Supplemental Web Materials

A Regularization Corrected Score Method for Nonlinear Regression Models with Covariate Error

David M. Zucker

Department of Statistics, Hebrew University

Mount Scopus, Jerusalem, Israel

email: mszucker@mscc.huji.ac.il

Malka Gorfine

Faculty of Industrial Engineering and Management

Technion - Israel Institute of Technology, Haifa, Israel

Yi Li

Department of Biostatistics

University of Michigan, Ann Arbor MI, USA

Mahlet G. Tadesse

Department of Mathematics and Statistics

Georgetown University, Washington DC, USA

and

Donna Spiegelman

Departments of Epidemiology and Biostatistics

Harvard School of Public Health, Boston MA, USA

Web Appendix A.1: Evaluation of Integrals of the Form $A_i g(x)$

As noted in the main text, the RECS procedure requires evaluation of integrals of the form

$$E[g(W_i)|X_i = x] = A_i g(x) = \int a_i(x, w) g(w) dw. \tag{1}$$

In our simulation studies, we work with measurement error models of the form $W = X + \sigma(X)\varepsilon$, where ε is a random variable with density f_ε, independent of X. In this case, we can write

$$A_i g(x) = \int f_\varepsilon(v) g(x + \sigma(x) v) dv.$$

This integral can be evaluated using a quadrature approximation of the form

$$A_i g(x) \doteq \sum_{k=1}^{K'} r_k g(x + \sigma(x) v_k)$$

for suitable quadrature points and weights v_k and r_k. One case we consider is $\varepsilon \sim N(0, 1)$, and in this case we use (modified) Gauss-Hermite points and weights. Another case we consider is the case where ε is the double-exponential (Laplace) distribution with variance 1, and in this case we use Gauss-Laguerre type points and weights (with the points and weights on the negative side of the real axis being the mirror image of those on the positive side).

In cases where specialized methods of the above sort are not readily available, a more general numerical approach can be used. In particular, if we let $F_{W|X}$ be the conditional distribution function of W given X, then $A_i g(x)$ can be evaluated as

$$A_i g(x) \doteq \frac{1}{K'} \sum_{k=1}^{K'} g(F_{W|X}^{-1}(v_k|x))$$

with $v_k = (k - \frac{1}{2})/K'$.

For the MODCHI distribution used in the simulation study, integrals of the form

$A_i g(x)$ were evaluated using the representation

$$A_i g(x) = \frac{2}{\Pr(\chi_1^2 \leq 5)} \int_0^{\sqrt{5}} \varphi(v) g\left(x + \gamma \left[\frac{v^2 - \mu_{MC}}{\sigma_{MC}}\right]\right) dv$$

in conjunction with Gauss-Hermite quadrature, where μ_{MC} and σ_{MC} denote, respectively, the mean of the chi-square distribution truncated at 5.

Web Appendix A.2: Theorem of Section 3.2 - Regularity Conditions and Proof

Regularity Conditions

A1. The parameter space $\mathbf{\Theta}$ is compact with a nonempty interior which includes the true value $\boldsymbol{\theta}_0$.

A2. The function $\mathbf{u}_i(x, \boldsymbol{\theta})$ is continuously differentiable in $\boldsymbol{\theta}$ over $\mathbf{\Theta}$ for every x, with derivative that is bounded over x by an L^2 function of x.

A3. The matrix $\mathbf{D}_E(\boldsymbol{\theta})$ defined by $(D_E)_{rs}(\boldsymbol{\theta}) = -E[(\partial^2/\partial\theta_r\partial\theta_s)\log f(Y|X, \mathbf{Z}, \boldsymbol{\theta})]$ is positive definite over $\mathbf{\Theta}$.

A4. The null space $\mathcal{N}(A_i^*)$ consists only of the zero function, i.e., the only solution to $A_i^* h = 0$ is $h = 0$.

Proof

In proving Part (a) of the theorem, we rely on the L^2 theory of integral operators as set forth, for example, in Kress (1989). We recall that, for a general integral operator B, the range of B is defined as $\mathcal{R}(B) = \{h \in L^2 : h = Bg \text{ for some } g \in L^2\}$ and the null space of B is defined $\mathcal{N}(B) = \{g \in L^2 : Bg = 0\}$. We use a superscript $\perp$ to denote orthogonal complement and the notation $cl(C)$ to denote the L^2 closure of a set $C \subset L^2$.

Kress (1989, Theorem 15.8) states that for a bounded linear operator $B : L^2 \to L^2$ with adjoint B^*, we have $\mathcal{N}(B^*)^\perp = cl(\mathcal{R}(B))$. Since $a_i(x, w)$ is a conditional density, the operator A_i is a bounded linear operator with norm 1. Assumption A4 specifies that $\mathcal{N}(A_i^*)$ consists only of the zero function. It follows that the L^2 closure of $\mathcal{R}(A_i)$ is equal to the whole of L^2. This, in turn, implies that $\inf_{\bar{\delta}\in L^2} \|A_i\bar{\delta} - g\| = 0 \; \forall \, g \in L^2$, although the infimum is not necessarily attained, which, in our setting, corresponds to the fact that an exact corrected score may not exist.

Now, for a given L^2 function δ, let $\bar{\delta}(\alpha)$ denote the minimizer of $\mathcal{L}(\bar{\delta}; A_i, \delta, \alpha) = \|A_i\bar{\delta} - \delta\|^2 + \alpha\|\bar{\delta}\|^2$. We claim that $\lim_{\alpha\to 0} \|\bar{\delta}(\alpha) - \delta\|$=0. The proof is simple, and is implicit in Kress (1989, Chapter 16), but we give it for completeness. We have

$$\|A_i\bar{\delta}(\alpha) - \delta\|^2 \leq \|A_i\bar{\delta}(\alpha) - \delta\|^2 + \alpha\|\bar{\delta}(\alpha)\|^2 \leq \|A_i g - \delta\|^2 + \alpha\|g\|^2$$

for any $g \in L^2$. Letting $\alpha \to 0$, we get $\lim_{\alpha\to 0} \|A_i\bar{\delta}(\alpha) - \delta\|^2 \leq \|A_i g - \delta\|^2$. Since g was arbitrary, we get $\lim_{\alpha\to 0} \|A_i\bar{\delta}(\alpha) - \delta\|^2 \leq \inf_g \|A_i g - \delta\|^2$, but the infimum on the right side, as we just saw, is equal to zero. We have thus proved the claim.

In the context of our RECS estimator, we obtain the following result: defining $r_{ij}(x, \boldsymbol{\theta}, \alpha) = A_i\bar{u}_{ij}(x, \boldsymbol{\theta}, \alpha) - u_{ij}(x, \boldsymbol{\theta})$, we have $\|r_{ij}(\cdot, \boldsymbol{\theta}, \alpha)\| \to 0$ as $\alpha \to 0$. At this point, we have this convergence only at a fixed value of $\boldsymbol{\theta}$. However, since $\boldsymbol{\Theta}$ is assumed compact and $u_{ij}(x, \boldsymbol{\theta})$ is continuous in $\boldsymbol{\theta}$, pointwise convergence implies uniform convergence. This yields the desired result.

Let us now turn to Part (b). Define $\Delta_{ijs}(x) = u_{ijs}(x, \boldsymbol{\theta}) - A_i u_{ijs}(x, \boldsymbol{\theta})$ and $\bar{\Delta}_{ijs}(w, \boldsymbol{\theta}, \alpha) = \bar{u}_{ijs}(w, \boldsymbol{\theta}, \alpha) - u_{ijs}(w, \boldsymbol{\theta})$. As stated near the end of Section 2 of the main paper, the minimizer of $\mathcal{L}(\bar{\delta}; A_i, \delta, \alpha)$ can be expressed as $\bar{\delta} = (A_i^* A_i + \alpha\mathcal{I})^{-1} A_i^* \delta$. This is a linear function of δ. It follows that, just as the function $\bar{\Delta}_{ij}(\cdot, \boldsymbol{\theta}, \alpha)$ is the minimizer of $\|A_i\bar{\delta} - \Delta_{ij}\|^2 + \alpha\|\bar{\delta}\|^2$, so, too, the function $\bar{\Delta}_{ijs}(\cdot, \boldsymbol{\theta}, \alpha)$ is the minimizer of $\|A_i\bar{\delta} - \Delta_{ijs}\|^2 + \alpha\|\bar{\delta}\|^2$. We can therefore apply the arguments just used in the proof of Part (a) to prove Part (b).

Given the results in Parts (a) and (b) of the theorem, the results in Parts (c) and

(e) follow from standard estimating equations theory, as in, for example, Huber (1967), White (1982), and van der Vaart (1998, Ch. 5).

Finally, we turn to Part (d). From the development in Part (a), we have

$$\bar{\mathbf{u}}_E(\tilde{\boldsymbol{\theta}}^{(\alpha)}, \alpha) = E[\bar{\mathbf{u}}_i(W_i, \tilde{\boldsymbol{\theta}}^{(\alpha)})] = E[\mathbf{u}_i(X_i, \tilde{\boldsymbol{\theta}}^{(\alpha)})] + E[\mathbf{r}_i(X_i, \tilde{\boldsymbol{\theta}}^{(\alpha)}, \alpha)].$$

The left side of the above equation, by definition, is zero. Expanding the first term of the right side in a first-order Taylor series around $\boldsymbol{\theta}_0$, we get

$$-\mathbf{D}_E(\boldsymbol{\theta}^{\#})(\tilde{\boldsymbol{\theta}}^{(\alpha)} - \boldsymbol{\theta}_0) + E[\mathbf{r}_i(X_i, \tilde{\boldsymbol{\theta}}^{(\alpha)}, \alpha)] = 0.$$

where $\mathbf{D}_E(\boldsymbol{\theta})$ is as defined in Assumption A3 and $\boldsymbol{\theta}^{\#}$ is some value between $\boldsymbol{\theta}_0$ and $\tilde{\boldsymbol{\theta}}^{(\alpha)}$. Part (a) implies that the second term on the right side tends to zero as $\alpha \to 0$. Given this, along with the nonsingularity condition A3, we obtain the desired conclusion $\tilde{\boldsymbol{\theta}}^{(\alpha)} \to \boldsymbol{\theta}_0$ as $\alpha \to 0$.

Web Appendix A.3: Correcting the Covariance for Estimation of $\mathbf{A}^{(i)}$

We describe here how to correct the covariance of $\hat{\boldsymbol{\theta}}$ for estimation of $\mathbf{A}^{(i)}$. In the development below, we generally suppress the dependence of various quantities on the penalty parameter α.

The parameter $\boldsymbol{\xi}$ entering into $a_i(x, w, \boldsymbol{\xi})$ is estimated on the basis of an external or internal validation sample, or a replicate measures study, of size denoted by m. Let $\boldsymbol{\xi}_0$ denote the true value of $\boldsymbol{\xi}$. We assume that the estimator $\hat{\boldsymbol{\xi}}$ has an approximate normal distribution with mean $\boldsymbol{\xi}_0$ and covariance matrix $m^{-1}\boldsymbol{\Gamma}$, along with an estimator of the matrix $\boldsymbol{\Gamma}$. This setup is a typical one in practice. For the asymptotics we assume that m and n are of the same order of magnitude, i.e., $m/n \to \zeta$ for some constant ζ as $n \to \infty$. Otherwise the error in $\hat{\boldsymbol{\xi}}$ will either be dominated by or will dominate the error in $\boldsymbol{\theta}$ due to the variation in the outcome data. Typically ζ will be between 0 and 1. The asymptotic covariance matrix of $\boldsymbol{\xi}$ may then be expressed as $n^{-1}\zeta^{-1}\boldsymbol{\Gamma}$. To emphasize the dependence of the corrected score on $\boldsymbol{\xi}$, we write $\bar{\mathbf{U}}(\boldsymbol{\theta}, \boldsymbol{\xi})$. The estimated asymptotic covariance matrix of $\bar{\mathbf{U}}(\tilde{\boldsymbol{\theta}}^{(\alpha)}, \boldsymbol{\xi}_0)$ is $n^{-1}\mathbf{F}^{(\alpha)}(\hat{\boldsymbol{\theta}})$ where $\mathbf{F}^{(\alpha)}(\boldsymbol{\theta})$ is as defined in (8). We denote the asymptotic covariance between $\bar{\mathbf{U}}(\boldsymbol{\theta})$ and $\hat{\boldsymbol{\xi}}$ by $n^{-1}\boldsymbol{\Upsilon}$. The form of $\boldsymbol{\Upsilon}$ depends on the type of data used to estimate $\boldsymbol{\xi}$, and will be discussed shortly.

Let $\bar{\mathbf{U}}'(\boldsymbol{\theta}, \boldsymbol{\xi})$ denote the matrix whose (r, ν) element is the partial derivative of $\bar{U}_r(\boldsymbol{\theta}, \boldsymbol{\xi})$ with respect to ξ_ν. By Taylor expansion, we have

$$\mathbf{0} = \bar{\mathbf{U}}(\hat{\boldsymbol{\theta}}^{(\alpha)}, \hat{\boldsymbol{\xi}}) = \bar{\mathbf{U}}(\tilde{\boldsymbol{\theta}}^{(\alpha)}, \boldsymbol{\xi}_0)\bar{\mathbf{D}}(\tilde{\boldsymbol{\theta}}^{(\alpha)}, \boldsymbol{\xi}_0)(\hat{\boldsymbol{\theta}}^{(\alpha)} - \tilde{\boldsymbol{\theta}}^{(\alpha)}) + \bar{\mathbf{U}}'(\tilde{\boldsymbol{\theta}}^{(\alpha)}, \boldsymbol{\xi}_0)(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0) + o_p(1/n),$$

leading to

$$-\sqrt{n}(\hat{\boldsymbol{\theta}}^{(\alpha)} - \tilde{\boldsymbol{\theta}}^{(\alpha)}) = \bar{\mathbf{D}}(\tilde{\boldsymbol{\theta}}^{(\alpha)}, \boldsymbol{\xi}_0)^{-1}\left[\{\sqrt{n}\,\bar{\mathbf{U}}(\tilde{\boldsymbol{\theta}}^{(\alpha)}, \boldsymbol{\xi}_0)\} + \bar{\mathbf{U}}'(\tilde{\boldsymbol{\theta}}^{(\alpha)}, \boldsymbol{\xi}_0)\{\sqrt{n}\,(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0)\}\right] + o_p(1/n).$$

Accordingly, the estimated asymptotic covariance matrix of $\sqrt{n}(\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}^{(\alpha)})$ is $\hat{\mathbf{V}}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\xi}})$, where $\hat{\mathbf{V}}(\boldsymbol{\theta}, \boldsymbol{\xi})$ is now defined as

$$\hat{\mathbf{V}}(\boldsymbol{\theta}, \boldsymbol{\xi}) = \bar{\mathbf{D}}(\boldsymbol{\theta}, \boldsymbol{\xi})^{-1}\left[\mathbf{F}^{(\alpha)}(\boldsymbol{\theta}) + \zeta^{-1}\bar{\mathbf{U}}'(\boldsymbol{\theta}, \boldsymbol{\xi})\hat{\boldsymbol{\Gamma}}\bar{\mathbf{U}}'(\boldsymbol{\theta}, \boldsymbol{\xi})^T + \hat{\boldsymbol{\Upsilon}}\bar{\mathbf{U}}'(\boldsymbol{\theta}, \boldsymbol{\xi})^T\right]\bar{\mathbf{D}}(\boldsymbol{\theta}, \boldsymbol{\xi})^{-1}.$$

If $\boldsymbol{\xi}$ is estimated from an external validation or replicate measures study, then $\hat{\boldsymbol{\xi}}$ is obviously independent of the corrected score function $\bar{\mathbf{U}}(\boldsymbol{\theta})$, and thus $\boldsymbol{\Upsilon} = \mathbf{0}$. When $\boldsymbol{\xi}$ is estimated from an internal validation or replicate measures study, $\boldsymbol{\Upsilon}$ is nonzero and must be estimated. We consider the setting where the validation/replicate data are i.i.d. across individuals, and $\boldsymbol{\xi}$ is estimated by maximum likelihood. Let $g_i(\boldsymbol{\xi})$ denote the log likelihood function for the validation/replicate data on individual i. The overall normalized log likelihood for $\boldsymbol{\xi}$ is then $g(\boldsymbol{\xi}) = m^{-1}\sum_{i\in\mathcal{R}} g_i(\boldsymbol{\xi})$, where $\mathcal{R}$ denotes the set of individuals in the internal validation/replicate sample. Let $\mathbf{g}'(\boldsymbol{\xi})$ and $\mathbf{g}''(\boldsymbol{\xi})$ denote the gradient vector and Hessian matrix, respectively, of $g(\boldsymbol{\xi})$, and let $\mathbf{g}_i'(\boldsymbol{\xi})$ denote the gradient of $g_i(\boldsymbol{\xi})$. We can then express $\hat{\boldsymbol{\xi}}$ in terms of the classic asymptotic approximation $\hat{\boldsymbol{\xi}} \doteq -\mathbf{g}''(\boldsymbol{\xi}_0)^{-1}\mathbf{g}'(\boldsymbol{\xi}_0)$.

Define $\boldsymbol{\Omega} = \text{Cov}(\bar{\mathbf{u}}_i(\boldsymbol{\theta}_0, \boldsymbol{\xi}_0), \mathbf{g}'_i(\boldsymbol{\xi}_0))$. The matrix $\boldsymbol{\Omega}$ can be estimated empirically by

$$\hat{\boldsymbol{\Omega}} = \frac{1}{m} \sum_{i \in \mathcal{R}} \bar{\mathbf{u}}_i(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\xi}}) \mathbf{g}'_i(\hat{\boldsymbol{\xi}})^T. \tag{2}$$

We then estimate $\boldsymbol{\Upsilon}$ by $\hat{\boldsymbol{\Upsilon}} = \hat{\boldsymbol{\Omega}} \mathbf{g}''(\hat{\boldsymbol{\xi}})^{-1}$. In addition, in the present setup we have $\hat{\boldsymbol{\Gamma}} = \mathbf{g}''(\hat{\boldsymbol{\xi}})^{-1}$.

In principle, expressions can be worked out for the partial derivatives that make up the matrix $\bar{\mathbf{U}}'(\boldsymbol{\theta}, \boldsymbol{\xi})$, but the algebra is cumbersome. Therefore, in our practical implementation, we use numerical partial derivatives.

Web Appendix B: Description of Bayesian MCMC Method

Richardson et al. (2002) proposed a semiparametric Bayesian measurement error method that uses a mixture of normal distributions to model the distribution of the error prone covariates X. They considered the case in which the main study sample consists of n units on which only W and Y are observed and a validation sample of size m is available with observations on X, W, Y. All $n^* = n + m$ observations are used to formulate a hierarchical model for the measurement error model, as follows ($i = 1, \ldots, n^*$):

$$\begin{aligned}
Y_i | X_i, \beta_0, \beta_1 &\sim \text{Bernoulli}(p_i) \quad \text{with} \quad p_i = \frac{\exp(\beta_0 + \beta_1 X_i)}{1 + \exp(\beta_0 + \beta_1 X_i)} \\
W_i | X_i, \alpha_0, \alpha_1, \gamma &\sim N(\alpha_0 + \alpha_1 X_i, \gamma) \\
X_i | \boldsymbol{\omega}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2, k &\sim \sum_{j=1}^{k} \omega_j N(\mu_j, \sigma_j^2).
\end{aligned}$$

Conjugate priors are specified for the model parameters with hyperparameters defined to lead to weakly informative priors centered around estimates from the validation group. Both the validation set and the main study contribute to the joint posterior distribution which is given by:

$$p(\boldsymbol{\beta})\ p(\boldsymbol{\alpha})\ p(\gamma)\ p(\boldsymbol{\omega}|k)\ p(\boldsymbol{\mu}|k)\ p(\boldsymbol{\sigma}^2|k)\ p(k) \\ \times \prod_{i=1}^{n^*} \left[p(Y_i|X_i, \boldsymbol{\beta})\ p(W_i|X_i, \boldsymbol{\alpha}, \gamma)\ p(X_i|c_i, \boldsymbol{\mu}, \boldsymbol{\sigma}^2)\ p(c_i|\boldsymbol{\omega}, k)\right], \tag{3}$$

where c_i is the component allocation indicator for X_i. The model is fit using a Markov chain Monte Carlo (MCMC) algorithm that iterates over the following steps:

1. update the regression coefficients and measurement error parameters using a Metropolis-Hastings algorithm
2. update each X_i $(i = 1, \ldots, n)$ in the main study using a random walk Metropolis step
3. update each of the mixture component parameters using Gibbs sampling
4. update the number of mixture components, k, and consequently the relevant mixture parameters using split/merge and birth/death moves.

In the scenarios we are considering, the validation sample includes data on (X, W) only and not on Y (for Scenario A3 [differential error], the measurement error parameters were fixed at their true values and assumed known in the MCMC analysis). In addition, we consider measurement error models of the form $W_i = X_i + \epsilon_i$ with different parametric distributions for ϵ_i as described in the various simulation scenarios. Some slight modifications in the Richardson et al. (2002) method are required to accommodate these features. In particular, an additional step in the MCMC algorithm is needed to update the unobserved Y_i $(i = 1, \ldots, m)$ in the validation sample. In addition, the prior settings for the regression coefficients need to be modified. Since we can no longer obtain regression coefficient estimates from the validation set, we center the prior for β_j around the naive estimates $\tilde{\beta}_j$ $(j = 0, 1)$. An important implication of not observing Y in the validation sample is that the prior for the β_j needs to be relatively informative; we set the prior variance to be proportional to the variance of the naive estimate. For the updates of the measurement error parameters and the unobserved X_i's, the assumed parametric form of the measurement error model is used in the corresponding MCMC steps. The prior specifications we used are as follows:

$$
\begin{aligned}
\beta &\sim N(\tilde{\beta},\ \eta\ \widehat{\mathrm{Var}}(\tilde{\beta})) \\
k &\sim \text{Truncated-Poisson}(\lambda = 3) \qquad k = 1, \ldots, 30 \\
c_i|k &\sim \text{Multinomial}(1; \omega_1, \ldots, \omega_k) \qquad \boldsymbol{\omega} \sim \text{Dirichlet}(1, \ldots, 1) \\
\mu_j &\sim N(\xi, R^2) \qquad \sigma_j^{-2} \sim \text{Gamma}(2, b_\sigma) \qquad b_\sigma \sim \text{Gamma}(0.2, 10/R^2),
\end{aligned}
$$

where ξ and R are respectively set to the midpoint and the range of the initial values for X_i, taken to be $\hat{X}_i \sim N(W_i, \widehat{\mathrm{Var}}(\hat{\epsilon}_i))$ with $\widehat{\mathrm{Var}}(\hat{\epsilon}_i)$ estimated according to the assumed error model. The hyperparameter η controls the strength of the prior information for the regression coefficients β_j with larger values leading to less informative priors. In our analyses, we consider $\eta = 10$ and 100.

Web Appendix C.1: Additional Results for Simulation Scenarios A-C

Tables 1, 2, and 3-5 provide results for Simulation Scenarios A, B, and C, respectively. These tables include the results provided in the main paper along with additional results beyond those provided in the main paper. For Scenario C we add results on the slope β_2 of the second error-prone covariate. For all scenarios we add results on the intercept β_0 and results for the SIMEX-L method, the SIMEX-NL method, and the MCMC method with $\eta = 10$.

Web Appendix C.2: Simulation Scenario D

Simulation Set D examined, in the setting of Simulation Sets A and B, the effect of misspecifying the error distribution. We generated the errors according to one of two possible non-normal distributions, but implemented our method assuming the errors are normal. The non-normal error distributions used were the MODCHI(γ) distribution and a modified version of Azzalini's (1985) skewed normal distribution. Azzalini's skewed normal distribution $SN(\lambda)$ has density $2\phi(y)\Phi(\lambda y)$, $y \in \mathbb{R}$, where ϕ and Φ denote the standard normal density and distribution function, respectively, and λ is a parameter that regulates the skewness ($\lambda = 0$ gives the standard normal). Our modified version recenters to mean zero and then rescales to the specified variance. We took $W_i = X_i + \epsilon_i$, with X_i taken to be either $N(0,1)$, $SN(50)$, or MODCHI, and the distribution of ϵ_i taken to be either the MODCHI(γ) or the the modified $SN(\lambda)$ with $\lambda = 50$. The skewness of the $SN(50)$ distribution is 1, and that of the MODCHI is 1.7.

Web Appendix D: Computer Code

The computer code for the numerical results presented in this paper are contained in a zip file called `RECS Paper Codes.zip`. This file contains five folders labeled as follows: `RECS Method`, `N-S Method`, `SIMEX Method`, `H-W Method`,and `MCMC Method`. In addition, it contains a file called `readme.txt` with some explanatory notes.

The `RECS Method` folder contains three zip files with the codes for the RECS method, as follows:

R code for Simulation Studies A, B, and D: `recs-siml-univ.zip`

R code for Simulation Study C: `recs-siml-multiv.zip`

R code for running the example: `recs-example.zip`

The codes were written in R, with some subroutines written in Fortran to increase the speed of the computation. The codes are also available on the first author's website

`http://pluto.huji.ac.il/~mszucker`

under the same file names (add `/recs-siml-univ.zip` to the URL for the first file, and similarly for the other two). In the event that the codes are updated, the updated versions will be posted on the above website. The other folders contain the codes for the other methods. More details are provided in the `readme.txt` file.

Additional References

Huber, P. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, pp. 221–233. University of California Press: Berkeley.

van der Vaart, A.W. (1998). *Asymptotic Statistics.* Cambridge University Press: Cambridge.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica* **50**:1–25.

Web Table 1

Simulation Study A

$X \sim N(0,1)$ and $\epsilon \sim N(0, \gamma_1 + \gamma_2|X| + \gamma_3 Y)$

Sample Size n=200, Validation Sample Size m=70

	β_0					β_1					
	M	MD	Emp-SD	IQ-SD	95% CI	M	MD	Emp-SD	IQ-SD	95% CI	F
				$\gamma_1 = 0.5$ $\gamma_2 = 0$ $\gamma_3 = 0$							
Naive	0.9526	0.9459	0.1737	0.1695	0.930	0.6469	0.6379	0.1512	0.1520	0.360	0
RECS	1.0323	1.0217	0.2122	0.1983	0.959	1.0782	1.0314	0.3373	0.3013	0.958	0
N&S	1.0085	1.0004	0.1902	0.1881	0.958	1.0368	1.0257	0.2740	0.2867	0.907	0
SIMEX - Q	0.9942	0.9840	0.1890	0.1890	0.953	0.8940	0.8737	0.2245	0.2157	0.877	0
SIMEX - L	0.9718	0.9620	0.1800	0.1764	0.942	0.7619	0.7501	0.1803	0.1794	0.663	0
SIMEX - NL	1.0548	1.0252	0.3080	0.2261	0.934	1.0681	1.0161	0.3312	0.3032	0.977	61
H&W	0.9470	0.9752	0.6279	0.2387	0.975	1.0600	0.9514	0.5264	0.3818	0.951	4
MCMC η =10	1.0209	1.0159	0.1965	0.2011	0.93	1.0278	1.0194	0.2635	0.2867	0.970	0
MCMC η =100	1.1000	1.1028	0.2299	0.2360	0.96	1.2797	1.2364	0.4072	0.4324	0.950	0
				$\gamma_1 = 1$ $\gamma_2 = 0$ $\gamma_3 = 0$							
Naive	0.9231	0.9153	0.1694	0.1665	0.909	0.4713	0.4660	0.1262	0.1264	0.027	0
RECS	1.0613	1.0346	0.2683	0.2283	0.963	1.1567	1.0231	0.5421	0.4266	0.933	2
N&S	0.9835	0.9774	0.1873	0.1926	0.942	0.9487	0.9859	0.2283	0.2365	0.897	8
SIMEX - Q	0.9645	0.9515	0.1855	0.1809	0.932	0.7182	0.6989	0.2048	0.2039	0.626	0
SIMEX - L	0.9390	0.9270	0.1748	0.1675	0.922	0.5663	0.5595	0.1525	0.1534	0.150	0
SIMEX - NL	1.0564	1.0068	0.4722	0.2328	0.799	1.1431	0.9997	0.5593	0.4151	0.939	63
H&W	0.9065	0.9495	0.6558	0.2483	0.975	0.8840	0.7851	0.5249	0.3751	0.917	4
MCMC η = 10	0.9802	0.9705	0.1899	0.1939	0.95	0.8551	0.8296	0.2428	0.2878	0.870	0
MCMC η =100	1.0910	1.0840	0.2366	0.2477	0.95	1.2304	1.1671	0.4350	0.4848	0.970	0
				$\gamma_1 = 0.3$ $\gamma_2 = 0.25$ $\gamma_3 = 0$							
Naive	0.9522	0.9446	0.1734	0.1728	0.932	0.6573	0.6494	0.1553	0.1573	0.390	0
RECS	1.0233	1.0121	0.2057	0.1953	0.958	1.0429	1.0054	0.3102	0.2784	0.949	0
N&S	1.0050	0.9963	0.1861	0.1827	0.958	1.0929	1.0881	0.3001	0.3208	0.899	1
SIMEX - Q	0.9933	0.9825	0.1884	0.1913	0.950	0.9169	0.8926	0.2368	0.2313	0.891	0
SIMEX - L	0.9713	0.9615	0.1796	0.1779	0.943	0.7750	0.7620	0.1857	0.1868	0.682	0
SIMEX - NL	1.0427	1.0199	0.2912	0.2209	0.932	1.1265	1.0607	0.3867	0.3321	0.972	74
H&W	0.9408	0.9680	0.5539	0.2150	0.973	1.0370	0.9679	0.4451	0.2980	0.949	1
MCMC η =10	1.0415	1.0487	0.1988	0.1990	0.93	1.0953	1.0994	0.2649	0.2854	0.970	0
MCMC η =100	1.0946	1.1061	0.2189	0.2245	0.95	1.2539	1.2292	0.3691	0.3750	0.950	0
				$\gamma_1 = 0.7$ $\gamma_2 = 0.35$ $\gamma_3 = 0$							
Naive	0.9238	0.9145	0.1695	0.1664	0.905	0.4812	0.4749	0.1284	0.1279	0.036	0
RECS	1.0363	1.0161	0.2423	0.2153	0.960	1.0658	0.9797	0.4673	0.3663	0.912	1
N&S	0.9763	0.9696	0.1845	0.1875	0.941	0.9926	1.0305	0.2402	0.2348	0.890	7
SIMEX - Q	0.9650	0.9536	0.1855	0.1861	0.933	0.7358	0.7167	0.2115	0.2053	0.660	0
SIMEX - L	0.9397	0.9285	0.1748	0.1668	0.919	0.5780	0.5685	0.1552	0.1542	0.177	0
SIMEX - NL	1.0484	1.0065	0.4375	0.2305	0.807	1.2113	1.0383	0.6183	0.4470	0.955	63
H&W	0.9388	0.9663	0.6524	0.2431	0.972	1.0310	0.9134	0.5109	0.3566	0.941	4
MCMC η =10	1.0035	0.9925	0.1937	0.2040	0.94	0.9550	0.9292	0.2552	0.2945	0.940	0
MCMC η =100	1.1025	1.1115	0.2345	0.2208	0.96	1.2866	1.2396	0.4414	0.4678	0.940	0
				$\gamma_1 = 0.15$ $\gamma_2 = 0.25$ $\gamma_3 = 0.25$							
Naive	0.9524	0.9433	0.1715	0.1717	0.935	0.6512	0.6437	0.1513	0.1488	0.373	0
RECS	1.0274	1.0114	0.2066	0.1964	0.959	1.0782	1.0316	0.3477	0.2921	0.950	0
N&S	1.0140	1.0116	0.1847	0.1848	0.961	1.1099	1.1037	0.2987	0.3215	0.896	2
SIMEX - Q	0.9891	0.9783	0.1838	0.1838	0.955	0.8718	0.8545	0.2241	0.2194	0.840	0
SIMEX - L	0.9704	0.9600	0.1769	0.1794	0.945	0.7587	0.7491	0.1799	0.1735	0.657	0
SIMEX - NL	1.0223	1.0131	0.2497	0.2053	0.944	1.0206	0.9698	0.3251	0.2758	0.942	102
H&W	1.0370	1.0350	0.4937	0.2231	0.985	1.0410	0.9682	0.4373	0.3121	0.957	2
MCMC η =10	1.0271	1.0294	0.1981	0.2182	0.95	0.9768	0.9642	0.2461	0.2941	0.94	0
MCMC η =100	1.0687	1.0769	0.2165	0.2251	0.93	1.0950	1.0748	0.3028	0.3490	0.99	0
				$\gamma_1 = 0.35$ $\gamma_2 = 0.25$ $\gamma_3 = 0.50$							
Naive	0.9292	0.9207	0.1676	0.1657	0.914	0.5091	0.5029	0.1282	0.1285	0.060	0
RECS	1.0439	1.0182	0.2389	0.2105	0.962	1.1123	0.9989	0.4765	0.3793	0.929	0
N&S	0.9982	0.9929	0.1835	0.1839	0.945	1.0293	1.0674	0.2451	0.2514	0.858	12
SIMEX - Q	0.9662	0.9564	0.1797	0.1757	0.942	0.7259	0.7125	0.1986	0.1868	0.605	0
SIMEX - L	0.9450	0.9362	0.1721	0.1720	0.928	0.6028	0.5954	0.1534	0.1520	0.230	0
SIMEX - NL	1.0190	0.9887	0.3647	0.2187	0.922	0.9651	0.8793	0.4155	0.2943	0.874	67
H&W	1.1080	1.0940	0.6033	0.2595	0.981	1.0600	0.9600	0.5004	0.3425	0.942	3
MCMC η =10	1.0072	1.0051	0.1939	0.2109	0.93	0.8949	0.8863	0.2432	0.2653	0.86	0
MCMC η =100	1.0789	1.0910	0.2335	0.2403	0.95	1.1408	1.0934	0.3828	0.3606	0.98	0

N&S - Novick and Stefanski (2002); H&W - Huang and Wang (2001)
SIMEX results are based on $B = 100$, $\lambda = (0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9)$ and the simex R library
L - linear extrapolation; Q - quadratic extrapolation; NL - non-linear extrapolation of Cook and Stefanski (1994)
M - empirical mean; MD - empirical median; Emp-SD - empirical standard deviation; IQ-SD - inter-quartile dispersion; 95% CI - empirical coverage rate of 95% Wald confidence interval (or HPD interval, for MCMC) ; F - number of samples with no solution

Web Table 2

Simulation Study B
Non-Normal Measurement Error with Normal or Non-Normal True Covariate
Sample Size n=200, Validation Sample Size m=70

	β_0					β_1					
	M	MD	Emp-SD	IQ-SD	95% CI	M	MD	Emp-SD	IQ-SD	95% CI	F
$X \sim N(0,1)$; $\epsilon \sim$ Double exponential with $\text{Var}(\epsilon) = 0.5$											
Naive	0.9467	0.9350	0.1775	0.1804	0.925	0.6378	0.6287	0.1609	0.1517	0.334	0
RECS	1.0288	1.0048	0.2192	0.1989	0.952	1.0684	0.9986	0.3611	0.3264	0.948	0
N&S	1.0040	0.9978	0.1883	0.1758	0.969	1.0471	1.0126	0.2970	0.3293	0.888	0
SIMEX - Q	0.9884	0.9763	0.1931	0.1883	0.952	0.8905	0.8641	0.2438	0.2276	0.871	0
SIMEX - L	0.9654	0.9529	0.1836	0.1772	0.940	0.7519	0.7381	0.1924	0.1831	0.652	0
SIMEX - NL	1.0471	1.0189	0.2754	0.2135	0.945	1.1111	1.0256	0.4207	0.3336	0.967	68
H&W	0.9980	0.9919	0.4298	0.2157	0.982	1.0770	0.9848	0.4523	0.3202	0.951	0
MCMC η =10	1.0208	1.0349	0.1957	0.2084	0.94	1.0349	1.0186	0.2258	0.2753	0.990	0
MCMC η =100	1.1028	1.0933	0.2216	0.2410	0.97	1.2987	1.2613	0.3796	0.4360	0.970	0
$X \sim N(0,1)$; $\epsilon \sim$ Double exponential with $\text{Var}(\epsilon) = 1$											
Naive	0.9174	0.9078	0.1731	0.1787	0.911	0.4659	0.4598	0.1370	0.1349	0.039	0
RECS	1.0545	1.0167	0.2735	0.2291	0.962	1.1362	1.0089	0.5714	0.4335	0.933	6
N&S	0.9700	0.9647	0.1851	0.1758	0.934	0.9625	1.0099	0.2794	0.2872	0.845	11
SIMEX - Q	0.9583	0.9474	0.1891	0.1861	0.933	0.7205	0.6970	0.2289	0.2165	0.629	0
SIMEX - L	0.9327	0.9228	0.1782	0.1772	0.912	0.5599	0.5514	0.1659	0.1601	0.197	0
SIMEX - NL	1.0032	0.9982	0.5742	0.2513	0.790	1.2130	1.0485	0.7743	0.4804	0.909	106
H&W	0.9353	0.9704	0.6913	0.2543	0.973	0.9886	0.8985	0.4636	0.3336	0.942	1
MCMC η=10	0.9779	0.9699	0.1876	0.2032	0.92	0.8304	0.8075	0.2289	0.2788	0.840	0
MCMC η=100	1.0711	1.0612	0.2221	0.2204	0.96	1.1476	1.0940	0.3785	0.4483	0.990	0
$X \sim N(0,1)$; $\epsilon \sim$ Modified $\chi^2_{(1)}$ with $\text{Var}(\epsilon) = 0.5$											
Naive	0.9732	0.9745	0.1785	0.1692	0.932	0.6822	0.6793	0.1706	0.1672	0.467	0
RECS	1.0335	1.0295	0.2091	0.1929	0.951	1.0555	1.0348	0.2981	0.2775	0.959	0
N&S	1.0636	1.0442	0.2100	0.2085	0.946	1.1449	1.1451	0.3230	0.3808	0.857	4
SIMEX - Q	1.0340	1.0277	0.2044	0.1846	0.949	0.9678	0.9500	0.2624	0.2617	0.890	0
SIMEX - L	0.9980	0.9973	0.1878	0.1727	0.942	0.8080	0.8028	0.2052	0.2024	0.720	0
SIMEX - NL	1.1297	1.0882	0.3964	0.2483	0.912	1.2028	1.1617	0.3912	0.3573	0.975	45
H&W	0.7934	0.8692	0.5560	0.1964	0.963	1.0260	0.9809	0.3822	0.2861	0.956	2
MCMC η =10	1.0813	1.0884	0.2050	0.2144	0.93	1.0596	1.0267	0.2486	0.2244	0.97	0
MCMC η =100	1.1501	1.1402	0.2368	0.2504	0.93	1.2154	1.1673	0.3372	0.2999	0.92	0
$X \sim N(0,1)$; $\epsilon \sim$ Modified $\chi^2_{(1)}$ with $\text{Var}(\epsilon) = 1$											
Naive	0.9510	0.9502	0.1765	0.1667	0.926	0.5196	0.5146	0.1547	0.1552	0.126	0
RECS	1.0519	1.0315	0.2444	0.2116	0.961	1.0869	1.0345	0.3970	0.3546	0.966	0
N&S	1.0914	1.0811	0.2151	0.2102	0.909	1.1232	1.1779	0.2711	0.2319	0.805	18
SIMEX - Q	1.0262	1.0140	0.2135	0.1890	0.933	0.8320	0.8234	0.2711	0.2765	0.747	0
SIMEX - L	0.9734	0.9703	0.1857	0.1690	0.931	0.6269	0.6187	0.1880	0.1853	0.316	0
SIMEX - NL	1.1226	1.0980	1.0702	0.4804	0.435	1.6444	1.4262	0.9090	0.7413	0.958	59
H&W	0.5574	0.7494	0.8302	0.2506	0.934	0.9640	0.8949	0.4753	0.3069	0.927	3
MCMC η =10	1.0585	1.0639	0.1975	0.2187	0.95	0.9402	0.9125	0.2241	0.2299	0.940	0
MCMC η =100	1.1556	1.1341	0.2378	0.2447	0.94	1.1502	1.1252	0.2996	0.2924	0.980	0
$X \sim$ Modified $\chi^2_{(1)}$ with $\text{Var}(X) = 1$; $\epsilon \sim$ Modified $\chi^2_{(1)}$ with $\text{Var}(\epsilon) = 0.5$											
Naive	0.9059	0.9015	0.1689	0.1632	0.903	0.5461	0.5348	0.1709	0.1662	0.250	0
RECS	1.0862	1.0356	0.2953	0.2626	0.967	1.1279	1.0479	0.4989	0.4415	0.938	0
N&S	1.2102	1.1694	0.3294	0.3620	0.813	1.2750	1.3552	0.4993	0.6506	0.672	51
SIMEX - Q	0.9703	0.9592	0.1968	0.1913	0.937	0.8175	0.7884	0.2870	0.2713	0.791	0
SIMEX - L	0.9253	0.9166	0.1750	0.1690	0.924	0.6492	0.6354	0.2056	0.1979	0.459	0
SIMEX - NL	1.0547	1.0083	0.9484	0.3625	0.495	1.2668	1.1121	0.6949	0.5137	0.945	43
H&W	0.6861	0.8618	0.8291	0.2053	0.944	0.9610	0.8409	0.5507	0.3729	0.912	6
MCMC η =10	1.0400	1.0104	0.2344	0.1886	0.90	0.7529	0.7365	0.3067	0.3059	0.900	0
MCMC η =100	1.1630	1.1179	0.3110	0.2506	0.91	1.0279	1.0105	0.4865	0.4735	0.970	0
$X \sim$ Modified $\chi^2_{(1)}$ with $\text{Var}(X) = 1$; $\epsilon \sim$ Modified $\chi^2_{(1)}$ with $\text{Var}(\epsilon) = 1$											
Naive	0.8784	0.8740	0.1661	0.1604	0.866	0.3917	0.3824	0.1448	0.1402	0.063	0
RECS	1.1134	1.0370	0.3589	0.2897	0.967	1.1272	1.0319	0.5784	0.5316	0.932	0
N&S	1.1288	1.0985	0.2997	0.2986	0.651	1.0147	1.1508	0.4306	0.5522	0.545	212
SIMEX - Q	0.9337	0.9243	0.1926	0.1794	0.917	0.6396	0.6176	0.2602	0.2498	0.490	0
SIMEX - L	0.8923	0.8876	0.1709	0.1616	0.890	0.4720	0.4604	0.1749	0.1683	0.099	0
SIMEX - NL	0.7096	0.8059	0.8610	0.4040	0.451	1.4470	1.1876	1.0740	0.8718	0.827	146
H&W	0.6004	0.7683	0.7942	0.2090	0.917	0.8084	0.6995	0.5894	0.4255	0.846	13
MCMC η =10	0.9865	0.9440	0.2132	0.2141	0.93	0.5704	0.5703	0.2674	0.2727	0.700	0
MCMC η =100	1.1116	1.0870	0.2817	0.2235	0.96	0.8416	0.8245	0.4370	0.4278	0.930	0

N&S - Novick and Stefanski (2002); H&W - Huang and Wang (2001)
SIMEX results are based on $B = 100$, $\lambda = (0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9)$ and the simex R library
L - linear extrapolation; Q - quadratic extrapolation; NL - non-linear extrapolation of Cook and Stefanski (1994)
M - empirical mean; MD - empirical median; Emp-SD - empirical standard deviation; IQ-SD - inter-quartile dispersion; 95% CI - empirical coverage rate of 95% Wald confidence interval (or HPD interval, for MCMC); F - number of samples with no solution

Web Table 3

Simulation Study C
Two Error-Prone Covariates and One Error-Free Covariate
Scenario C1: $X_1, X_2, Z \sim N(0,1)$ and $\epsilon_1, \epsilon_2 \sim N(0,1)$
Sample Size n=500

	β_1					β_2				
	M	MD	Emp-SD	IQ-SD	95% CI	M	MD	Emp-SD	IQ-SD	95% CI
Naive	0.4296	0.4242	0.0804	0.0781	0.000	0.4282	0.4235	0.0853	0.0857	0.000
RECS	1.0817	1.0341	0.3116	0.2595	0.972	1.0784	1.0305	0.3333	0.2764	0.956
N&S	0.8286	0.8288	0.1316	0.1319	0.826	0.8230	0.8259	0.1444	0.1454	0.775
SIMEX - Q	0.8113	0.8038	0.1377	0.1333	0.571	0.8096	0.8006	0.1450	0.1378	0.580
SIMEX - L	0.5235	0.5189	0.0917	0.0911	0.003	0.5219	0.5150	0.0969	0.0953	0.000
SIMEX - NL	-2.6768	-2.5286	1.1003	1.0482	0.000	-2.5988	-2.4974	1.1896	1.0356	0.000
H&W	0.6678	0.6675	0.5038	0.3360	0.800	0.6975	0.6744	0.5318	0.3407	0.786
MCMC η =10	0.7675	0.7560	0.1670	0.1550	0.550	0.7695	0.7706	0.1550	0.1188	0.640
MCMC η =100	1.0608	1.0335	0.2765	0.2508	0.870	1.0688	1.0593	0.2804	0.1833	0.890

	β_0					β_3					
	M	MD	Emp-SD	IQ-SD	95% CI	M	MD	Emp-SD	IQ-SD	95% CI	F
Naive	0.8587	0.8575	0.1163	0.1129	0.734	0.8626	0.8599	0.1201	0.1223	0.769	0
RECS	1.0500	1.0294	0.2112	0.1791	0.972	1.0525	1.0286	0.2079	0.1907	0.979	0
N&S	0.9432	0.9433	0.1379	0.1383	0.947	0.9483	0.9448	0.1382	0.1412	0.951	52
SIMEX - Q	0.9553	0.9502	0.1433	0.1394	0.886	0.9587	0.9524	0.1452	0.1519	0.920	0
SIMEX - L	0.8821	0.8803	0.1215	0.1214	0.783	0.8857	0.8819	0.1247	0.1278	0.809	0
SIMEX - NL	0.3396	0.3544	1.1798	0.5785	0.051	0.3642	0.4221	1.1742	0.5205	0.059	285
H&W	0.7535	0.8516	0.7885	0.2376	0.887	1.0277	0.9679	0.4669	0.3039	0.918	14
MCMC η =10	0.9101	0.9060	0.1349	0.1290	0.840	0.9172	0.9113	0.1281	0.1475	0.870	0
MCMC η =100	1.0407	1.0328	0.1964	0.1728	0.930	1.0445	1.0418	0.1622	0.1834	0.970	0

N&S - Novick and Stefanski (2002); H&W - Huang and Wang (2001)
SIMEX results are based on $B = 100$, $\lambda = (0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9)$ and the simex R library
L - linear extrapolation; Q - quadratic extrapolation; NL - non-linear extrapolation of Cook and Stefanski (1994)
M - empirical mean; MD - empirical median; Emp-SD - empirical standard deviation; IQ-SD - inter-quartile dispersion; 95% CI - empirical coverage rate of 95% Wald confidence interval (or HPD interval, for MCMC); F - number of samples with no solution

Web Table 4

Simulation Study C
Two Error-Prone Covariates and One Error-Free Covariate
Scenario C2: $X_1, X_2, Z \sim N(0,1)$ and $\epsilon_1, \epsilon_2 \sim N(0, 0.4 + 0.25(|X_1| + |X_2| + |Z|))$
Sample Size n=500

	β_1					β_2				
	M	MD	Emp-SD	IQ-SD	95% CI	M	MD	Emp-SD	IQ-SD	95% CI
Naive	0.4386	0.4334	0.0824	0.0792	0.001	0.4364	0.4313	0.0868	0.0857	0.000
RECS	1.1884	1.1122	0.3800	0.3126	0.986	1.1822	1.1151	0.4044	0.3313	0.977
N&S	0.8606	0.8619	0.1329	0.1312	0.842	0.8521	0.8559	0.1468	0.1454	0.834
SIMEX - Q	0.8345	0.8242	0.1431	0.1394	0.819	0.8315	0.8231	0.1497	0.1439	0.767
SIMEX - L	0.5339	0.5265	0.0940	0.0903	0.009	0.5314	0.5257	0.0985	0.0973	0.011
SIMEX - NL	-2.3990	-2.2229	0.9623	0.9407	0.000	-2.3830	-2.1980	1.0071	0.9666	0.000
H&W	0.6165	0.6176	0.4809	0.3277	0.733	0.6455	0.6072	0.4809	0.3271	0.766
MCMC η =10	0.7713	0.7621	0.1663	0.1656	0.600	0.7716	0.7630	0.1606	0.1188	0.620
MCMC η =100	1.0338	1.0084	0.2709	0.2637	0.890	1.0378	1.0093	0.2576	0.1849	0.900

	β_0					β_3					
	M	MD	Emp-SD	IQ-SD	95% CI	M	MD	Emp-SD	IQ-SD	95% CI	F
Naive	0.8593	0.8586	0.1162	0.1153	0.740	0.8712	0.8694	0.1223	0.1223	0.791	0
RECS	1.0787	1.0475	0.2359	0.1896	0.981	1.0828	1.0482	0.2405	0.2007	0.983	1
N&S	0.9353	0.9350	0.1360	0.1387	0.939	0.9668	0.9655	0.1439	0.1465	0.966	46
SIMEX - Q	0.9569	0.9531	0.1434	0.1394	0.907	0.9793	0.9746	0.1516	0.1562	0.933	0
SIMEX - L	0.8827	0.8824	0.1214	0.1219	0.808	0.8961	0.8921	0.1276	0.1314	0.834	0
SIMEX - NL	0.3789	0.3653	1.1233	0.5524	0.035	0.3096	0.3920	0.9646	0.4150	0.039	185
H&W	0.7637	0.8325	0.9796	0.2280	0.887	0.9627	0.9279	0.4248	0.2837	0.905	9
MCMC η =10	0.9148	0.9051	0.1358	0.1285	0.830	0.9237	0.9205	0.1262	0.1474	0.890	0
MCMC η =100	1.0341	1.0186	0.1903	0.1847	0.920	1.0358	1.0339	0.1626	0.1701	0.950	0

N&S - Novick and Stefanski (2002); H&W - Huang and Wang (2001)
SIMEX results are based on $B = 100$, $\lambda = (0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9)$ and the simex R library
L - linear extrapolation; Q - quadratic extrapolation; NL - non-linear extrapolation of Cook and Stefanski (1994)
M - empirical mean; MD - empirical median; Emp-SD - empirical standard deviation; IQ-SD - inter-quartile dispersion; 95% CI - empirical coverage rate of 95% Wald confidence interval (or HPD interval, for MCMC); F - number of samples with no solution

Web Table 5

Simulation Study C
Two Error-Prone Covariates and One Error-Free Covariate
Scenario C3: $X_1, X_2 \sim$ Modified χ^2_1, $\mathrm{Var}(X_1) = \mathrm{Var}(X_2) = 1$, $Z \sim N(0, 1)$, and $\epsilon_1, \epsilon_2 \sim$ Modified χ^2_1, $\mathrm{Var}(\epsilon_1) = \mathrm{Var}(\epsilon_2) = 1$
Sample Size n=500

	β_1					β_2					
	M	MD	Emp-SD	IQ-SD	95% CI	M	MD	Emp-SD	IQ-SD	95% CI	
Naive	0.3853	0.3814	0.0887	0.0915	0.000	0.3830	0.3811	0.0843	0.0856	0.000	
RECS	1.0430	0.9841	0.3726	0.3328	0.942	1.0322	0.9850	0.3319	0.3016	0.946	
N&S	0.8656	0.8664	0.2576	0.2880	0.782	0.8659	0.8750	0.2516	0.2872	0.784	
SIMEX - Q	0.7463	0.7315	0.1586	0.1598	0.445	0.7431	0.7317	0.1493	0.1488	0.450	
SIMEX - L	0.4672	0.4624	0.1009	0.1021	0.001	0.4647	0.4602	0.0958	0.0989	0.000	
SIMEX - NL	-1.5004	-1.2992	0.7297	0.5271	0.000	-1.5046	-1.3341	0.6896	0.5078	0.000	
H&W	0.4723	0.4276	0.2733	0.2108	0.654	0.4625	0.4216	0.2849	0.1992	0.606	
MCMC η =10	0.7216	0.7191	0.1696	0.2000	0.640	0.7054	0.7066	0.1493	0.1446	0.600	
MCMC η =100	1.1462	1.1199	0.3153	0.3259	0.910	1.1100	1.0941	0.3072	0.2442	0.950	
	β_0					β_3					
	M	MD	Emp-SD	IQ-SD	95% CI	M	MD	Emp-SD	IQ-SD	95% CI	F
Naive	0.8007	0.7971	0.1144	0.1113	0.548	0.8883	0.8893	0.1206	0.1207	0.809	0
RECS	1.0595	1.0225	0.2447	0.2084	0.963	1.0413	1.0232	0.1856	0.1667	0.971	0
N&S	1.0389	1.0389	0.1554	0.1553	0.945	0.9261	0.9201	0.1396	0.1415	0.946	34
SIMEX - Q	0.9049	0.8978	0.1456	0.1383	0.823	0.9626	0.9588	0.1402	0.1399	0.915	0
SIMEX - L	0.8199	0.8146	0.1191	0.1133	0.598	0.9059	0.9073	0.1243	0.1244	0.893	0
SIMEX - NL	0.5453	0.5554	0.1777	0.1286	0.027	0.4842	0.5551	0.9577	0.3875	0.122	60
H&W	0.3332	0.4979	0.6451	0.2517	0.681	0.9155	0.9023	0.1798	0.1638	0.971	1
MCMC η =10	0.9526	0.9413	0.1204	0.0946	0.900	0.9125	0.8945	0.1388	0.1345	0.840	0
MCMC η =100	1.2427	1.2461	0.1978	0.2014	0.830	1.0167	0.9887	0.1713	0.1519	0.910	0

N&S - Novick and Stefanski (2002); H&W - Huang and Wang (2001)
SIMEX results are based on $B = 100$, $\lambda = (0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9)$ and the simex R library
L - linear extrapolation; Q - quadratic extrapolation; NL - non-linear extrapolation of Cook and Stefanski (1994)
M - empirical mean; MD - empirical median; Emp-SD - empirical standard deviation; IQ-SD - inter-quartile dispersion; 95% CI - empirical coverage rate of 95% Wald confidence interval (or HPD interval, for MCMC); F - number of samples with no solution

Web Table 6

Simulation Study D
Performance of RECS Under Misspecified Error Model
Assumed Error Model: $\epsilon \sim N(0, \sigma^2)$
Results for Estimate of β_1

		Naive Estimate				RECS Estimate			
Distn of X	True Distn of ϵ	Mean	Median	Emp-SD	IQ-SD	Mean	Median	Emp-SD	IQ-SD
Normal	Skewed Normal	0.6621	0.6569	0.1655	0.1668	1.1885	1.1028	0.4790	0.3466
Normal	MODCHI	0.6821	0.6791	0.1706	0.1672	1.2778	1.1592	0.5621	0.4267
Skewed Normal	Skewed Normal	0.5992	0.5895	0.1620	0.1589	1.1983	1.0923	0.5506	0.4355
Skewed Nomal	MODCHI	0.6058	0.6043	0.1669	0.1698	1.3466	1.1862	0.6752	0.5526
MODCHI	Skewed Normal	0.5250	0.5209	0.1573	0.1532	1.3280	1.1335	0.7406	0.5932
MODCHI	MODCHI	0.5460	0.5347	0.1710	0.1661	1.6553	1.3567	1.0210	0.9698