Statistica Sinica **30** (2020), 955-975
doi:https://doi.org/10.5705/ss.202017.0457

LPRE CRITERION BASED ESTIMATING EQUATION APPROACHES FOR THE ERROR-IN-COVARIABLES MULTIPLICATIVE REGRESSION MODELS

Qihua Wang[1,2] and Dahai Hu[3]

[1] *Zhejiang Gongshang University,* [2] *Chinese Academy of Sciences and* [3] *University of Science and Technology of China*

Abstract: In this paper, we propose two estimating equation-based methods for estimating the regression parameter vector in a multiplicative regression model when a subset of covariates is subject to a measurement error, but replicate measurements of their surrogates are available. Both methods allow the number of replicate measurements to vary between subjects. No parametric assumption is imposed on the measurement error term or the true covariates, which are not observed in the data set. Under some regularity conditions, the asymptotic normality is proved for both proposed estimators. Furthermore, the estimators are compared theoretically when the distribution of the measurement error follows a normal distribution. Simulation studies are conducted to assess the performance of the proposed methods. A real-data analysis is used to illustrate our methods.

Key words and phrases: Estimating equations, measurement error, multiplicative regression model, product form, relative error, replicate measurement.

1. Introduction

Positive responses are common in many practical problems, such those related to economics or survival analysis. To handle positive responses, it is natural to consider the following multiplicative regression model:

$$Y_i = \exp(Z_i^T \beta_0)\varepsilon_i, \quad i = 1, \ldots, n, \tag{1.1}$$

where Y_i is a scalar response variable, Z_i is a random covariate vector with its first component equal to one (intercept), β_0 is the true regression parameter vector, and the error term ε is strictly positive. When the response variable Y_i is a failure time, model (1.1) is called the accelerated failure time (AFT) model in survival analysis, see Wei (1992) and Jin et al. (2003), for example. The multiplicative regression model also has an important application in economic theory; see Teekens and Koerts (1972).

For the positive response variable, there are many situations in which the relative errors, rather than the absolute errors, are of major concern. For example, consider the problem of predicting a person's income. Assume that the two true values are {100,000, 10,000}. Further, assume there are two results from the prediction: (1) {150,000, 11,000}, (2) {101,000, 60,000}. Predictors (1) and (2) have absolute errors {50,000, 1,000} and {1,000, 50,000}, respectively. Thus, criteria based on the absolute error cannot determine which of the two predictors is more exact. However, if we consider the relative error, predictor (1) has relative error {0.5, 0.1}, whereas predictor (2) has relative error {0.01, 5}. The relative error criterion suggests choosing predictor (1), which seems to reflect the two persons' incomes more realistically.

In the literature, the relative error criterion has been applied to the standard linear model and nonlinear regression model. See, for example, Narula and Wellington (1977), Makridakis (1986), Khoshgoftaar, Bhattacharyya and Richardson (1992) and Park and Stefanski (1998). However, the theoretical justifications of the relative least squares (RLS) and absolute relative error (ARE) criteria are quite challenge for the linear and non-linear models. As pointed out by Chen et al. (2010), the consistency and asymptotic normality of the RLS and ARE estimators are not established under general regularity conditions, even for the standard linear regression models. Chen et al. (2010) took into account the following two types of relative errors: $|Y_i - \exp(Z_i^T\beta)|/Y_i$, which is relative to the response, and $|Y_i - \exp(Z_i^T\beta)|/\exp(Z_i^T\beta)$, which is relative to the predictor of the response. Then, they developed the least absolute relative error criterion (LARE) for multiplicative models (1.1). This criterion minimizes the following objective function:

$$\sum_{i=1}^{n}\left\{\frac{|Y_i - \exp(Z_i^T\beta)|}{Y_i} + \frac{|Y_i - \exp(Z_i^T\beta)|}{\exp(Z_i^T\beta)}\right\}.$$

Some variable selection methods were proposed based on the loss function of LARE (Xia, Liu and Yang (2016); Liu, Lin and Wang (2016)). In order to capture more complex models, Zhang and Wang (2013) extended the LARE criterion to the partially linear multiplicative regression model, using local smoothing techniques for estimation and variable selection.

In spite of its robustness and being scale-free, the LARE criterion function is nonsmooth and the asymptotic variance of the LARE estimator involves the unknown density of the error term. To avoid the density estimation, Li et al. (2014) proposed a novel empirical likelihood approach for constructing confidence intervals/regions for the regression parameters of the multiplicative regression models. To obtain a differentiable criterion function, Chen et al. (2016) studied the least product relative error (LPRE) criterion, which minimizes the following objective function:

$$LPRE_n(\beta) = \sum_{i=1}^{n} \left\{ \frac{|Y_i - \exp(Z_i^T\beta)|}{Y_i} \right\} \times \left\{ \frac{|Y_i - \exp(Z_i^T\beta)|}{\exp(Z_i^T\beta)} \right\}. \tag{1.2}$$

The most attractive property of the LPRE objective function is that it is infinitely differentiable and strictly convex. Using this funciton, Wang, Liu and Lin (2015) developed a procedure to detect the existence of the unknown change point, and discussed a relative-based estimation of the change point.

To the best of our knowledge, the aforementioned LARE and LPRE methods commonly assume that covariates are observed precisely. However, we often encounter corrupt data in practice, where the covariate measurements include errors. Sometimes the covariates of interest may be difficult to obtain accurately owing to physical location or cost. More commonly, it is not possible to measure them precisely owing to the nature of the covariates or the imprecision of the instrument. As a result, only replicate measurements of their surrogate variables are available. A good example of the latter situation is the AIDS Clinical Trials Group (ACTG) 175 study (Hammer et al. (1996)), which investigated the effects of four types of HIV treatments: zidovudine only, zidovudine and didanosine, zidovudine and zalcitabine, and didanosine only. In the ACTG 175 study, the baseline measurements on CD4 counts were collected before treatment. CD4 counts can never be measured precisely owing to the imprecision of the instrument. Hence, most subjects have two replicate baseline measurements of CD4 counts.

It is well known that misleading results may be obtained by naively applying the aforementioned methods to the corrupt data. Hence, it is important that we develop methods to handle such errors. There is a large body of literature for this topic for other models. Hu and Lin (2004) introduced a modified score equation for multivariate failure time data. Recently, Sinha and Ma (2016) proposed a semiparametric method to treat errors in covariates in the censored proportional odds model when replicated measurements of their surrogates are available and the number of replicate measurements is fixed. For a censored quantile regression with a measurement error, Wu, Ma and Yin (2015) developed a corrected estimating equation method based on a kernel smoothing approximation. They considered two types of measurement errors: those following a Laplace distribution and those following a normal distribution. For additive hazard models in survival analysis, Yan and Yi (2016) developed a class of correction methods for error-contaminated survival data with replicate measurements. Comprehensive discussions on measurement errors can be found in Carroll et al. (2006) and Buonaccorsi (2010), and the references therein.

In this paper, we propose two estimating equation approaches, based on the LPRE criterion, to estimate the regression parameter vector in a multiplicative regression model when a subset of covariates is subject to a measurement error, but replicate measurements of their surrogates are available. The first method constructs an unbiased estimating equation based on the conditional mean score. The second method corrects the naive method to obtain an unbiased estimating equation. A similar idea to that used in the first method is used in Hu and Lin (2004). Both the methods allow the study subjects to have unequal numbers of surrogate measurements. Furthermore, no parameter model is imposed on the measurement error term and the true covariates, which are not observed in the data.

The remainder of this article is organized as follows. In Section 2, we describe the framework of the multiplicative model with covariates measured with errors. In Section 3, we propose a conditional mean score-based estimating equation method. In Section 4, a corrected estimating equation method is suggested. For further discussion on the effect of the measurement error, we compare our proposed estimators theoretically in Section 5. Simulation studies are conducted in Section 6 to assess the performances of the proposed methods. An example from ACTG315 data is presented in Section 7 to illustrate the proposed methods.

2. Model Framework

Assume that the aforementioned covariates $Z_i = (V_i^T, X_i^T)^T$ form a (p+q)-vector of explanatory variables, where V_i is a q-vector of explanatory variables that are precisely measured with the first component being one (intercept), and X_i is a p-vector of error-prone explanatory variables. Then, model (1.1) turns into

$$Y_i = \exp(V_i^T \alpha_0 + X_i^T \gamma_0)\varepsilon_i, \quad i = 1, \ldots, n, \tag{2.1}$$

where $(\alpha_0^T, \gamma_0^T)^T = \beta_0$ is the corresponding regression parameter vector.

Suppose that X_i is measured repeatedly n_i times ($n_i \geq 1$) by the surrogates $W_{i,r}$, $r = 1, \ldots, n_i$. We consider the classical additive measurement error model:

$$W_{i,j} = X_i + U_{i,j}, \quad j = 1, \ldots, n_i, i = 1, \ldots, n, \tag{2.2}$$

where $U_{i,j}$ is an independent and identically distributed (i.i.d.) copy of the random variable vector U whose distribution is symmetric; that is, U and $-U$ are from the same distribution. In addition, $U_{i,j}$ is independent of Z_i and ε_i.

3. Conditional Mean Score-Based Estimating Equation Approach

3.1. Review: Estimation without measurement error

If X_i is accurately observed, the estimation of β can be obtained by minimizing the LPRE objective function (1.2). A simple algebraic manipulation of the LPRE objective function (1.2) yields

$$LPRE_n(\beta) = \sum_{i=1}^{n} \left\{ Y_i \exp(-Z_i^T\beta) + Y_i^{-1} \exp(Z_i^T\beta) - 2 \right\}.$$

Owing to the fact that the $LPRE$ objective function is strictly convex, minimizing $LPRE_n(\beta)$ is equivalent to solving the estimating equation $U_n(\beta) = 0$, where $U_n(\beta) = \partial LPRE_n(\beta)/\partial\beta$. Define $\psi(Z_i, Y_i, \beta) = \{Y_i^{-1} \exp(Z_i^T\beta) - Y_i \exp(-Z_i^T\beta)\}Z_i$. Then,

$$U_n(\beta) = \sum_{i=1}^{n} \psi(Z_i, Y_i, \beta). \tag{3.1}$$

With the condition $E(1/\varepsilon - \varepsilon|Z) = 0$ (Chen et al. (2016)), it is easy to obtain that $E[U_n(\beta_0)] = 0$; that is, $U_n(\beta)$ is an unbiased estimating function of β. By the theory of the estimating equation, under regularity conditions, solving $U_n(\beta) = 0$ yields a consistent estimator of β (Chen et al. (2016)).

3.2. Estimation with measurement error and asymptotic properties

For simplification, denote the observed data $\mathcal{O}_{i,r} = (Y_i, V_i, W_{i,r})$ and let $\mathcal{U}_i = (Y_i, V_i, X_i)$ for $i = 1, \ldots, n$ and $r = 1, \ldots, n_i$. Recall that $\psi(Z_i, Y_i, \beta)$ is the summand of the unbiased estimating function $U_n(\beta)$ in (3.1). We can find a function $T^*(\mathcal{O}_{i,r}, \beta)$ such that

$$E[T^*(\mathcal{O}_{i,r}, \beta)|\mathcal{U}_i] = \psi(Z_i, Y_i, \beta),$$

which leads to the following unbiased estimating equation,

$$\sum_{i=1}^{n} n_i^{-1} \sum_{r=1}^{n_i} T^*(\mathcal{O}_{i,r}, \beta) = 0. \tag{3.2}$$

Next, let us construct $T^*(\mathcal{O}_{i,r}, \beta)$. Take $\hat{Z}_{i,r} = (V_i^T, W_{i,r}^T)^T$ and $J = (0_{p\times q}, I_{p\times p})^T$. Then, $\hat{Z}_{i,r} = Z_i + JU_{i,r}$. For simplicity, define $\varphi_0(\gamma) = E[\exp(U^T\gamma)]$ and $\varphi_1(\gamma) = E[U \exp(U^T\gamma)]$. The independence between the error $U_{i,r}$ and the true covariate Z_i implies

$$E[\exp(\hat{Z}_{i,r}^T\beta)\hat{Z}_{i,r}|Z_i] = \varphi_0(\gamma)\exp(Z_i^T\beta)Z_i + \exp(Z_i^T\beta)J\varphi_1(\gamma), \tag{3.3}$$

$$E[\exp(\hat{Z}_{i,r}^T\beta)|Z_i] = \varphi_0(\gamma)\exp(Z_i^T\beta). \tag{3.4}$$

For simplification, take

$$R_{i,r}^{(0)}(\beta) = \varphi_0^{-1}(\gamma)\exp(\hat{Z}_{i,r}^T\beta),$$
$$R_{i,r}^{(1)}(\beta) = \varphi_0^{-1}(\gamma)\exp(\hat{Z}_{i,r}^T\beta)\{\hat{Z}_{i,r} - J\varphi_0^{-1}(\gamma)\varphi_1(\gamma)\}.$$

A simple algebraic manipulation of (3.3) and (3.4) yields

$$\exp(Z_i^T\beta)Z_i = E\big[R_{i,r}^{(1)}(\beta)|\mathcal{U}_i\big], \tag{3.5}$$

$$\exp(Z_i^T\beta) = E\big[R_{i,r}^{(0)}(\beta)|\mathcal{U}_i\big]. \tag{3.6}$$

Recalling the definition of $\psi(Z_i, Y_i, \beta)$ in (3.1), the desired function $T^*(\mathcal{O}_{i,r}, \beta)$ can be defined as

$$T^*(\mathcal{O}_{i,r}, \beta) = Y_i^{-1}R_{i,r}^{(1)}(\beta) - Y_iR_{i,r}^{(1)}(-\beta).$$

By (3.5), $E[T^*(\mathcal{O}_{i,r}, \beta)|\mathcal{U}_i] = \psi(Z_i, Y_i, \beta)$. However, $\varphi_0(\gamma)$ and $\varphi_1(\gamma)$ in $T^*(\mathcal{O}_{i,r}, \beta)$ are unknown. We must define their estimation. Observing that $W_{i,r} - W_{i,s} = U_{i,r} - U_{i,s} (r \neq s)$ and that the errors $U_{i,r}$ are i.i.d and symmetric, we have

$$E[\exp\{\gamma^T(W_{i,r} - W_{i,s})\}] = \varphi_0^2(\gamma)$$

and

$$E[(W_{i,r} - W_{i,s})\exp\{\gamma^T(W_{i,r} - W_{i,s})\}] = 2\varphi_0(\gamma)\varphi_1(\gamma).$$

Denote $\xi_i = I(n_i > 1)$ and $\tilde{n} = \sum_{i=1}^n \xi_i$. Then, $\varphi_k(\gamma), (k = 0, 1)$ can be estimated by

$$\hat{\varphi}_0(\gamma) = \left[\frac{1}{\tilde{n}}\sum_{i=1}^n \frac{\xi_i}{\{n_i(n_i-1)\}}\sum_{r\neq s}\exp(\gamma^T(W_{i,r} - W_{i,s}))\right]^{1/2}$$

and

$$\hat{\varphi}_1(\gamma) = \{2\tilde{n}\hat{\varphi}_0(\gamma)\}^{-1}\sum_{i=1}^n \frac{\xi_i}{\{n_i(n_i-1)\}}\sum_{r\neq s}(W_{i,r} - W_{i,s})\exp(\gamma^T(W_{i,r} - W_{i,s})),$$

where, for every i with $n_i > 1$, (r, s) runs through all possible combinations of numbers in $\{1, \ldots, n_i\}$. When $n_i = 1$, both ξ_i and $n_i - 1$ equal zero, and we define the fraction $\xi_i/(n_i - 1)$ to be zero for convenience.

Let $\hat{R}_{i,r}^{(0)}(\beta)$ and $\hat{R}_{i,r}^{(1)}(\beta)$ be $R_{i,r}^{(0)}(\beta)$ and $R_{i,r}^{(1)}(\beta)$, respectively, with $\varphi_0(\gamma)$ and $\varphi_1(\gamma)$ replaced by $\hat{\varphi}_0(\gamma)$ and $\hat{\varphi}_1(\gamma)$, respectively. Thereafter, the resulting estimating equation is given by

$$\sum_{i=1}^{n} n_i^{-1} \sum_{r=1}^{n_i} \hat{T}^*(\mathcal{O}_{i,r}, \beta) = 0,$$

where $\hat{T}^*(\mathcal{O}_{i,r}, \beta) = Y_i^{-1} \hat{R}_{i,r}^{(1)}(\beta) - Y_i \hat{R}_{i,r}^{(1)}(-\beta)$. The solution of the above equation can be defined as an estimator of β, denoted as $\hat{\beta}_{cms}$.

For notational simplicity, we assume that $(Z^T, Y, \varepsilon)^T$, $(Z_i^T, Y_i, \varepsilon_i)^T$, for $i = 1, \ldots, n$ are i.i.d.. To describe the asymptotic properties of the proposed estimator, we first present some notations. For any vector or matrix a, we denote $a^{\otimes 2} = aa^T$. Define $\mathcal{A}_k = \{i : n_i = k, i = 1, \ldots, n\}$, for $k = 1, \ldots m$, and let $|\mathcal{A}_k|$ be the number of members of $\mathcal{A}_k$. Define $R_i^{(0)}(\beta) = n_i^{-1} \sum_{r=1}^{n_i} R_{i,r}^{(0)}(\beta)$, $\hat{R}_i^{(0)}(\beta) = n_i^{-1} \sum_{r=1}^{n_i} \hat{R}_{i,r}^{(0)}(\beta)$, $R_i^{(1)}(\beta) = n_i^{-1} \sum_{r=1}^{n_i} R_{i,r}^{(1)}(\beta)$, and $\hat{R}_i^{(1)}(\beta) = n_i^{-1} \sum_{r=1}^{n_i} \hat{R}_{i,r}^{(1)}(\beta)$.

Then, $\sum_{i=1}^{n} n_i^{-1} \sum_{r=1}^{n_i} \hat{T}^*(\mathcal{O}_{i,r}, \beta) = \sum_{i=1}^{n} [Y_i^{-1} \hat{R}_i^{(1)}(\beta) - Y_i \hat{R}_i^{(1)}(-\beta)]$. Take

$$\begin{aligned} v_i &= Y_i^{-1} R_i^{(1)}(\beta_0) - Y_i R_i^{(1)}(-\beta_0), \\ r_i &= E\Big(\frac{1}{\varepsilon} + \varepsilon\Big) \{2(1-\rho_1)\varphi_0^2(\gamma_0)\}^{-1} \{h_i^{(1)}(\gamma_0) - 2\varphi_0^{-1}(\gamma_0)\varphi_1(\gamma_0) h_i^{(0)}(\gamma_0)\}, \end{aligned}$$

where $\rho_1 = \lim |\mathcal{A}_1|/n$, $h_i^{(0)}(\gamma) = \{n_i(n_i-1)\}^{-1} \sum_{r \neq s} \exp\{\gamma^T (W_{i,r} - W_{i,s})\}$ and $h_i^{(1)}(\gamma) = \{n_i(n_i-1)\}^{-1} \sum_{r \neq s} (W_{i,r} - W_{i,s}) \exp\{\gamma^T (W_{i,r} - W_{i,s})\}$. Here, if $n_i = 1$, define $h_i^{(0)}(\gamma) = 0$ and $h_i^{(1)}(\gamma) = 0$, for convenience. Further, define $V_0 = E[(1/\varepsilon + \varepsilon) Z Z^T]$. The asymptotic normality of $\hat{\beta}_{cms}$ is established in the following theorem.

Theorem 1. *Under Conditions C1–C6 in the Appendix, $\hat{\beta}_{cms}$ exists and is unique in a neighbourhood of β_0 with probability converging to 1 as $n \to \infty$, and*

$$\sqrt{n}(\hat{\beta}_{cms} - \beta_0) \xrightarrow{D} N(0, \Gamma_{cms}),$$

where $\Gamma_{cms} = V_0^{-1} \Sigma_{cms} V_0^{-1}$ and $\Sigma_{cms} = \lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} E(v_i - \xi_i J r_i)^{\otimes 2}$.

To estimate Γ_{cms}, we define

$$\begin{aligned} \hat{v}_i &= Y_i^{-1} \hat{R}_i^{(1)}(\hat{\beta}_{cms}) - Y_i \hat{R}_i^{(1)}(-\hat{\beta}_{cms}), \\ \hat{r}_i &= \{2n(1-\hat{\rho}_1)\hat{\varphi}_0^2(\hat{\gamma}_{cms})\}^{-1} \sum_{j=1}^{n} \{Y_j^{-1} \hat{R}_j^{(0)}(\hat{\beta}_{cms}) + Y_j \hat{R}_j^{(0)}(-\hat{\beta}_{cms})\} \\ &\quad \times \{h_i^{(1)}(\hat{\gamma}_{cms}) - 2\hat{\varphi}_0^{-1}(\hat{\gamma}_{cms})\hat{\varphi}_1(\hat{\gamma}_{cms}) h_i^{(0)}(\hat{\gamma}_{cms})\}, \end{aligned}$$

where $\hat{\rho}_1 = |\mathcal{A}_1|/n$ and $\hat{\gamma}_{cms}$ denotes the last p components of $\hat{\beta}_{cms}$. Then, take $\hat{\Sigma}_{cms} = n^{-1} \sum_{i=1}^{n} \{\hat{v}_i - \xi_i J \hat{r}_i\}^{\otimes 2}$ and $\hat{V}_0 = n^{-1} \sum_{i=1}^{n} \{Y_i^{-1} \hat{R}_i^{(2)}(\hat{\beta}_{cms}) + Y_i \hat{R}_i^{(2)}(-\hat{\beta}_{cms})\}$, where $\hat{R}_i^{(2)}(\beta) = \partial \hat{R}_i^{(1)}(\beta)/\partial \beta^T$. Denote $\hat{\Gamma}_{cms} = \hat{V}_0^{-1} \hat{\Sigma}_{cms} \hat{V}_0^{-1}$.

Γ_{cms} can then be estimated by $\hat{\Gamma}_{cms}$.

4. Corrected Estimating Equation Method

4.1. Naive method and bias

Define $\bar{W}_{i,\cdot} = n_i^{-1}\sum_{r=1}^{n_i} W_{i,r}$, and $\hat{Z}_i = (V_i^T, \bar{W}_{i,\cdot}^T)^T = Z_i + J\bar{U}_{i,\cdot}$, where $\bar{U}_{i\cdot} = n_i^{-1}\sum_{r=1}^{n_i} U_{i,r}$. A naive computable estimating function $U_{nv}(\beta)$ can be obtained as

$$U_{nv}(\beta) = \sum_{i=1}^{n} \left\{Y_i^{-1}\exp(\hat{Z}_i^T\beta) - Y_i\exp(-\hat{Z}_i^T\beta)\right\}\hat{Z}_i = \sum_{i=1}^{n}\psi(\hat{Z}_i, Y_i, \beta) \tag{4.1}$$

by replacing Z_i in (3.1) with $\hat{Z}_i$. Let $\hat{\beta}_{nv}$ be the solution of $U_{nv}(\beta) = 0$. Here $\hat{\beta}_{nv}$ is known as the naive-LPRE estimator of β_0.

Recall the definition of $\hat{Z}_i$ and J, then, $\psi(\hat{Z}_i, Y_i, \beta)$ can be written as

$$\left\{Y_i^{-1}\exp(Z_i^T\beta)\exp(\bar{U}_{i,\cdot}^T\gamma) - Y_i\exp(-Z_i^T\beta)\exp(-\bar{U}_{i,\cdot}^T\gamma)\right\}(Z_i + J\bar{U}_{i,\cdot}). \tag{4.2}$$

Owing to the symmetry of U and the independence between $U_{i,r}$ and (Z_i, Y_i), we have

$$\begin{aligned} E[\psi(\hat{Z}_i, Y_i, \beta)|Y_i, Z_i] = {} & \varphi_0^{n_i}\left(\frac{\gamma}{n_i}\right)\psi(Z_i, Y_i, \beta) + J\{Y_i^{-1}\exp(Z_i^T\beta) + \\ & Y_i\exp(-Z_i^T\beta)\} \times \varphi_0^{n_i-1}\left(\frac{\gamma}{n_i}\right)\varphi_1\left(\frac{\gamma}{n_i}\right). \end{aligned} \tag{4.3}$$

Comparing (4.3) and $\psi(Z_i, Y_i, \beta)$ in (3.1), we find two main differences. Take $I_{1n}(\beta) = \varphi_0^{n_i}(\gamma/n_i)\psi(Z_i, Y_i, \beta)$ and $I_{2n}(\beta) = J\{Y_i^{-1}\exp(Z_i^T\beta) + Y_i\exp(-Z_i^T\beta)\} \times \varphi_0^{n_i-1}(\gamma/n_i)\varphi_1(\gamma/n_i)$. On the one hand, the term $I_{1n}(\beta)$ is equal to the product of the factor $\varphi_0^{n_i}(\gamma/n_i)$ and $\psi(Z_i, Y_i, \beta)$ in (3.1). On the other hand, the term $I_{2n}(\beta)$ is an extra term. With the assumption that $E[\varepsilon - 1/\varepsilon|Z] = 0$, we obtain that $E[I_{1n}(\beta_0)] = 0$. Therefore, $E[\psi(\hat{Z}_i, Y_i, \beta_0)] = E[I_{1n}(\beta_0) + I_{2n}(\beta_0)] = E[I_{2n}(\beta_0)]$. However, $E[I_{2n}(\beta_0)] = E(1/\varepsilon + \varepsilon)J\varphi_0^{n_i-1}(\gamma_0/n_i)\varphi_1(\gamma_0/n_i)$. Thus, $I_{2n}(\beta_0)$ may be the term causing bias of the naive estimator $\hat{\beta}_{nv}$. It's obvious that $\varphi_0(\gamma) > 0$ from the definition of $\varphi_0(\gamma)$. In general, $\varphi_1(\gamma)$ is also a nonzero vector. In the remark, we discuss some common scenarios.

Remark 1.

- When U is a scalar variable, $\partial\varphi_1(\gamma)/\partial\gamma = E[U^2\exp(U^T\gamma)] > 0$ unless U is zero almost surely. Consequently, $\varphi_1(\gamma)$ increases strictly as γ increases with $\varphi_1(0) = 0$. As a result, $\varphi_1(\gamma)$ departs from zero when $\gamma \neq 0$.

- Denote $a^{(i)}$ as the ith component of the vector a. Assume that the measurement error $U = (U^{(1)}, \ldots, U^{(p)})^T$ and $U^{(1)}, \ldots, U^{(p)}$ are independent of each other. Some simple algebraic manipulation yields that $\varphi_1^{(i)}(\gamma) = E[U^{(i)} \exp(U^{(i)}\gamma^{(i)})] \prod_{j\neq i} E[\exp(U^{(j)}\gamma^{(j)})]$. As discussed above, $\varphi_1(\gamma)$ does not equal zero when $\gamma \neq 0$.

- Assume that $U \sim N(0, \Sigma_{p\times p})$. Some basic calculation yields that $\varphi_0(\gamma) = \exp(\gamma^T \Sigma\gamma/2)$ and $\varphi_1(\gamma) = \Sigma\gamma \exp(\gamma^T \Sigma\gamma/2)$. Because Σ is positive definite, $\varphi_1(\gamma) = 0$ if and only if $\gamma = 0$.

In the above discussion, for the three commonly used cases, we conclude that $\varphi_1(\gamma) = 0$ if and only if $\gamma = 0$. However, γ_0, the true value of γ, is not zero for the measurement error model considered here. Otherwise, the estimating problem reduces to that without a measurement error. Combined with the fact that $\varphi_0(\gamma) > 0$ and $E[1/\varepsilon + \varepsilon] > 0$, we have that $E[I_{2n}(\beta_0)] \neq 0$. Consequently, $E[\psi(\hat{Z}_i, Y_i, \beta_0)] \neq 0$ and $U_{nv}(\beta)$ is a biased estimating function. The resultant estimator $\hat{\beta}_{nv}$ does not converge to the true parameter β_0.

In addition, note that $I_{1n}(\beta)$ is the unbiased estimating function $\psi(Z_i, Y_i, \beta)$ multiplied by $\varphi_0^{n_i}(\gamma/n_i)$. This factor $\varphi_0^{n_i}(\gamma/n_i)$ may lead to a loss of efficiency for the naive estimator $\hat{\beta}_{nv}$. Based on the above two considerations, we develop a corrected estimating equation approach in the following subsection.

4.2. Corrected estimation and asymptotic properties

To eliminate the bias of the naive estimator and to obtain a more reasonable estimator, we can construct an unbiased estimating function as follows

$$U^*(\beta) = \sum_{i=1}^{n} \tilde{\psi}_i,$$

where

$$\begin{aligned}\tilde{\psi}_i = &\Big\{\varphi_0^{n_i}\Big(\frac{\gamma}{n_i}\Big)\Big\}^{-1}\Big[\psi(\hat{Z}_i, Y_i, \beta) - J\{Y_i^{-1}\exp(Z_i^T\beta) \\ &+ Y_i \exp(-Z_i^T\beta)\}\varphi_0^{n_i-1}\Big(\frac{\gamma}{n_i}\Big)\varphi_1\Big(\frac{\gamma}{n_i}\Big)\Big].\end{aligned}$$

Recalling (4.3), we can see that $E[\tilde{\psi}_i|\mathcal{U}_i] = \psi(Z_i, Y_i, \beta)$. However, Z_i in $\tilde{\psi}_i$ cannot be observed. Note that

$$\begin{aligned}&E[Y_i^{-1}\exp(\hat{Z}_i^T\beta) + Y_i\exp(-\hat{Z}_i^T\beta)|\mathcal{U}_i] \\ &= \{Y_i^{-1}\exp(Z_i^T\beta) + Y_i\exp(-Z_i^T\beta)\}\varphi_0^{n_i}\Big(\frac{\gamma}{n_i}\Big).\end{aligned} \tag{4.4}$$

From (4.4), we have

$$Y_i^{-1}\exp(Z_i^T\beta) + Y_i\exp(-Z_i^T\beta)$$

$$= E\left[\varphi_0^{-n_i}\Big(\frac{\gamma}{n_i}\Big)\{Y_i^{-1}\exp(\hat{Z}_i^T\beta) + Y_i\exp(-\hat{Z}_i^T\beta)\}|\mathcal{U}_i\right]. \tag{4.5}$$

Therefore, we can define ψ_i^* as follow,

$$\begin{aligned}\psi_i^* = &\Big\{\varphi_0^{n_i}\Big(\frac{\gamma}{n_i}\Big)\Big\}^{-1}\Big[\psi(\hat{Z}_i, Y_i, \beta)\\ &- J\{Y_i^{-1}\exp(\hat{Z}_i^T\beta) + Y_i\exp(-\hat{Z}_i^T\beta)\}\varphi_1\Big(\frac{\gamma}{n_i}\Big)\varphi_0^{-1}\Big(\frac{\gamma}{n_i}\Big)\Big],\end{aligned}$$

by replacing the term $Y_i^{-1}\exp(Z_i^T\beta) + Y_i\exp(-Z_i^T\beta)$ in $\tilde{\psi}_i$ with the expression $\varphi_0^{-n_i}(\gamma/n_i)\{Y_i^{-1}\exp(\hat{Z}_i^T\beta) + Y_i\exp(-\hat{Z}_i^T\beta)\}$. However, $\varphi_0(\gamma)$ and $\varphi_1(\gamma)$ in ψ_i^* are unknown. Define

$$\begin{aligned}\hat{\psi}_i^* = &\Big\{\hat{\varphi}_0^{n_i}\Big(\frac{\gamma}{n_i}\Big)\Big\}^{-1}\Big[\psi(\hat{Z}_i, Y_i, \beta)\\ &- J\{Y_i^{-1}\exp(\hat{Z}_i^T\beta) + Y_i\exp(-\hat{Z}_i^T\beta)\}\hat{\varphi}_1\Big(\frac{\gamma}{n_i}\Big)\hat{\varphi}_0^{-1}\Big(\frac{\gamma}{n_i}\Big)\Big],\end{aligned}$$

by replacing $\varphi_0(\gamma/n_i)$ and $\varphi_1(\gamma/n_i)$ in ψ^* with $\hat{\varphi}_0(\gamma/n_i)$ and $\hat{\varphi}_1(\gamma/n_i)$, respectively, given in the previous section, and we obtain a resultant estimating equation for β_0 as follows

$$\sum_{i=1}^n \hat{\psi}_i^* = 0.$$

Let $\hat{\beta}_{cee}$ be the solution to the above estimating equation. For simplicity, denote $\eta_0(k,\gamma) = E[\exp\{k^{-1}\gamma^T(U_1 + \cdots + U_k)\}]$ and $\eta_1(k,\gamma) = \partial\eta_0(k,\gamma)/\partial\gamma$, for any positive integer k. Clearly, $\eta_0(1,\gamma) = \varphi_0(\gamma)$, $\eta_0(k,\gamma) = \varphi_0^k(\gamma/k)$ and $\eta_1(k,\gamma) = \varphi_0^{k-1}(\gamma/k)\varphi_1(\gamma/k)$. Then, denote

$$\begin{aligned}\tilde{R}_i^{(0)}(\beta) &= \eta_0^{-1}(n_i,\gamma)\exp(\hat{Z}_i^T\beta),\\ \check{R}_i^{(0)}(\beta) &= \hat{\eta}_0^{-1}(n_i,\gamma)\exp(\hat{Z}_i^T\beta),\\ \tilde{R}_i^{(1)}(\beta) &= \eta_0^{-1}(n_i,\gamma)\exp(\hat{Z}_i^T\beta)\{\hat{Z}_i - J\eta_1(n_i,\gamma)\eta_0^{-1}(n_i,\gamma)\},\\ \check{R}_i^{(1)}(\beta) &= \hat{\eta}_0^{-1}(n_i,\gamma)\exp(\hat{Z}_i^T\beta)\{\hat{Z}_i - J\hat{\eta}_1(n_i,\gamma)\hat{\eta}_0^{-1}(n_i,\gamma)\},\end{aligned}$$

where $\hat{\eta}_0(k,\gamma) = \hat{\varphi}_0^k(\gamma/k)$ and $\hat{\eta}_1(k,\gamma) = \hat{\varphi}_0^{k-1}(\gamma/k)\hat{\varphi}_1(\gamma/k)$, and $\hat{\varphi}_k(\cdot)(k = 0, 1)$ is defined as in the previous section. By a simple calculation, we have

$$\begin{aligned}\psi_i^* &= Y_i^{-1}\tilde{R}_i^{(1)}(\beta) - Y_i\tilde{R}_i^{(1)}(-\beta),\\ \hat{\psi}_i^* &= Y_i^{-1}\check{R}_i^{(1)}(\beta) - Y_i\check{R}_i^{(1)}(-\beta).\end{aligned}$$

Let

$$\tilde{v}_i = Y_i^{-1}\tilde{R}_i^{(1)}(\beta_0) - Y_i\tilde{R}_i^{(1)}(-\beta_0),$$

$$\tilde{r}_{i,k} = E\Big(\frac{1}{\varepsilon}+\varepsilon\Big)\Big\{2(1-\rho_1)\varphi_0^2\Big(\frac{\gamma_0}{k}\Big)\Big\}^{-1} \\ \times\Big\{h_i^{(1)}\Big(\frac{\gamma_0}{k}\Big) - 2\varphi_0^{-1}\Big(\frac{\gamma_0}{k}\Big)\varphi_1\Big(\frac{\gamma_0}{k}\Big)h_i^{(0)}\Big(\frac{\gamma_0}{k}\Big)\Big\},$$

where $h_i^{(k)}(\gamma)(k=0,1)$ are defined in Section 3.2. Let $\rho_k = \lim_{n\to\infty}|\mathcal{A}_k|/n$. Further, recall that $V_0 = E[(1/\varepsilon+\varepsilon)ZZ^T]$, which is defined in Theorem 1. The asymptotic normality of $\hat{\beta}_{cee}$ is established in the following theorem.

Theorem 2. *Under Conditions $C1$-$C6$ in Appendix, $\hat{\beta}_{cee}$ exists and is unique in a neighbourhood of β_0 with probability converging to 1 as $n\to\infty$, and*

$$\sqrt{n}(\hat{\beta}_{cee}-\beta_0) \xrightarrow{D} N(0,\Gamma_{cee}),$$

where $\Gamma_{cee} = V_0^{-1}\Sigma_{cee}V_0^{-1}$ and $\Sigma_{cee} = \lim n^{-1}\sum_{i=1}^n E\{\tilde{v}_i - \xi_i J\sum_{k=1}^m \rho_k\tilde{r}_{i,k}\}^{\otimes 2}$.

To estimate Γ_{cee}, we define

$$\check{v}_i = Y_i^{-1}\check{R}_i^{(1)}(\hat{\beta}_{cee}) - Y_i\check{R}_i^{(1)}(-\hat{\beta}_{cee}),$$

$$\check{r}_{i,k} = \Big\{2n(1-\hat{\rho}_1)\hat{\varphi}_0^2\Big(\frac{\hat{\gamma}_{cee}}{k}\Big)\Big\}^{-1}\sum_{j=1}^n\Big\{Y_j^{-1}\check{R}_j^{(0)}(\hat{\beta}_{cee}) + Y_j\check{R}_j^{(0)}(-\hat{\beta}_{cee})\Big\} \\ \times\Big\{h_i^{(1)}\Big(\frac{\hat{\gamma}_{cee}}{k}\Big) - 2\hat{\varphi}_0^{-1}\Big(\frac{\hat{\gamma}_{cee}}{k}\Big)\hat{\varphi}_1\Big(\frac{\hat{\gamma}_{cee}}{k}\Big)h_i^{(0)}\Big(\frac{\hat{\gamma}_{cee}}{k}\Big)\Big\},$$

where $\hat{\gamma}_{cee}$ is the last p components of $\hat{\beta}_{cee}$. Let $\hat{\rho}_k = |\mathcal{A}_k|/n$. Then take $\hat{\Sigma}_{cee} = n^{-1}\sum_{i=1}^n\{\check{v}_i - \xi_i J\sum_{k=1}^m\hat{\rho}_k\check{r}_{i,k}\}^{\otimes 2}$ and $\tilde{V}_0 = n^{-1}\sum_{i=1}^n\{Y_i^{-1}\check{R}_i^{(2)}(\hat{\beta}_{cee}) + Y_i\check{R}_i^{(2)}(-\hat{\beta}_{cee})\}$, where $\check{R}_i^{(2)}(\beta) = \partial\check{R}_i^{(1)}(\beta)/\partial\beta^T$. Denote $\hat{\Gamma}_{cee} = \tilde{V}_0^{-1}\hat{\Sigma}_{cee}\tilde{V}_0^{-1}$. Γ_{cee} can then be estimated by $\hat{\Gamma}_{cee}$.

5. Comparison of the Two Methods

When the distribution of the measurement error U is unknown, $\varphi_s(\gamma)(s=0,1)$ must be estimated with the sample. This makes the asymptotic covariance structures complex, and hence it is hard to compare the asymptotic efficiency of the two proposed methods. However, we may compare the two methods for a special case where the distribution of the measurement error U is known to be normal. For simplicity, take $n_i = k$. Hence, the first estimator $\hat{\beta}_{cms}$ reduces to the solution of the following estimating equation:

$$\sum_{i=1}^n k^{-1}\sum_{r=1}^k T^*(\mathcal{O}_{i,r},\beta) = 0,$$

denoted as $\hat{\beta}_{cms}^*$. Similarly, the second estimator $\hat{\beta}_{cee}$ reduces to the solution of the following estimating equation:

$$\sum_{i=1}^{n} \psi_i^* = 0,$$

denoted as $\hat{\beta}_{cee}^*$. Thus, we have the following results.

Theorem 3. *Under Condition C1-C3 and C5 in the Appendix, both $\hat{\beta}_{cms}^*$ and $\hat{\beta}_{cee}^*$ exist and are unique in a neighbourhood of β_0 with probability converging to 1 as $n \to \infty$. In addition,*

$$\sqrt{n}(\hat{\beta}_{cms}^* - \beta_0) \xrightarrow{D} N(0, \Gamma_{cms}^*) \quad \text{and} \quad \sqrt{n}(\hat{\beta}_{cee}^* - \beta_0) \xrightarrow{D} N(0, \Gamma_{cee}^*),$$

where $\Gamma_{cms}^* = V_0^{-1}\Sigma_{cms}^* V_0^{-1}$ and $\Gamma_{cee}^* = V_0^{-1}\Sigma_{cee}^* V_0^{-1}$, with $\Sigma_{cms}^* = E[v_i^{\otimes 2}]$, and $\Sigma_{cee}^* = E[\tilde{v}_i^{\otimes 2}]$, with V_0, v_i, and $\tilde{v}_i$ defined in Section 3.2 and 4.2.

In order to compare the asymptotic covariances of the two proposed estimators, we only need to compare Σ_{cms}^* and Σ_{cee}^*. In the following lemma, we establish the expressions of Σ_{cms}^* and Σ_{cee}^*.

Lemma 1. *Assume the conditions of Theorem 3. If $EZ = 0$, ε is independent of Z, $E(U) = 0$ and $cov(U) = \Sigma_u$, we then have*

$$\begin{aligned}
\Sigma_{cms}^* =& k^{-1}\Bigg\{E\left(\frac{1}{\varepsilon^2} + \varepsilon^2\right)\varphi_0(2\gamma_0)\varphi_0^{-2}(\gamma_0)\big[E(Z^{\otimes 2}) \\
&+ J\varphi_2(2\gamma_0)\varphi_0^{-1}(2\gamma_0)J^T - J\varphi_1(2\gamma_0)\varphi_0^{-1}(2\gamma_0)\varphi_1^T(\gamma_0)\varphi_0^{-1}(\gamma_0)J^T \\
&- J\varphi_1(\gamma_0)\varphi_0^{-1}(\gamma_0)\varphi_1^T(2\gamma_0)\varphi_0^{-1}(2\gamma_0)J^T + \{J\varphi_1(\gamma_0)\varphi_0^{-1}(\gamma_0)\}^{\otimes 2}\big] \\
&- 2\varphi_0^{-2}(\gamma_0)\big[E(Z^{\otimes 2}) + J\Sigma_u J^T - \{J\varphi_1(\gamma_0)\varphi_0^{-1}(\gamma_0)\}^{\otimes 2}\big]\Bigg\} \\
&+ \frac{(k-1)}{kE(\varepsilon - 1/\varepsilon)^2 E(Z^{\otimes 2})}, \\
\Sigma_{cee}^* =& E\Big(\frac{1}{\varepsilon^2} + \varepsilon^2\Big)\varphi_0^k\Big(\frac{2\gamma_0}{k}\Big)\varphi_0^{-2k}\Big(\frac{\gamma_0}{k}\Big) \\
&\times \Big[E(Z^{\otimes 2}) + k^{-1}J\varphi_2\Big(\frac{2\gamma_0}{k}\Big)\varphi_0^{-1}\Big(\frac{2\gamma_0}{k}\Big)J^T \\
&+ \frac{(k-1)}{k\big\{J\varphi_1\big(2\gamma_0/k\big)\varphi_0^{-1}\big(2\gamma_0/k\big)\big\}^{\otimes 2}} \\
&- J\varphi_1\Big(\frac{2\gamma_0}{k}\Big)\varphi_0^{-1}\Big(\frac{2\gamma_0}{k}\Big)\varphi_1^T\Big(\frac{\gamma_0}{k}\Big)\varphi_0^{-1}\Big(\frac{\gamma_0}{k}\Big)J^T \\
&- J\varphi_1\Big(\frac{\gamma_0}{k}\Big)\varphi_0^{-1}\Big(\frac{\gamma_0}{k}\Big)\varphi_1^T\Big(\frac{2\gamma_0}{k}\Big)\varphi_0^{-1}\Big(\frac{2\gamma_0}{k}\Big)J^T \\
&+ \Big\{J\varphi_1\Big(\frac{\gamma_0}{k}\Big)\varphi_0^{-1}\Big(\frac{\gamma_0}{k}\Big)\Big\}^{\otimes 2}\Big]
\end{aligned}$$

$$- 2\varphi_0^{-2k}\Big(\frac{\gamma_0}{k}\Big)\Big[E(Z^{\otimes 2}) + k^{-1}J\Sigma_u J^T - \Big\{J\varphi_1\Big(\frac{\gamma_0}{k}\Big)\varphi_0^{-1}\Big(\frac{\gamma_0}{k}\Big)\Big\}^{\otimes 2}\Big].$$

It is also difficult to compare Σ^*_{cms} and Σ^*_{cee} directly. Therefore, we compare them in some special cases. We assume that U is from $N(0,\Sigma)$, where Σ is known. Then, $\varphi_0(\gamma) = \exp(\gamma^T\Sigma\gamma/2)$, and $\varphi_1(\gamma) = \exp(\gamma^T\Sigma\gamma/2)\Sigma\gamma$, $\varphi_2(\gamma) = \exp(\gamma^T\Sigma\gamma/2)\{\Sigma + (\Sigma\gamma)^{\otimes 2}\}$. Using some simple algebraic manipulation, we have

$$\begin{aligned}\Sigma^*_{cms} = & k^{-1}\big[E(\varepsilon^2+\varepsilon^{-2})\exp(\gamma_0^T\Sigma\gamma_0)\big\{E[Z^{\otimes 2}] + (J\Sigma\gamma_0)^{\otimes 2} + J\Sigma J^T\big\} \\ & - 2\exp(-\gamma_0^T\Sigma\gamma_0)\big\{E[Z^{\otimes 2}] - (J\Sigma\gamma_0)^{\otimes 2} + J\Sigma J^T\big\}\big] \\ & + \frac{(k-1)}{k\big\{E(\varepsilon^2+\varepsilon^{-2}) - 2\big\}E[Z^{\otimes 2}]}.\end{aligned}$$

Similarly, it follows that

$$\begin{aligned}\Sigma^*_{cee} = & E(\varepsilon^2+\varepsilon^{-2})\exp(k^{-1}\gamma_0^T\Sigma\gamma_0)\big\{E[Z^{\otimes 2}] + k^{-2}(J\Sigma\gamma_0)^{\otimes 2} + k^{-1}J\Sigma J^T\big\} \\ & - 2\exp(-k^{-1}\gamma_0^T\Sigma\gamma_0)\big\{E[Z^{\otimes 2}] - k^{-2}(J\Sigma\gamma_0)^{\otimes 2} + k^{-1}J\Sigma J^T\big\}.\end{aligned}$$

Theorem 4. *Assume the conditions of Lemma* 1*. If* $U \sim N(0,\Sigma)$*, we then have* $\Sigma^*_{cms} \geq \Sigma^*_{cee}$*.*

Theorem 4 shows that $\hat{\beta}^*_{cee}$ outperforms $\hat{\beta}^*_{cms}$ under the normal assumption of the measurement error. This result implies that $\hat{\beta}_{cee}$ may outperform $\hat{\beta}_{cms}$ under the normal assumption of the measurement error U, which is also verified by our simulation studies. For other familiar distributions of U, the covariance matrices do not have a simple form, which makes comparisons difficult. Instead, various simulations have been conducted to compare the two methods.

6. Simulation Studies

In this section, various simulation studies were conducted to assess the finite-sample performance of the proposed estimators. Response variable Y was generated from the multiplicative regression model,

$$Y = \exp(c_0 + \alpha_0 V_1 + \gamma_0 X)\varepsilon, \tag{6.1}$$

where V_1 and X are two covariates generated from the bivariate normal distribution with $Var(X) = Var(V_1) = 1$ and $Cov(X, V_1) = 0.5$, and $(c_0, \alpha_0, \gamma_0) = (1, 1, 2)$. We considered two model error distributions: $\log\varepsilon \sim Uniform(-2, 2)$ and $\log\varepsilon \sim N(0, 0.25)$. Both cases are usually considered in some literatures on the relative error; see, for example, Chen et al. (2010), Zhang and Wang (2013), and Chen et al. (2016), among others. The covariate V_1 was measured precisely, whereas X was measured with an error. The surrogate W of X was generated from the classical error model $W = X + U$, where U is the measurement error term. In order to show that the proposed error-corrected methods can handle many symmetric measurement error distributions, we considered two different distributions for U, namely, $N(0, 0.25)$ and $Uniform(-\sqrt{3}/2, \sqrt{3}/2)$, and in both cases where the error variance is 0.25. For each parameter configuration, every subject has three replicates of the surrogate ($n_i = 3$). To assess the finite-sample performance, we calculated the bias (Bias), empirical standard errors (SE) and standard error estimators (SEE). The sample size n is taken to be 200 and 500, respectively, and the simulation results are based on 2,000 replications.

We analyzed the simulated data sets using seven methods: the **LPRE**-based **f**ull data (**LPREf**) method, the **l**east **s**quare-based **f**ull data (**LSf**) method, the **LPRE**-based **n**ai**v**e method (**LPREnv**) (given in Section 4.1), the **l**east **s**quare-based **n**ai**v**e (**LSnv**) method, the **c**orrected **l**east **s**quare (**CLS**) method (Carroll et al. (2006); Buonaccorsi (2010)), the **c**onditional **m**ean **s**core (**CMS**) method proposed in Section 3.2, and the **c**orrected **e**stimating **e**quation (**CEE**) method proposed in Section 4.2. The LPREf and LSf methods are treated as gold standards. For the LPREf method, the LPREf estimator is obtained by minimizing the LPRE criterion suing the true values of X for all subjects. In order to implement the least square based methods, we converted model (6.1) into the following linear model by taking a logarithmic transformation,

$$Y^* = c_0 + \alpha_0 V_1 + \gamma_0 X + \varepsilon^*, \tag{6.2}$$

where $Y^* = \log Y$ and $\varepsilon^* = \log \varepsilon$. The LSf estimator is just the least square estimator of (6.2) using the true covariates. The LSnv estimator is the LSf estimator, but with X replaced by the average of its surrogates. For the CLS method, we implemented the corrected least square method for the linear model (6.2). For the LPRE based methods, we used the Newton-Raphson procedure to solve the estimating equations by taking (0, 0, 0) as the initial value of (c, α, γ). The results was reported in Tables 1 and 2.

Table 1 was conducted with $\log \varepsilon \sim Uniform(-2, 2)$ whereas Table 2 was carried out with $\log \varepsilon \sim N(0, 0.25)$. From Tables 1 and 2, we have the following observations. Both naive estimators (LPREnv and LSnv) for α_0 and γ_0 suffer bias, and the bias does not decrease as the sample size increases. This implies that naive methods may define inconsistent estimators for α_0 and γ_0. All estimators except for the naive estimators exhibit very small bias, and the bias decreases as the sample size increases, as expected. Hence, both the proposed methods and

Table 1. Simulation results for $\log \varepsilon \sim Uniform(-2, 2)$. LPREf, LSf, LPREnv, LSnv, CLS, CMS, and CEE stand for the full LPRE, full least square(LS), naive LPRE, naive LS, classical corrected LS, proposed conditional mean score, and proposed corrected estimating equation estimators. All entries are multiplied by 100.

	$\hat{\alpha}$			$\hat{\gamma}$			$\hat{c}$		
method	Bias	SE	SEE	Bias	SE	SEE	Bias	SE	SEE
				$n = 200$, $U \sim Uniform(-\sqrt{3}/2, \sqrt{3}/2)$					
LPREf	-0.17	7.76	7.69	0.14	7.70	7.68	0.14	6.68	6.65
LSf	-0.18	9.37	9.44	0.13	9.32	9.42	0.18	8.09	8.13
LPREnv	9.63	9.42	8.88	-19.38	8.57	8.51	0.16	7.95	7.84
LSnv	9.72	10.53	10.31	-19.65	9.72	9.90	0.22	8.97	9.00
CLS	-0.60	11.02	10.76	0.85	11.14	11.34	0.19	9.09	9.17
CMS	-0.48	9.41	8.96	0.73	9.31	9.19	0.15	7.77	7.71
CEE	-0.71	9.92	9.32	1.14	10.00	9.81	0.15	8.13	7.97
				$n = 200$, $U \sim N(0, 0.25)$					
LPREf	-0.08	7.86	7.68	0.17	7.98	7.69	0.11	6.74	6.66
LSf	-0.15	9.38	9.43	0.22	9.70	9.43	0.12	8.14	8.14
LPREnv	9.90	8.97	8.90	-19.84	9.09	8.55	0.06	8.13	7.89
LSnv	9.80	9.96	10.30	-19.80	10.12	9.90	0.13	9.06	9.01
CLS	-0.40	10.41	10.76	0.66	11.58	11.39	0.11	9.16	9.18
CMS	-0.88	10.75	10.13	1.45	12.81	11.48	0.08	8.73	8.42
CEE	-0.51	9.50	9.37	1.01	10.58	9.93	0.06	8.31	8.03
				$n = 500$, $U \sim Uniform(-\sqrt{3}/2, \sqrt{3}/2)$					
LPREf	0.09	4.98	4.86	-0.14	4.99	4.87	-0.01	4.27	4.21
LSf	0.10	6.08	5.96	-0.15	6.10	5.97	-0.01	5.22	5.15
LPREnv	10.14	5.88	5.68	20.00	5.71	5.45	-0.12	5.10	5.01
LSnv	10.27	6.68	6.52	-20.28	6.45	6.26	-0.10	5.82	5.70
CLS	0.17	7.01	6.79	-0.10	7.38	7.15	-0.07	5.90	5.78
CMS	0.17	5.88	5.70	-0.16	6.09	5.85	-0.08	4.96	4.90
CEE	0.12	6.21	5.96	-0.01	6.54	6.27	-0.08	5.18	5.08
				$n = 500$, $U \sim N(0, 0.25)$					
LPREf	-0.10	4.87	4.87	-0.10	4.94	4.86	0.20	4.25	4.21
LSf	-0.08	5.94	5.97	-0.17	6.01	5.96	0.23	5.19	5.16
LPREnv	9.75	5.79	5.73	-20.06	5.65	5.50	0.23	5.12	5.05
LSnv	9.84	6.55	6.53	-20.18	6.31	6.26	0.24	5.75	5.71
CLS	-0.26	6.84	6.82	-0.03	7.24	7.19	0.24	5.81	5.79
CMS	-0.56	6.74	6.56	0.44	7.96	7.51	0.24	5.49	5.42
CEE	-0.43	6.17	6.03	0.24	6.65	6.39	0.22	5.22	5.14

CLS can effectively correct the biases caused by measurement errors and define consistent estimators. It is noted that the bias for all of the estimators, including both the naive estimators of c_0, are very small, and that the SEE and SE of all the

Table 2. Simulation results for $\log\varepsilon \sim N(0, 0.25)$. LPREf, LSf, LPREnv, LSnv, CLS, CMS, and CEE stand for the full LPRE, full least square(LS), naive LPRE, naive LS, classical corrected LS, proposed conditional mean score, and proposed corrected estimating equation estimators. All entries are multiplied by 100.

	$\hat{\alpha}$			$\hat{\gamma}$			$\hat{c}$		
method	Bias	SE	SEE	Bias	SE	SEE	Bias	SE	SEE
$n = 200$, $U \sim Uniform(-\sqrt{3}/2, \sqrt{3}/2)$									
LPREf	-0.08	4.16	4.05	-0.10	4.09	4.04	0.10	3.51	3.53
LSf	-0.08	4.13	4.09	-0.10	4.07	4.09	0.09	3.50	3.53
LPREnv	9.75	6.17	5.90	-19.82	5.59	5.62	0.20	5.18	5.21
LSnv	9.91	6.12	5.97	-20.11	5.55	5.74	0.17	5.14	5.22
CLS	-0.31	6.65	6.44	0.35	6.94	6.93	0.14	5.24	5.43
CMS	-0.15	6.08	5.81	-0.04	6.01	6.01	0.13	4.88	4.99
CEE	-0.34	6.69	6.30	0.38	6.91	6.73	0.15	5.25	5.36
$n = 200$, $U \sim N(0, 0.25)$									
LPREf	0.16	4.29	4.04	-0.07	4.25	4.04	0.08	3.61	3.53
LSf	0.15	4.27	4.08	-0.07	4.24	4.09	0.08	3.59	3.53
LPREnv	10.06	6.29	5.93	-20.07	5.97	5.72	-0.05	5.39	5.27
LSnv	10.03	6.22	5.97	-20.04	5.83	5.75	-0.03	5.31	5.23
CLS	-0.19	6.72	6.44	0.48	7.13	7.02	-0.01	5.44	5.44
CMS	-0.30	7.99	7.05	0.80	11.06	8.27	0.04	6.00	5.78
CEE	-0.21	6.83	6.38	0.56	7.34	6.89	-0.05	5.55	5.45
$n = 500$, $U \sim Uniform(-\sqrt{3}/2, \sqrt{3}/2)$									
LPREf	0.00	2.66	2.59	0.03	2.62	2.59	-0.05	2.18	2.24
LSf	0.00	2.66	2.59	0.03	2.63	2.59	-0.05	2.17	2.24
LPREnv	9.96	3.76	3.79	-19.77	3.62	3.61	-0.07	3.24	3.34
LSnv	10.09	3.75	3.79	-20.07	3.58	3.63	-0.06	3.18	3.31
CLS	-0.04	4.07	4.07	0.12	4.43	4.36	-0.04	3.29	3.42
CMS	0.02	3.75	3.72	0.02	3.95	3.84	-0.04	3.10	3.18
CEE	-0.03	4.09	4.05	0.11	4.41	4.30	-0.07	3.33	3.43
$n = 500$, $U \sim N(0, 0.25)$									
LPREf	-0.07	2.52	2.58	0.04	2.58	2.58	0.10	2.25	2.24
LSf	-0.07	2.50	2.58	0.04	2.58	2.58	0.10	2.24	2.23
LPREnv	9.92	3.77	3.81	-19.99	3.69	3.68	0.03	3.40	3.37
LSnv	9.90	3.72	3.78	-19.97	3.59	3.63	0.07	3.31	3.31
CLS	-0.19	4.03	4.07	0.17	4.38	4.41	0.06	3.40	3.43
CMS	-0.25	4.69	4.60	0.24	5.69	5.46	0.02	3.71	3.70
CEE	-0.18	4.11	4.10	0.17	4.52	4.42	0.02	3.48	3.48

estimators are quite close to each other. When both the model error $\log\varepsilon$ and the measurement error U follow the normal distribution, the classical CLS method is of slightly smaller SE than the proposed CEE and CMS methods. However, when both $\log \varepsilon$ and U follow the uniform distribution, both the proposed CEE and CMS methods perform better than the CLS method in terms of SE. When $\log \varepsilon$ follows a normal distribution, but the measurement error U follows a uniform distribution, the proposed CMS method outperforms CEE and CLS in terms of SE. When $\log \varepsilon$ is from the uniform distribution, but the measurement error U is from the normal distribution, the proposed CEE estimator is of smaller SE than CMS and CLS.

7. Data Analysis

As an illustration, we apply the proposed methods to an AIDS clinic study conducted by the AIDS Clinical Trial Group (ACTG) 315 (Lederman et al. (1998); Wu and Ding (1999); Liang, Wu and Carroll (2003)). In this study, patients with evaluable HIV-1 infection were treated with potent antiviral drugs consisting of ritonavir, 3TC, and AZT. Both plasma HIV RNA copies (viral load) and CD4+ cell counts were repeatedly quantified at treatment days 0, 2, 7, 10, 14, 28, 56, 84, 168, and 336 after initiation of treatment. Because plasma HIV RNA copies (viral load) and CD4+ cell counts are two crucial medical index in AIDS clinical research, it is necessary to study the relationship during HIV/AIDS treatment. The data of 46 evaluable patients in the study are available at `https://www.urmc.rochester.edu/biostat/people/faculty/wusite/datasets/ACTG315LongitudinalDataViralLoad.cfm`. In this example, we focus on the data for the first 2 days of treatment. Among the 45 evaluable patients, 33 patients have two measurements of day 0 and day 2, 10 patients with just one measurement on day 0, and two patients with just one measurement on day 2. We are interested in the relationship between the average viral load and the average CD4+ cell counts of the first two days of treatment. However, both viral load and CD4+ cell counts are subject to measurement errors. Adjusting the measurement error usually requires replication, validation data, or other information to estimate the error structure.

Inspired by a referee's suggestion, paired sample t-tests were used to test whether the measured values (viral load and CD4+ cell counts) of day 0 and day 2 can be treated as replicate measurements of the average values for the first two days of treatment. The p-values are 0.347 and 0.128 for viral load and CD4+ cell counts, respectively. This implies that the viral loads and the CD4+ cell counts in day 0 and day 2 can be treated as the replicate surrogates of the average viral load and CD4+ cell counts for the first two days of treatment respectively.

Table 3. Analysis of the ACTG315 data.

method	c_0		γ_0	
	Est	SEE	Est	SEE
LSnv	12.217	0.436	-0.377	0.216
CLS	12.282	0.463	-0.412	0.237
LPREnv	12.309	0.374	-0.453	0.198
CMS	12.383	0.401	-0.491	0.212
CEE	12.424	0.416	-0.514	0.220

Thereafter, we considered the following additive measurement error model to link the underlying CD4+ cell counts to its surrogate measurements:

$$W_{i,r} = X_i + U_{i,r}, \quad r = 1, \dots, n_i, \quad i = 1, \dots, 45,$$

where $n_i = 2$ for subjects with two measurements in day 0 and day 2; otherwise, $n_i = 1$. We take the average of the viral loads for day 0 and day 2 as the response variable Y_i for each patient. Clearly, the response is positive. Hence, it is natural to use the following multiplicative regression model

$$Y_i = \exp(c_0 + \gamma_0 X_i)\varepsilon_i,$$

to fit the data set, where c_0 is the intercept and γ_0 is the regression parameter. Here, we could treat X_i and Y_i as the average CD4+ cell counts and viral loads of the first two days, respectively. The absolute error criterion cannot be applied to the multiplicative model directly, otherwise an inconsistent estimator is defined. In order to make a comparison with the least square based approach (an absolute error criterion), we also consider the linear model by taking the logarithmic transformation. We analyzed the data set using the five methods, LSnv, CLS, LPREnv, CMS and CEE methods, respectively. Table 3 calculated estimate values of all the five methods and standard error values.

First, the proposed CMS and CEE estimators for γ_0 and c_0 have larger absolute values than LPREnv does, and they are close to each other; the classical corrected least square (CLS) estimator also is of larger absolute value than the naive least square (LSnv) estimator. This implies that ignoring the measurement error can attenuate the estimate considerably.

Secondly, it also can be observed that the relative error based LPREnv, CMS and CEE estimators of γ_0 are of bigger absolute values than the least square based LSnv and CLS estimators. Hence, the proposed methods show that the average CD4+ cell counts are more closely related to the HIV viral loads. From Table 3, it is seen that the estimated value of γ_0 based on the CLS is -0.412, and the estimated values of γ_0 based on the two proposed methods are -0.491 and -0.514, respectively. The relative differences between the CLS estimator and the proposed CMS and CEE estimators are 0.1912 and 0.2475, respectively. Thus we conclude that the proposed methods perform similarly to the CLS estimator in some cases, and perform better in some other cases. This suggests that the proposed methods are useful for practical settings, and the criterion based on the relative error is more reasonable. One of the reasons may be that the distribution of $\log \epsilon$ or the distribution of the measurement error U are not normal. Another reason may be that the proposed criterion is scale free, which can use information of the subjects with small values effectively. If we use the log-linear model with least square loss, the large values of some subjects can overwhelm the effect of the small values of some subjects.

Supplementary Material

The online Supplementary Material contains an additional simulation with the assumptions violated, as well as the detailed proofs of the theorems.

Acknowledgements

Wang's research was supported by the National Natural Science Foundation of China (General program 11871460, Key program 11331011, and program for Creative Research Group in China 61621003), a grant from the Key Lab of Random Complex Structure and Data Science, CAS.

Appendix

Regularity Conditions:

Condition C1. $E[(1/\varepsilon+\varepsilon)^2\exp(\delta\|Z\|)]<\infty$ for some $\delta>0$.

Condition C2. $E[(1/\varepsilon+\varepsilon)ZZ^T]$ is positive definite.

Condition C3. The model error ε satisfies $E(\varepsilon-1/\varepsilon|Z)=0$.

Condition C4. The measurement errors $U_{i,r}, r=1,\ldots,n_i$ are independent and identically distributed (i.i.d.), symmetrically distributed and independent of (Z_i,Y_i) for $i=1,\ldots,n$.

Condition C5. $E(|U|^2)<\infty$. In addition, there exists a compact neighborhood $\mathcal{B}$ of γ_0 such that

$$E[\sup_{\gamma\in\mathcal{B}}|U|^2\exp(U^T\gamma)]<\infty \quad \text{and} \quad E[\sup_{\gamma\in\mathcal{B}}|U|^2\exp(2U^T\gamma)]<\infty.$$

Condition C6. The repeated times n_i has an upper bound m, namely, $1 \leq n_i \leq m$. In addition, the limit of $|\mathcal{A}_k|/n$ exist, denoted by ρ_k, where $k = 1, \ldots, m$. Conditions C1-C3 are almost minimal for the asymptotic normality to hold in LPRE with the covariates measured precisely. Condition C4-C6 are the regular conditions to deal with the measurement error in the covariates.

References

Buonaccorsi, J. P. (2010). *Measurement Error: Models, Methods, and Applications.* Chapman and Hall/CRC.

Carroll, R. J., Ruppert, D., Stefanski, L. A. and Crainiceanu, C. M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective.* Chapman and Hall/CRC.

Chen, K., Guo, S., Lin, Y. and Ying, Z. (2010). Least absolute relative error estimation. *Journal of the American Statistical Association* **105**, 1104–1112.

Chen, K., Lin, Y., Wang, Z. and Ying, Z. (2016). Least product relative error estimation. *Journal of Multivariate Analysis* **144**, 91–98.

Hammer, S. M., Katzenstein, D., Hughes, M. D., Gundacker, H., Schooley, R. T., Haubrich, R., Henry, W. K., Lederman, M. M., Phair, J. P., Niu, M. and et al. (1996). A trial comparing nucleoside monotherapy with combination therapy in hiv-infected adults with cd4 cell counts from 200 to 500 per cubic millimeter. *The New England Journal of Medicine* **335**, 1081–1090.

Hu, C. and Lin, D. Y. (2004). Semiparametric failure time regression with replicates of mismeasured covariates. *Journal of the American Statistical Association* **99**, 105–118.

Jin Z., Lin D.Y., Wei L.J. and Ying Z. (2003). Rank-based inference for the accelerated failure time model. *Biometrika*, **90**, 341–353.

Khoshgoftaar, T. M., Bhattacharyya, B. B. and Richardson, G. D. (1992). Predicting software errors, during development, using nonlinear regression models: a comparative study. *IEEE Transactions on Reliability* **41**, 390–395.

Lederman, M. M., Connick, E., Landay, A. L., Kuritzkes, D. R., Spritzler, J., Clair, M. S., Kotzin, B. L., Fox, L., Chiozzi, M. H., Leonard, J. M. and et al. (1998). Immunologic responses associated with 12 weeks of combination antiretroviral therapy consisting of zidovudine, lamivudine, and ritonavir: Results of aids clinical trials group protocol 315. *The Journal of Infectious Diseases* **178**, 70–79.

Li, Z., Lin, Y., Zhou, G. and Zhou, W. (2014). Empirical likelihood for least absolute relative error regression. *Test* **23**, 86–99.

Liang, H. Y., Wu, H. and Carroll, R. J. (2003). The relationship between virologic and immunologic responses in aids clinical research using mixed-effects varying-coefficient models with measurement error. *Biostatistics* **4**, 297–312.

Liu, X., Lin, Y. and Wang, Z. (2016). Group variable selection for relative error regression. *Journal of Statistical Planning and Inference* **175**, 40–50.

Makridakis, S. (1986). The forecasting accuracy of major time series methods. *Journal of the American Statistical Association* **81**, 262.

Narula, S. C. and Wellington, J. F. (1977). Prediction, linear regression and the minimum sum of relative errors. *Technometrics* **19**, 185–190.

Park H., Stefanski L.A. (1998). Relative-error prediction. *Statistics* & *Probability Letters* **40**, 227–236.

Sinha, S. and Ma, Y. (2016). Analysis of proportional odds models with censoring and errors-in-covariates. *Journal of the American Statistical Association* **111**, 1301–1312.

Teekens, R. and Koerts, J. (1972). Some statistical implications of the log transformation of multiplicative models. *Econometrica* **40**, 793–819.

Wang, Z., Liu, W. and Lin, Y. (2015). A change-point problem in relative error-based regression. *Test* **24**, 835–856.

Wei, L. J. (1992). The accelerated failure time model: A useful alternative to the cox regression model in survival analysis. *Statistics in Medicine* **11**, 1871–1879.

Wu, H. and Ding, A. A. (1999). Population hiv-1 dynamics in vivo: applicable models and inferential tools for virological data from aids clinical trials. *Biometrics* **55**, 410–418.

Wu, Y., Ma, Y. and Yin, G. (2015). Smoothed and corrected score approach to censored quantile regression with measurement errors. *Journal of the American Statistical Association* **110**, 1670–1683.

Xia, X., Liu, Z. and Yang, H. (2016). Regularized estimation for the least absolute relative error models with a diverging number of covariates. *Computational Statistics & Data Analysis* **96**, 104–119.

Yan, Y. and Yi, G. Y. (2016). A class of functional methods for error-contaminated survival data under additive hazards models with replicate measurements. *Journal of the American Statistical Association* **111**, 684–695.

Zhang, Q. and Wang, Q. (2013). Local least absolute relative error estimating approach for partially linear multiplicative model. *Statistica Sinica* **23**, 1091–1116.

School of Statistics and Mathematics, Zhejiang Gongshang University, Zhejiang, 310018, China.
Academy of Mathematics and Systems, Chinese Academy of Science, Beijing, 100190, China.

E-mail: qhwang@amss.ac.cn

Department of Statistics and Finance, University of Science and Technology of China, Hefei, Anhui, 230026, China.

E-mail: dhh@mail.ustc.edu.cn

(Received August 2016; accepted June 2018)