement, and percentage of sperm with normal morphological shape. The purpose of the study was to assess the effect of pesticide exposure on overall semen quality. Investigators were also interested in knowing whether all three semen measures were affected to a similar degree or whether they varied in their sensitivity and responsiveness to exposure. Because the three outcomes were measured on very different scales, standard methods could not easily be applied to address these questions.

In this paper, we propose a scaled linear mixed model for analyzing multiple continuous outcomes. In its most general form, the model allows a different exposure effect for each outcome. An important special case is a model that represents the exposure effects using a common global measure that can be characterized in terms of effect sizes. Correlations among different outcomes within the same subject are accommodated using random effects. Two methods are proposed for parameter estimation. They include the maximum likelihood method and the working parameter method. A key feature of both methods is that they can easily be implemented by repeatedly calling software for fitting standard linear mixed models, such as SAS PROC MIXED. Compared to the maximum likelihood method, the working parameter method is easier to implement yet maintains full efficiency with respect to the parameters of interest. We illustrate the proposed methods with the analysis of the Chinese semen data.

2. The Scaled Linear Mixed Model

Suppose that, for the ith of n subjects, we observe M continuous outcomes $\mathbf{y}_i = (y_{i1}, \ldots, y_{iM})^{\mathrm{T}}$, as well as a vector of covariates $\mathbf{x}_i$ and an exposure variable w_i. Our main interest is to assess an overall exposure effect using the information from the M outcomes. Since the M outcomes are often measured on different scales, we consider a scaled linear mixed model

$$\frac{y_{ij}}{\sigma_j} = \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}_j + w_i\alpha_j + \mathbf{z}_{ij}^{\mathrm{T}}\mathbf{b}_i + \epsilon_{ij}, \tag{1}$$

where $j = 1, \ldots, M$, σ_j is a scale parameter for the jth outcome, $(\boldsymbol{\beta}_j, \alpha_j)$ are regression coefficients, $\mathbf{z}_{ij}$ is a design vec-

tor, $\mathbf{b}_i$ is a $q \times 1$ vector of random effects, and ϵ_{ij} are independent random errors following $N(0,1)$. The $\mathbf{b}_i$ are distributed as $N(0, \mathbf{D}(\theta))$ and are used to model correlation among different outcomes of the same subject and θ is a $c \times 1$ vector of variance components to be estimated. We allow the design vector $\mathbf{z}_{ij}$ to be outcome specific so that different correlations among different outcomes can be assumed, e.g., one could assume a factor-analytic type correlation.

Model (1) is appealing in that it is simply a linear mixed model for outcomes standardized by error standard deviations. This most general formulation will yield results similar to those obtained by fitting a separate model for each outcome. Slight differences are to be expected, however, due to the fact that the model allows outcomes to be correlated. The real value of the formulation can be seen, however, by considering restricted submodels. In particular, imposing the restriction $\alpha_1 = \alpha_2 = \cdots = \alpha_M = \alpha$ allows one to test for a common exposure effect on the standardized (by error standard deviations) outcomes. From an applied perspective, the parameter α can be thought of as the common effect size, an attractive interpretation to practitioners. Another feature is that the exposure effect can be detected with good statistical power using a 1 d.f. global test statistic.

Let $\Psi = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_M^2)$, $\mathbf{X}_i = \{\mathbf{x}_i^{\mathrm{T}} \otimes \mathbf{I}, w_i\mathbf{I}\}$, $\gamma = (\boldsymbol{\beta}_1^{\mathrm{T}}, \ldots, \boldsymbol{\beta}_M^{\mathrm{T}}, \alpha^{\mathrm{T}})^{\mathrm{T}}$, $\alpha = (\alpha_1, \ldots, \alpha_M)^{\mathrm{T}}$, $\mathbf{Z}_i = (\mathbf{z}_{i1}, \ldots, \mathbf{z}_{iM})^{\mathrm{T}}$, $\mathbf{y}_i = (y_{i1}, \ldots, y_{iM})^{\mathrm{T}}$, and $\epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{iM})^{\mathrm{T}}$, where $\otimes$ denotes a direct product and $\mathbf{I}$ is an $M \times M$ identity matrix. Model (1) can be succinctly written in a matrix form as

$$\Psi^{-1/2}\mathbf{y}_i = \mathbf{X}_i\gamma + \mathbf{Z}_i\mathbf{b}_i + \epsilon_i. \tag{2}$$

It follows that the marginal distribution of $\mathbf{y}_i$ is $N(\Psi^{1/2}\mathbf{X}_i\gamma, \Psi^{1/2}\mathbf{V}_i(\theta)\Psi^{1/2})$, where $\mathbf{V}_i(\theta) = \mathbf{Z}_i\mathbf{D}(\theta)\mathbf{Z}_i^{\mathrm{T}} + \mathbf{I}$. The log-likelihood function of $(\gamma, \theta, \sigma^2)$, where $\sigma^2 = (\sigma_1^2, \ldots, \sigma_M^2)^{\mathrm{T}}$, is

$$\sum_{i=1}^{n} \left\{ -\frac{M}{2}\ln 2\pi - \frac{1}{2}\ln|\Psi| - \frac{1}{2}\ln|\mathbf{V}_i| \right.$$
$$\left. -\frac{1}{2}\left(\Psi^{-1/2}\mathbf{y}_i - \mathbf{X}_i\gamma\right)^{\mathrm{T}}\mathbf{V}_i(\theta)^{-1}\left(\Psi^{-1/2}\mathbf{y}_i - \mathbf{X}_i\gamma\right) \right\}. \tag{3}$$

Since the scale parameter matrix Ψ enters into both the marginal mean and covariance of $\mathbf{y}_i$, standard linear mixed model machinery cannot be used directly to fit the scaled linear mixed model (1). A new estimation procedure is needed.

3. The Estimation Procedures

In this section, we develop two simple iterative algorithms using the maximum likelihood method (Section 3.1) and the working parameter method (Section 3.2) to fit the general model in (1). We will show how these algorithms can be easily implemented by repeatedly calling existing software, such as SAS PROC MIXED, for fitting standard linear mixed models. The latter feature makes the approach particularly attractive to practitioners.

3.1 *The Maximum Likelihood Method*

Examination of equation (2) suggests that, if σ^2 were known, one could easily calculate the maximum likelihood estimates (MLEs) of γ and θ by fitting the linear mixed model

$$\mathbf{y}_i^* = \mathbf{X}_i\gamma + \mathbf{Z}_i\mathbf{b}_i + \epsilon_i, \qquad (4)$$

where $\mathbf{y}_i^* = \Psi^{-1/2}\mathbf{y}_i$ denotes the standardized (by error standard deviations) $\mathbf{y}_i$ and $\epsilon_i \sim N(0, \mathbf{I})$. Hence, the difficulty in fitting the scaled linear mixed model (2) mainly lies in calculating the MLE of σ^2. These observations are supported by deriving the score equations of $(\gamma, \theta, \sigma^2)$.

Specifically, differentiation of the log-likelihood function (3) with respect to $(\gamma, \theta, \sigma^2)$ shows that the maximum likelihood estimators $(\hat{\gamma}, \hat{\theta}, \hat{\sigma}^2)$ solve the score equations

$$\sum_{i=1}^n \mathbf{X}_i^{\mathrm{T}} \mathbf{V}_i^{-1}(\mathbf{y}_i^* - \mathbf{X}_i\gamma) = 0 \quad (5)$$

$$\sum_{i=1}^n \left\{ -\frac{1}{2}\mathrm{tr}\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \right.$$
$$\left. +\frac{1}{2}(\mathbf{y}_i^* - \mathbf{X}_i\gamma)^{\mathrm{T}}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_k}\mathbf{V}_i^{-1}(\mathbf{y}_i^* - \mathbf{X}_i\gamma) \right\} = 0 \quad (6)$$

$$\sum_{i=1}^n \left\{ -\frac{1}{2\sigma_j^2} + \frac{1}{2\sigma_j^2}\mathbf{y}_i^{*\mathrm{T}}\Delta_j\mathbf{V}_i^{-1}(\mathbf{y}_i^* - \mathbf{X}_i\gamma) \right\} = 0, \quad (7)$$

where Δ_j is a diagonal matrix with the jth diagonal element equal to one and the other diagonal elements equal to zero and $\mathbf{y}_i^*$ in (5)–(7) contains the unknown parameters σ_j^2. It can be easily seen that, for given σ^2, the score equations for γ and θ in (5) and (6) coincide with those under model (4). Using the identity $\mathbf{V}_i^{-1}(\mathbf{y}_i^* - \mathbf{X}_i\hat{\gamma}) = (\mathbf{y}_i^* - \mathbf{X}_i\hat{\gamma} - \mathbf{Z}_i\hat{\mathbf{b}}_i)$ (Harville, 1977), where $\hat{\mathbf{b}}_i$ is the best linear unbiased predictor (BLUP) of $\mathbf{b}_i$ from fitting model (4), we can rewrite the score equation for σ_j^2 in (7) as $\sigma_j^2 = \sigma_j^2(n^{-1}\Sigma_{i=1}^n y_{ij}^* r_{ij})$, where r_{ij} is the jth component of the residual vector $\mathbf{r}_i = \mathbf{y}_i^* - \mathbf{X}_i\hat{\gamma} - \mathbf{Z}_i\hat{\mathbf{b}}_i$ for the ith subject and can be easily calculated after fitting model (4).

These results suggest that the MLEs $(\hat{\gamma}, \hat{\theta}, \hat{\sigma}^2)$ can be calculated using the following iterative algorithm, which repeatedly calls existing software for fitting standard linear mixed models. This algorithm can be viewed as solving the score equations (5)–(7) using a Gauss–Seidel algorithm (Lange, 1999).

(1) Set initial values for σ^2, e.g., σ_j^2 = sample variance of y_{ij}.
(2) Calculate $y_{ij}^* = y_{ij}/\sigma_j$.
(3) Estimate γ and θ by fitting the linear mixed model (4), with the variances of the error terms ϵ_i constrained to be one, e.g., using SAS PROC MIXED with the PARM statement.
(4) Update σ_j^2 using $\sigma_{j,\mathrm{new}}^2 = \sigma_{j,\mathrm{old}}^2(n^{-1}\Sigma_{i=1}^n y_{ij}^* r_{ij})$, where the residuals $\mathbf{r}_i$ are by-products of step 3 from fitting model (4) and can be easily obtained from SAS PROC MIXED.
(5) Check for convergence. If $|n^{-1}\Sigma_{i=1}^n y_{ij}^* r_{ij} - 1|$ is less than a prespecified convergence criterion, then stop; otherwise, go back to step 2.

It is of interest to study the properties of the MLEs $(\hat{\gamma}, \hat{\theta}, \hat{\sigma}^2)$. Unlike the MLEs in standard linear mixed models, the MLEs of the regression coefficients γ and the variance components (θ, σ^2) in the scaled linear mixed model (2) are not asymptotically orthogonal. In other words,

$$\lim_{n\to\infty} n\,\mathrm{cov}(\hat{\gamma}, \hat{\theta}) \neq 0$$

and

$$\lim_{n\to\infty} n\,\mathrm{cov}(\hat{\gamma}, \hat{\sigma}^2) \neq 0.$$

This can be shown by explicitly calculating the asymptotic covariance matrix of the MLE $(\hat{\gamma}, \hat{\theta}, \hat{\sigma}^2)$ (see Appendix A.1). An attractive feature is that $\mathrm{cov}(\hat{\gamma})$ and $\mathrm{cov}(\hat{\theta})$ do not depend on σ^2 (Appendix A.3). However, $\mathrm{cov}(\hat{\gamma}) \neq (\Sigma_{i=1}^n \mathbf{X}_i\mathbf{V}_i^{-1}\mathbf{X}_i)^{-1}$, the usual covariance of the fixed effects from a standard mixed model, but involves an adjustment factor. This can be easily seen from equation (A.1).

3.2 The Working Parameter Method

A key property of the MLEs $(\hat{\gamma}, \hat{\theta}, \hat{\sigma}^2)$ is that they are asymptotically most efficient. However, its implementation could be complicated. For example, when using the iterative procedure in Section 3.1, one needs to constrain the variances of ϵ_i. This may not be allowed in some software, such as the S-plus function LME. Another complication is that some programming is needed when we update σ^2. In this section, we propose a simpler iterative algorithm by introducing working parameters. We first describe the working parameter algorithm and then show that it yields consistent estimates of $(\gamma, \theta, \sigma^2)$.

The main idea of this approach is to fit model (4) without constraining the variances of the ϵ_i to be ones and then to properly update the scale parameters σ^2 using the estimated measurement error variances, which are used as working parameters. This algorithm is given as follows.

(1) Set initial values of σ^2, e.g., σ_j^2 = sample variance of y_{ij}.
(2) Calculate $y_{ij}^* = y_{ij}/\sigma_j$.
(3) Fit the linear mixed model

$$\mathbf{y}_i^* = \mathbf{X}_i\gamma + \mathbf{Z}_i\mathbf{b}_i + \mathbf{e}_i, \qquad (8)$$

where $\mathbf{e}_i \sim N(0, \mathrm{diag}(\tau))$ and $\tau = (\tau_1, \ldots, \tau_M)^{\mathrm{T}}$. Note that no constraint is placed on τ when we fit model (8). We refer to τ as working parameters. This fitting gives estimators of (γ, θ, τ).
(4) Update σ_j^2 using $\sigma_{j,\mathrm{new}}^2 = \sigma_{j,\mathrm{old}}^2\tau_j$ $(j = 1, \ldots, M)$.
(5) Check for convergence. If $|\tau_j - 1|$ is less than a prespecified convergence criterion for all $j = 1, \ldots, M$, then stop; otherwise, go back to step 2.

Compared to the MLE iterative algorithm discussed in Section 3.1, the above algorithm is simpler and requires less programming. We now investigate the properties of these estimators. This can be done by studying the estimating equations for (γ, θ, τ) that are implicitly being solved by the above procedure at convergence, i.e., when $\tau_j = 1$ $(j = 1, \ldots, M)$,

$$\sum_{i=1}^n \mathbf{X}_i^{\mathrm{T}} \mathbf{V}_i^{-1}(\mathbf{y}_i^* - \mathbf{X}_i\gamma) = 0 \quad (9)$$

$$\sum_{i=1}^n \left\{ -\frac{1}{2}\mathrm{tr}\left(\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \right.$$
$$\left. +\frac{1}{2}(\mathbf{y}_i^* - \mathbf{X}_i\gamma)^{\mathrm{T}}\mathbf{V}_i^{-1}\frac{\partial \mathbf{V}_i}{\partial \theta_k}\mathbf{V}_i^{-1}(\mathbf{y}_i^* - \mathbf{X}_i\gamma) \right\} = 0 \quad (10)$$

$$\sum_{i=1}^{n}\left\{-\frac{1}{2}\mathrm{tr}\left(\mathbf{V}_i^{-1}\Delta_j\right)\right.$$

$$\left.+\frac{1}{2}(\mathbf{y}_i^* - \mathbf{X}_i\gamma)^{\mathrm{T}}\mathbf{V}_i^{-1}\Delta_j\mathbf{V}_i^{-1}(\mathbf{y}_i^* - \mathbf{X}_i\gamma)\right\} = 0. \quad (11)$$

Note that the above equations are derived by letting τ equal one in the standard score equations under the linear mixed model (8) with τ as free parameters (Harville, 1977). The estimating equations for γ and θ in (9) and (10) are identical to those used to derive the MLEs (equations (5) and (6)). However, the estimating equation of σ^2 in (11) is different from its MLE counterpart in (7). Although our working parameter method does not yield the MLEs of $(\gamma, \theta, \sigma^2)$, it can be easily shown that equation (11) is an unbiased estimating equation for σ^2 under model (1). Straightforward application of the standard estimating equation theory (Foutz, 1977) then shows that this algorithm yields consistent and asymptotically normal estimators of $(\gamma, \theta, \sigma^2)$.

Specifically, denote the set of the estimating equations (9)–(11) by $\mathbf{U} = (\mathbf{U}_\gamma^{\mathrm{T}}, \mathbf{U}_\theta^{\mathrm{T}}, \mathbf{U}_{\sigma^2}^{\mathrm{T}})^{\mathrm{T}}$ and their solution by $\tilde{\xi} = (\tilde{\gamma}^{\mathrm{T}}, \tilde{\theta}^{\mathrm{T}}, \tilde{\sigma^2}^{\mathrm{T}})^{\mathrm{T}}$. Then we have $\tilde{\xi}$ is asymptotically normally distributed with mean equal to the true value of $\xi = (\gamma^{\mathrm{T}}, \theta^{\mathrm{T}}, \sigma^{2\mathrm{T}})^{\mathrm{T}}$ and a sandwich covariance matrix $[\mathrm{E}(-\partial\mathbf{U}/\partial\xi^{\mathrm{T}})]^{-1}\mathrm{cov}(\mathbf{U})[\mathrm{E}(-\partial\mathbf{U}^{\mathrm{T}}/\partial\xi)]^{-1}$. A detailed expression of this covariance matrix is given in Appendix A.2. Like the MLEs $(\hat{\gamma}, \hat{\theta}, \hat{\sigma^2})$, $\tilde{\gamma}$ and $(\tilde{\theta}, \tilde{\sigma^2})$ are not asymptotically orthogonal, i.e., $\lim_{n\to\infty} n\,\mathrm{cov}[\tilde{\gamma}, (\tilde{\theta}, \tilde{\sigma^2})] \neq 0$ and $\mathrm{cov}(\tilde{\gamma}) \neq [\Sigma_{i=1}^{n}\mathbf{X}_i\mathbf{V}_i^{-1}\mathbf{X}_i]^{-1}$. However, $\mathrm{cov}(\tilde{\gamma})$ and $\mathrm{cov}(\tilde{\theta})$ do not depend on σ^2 (Appendix A.3).

3.3 *Study of the Asymptotic Relative Efficiency*

Although the estimator $\tilde{\xi} = (\tilde{\gamma}, \tilde{\theta}, \tilde{\sigma^2})$ yielded by the working parameter (WP) approach discussed in Section 3.2 is consistent, it is not the MLE and therefore may be inefficient. Hence, it is of interest to study the loss of information and its asymptotic relative efficiency (ARE) with respect to the MLE $\hat{\xi} = (\hat{\gamma}, \hat{\theta}, \hat{\sigma^2})$.

Following Bhapkar (1991), we define the information matrix of the working parameter estimator $\tilde{\xi}$ as $\tilde{\mathcal{I}} = \mathrm{E}(-\partial\mathbf{U}^{\mathrm{T}}/\partial\xi)\mathrm{cov}^{-1}(\mathbf{U})\mathrm{E}(-\partial\mathbf{U}/\partial\xi^{\mathrm{T}})$. The results in Appendices A.1 and A.2 show that the information matrix $\mathcal{I}$ of the MLE $\hat{\xi}$ and the information matrix $\tilde{\mathcal{I}}$ of $\tilde{\xi}$ are

$$\mathcal{I} = \begin{bmatrix} \mathcal{I}_{11} & 0 & \mathcal{I}_{13} \\ 0 & \mathcal{I}_{22} & \mathcal{I}_{23} \\ \mathcal{I}_{13}^{\mathrm{T}} & \mathcal{I}_{23}^{\mathrm{T}} & \mathcal{I}_{33} \end{bmatrix}$$

and

$$\tilde{\mathcal{I}} = \begin{bmatrix} \mathcal{I}_{11} & 0 & \mathcal{I}_{13} \\ 0 & \mathcal{I}_{22} & \mathcal{I}_{23} \\ \mathcal{I}_{13}^{\mathrm{T}} & \mathcal{I}_{23}^{\mathrm{T}} & \tilde{\mathcal{I}}_{33} \end{bmatrix}, \quad (12)$$

where the expressions of $\mathcal{I}_{jk}$ and $\tilde{\mathcal{I}}_{33}$ are given in Appendices A.1 and A.2. A comparison of $\mathcal{I}$ and $\tilde{\mathcal{I}}$ shows that the loss of information of the working parameter estimator $\tilde{\xi}$ lies only in estimation of σ^2, i.e., $\mathcal{I} - \tilde{\mathcal{I}}$ is zero except for the last block diagonal element, which equals $\mathcal{I}_{33} - \tilde{\mathcal{I}}_{33}$ and is a positive definite matrix. Since $\mathcal{I}_{13}$ and $\mathcal{I}_{23}$ do not equal zero, the estimators of γ and θ are not asymptotically orthogonal to

the estimators of σ^2. This means the loss of information in estimating σ^2 could affect the AREs of the WP estimators $(\tilde{\gamma}, \tilde{\theta})$ with respect to the MLEs $(\hat{\gamma}, \hat{\theta})$, i.e., $(\tilde{\gamma}, \tilde{\theta})$ may not be fully efficient relative to the MLEs $(\hat{\gamma}, \hat{\theta})$.

Rather than studying the global ARE of $\tilde{\xi}$ with respect to $\hat{\xi}$, we are more interested in studying the AREs of each of the three parameter estimators $(\gamma, \theta, \sigma^2)$ separately, which are defined as (Serfling, 1980, Section 4.1.2)

$$\mathrm{ARE}(\tilde{\eta}, \hat{\eta}) = \left[\lim_{n\to\infty} \frac{|\mathrm{cov}(\hat{\eta})|}{|\mathrm{cov}(\tilde{\eta})|}\right]^{1/d}, \quad (13)$$

where η equals γ, θ or σ^2; d is the dimension of η and equals p, c, M for γ, θ, σ^2, respectively; and p is the dimension of γ. The covariances $\mathrm{cov}(\hat{\gamma}), \mathrm{cov}(\hat{\theta}), \mathrm{cov}(\hat{\sigma^2})$ can be obtained using the partitioned inverse information matrix $\mathcal{I}$ in Appendix A.1, while $\mathrm{cov}(\tilde{\gamma}), \mathrm{cov}(\tilde{\theta}), \mathrm{cov}(\tilde{\sigma^2})$ can be obtained using the partitioned sandwich covariance matrix in Appendix A.2. Their forms are given in Appendix A.3. Using the results in Appendices A.1 and A.2, we can show that AREs in (13) have the following attractive property.

PROPOSITION 1: $\mathrm{ARE}(\tilde{\gamma}, \hat{\gamma})$, $\mathrm{ARE}(\tilde{\theta}, \hat{\theta})$, *and* $\mathrm{ARE}(\tilde{\sigma^2}, \hat{\sigma^2})$ *are independent of* σ^2.

Proof. See Appendix A.3.

It is difficult to study the loss of efficiency of the WP estimators $(\tilde{\gamma}, \tilde{\theta}, \tilde{\sigma^2})$ relative to their MLE counterparts under the general model (1). We specialize to the scaled random intercept model with a common dose effect,

$$y_{ij}/\sigma_j = \mathbf{x}_i^{\mathrm{T}}\beta + w_i\alpha + b_i + \epsilon_{ij},$$

where $b_i \sim N(0, \theta)$ and ϵ_{ij} are independent $N(0, 1)$. Straightforward, though tedious, calculations show that, under this scaled random intercept model, the key parameters of interest under the working parameter method $(\tilde{\alpha}, \tilde{\theta})$ are fully efficient compared to their MLE counterparts for any $\mathbf{x}_i$ and a binary exposure variable w_i, but the remaining parameters $(\tilde{\beta}, \tilde{\sigma^2})$ under the working parameter method may not be fully efficient.

To numerically study the loss of efficiency of estimation of the remaining parameters (β, σ^2) under the working parameter method, we consider the simple model

$$y_{ij}/\sigma_j = \beta_{0j} + w_i\alpha + b_i + \epsilon_{ij}, \quad (14)$$

where w_i is a binary exposure variable taking the value zero for half of the n subjects and one for the other. Using the results in Appendices A.1 and A.2, we can easily derive closed-form expressions of $\mathrm{ARE}(\tilde{\beta}, \hat{\beta})$ and $\mathrm{ARE}(\tilde{\sigma^2}, \hat{\sigma^2})$ as functions of (γ, θ).

We provide in Table 1 numerical values of these AREs for a wide variety of configurations of (β, α, θ). Note that we do not need to specify the values of σ since Proposition 1 shows that the AREs do not depend on σ. To mimic the Chinese semen data, we assumed three outcomes $(M = 3)$. Our results suggest that the efficiency of $\tilde{\beta}$ and $\tilde{\sigma^2}$ depends on the parameter configurations and loss of efficiency is observed when (β, α) and θ are large. In the interest of space, we only report cases with positive values of β and α in Table 1. Similar results are obtained when β and α are negative. The table

Table 1

Asymptotic relative efficiency (ARE) of the WP estimates $(\tilde{\beta}, \tilde{\sigma}^2)$ with respect to the MLEs $(\hat{\beta}, \hat{\sigma}^2)$ under model (14)

			ARE	
θ	β	α	β	σ^2
0.25	$(0,1,2)$	0	0.99	0.98
0.25	$(0,1,2)$	1	0.96	0.91
0.25	$(0,1,2)$	5	0.53	0.39
0.25	$(3,4,5)$	0	0.99	0.98
0.25	$(3,4,5)$	1	0.92	0.91
0.25	$(3,4,5)$	5	0.43	0.39
1.00	$(0,1,2)$	0	0.97	0.88
1.00	$(0,1,2)$	1	0.93	0.83
1.00	$(0,1,2)$	5	0.54	0.34
1.00	$(3,4,5)$	0	0.89	0.88
1.00	$(3,4,5)$	1	0.85	0.83
1.00	$(3,4,5)$	5	0.44	0.40
5.00	$(0,1,2)$	0	0.89	0.52
5.00	$(0,1,2)$	1	0.81	0.51
5.00	$(0,1,2)$	5	0.51	0.34
5.00	$(3,4,5)$	0	0.61	0.52
5.00	$(3,4,5)$	1	0.59	0.51
5.00	$(3,4,5)$	5	0.40	0.34

does not include any results for the exposure effect α or variance components θ since estimates of these parameters are fully efficient under the working parameter method. In settings like ours where primary interest lies in estimating the exposure effect α, our results suggest that the simple working parameter approach provides an effective and efficient alternative to MLE estimation.

4. Application to the Chinese Semen Data

We fit scaled linear mixed models to the Chinese semen data introduced in Section 1 using the MLE method and the working parameter method discussed in Sections 3.1 and 3.2. The three outcome variables used to measure semen quality include sperm concentration (CONCEN), percentage of sperm with normal motility (MOTIL), and percentage of sperm with normal morphology (MORPH). Our main interest was to study the effect of occupational organophosphate pesticide exposure on overall semen quality with the aim of constructing a global measure of the exposure effect and

developing a global test. The exposure variable (EXP) was an indicator of whether a worker was employed at the pesticide factory. The covariates of interest were age (AGE) and sexual abstinence period (ABSTIN).

We first examined the data using descriptive statistics. Examination of the distributions of the three outcome variables showed that the distribution of CONCEN was somewhat skewed, while the other two variables appeared to be approximately normally distributed. Hence, we took a log transformation of CONCEN to make the normality assumption more plausible, calling the resulting variable LNCONCEN. Table 2 gives the exposure-specific mean and standard deviation of each outcome. One can easily see that the scales of the three outcomes differ substantially and hence the crude exposure effects represented by the mean differences vary dramatically among the three outcomes. To explore whether standardization could lead to a similar degree of the exposure effects on the three outcomes, we standardized each outcome by its sample standard deviation calculated by pooling the data over the two exposure groups, calling the resulting variables S_LNCONCEN, S_MOTIL, S_MORPH. Their means and the standard deviations are also given in Table 2. These results suggest that a common exposure effect on the scaled outcomes may well be plausible. This observation is further supported by the boxplots of the standardized outcomes, presented in Figure 1. Note that we have not adjusted for the possible confounding effects of AGE and ABSTIN.

We began our formal statistical analysis by fitting a scaled linear mixed model with different exposure effects for each outcome. For the jth outcome measured on the ith subject $(i = 1, \ldots, 43)$, the model can be written as

$$y_{ij}/\sigma_j = \beta_{0j} + \beta_{1j}\text{AGE}_i + \beta_{2j}\text{ABSTIN}_i + \alpha_j\text{EXP}_i + b_i + \epsilon_{ij},$$
(15)

where y_{ij} $(j = 1,2,3)$ denotes LNCONCEN, MOTIL, MORPH, respectively, $\text{EXP}_i = 1$ if a worker was exposed to pesticides and zero otherwise, the random intercept $b_i \sim N(0, \theta)$, and ϵ_{ij} are independent $N(0,1)$. For simplicity, we assume in model (15) a simple random intercept model, which assumes equal correlation among the three outcomes. To examine this assumption, we calculated the sample correlation matrix. The sample correlations among the three outcomes

Table 2

Exposure-specific summary statistics for the three semen measures

	Unexposed ($n = 23$)		Exposed ($n = 20$)		
	Mean	SD	Mean	SD	Mean difference
CONCEN	74.51	43.64	42.92	22.49	−31.59
LNCONCEN	4.14	0.60	3.58	0.66	−0.56
MOTIL	57.22	13.69	47.25	15.47	−9.97
MORPH	61.39	8.93	57.22	8.67	−4.17
S_LNCONCEN	6.07	0.87	5.25	0.97	−0.82
S_MOTIL	3.77	0.90	3.10	1.01	−0.67
S_MORPH	6.87	0.99	6.37	0.97	−0.50

S_LNCONCEN = LNCONCEN/pooled sample SD; S_MOTIL = MOTIL/pooled sample SD; S_MORPH = MORPH/pooled sample SD.
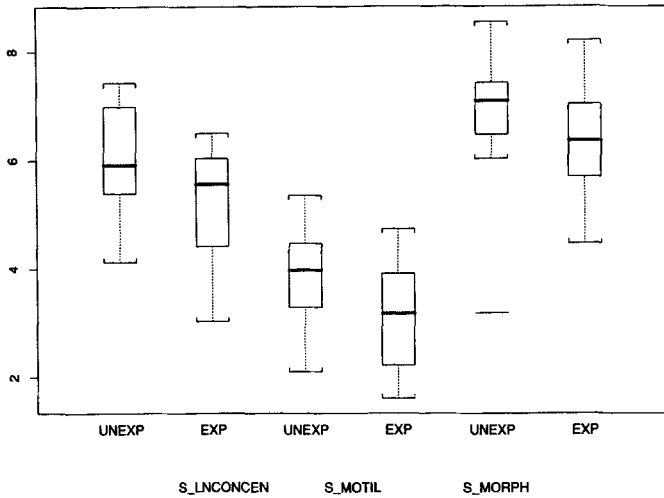
Figure 1. Boxplots of the three standardized semen measures (S_LNCONCEN, S_MOTIL, S_MORPH) stratified by exposure status, where a log transformation was taken for CONCEN and the sample standard deviation was used to standardize each measure.

(LNCONCEN, MOTIL, MORPH) are 0.18, 0.23, 0.40. This seems to indicate the same correlation between LNCONCEN and the other two variables, but the correlation between MOTIL and MORPH might be different. To examine this, we added another random effect shared only by the last two variables MOTIL and MORPH to (15) as

$$y_{ij}/\sigma_j = \beta_{0j} + \beta_{1j}\text{AGE}_i + \beta_{2j}\text{ABSTIN}_i$$
$$+ \alpha_j\text{EXP}_i + b_i + \tilde{z}_{ij}^{\mathrm{T}}\tilde{b}_i + \epsilon_{ij}, \qquad (16)$$

where $\tilde{z}_{ij} = (0,1,1)^{\mathrm{T}}$ and $\tilde{b}_i \sim N(0,\tilde{\theta})$ and is independent of b_i. This model captures the above observed correlation

structure. A likelihood ratio test was performed to test for H_0: $\tilde{\theta} = 0$ and gave a p-value 0.12. Note that the null hypothesis is on the boundary of the parameter space and the likelihood ratio test follows a 50:50 mixture of χ_0^2 and χ_1^2 (Self and Liang, 1987). This suggests that the equal-correlation assumption under (15) is plausible.

The MLEs of the individual exposure effects α_j are -0.97 (SE $= 0.37$), -0.92 (SE $= 0.37$), and -0.67 (SE $= 0.36$). We next fit a common exposure effect model by imposing the restriction $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$. Comparing the heterogeneous to the homogeneous exposure effect model yielded a likelihood ratio test statistic 0.54 (d.f. $= 2$, p-value $= 0.76$), suggesting quite strongly that a common exposure effect assumption is adequate.

Table 3 presents the estimates and the standard errors from fitting the common dose effect model using the ML method and the WP method. Both methods estimate the regression coefficient α as -0.85 (SE $= 0.25$), which measures the global exposure effect on the scaled outcomes and can be interpreted in terms of effect size. This result shows that subjects exposed to pesticides have significantly lower semen quality compared to those who are not exposed (p-value $= 0.001$). Specifically, the mean of each of the three semen quality measures of the exposure group is 0.85 error standard deviations less than that of the unexposed group. The exposure effects on the original scales of the semen outcomes LNCONCEN, MOTIL, MORPH can be estimated using $\alpha\sigma_j$ and are $-0.47, -10.18, -6.17$ using the ML method and $-0.48, -10.04, -6.14$ using the WP method. The coefficients of age and sexual abstinence period are not statistically significant and indicate these two variables do not have a significant impact on semen quality, at least for our data set.

A comparison of the MLEs and the WP method estimates in Table 2 suggests that the simple WP method yields virtually identical estimates of the regression coefficients β and α and the variance component θ. Their standard errors are also

Table 3
The MLEs and the estimates using the working parameter method (WPM) from application to the Chinese semen data

		MLE			WPM		
		Estimate	SE	Naive SE	Estimate	SE	Naive SE
Intercept	β_{01}	7.40	1.03	0.64	7.26	1.04	0.64
	β_{02}	5.58	0.88	0.64	5.66	0.90	0.64
	β_{03}	8.26	1.11	0.64	8.30	1.14	0.64
Age	β_{11}	-0.01	0.02	0.02	-0.01	0.02	0.02
	β_{12}	-0.02	0.02	0.02	-0.02	0.02	0.02
	β_{13}	0.02	0.02	0.02	0.02	0.02	0.02
Abstinence	β_{21}	0.03	0.03	0.03	0.02	0.03	0.03
	β_{22}	-0.05	0.03	0.03	-0.06	0.03	0.03
	β_{23}	-0.05	0.03	0.03	-0.05	0.03	0.03
Exposure	α	-0.85	0.25	0.24	-0.85	0.25	0.24
	θ	0.27	0.16	0.13	0.28	0.16	0.14
	σ_1^2	0.31	0.07		0.32	0.07	
	σ_2^2	143.92	32.27		139.43	32.18	
	σ_3^2	52.74	11.83		52.13	12.03	
Log likelihood		-362.93			-362.96		

almost identical. This result is consistent with our theoretical finding in Section 3.3 and suggests that the WP method estimates of the regression coefficients and the variance components are highly efficient compared to their MLE counterparts. The estimates of the scale parameters σ^2 using the two methods are slightly different. The standard errors of the σ^2 estimates using the WP method are slightly larger than those of the MLEs. This result is consistent with our theoretical finding. The log likelihood of the estimates using the WP method is slightly lower than that of the MLEs (Table 3).

Although one can estimate the model parameters under the two methods by repeatedly calling software, such as SAS PROC MIXED, for fitting linear mixed models, additional programming is needed to compute the correct standard errors. (SAS macros for implementing these two methods and calculating the standard errors are available from the authors on request.) It is hence of interest to study the behavior of the naive standard errors output from SAS at convergence. The naive covariance matrices of the MLEs of (γ, θ) are calculated as $\mathcal{I}_{11}^{-1}$ and $\mathcal{I}_{22}^{-1}$, while the naive covariance matrices of the working parameter estimates of (γ, θ) are calculated as the $(p + c) \times (p + c)$ upper block diagonal matrix of $\mathbf{B}^{-1}$, where $\mathcal{I}_{11}^{-1}$, $\mathcal{I}_{22}^{-1}$, and $\mathbf{B}$ are defined in Appendices A.1 and A.2. We present in Table 3 these naive standard error estimates for the Chinese semen data. It is interesting to note that these naive standard errors perform reasonably well except for the intercepts.

5. Discussion

We have proposed a scaled linear mixed model for multiple continuous outcomes. In its most general form, the model allows for a different exposure effect on each outcome. By comparing this model to one that specifies a common exposure effect, we can test for heterogeneity of the exposure effects. The common dose effects model provides an appealing global measure of the exposure effects that can be characterized in terms of effect sizes. In both cases, correlations among different outcomes measured on the same subject are accommodated using random effects. Our model is especially powerful for detecting and estimating the exposure effect when all outcomes affect the exposure to a similar degree.

Sometimes different outcomes are affected by the exposure to different degrees. One way to model this is to use our heterogeneous exposure effect model to report individual exposure effect estimates and use the Bonferroni adjustment. However, it has been found that the Bonferroni adjustment is often too conservative in analyzing multiple outcomes (Saviltz and Olshan, 1995). Alternatively, one can adopt other models, such as latent variable models (Sammel and Ryan, 1996). Our general model (1) allows specifying flexible correlation structures. In order to specify an appropriate correlation structure, it would be helpful to first examine the sample correlation matrix of the outcomes.

We scale the outcomes using the unknown error standard deviations σ_j^2. There are also other scaling methods. One approach is to fit standard linear mixed models assuming different exposure effects and then to standardize the estimated regression coefficients. In contrast to our method, this *ad hoc* method does not provide a global measure of the exposure effects and the interpretation of the resulting standardized coefficients is not clear. It is also less powerful for detecting

the global exposure effect than our method. An alternative approach is to standardize each outcome by its sample standard deviation before fitting a linear mixed model with a common exposure effect. A major drawback of this approach is that the sample standard deviation estimates are inappropriate since subjects are from heterogeneous populations (e.g., exposed and unexposed groups). They have different covariate values (e.g, different exposure status) and have different mean values.

We have proposed fitting this model using either the maximum likelihood method or the working parameter method. Both methods can be easily implemented by repeatedly calling software, such as SAS PROC MIXED, for fitting standard linear mixed models. Compared to the ML method, the working parameter method is easier to implement. Our results show that the estimators of the key regression coefficients and the variance components yielded by the working parameter method are highly efficient compared to their MLE counterparts. The estimators of the scale parameters using the working parameter method can be less efficient compared to the MLEs. Hence, if one is also interested in the scale parameters, it would be a better strategy to use the MLEs.

RÉSUMÉ

Nous proposons un modèle linéaire mixte standardisé pour évaluer les effets de l'exposition et des autres covariables sur des réponses multiples continues. La forme la plus générale du modèle permet un facteur d'exposition différent pour chaque réponse. Un cas particulier important est un modèle qui représente les effets de l'exposition en utilisant une mesure commune globale qui peut être caractérisée en terme d'effet taille. Les corrélations entre les différentes réponses pour un même sujet sont prises en compte par des effets aléatoires. Nous développons deux approches pour l'ajustement du modèle, incluant la méthode du maximum de vraisemblance et la méthode du paramètre de travail. Une propriété clé des deux méthodes est qu'elles peuvent être facilement implémentées en appelant de façon répétée un logiciel pour ajuster un modèle linéaire à effets mixtes standard, par exemple SAS PROC MIXED. Comparé à la méthode du maximum de vraisemblance, la méthode du paramètre de travail est plus facile à implémenter et aboutit à des estimateurs complètement efficients des paramètres d'intérêt. Nous illustrons les méthodes proposées en analysant les données d'une étude des effets de l'exposition professionnelle aux pesticides sur la qualité du sperme dans une cohorte d'hommes chinois.

REFERENCES

Bartholomew, D. J. (1987). *Latent Variable Models and Factor Analysis*. Oxford: Oxford University Press.

Bhapkar, V. P. (1991). Sufficiency, ancillarity, and information in estimating functions. In *Estimating Functions*, V. P. Godambe (ed), 242–254. Oxford: Oxford University Press.

Foutz, R. V. (1977). On the unique consistent solution to the likelihood equations. *Journal of the American Statistical Association* **72**, 147–148.

Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association* **72,** 320–340.

Lange, K. (1999). *Numerical Analysis for Statisticians.* New York: Springer.

Lefkopoulou, M., Moore, D., and Ryan, L. (1989). The analysis of multiple correlated binary outcomes: Application to rodent teratology experiments. *Journal of the American Statistical Association* **84,** 810–815.

Legler, J., Lefkopoulou, M., and Ryan, L. (1995). Efficiency and power of tests for multiple binary outcomes. *Journal of the American Statistical Association* **90,** 680–693.

O'Brien, P. C. (1984). Procedures for comparing samples with multiple endpoints. *Biometrics* **40,** 1079–1087.

Padungtod, C., Lasley, B. L., Christiani, D. C., Ryan, L., and Xu, X. (1999). Occupational pesticide exposure and semen quality among Chinese workers. *Journal of Occupational and Environmental Health* **40,** 1038–1047.

Pocock, S. T., Geller, N. L., and Tsiatis, A. A. (1987). The analysis of multiple endpoints in clinical trials. *Biometrics* **43,** 487–498.

Sammel, M. D. and Ryan, L. M. (1996). Latent variable models with fixed effects. *Biometrics* **52,** 650–663.

Sammel, M. D., Lin, X., and Ryan, L. M. (1999). Multivariate linear mixed models for multiple outcomes. *Statistics in Medicine* **18,** 2479–2492.

Saviltz, D. A. and Olshan, A. F. (1995). Multiple comparisons and related issues in the interpretation of epidemiologic data. *American Journal of Epidemiology* **142,** 904–908.

Self, S. and Liang, K. Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association* **82,** 605–10.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* New York: Wiley.

APPENDIX

A.1. Asymptotic Covariance Matrix of the MLE $(\hat{\gamma}, \hat{\theta}, \hat{\sigma}^2)$

Differentiating the score equations (5)–(7) with respect to $(\gamma, \theta, \sigma^2)$ and taking expectations of the resulting expressions gives the information matrix $\mathcal{I}$ given in equation (12), for $k, k' = 1, \ldots, c$ and $j, j' = 1, \ldots, M$, as follows:

$$\mathcal{I}_{11} = \sum_{i=1}^{n} \mathbf{X}_i^{\mathrm{T}} \mathbf{V}_i^{-1} \mathbf{X}_i$$

$$\mathcal{I}_{13}[\cdot, j] = \sum_{i=1}^{n} \frac{1}{2\sigma_j^2} \mathbf{X}_i^{\mathrm{T}} \mathbf{V}_i^{-1} \Delta_j \mathbf{X}_i \gamma$$

$$\mathcal{I}_{22}[k, k'] = \sum_{i=1}^{n} \frac{1}{2} \mathrm{tr} \left(\mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_{k'}} \right)$$

$$\mathcal{I}_{23}[k, j] = \sum_{i=1}^{n} \frac{1}{2\sigma_j^2} \mathrm{tr} \left(\mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} \Delta_j \right)$$

$$\mathcal{I}_{33}[j, j'] = \frac{n}{4\sigma_j^4} \delta[j = j'] + \sum_{i=1}^{n} \frac{1}{4\sigma_j^2 \sigma_{j'}^2} \left[\mathrm{tr}(\Delta_j \mathbf{V}_i^{-1} \Delta_{j'} \mathbf{V}_i) \right.$$
$$\left. + \gamma^{\mathrm{T}} \mathbf{X}_i^{\mathrm{T}} \Delta_j \mathbf{V}_i^{-1} \Delta_{j'} \mathbf{X}_i \gamma \right].$$

Here $\delta[j = j']$ equals one if $j = j'$ and zero otherwise. The asymptotic covariance matrix of the MLE $(\hat{\gamma}, \hat{\theta}, \hat{\sigma}^2)$ is $\mathcal{I}^{-1}$. Note that $\mathcal{I}_{13}[\cdot, j]$ denotes the jth row of $\mathcal{I}_{13}$.

A.2. Asymptotic Covariance Matrix of the WP Estimators $(\tilde{\gamma}, \tilde{\theta}, \tilde{\sigma}^2)$

Let $\xi = (\gamma^{\mathrm{T}}, \theta^{\mathrm{T}}, \sigma^{2\mathrm{T}})^{\mathrm{T}}$, $\mathbf{A} = \mathrm{E}(-\partial \mathbf{U}/\partial \xi^{\mathrm{T}})$, and $\mathbf{B} = \mathrm{cov}(\mathbf{U})$. Some calculations give

$$\mathbf{A} = \begin{bmatrix} \mathcal{I}_{11} & 0 & \mathcal{I}_{13} \\ 0 & \mathcal{I}_{22} & \mathcal{I}_{23} \\ 0 & \mathbf{A}_{32} & \mathbf{A}_{33} \end{bmatrix}$$

and

$$\mathbf{B} = \begin{bmatrix} \mathcal{I}_{11} & 0 & 0 \\ 0 & \mathcal{I}_{22} & \mathbf{A}_{32}^{\mathrm{T}} \\ 0 & \mathbf{A}_{32} & \mathbf{B}_{33} \end{bmatrix},$$

where, for $k = 1, \ldots, c$ and $j, j' = 1, \ldots, M$,

$$\mathbf{A}_{32}[k, j] = \sum_{i=1}^{n} \frac{1}{2} \mathrm{tr} \left(\mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} \mathbf{V}_i^{-1} \Delta_j \right)$$

$$\mathbf{A}_{33}[j, j'] = \sum_{i=1}^{n} \frac{1}{2\sigma_j^2} \mathrm{tr} \left(\Delta_j \mathbf{V}_i^{-1} \Delta_{j'} \right)$$

$$\mathbf{B}_{33}[j, j'] = \sum_{i=1}^{n} \frac{1}{2} \mathrm{tr} \left(\mathbf{V}_i^{-1} \Delta_j \mathbf{V}_i^{-1} \Delta_{j'} \right).$$

Then the asymptotic covariance matrix of $(\tilde{\gamma}, \tilde{\theta}, \tilde{\sigma}^2)$ is $\mathbf{A}^{-1} \mathbf{B} (\mathbf{A}^{-1})^{\mathrm{T}}$. The information matrix of $(\tilde{\gamma}, \tilde{\theta}, \tilde{\sigma}^2)$ can be defined as (Bhapkar, 1991) $\tilde{\mathcal{I}} = \mathbf{A}^{\mathrm{T}} \mathbf{B}^{-1} \mathbf{A}$. Some tedious calculations show $\tilde{\mathcal{I}}$ is given in (12), with

$$\tilde{\mathcal{I}}_{33} = \mathcal{I}_{23}^{\mathrm{T}} \mathcal{I}_{22}^{-1} \mathcal{I}_{23} + (\mathbf{A}_{33} - \mathcal{I}_{23}^{\mathrm{T}} \mathcal{I}_{22}^{-1} \mathbf{A}_{32}^{\mathrm{T}})$$
$$\times (\mathbf{B}_{33} - \mathbf{A}_{32} \mathcal{I}_{22}^{-1} \mathbf{A}_{32}^{\mathrm{T}})^{-1} (\mathbf{A}_{33} - \mathbf{A}_{32} \mathcal{I}_{22}^{-1} \mathcal{I}_{23}).$$

A.3. Proof of Proposition 1

Examination of the information matrix $\mathcal{I}$ in Appendix A.1 suggests that we can write $\mathcal{I}$ as

$$\mathcal{I} = \begin{bmatrix} \mathcal{I}_{11}^* & \mathcal{I}_{12}^* \Psi^{-1} \\ \Psi^{-1} \mathcal{I}_{21}^* & \Psi^{-1} \mathcal{I}_{22}^* \Psi^{-1} \end{bmatrix},$$

where $\mathcal{I}_{11}^* = \mathrm{diag}(\mathcal{I}_{11}, \mathcal{I}_{22})$ and $\mathcal{I}_{12}^*, \mathcal{I}_{21}^*, \mathcal{I}_{22}^*$ are defined accordingly. Note that $(\mathcal{I}_{11}^*, \mathcal{I}_{12}^*, \mathcal{I}_{21}^*, \mathcal{I}_{22}^*)$ in our formulation only depend on (γ, θ) but are free of σ^2. It follows that the covariance matrix of $(\hat{\gamma}, \hat{\theta})$ is

$$\mathrm{cov}(\hat{\gamma}, \hat{\theta}) = \left[\mathcal{I}_{11}^* - \mathcal{I}_{12}^* \Psi^{-1} (\Psi^{-1} \mathcal{I}_{22}^* \Psi^{-1})^{-1} \Psi^{-1} \mathcal{I}_{21}^* \right]^{-1}$$
$$= \left[\mathcal{I}_{11}^* - \mathcal{I}_{12}^* \mathcal{I}_{22}^{*-1} \mathcal{I}_{21}^* \right]^{-1}, \tag{A.1}$$

which is free of σ^2. The covariance matrix of $\hat{\sigma}^2$ is

$$\text{cov}(\hat{\sigma}^2) = \left[\Psi^{-1} \mathcal{I}_{22}^* \Psi^{-1} - \Psi^{-1} \mathcal{I}_{21}^* \mathcal{I}_{11}^{*}{}^{-1} \mathcal{I}_{12}^* \Psi^{-1} \right]^{-1}$$

$$= \Psi \left[\mathcal{I}_{22}^* - \mathcal{I}_{21}^* \mathcal{I}_{11}^{*}{}^{-1} \mathcal{I}_{12}^* \right]^{-1} \Psi. \qquad (A.2)$$

Using the results in Appendix A.2, we can write the matrices **A** and **B** as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11}^* & \mathbf{A}_{12}^* \\ \mathbf{A}_{21}^* & \mathbf{A}_{22}^* \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \Psi^{-1} \end{bmatrix}$$

and

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11}^* & \mathbf{B}_{12}^* \\ \mathbf{B}_{21}^* & \mathbf{B}_{22}^* \end{bmatrix},$$

where the partition of **A** and **B** is similar to that of $\mathcal{I}$ and $(\mathbf{A}_{ij}^*, \mathbf{B}_{ij}^*)$ $(i, j = 1, 2)$ are free of σ^2. The covariance matrix of $(\tilde{\gamma}, \tilde{\theta}, \hat{\sigma}^2)$ hence can be written as

$$\mathbf{A}^{-1} \mathbf{B} \left(\mathbf{A}^{-1} \right)^{\mathrm{T}} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \Psi \end{bmatrix} \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \Psi \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12}\Psi \\ \Psi\mathbf{C}_{21} & \Psi\mathbf{C}_{22}\Psi \end{bmatrix},$$

where

$$\begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11}^* & \mathbf{A}_{12}^* \\ \mathbf{A}_{21}^* & \mathbf{A}_{22}^* \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{B}_{11}^* & \mathbf{B}_{12}^* \\ \mathbf{B}_{21}^* & \mathbf{B}_{22}^* \end{bmatrix}$$

$$\times \left\{ \begin{bmatrix} \mathbf{A}_{11}^* & \mathbf{A}_{12}^* \\ \mathbf{A}_{21}^* & \mathbf{A}_{22}^* \end{bmatrix}^{-1} \right\}^{\mathrm{T}}.$$

It follows that $\text{cov}(\tilde{\gamma}, \tilde{\theta}) = \mathbf{C}_{11}$, which is free of σ^2, and $\text{cov}(\tilde{\sigma}^2) = \Psi\mathbf{C}_{22}\Psi$. Combining these results with those in equations (A.1) and (A.2) and using equation (13), we have $\text{ARE}(\tilde{\gamma}, \hat{\gamma})$, $\text{ARE}(\tilde{\theta}, \hat{\theta})$, and $\text{ARE}(\tilde{\sigma}^2, \hat{\sigma}^2)$ are independent of σ^2.