Journal of the Korean Statistical Society (2003), **32**: *1, pp 85–92*

SOME PROPERTIES OF SIMEX ESTIMATOR IN PARTIALLY LINEAR MEASUREMENT ERROR MODEL†

MEESEON JEONG[1] AND CHOONGRAK KIM[2]

ABSTRACT

We consider the partially linear model $E(Y) = \mathbf{X}^t\beta + \eta(Z)$ when the $\mathbf{X}$'s are measured with additive error. The semiparametric likelihood estimation ignoring the measurement error gives inconsistent estimator for both β and $\eta(\cdot)$. In this paper we suggest the SIMEX estimator for β to correct the bias induced by measurement error, and explore its properties. We show that the rational linear extrapolant is proper in extrapolation step in the sense that the SIMEX method under this extrapolant gives consistent estimator. It is also shown that the SIMEX estimator is asymptotically equivalent to the semiparametric version of the usual parametric correction for attenuation suggested by Liang *et al.* (1999). A simulation study is given to compare two variance estimating methods for SIMEX estimator.

AMS 2000 subject classifications. Primary 62G08; Secondary 62G20.
Keywords. Extrapolant, nonparametric regression, simulation-extrapolation.

1. INTRODUCTION

If predictors in a regression model are subject to measurement error and we use their surrogates to estimate parameters of interest, then the resulting estimates are usually inconsistent. For consistent estimation in measurement error model, several methods have been suggested such as regression calibration (Carroll and Stefanski, 1990) and simulation-extrapolation (SIMEX) (Cook and Stefanski, 1995). Roughly speaking, SIMEX is more robust to the distributional assumptions than the regression calibration, but SIMEX requires more computational amounts. Fuller (1987) and Carroll *et al.* (1995) give good reviews for

Received October 2002; accepted February 2003.
†Research of the first author was supported by KOSEF and also partially supported by the *BK21 Project*.
[1]Radiation Health Research Institute, KHNP, Seoul 132-703, Korea
[2]Department of Statistics, Pusan National University, Pusan 609-735, Korea

these measurement error analysis. In this paper, we consider the measurement error in a partially linear model,

$$Y_i = \mathbf{X}_i^t\beta + \eta(Z_i) + \varepsilon_i\ , \quad i = 1, \ldots, n \tag{1.1}$$

where $\mathbf{X} = (X_1, \ldots, X_k)^t$ are covariates related linearly to response Y with parametric coefficients $\beta = (\beta_1, \ldots, \beta_k)^t$, Z is a scalar covariate, and $\eta(\cdot)$ is unknown smooth function. Assume the errors (ε_i) are *iid* with mean zero and variance σ^2 and independent of covariates. The partially linear model was studied by Heckman (1986), Chen (1988), Speckman (1988), and Severini and Staniswalis (1994). We are interested in the estimation of parametric coefficients β, when the covariates $\mathbf{X}$ are measured with error and we observe the surrogates $\mathbf{W} = (W_1, \ldots, W_k)^t$ instead of $\mathbf{X}$, assuming

$$\mathbf{W}_i = \mathbf{X}_i + \mathbf{U}_i\ , \quad i = 1, \ldots, n \tag{1.2}$$

where $(\mathbf{U}_i)$ are *iid* measurement errors from $N(\mathbf{0}, \Sigma_{uu})$ and independent of $(Y, \mathbf{X}, Z)$. Here Σ_{uu} is usually assumed to be known or to be easily estimated. The assumption of normality and additivity of measurement errors are not crucial, see Chapter 4 of Carroll *et al.* (1995).

In this paper, we consider the SIMEX estimator for β under the so-called partially linear measurement error model (PLMEM), defined by (1.1) and (1.2), and study its asymptotic properties. In Section 2, we introduce the SIMEX method in PLMEM. In Section 3, we show that the SIMEX estimator with the rational linear extrapolant is consistent for β and that it is asymptotically equivalent to the estimator of Liang *et al.* (1999). Section 4 gives the comparison between two methods of variance estimation for SIMEX estimator *via* a Monte Carlo study.

2. SIMEX ESTIMATOR IN PLMEM

If the covariates $\mathbf{X}$ are not related to the measurement error, one method to estimate the parametric coefficients β in model (1.1) is to minimize

$$\sum_{i=1}^{n}\{\widetilde{Y}_i - \widetilde{\mathbf{X}}_i^t\beta\}^2 \tag{2.1}$$

where $\widetilde{Y}$ and $\widetilde{\mathbf{X}}$ are the residuals after regressing nonparametrically Y and $\mathbf{X}$ on Z, respectively. The resulting least squares estimator of β is the same as that of Severini and Staniswalis (1994) if kernel regression is used for nonparametric fit.

When there exists the measurement error in covariates $\mathbf{X}$, the naive estimator for β in PLMEM is obtained by ignoring measurement error and by minimizing (2.1) with $\mathbf{X}$ replaced by $\mathbf{W}$. Then we have

$$\widehat{\beta}_{naive} = \left\{ \sum_{i=1}^{n} \widetilde{\mathbf{W}}_i \widetilde{\mathbf{W}}_i^t \right\}^{-1} \sum_{i=1}^{n} \widetilde{\mathbf{W}}_i \widetilde{Y}_i \tag{2.2}$$

where $\widetilde{\mathbf{W}}$ are similarly defined to $\widetilde{\mathbf{X}}$. However, we can easily show that $\widehat{\beta}_{naive}$ is inconsistent for β under some distributional assumption. See Section 3.1 for the proof. Therefore, we consider the SIMEX method to reduce the bias by measurement error in PLMEM.

The first part of the SIMEX method is the simulation step. Let λ be a fixed positive constant. For $b = 1, \ldots, B$, we obtain new predictors $\{\mathbf{W}_{b,i}(\lambda)\}_{i=1}^{n}$ where $\mathbf{W}_{b,i}(\lambda) = \mathbf{W}_i + \sqrt{\lambda}\mathbf{U}_{b,i}$, $i = 1, \ldots, n$. Here $\{\mathbf{U}_{b,i}\}_{i=1}^{n}$ are *iid* additional errors generated from $N(\mathbf{0}, \Sigma_{uu})$. By nonparametric fit such as smoothing spline or kernel regression, we get the residuals $\widetilde{Y}_i$ and $\widetilde{\mathbf{W}}_{b,i}(\lambda)$, and by (2.2) we obtain the resulting naive estimator, $\widehat{\beta}_b(\lambda)$, $b = 1, \ldots, B$. Then we have the average $\widehat{\beta}(\lambda) = \sum_{b=1}^{B} \widehat{\beta}_b(\lambda)/B$ as the estimator when there exist the measurement errors with variance $(1 + \lambda)\Sigma_{uu}$. Repeat this process to get $\widehat{\beta}(\lambda_m)$ for $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$.

For the extrapolation step, let $g(\lambda, \Gamma)$ be a regression function with unknown coefficients Γ for response $\widehat{\beta}(\lambda)$ and covariate λ, called an extrapolant in the SIMEX method. We will discuss about the proper extrapolant in PLMEM later. The data $\{\lambda_m, \widehat{\beta}(\lambda_m)\}_{m=1}^{M}$ obtained in the simulation step are fitted to get $g(\lambda, \widehat{\Gamma})$ where $\widehat{\Gamma}$ is the estimate of Γ. Then we extrapolate the fitted model back to $\lambda = -1$ to obtain the SIMEX estimate of β in PLMEM, denoted by $\widehat{\beta}_{SIMEX} = g(-1, \widehat{\Gamma})$.

3. Some Properties of SIMEX Estimator in PLMEM

3.1. Proper extrapolating function

In SIMEX estimation, the following types of extrapolating function are often used:

(i) $g_L(\lambda, \Gamma) = \gamma_0 + \gamma_1 \lambda$, $\Gamma = (\gamma_0, \gamma_1)$: linear;

(ii) $g_Q(\lambda, \Gamma) = \gamma_0 + \gamma_1 \lambda + \gamma_2 \lambda^2$, $\Gamma = (\gamma_0, \gamma_1, \gamma_2)$: quadratic;

(iii) $g_{RL}(\lambda, \Gamma) = \gamma_0 + \gamma_1/(\gamma_2 + \lambda)$, $\Gamma = (\gamma_0, \gamma_1, \gamma_2)$: rational linear.

However, the choice of proper extrapolant is important to get consistent estimator. Carroll *et al.* (1995) gives the summary of proper extrapolants for some measurement error models. And Kim *et al.* (2000) considered a nonparametric method *via* the Bezier curve to extrapolate the simulated data. In this section, we obtain the proper extrapolant in PLMEM which gives consistent estimator for β under some distributional assumptions.

Assume that $(\mathbf{X}, Z)$ are jointly normally distributed with mean $(\boldsymbol{\mu}_x, \mu_z)$, $\mathrm{Var}(\mathbf{X}) = \Sigma_{xx}$, $\mathrm{Var}(Z) = \sigma_z^2$ and $\mathrm{Cov}(\mathbf{X}, Z) = \Sigma_{xz}$. Let $\Sigma_{x|z}$ be the conditional variance of $\mathbf{X}$ given Z. Then, under the model (1.1) and (1.2), we can easily show that

$$E(Y|\mathbf{W}, Z) = E(\mathbf{X}|\mathbf{W}, Z)^t\beta + \eta(Z) = \mathbf{W}^t\beta^* + \eta^*(Z) \tag{3.1}$$

where

$$\begin{aligned}\beta^* &= (\Sigma_{x|z} + \Sigma_{uu})^{-1}\Sigma_{x|z}^t\beta,\\ \eta^*(Z) &= \eta(Z) + (\boldsymbol{\mu}_x + \Sigma_{xz}(Z - \mu_z)/\sigma_z^2)^t(I - (\Sigma_{x|z} + \Sigma_{uu})^{-1}\Sigma_{x|z}^t)\beta.\end{aligned}$$

Since $E(Y|Z) = E(\mathbf{W}|Z)^t\beta^* + \eta^*(Z)$, we have $E[Y - E(Y|\mathbf{W}, Z)]^2 = E[(Y - E(Y|Z)) - (\mathbf{W} - E(\mathbf{W}|Z))^t\beta^*]^2$. This implies that $\widehat{\beta}_{naive}$, the minimizer of (2.1) using $\mathbf{W}$ instead of $\mathbf{X}$ is not consistent for β but for β^*. By using this fact, we can also see that $\widehat{\beta}(\lambda)$ for a fixed constant λ consistently estimates $\{\Sigma_{x|z} + (1 + \lambda)\Sigma_{uu}\}^{-1}\Sigma_{x|z}^t\beta$. Therefore, $\widehat{\beta}_{SIMEX}$ in PLMEM becomes consistent for β when we use the rational linear extrapolant $g_{RL}(\lambda, \Gamma)$ to fit the data $\{\lambda_m, \widehat{\beta}(\lambda_m)\}_{m=1}^M$.

3.2. Asymptotic equivalence to Liang's estimator

Under the model (1.1) and (1.2), Liang *et al.* (1999) suggested a method to estimate β consistently. They noticed that $E\{Y_i - \mathbf{W}_i^t\beta - \eta(Z_i)\}^2 = E\{Y_i - \mathbf{X}_i^t\beta - \eta(Z_i)\}^2 + \beta^t\Sigma_{uu}\beta$ and by minimizing $\sum_{i=1}^n\{Y_i - \mathbf{W}_i^t\beta - \eta(Z_i)\}^2 - n\beta^t\Sigma_{uu}\beta$, they obtained an estimator of β given by

$$\widehat{\beta}_L = \left\{\sum_{i=1}^n \widetilde{\mathbf{W}}_i\widetilde{\mathbf{W}}_i^t - n\Sigma_{uu}\right\}^{-1}\sum_{i=1}^n \widetilde{\mathbf{W}}_i\widetilde{Y}_i. \tag{3.2}$$

They also showed that $\widehat{\beta}_L$ is consistent for β and that it is asymptotically normally distributed with covariance $n^{-1}\mathbf{D}^{-1}\mathbf{G}\mathbf{D}^{-1}$, which is estimated by replacing $\mathbf{D}$ and $\mathbf{G}$ by $\widehat{\mathbf{D}} = (n-k)^{-1}\sum_{i=1}^n \widetilde{\mathbf{W}}_i\widetilde{\mathbf{W}}_i^t - \Sigma_{uu}$ and $\widehat{\mathbf{G}} = n^{-1}\sum_{i=1}^n\{\widetilde{\mathbf{W}}_i(\widetilde{Y}_i - \widetilde{\mathbf{W}}_i^t\widehat{\beta}_L) + \Sigma_{uu}\widehat{\beta}_L\}\{\widetilde{\mathbf{W}}_i(\widetilde{Y}_i - \widetilde{\mathbf{W}}_i^t\widehat{\beta}_L) + \Sigma_{uu}\widehat{\beta}_L\}^t$, respectively. For the definition of

D and **G**, see Liang *et al.* (1999). Then, it can be shown that the SIMEX estimator with the extrapolant $g_{RL}(\lambda, \Gamma)$ is asymptotically equal to $\widehat{\boldsymbol{\beta}}_L$. The proof is given in Appendix.

4. Simulation Study

To see the numerical behavior of the SIMEX estimator discussed in Section 2, and compare it with $\widehat{\boldsymbol{\beta}}_L$ in (3.2), we perform a small sample simulation study focused on the coverage probability of each estimator. To do this we need to estimate the variance of $\widehat{\boldsymbol{\beta}}_{SIMEX}$, and two methods are currently available. One is the simulation-extrapolation method suggested by Stefanski and Cook (1995). The conditions are easily checked to apply their method to the SIMEX estimator in PLMEM. The other is so-called the estimating equation method discussed by Carroll *et al.* (1996). In this section, $\widehat{\boldsymbol{\beta}}_{SIMEX}$ with these variance estimates is compared to $\widehat{\boldsymbol{\beta}}_L$ in terms of asymptotic coverage probability.

Under the model (1.1) and (1.2), it is assumed that $k = 1$, *i.e.*, X_i is scalar covariate and therefore all matrix notations in Section 2 are now changed into scalars, for example, Σ_{uu} by σ_u^2. We generate (X_i, Z_i) from the bivariate normal distribution with mean (0,0) and $\sigma_x^2 = 0.2$, $\sigma_z^2 = 0.1$ and $\sigma_{xz} = 0.05$. The model error ε_i and the measurement error U_i are generated from $N(0, 0.1)$ and $N(0, \sigma_u^2)$, respectively. Let $\eta(Z) = \exp(-Z^2)$, $\beta = 1$, $\sigma_u^2 = 0.1$ and $n = 100$. For the nonparametric fit to get the residuals $\widetilde{Y}_i$ and $\widetilde{W}_i$, the function *smooth.spline* of S-PLUS is used since it gives bandwidth automatically selected.

First, we calculate $\widehat{\beta}_{naive}$ and $\widehat{\beta}_L$. To obtain $\widehat{\beta}_{SIMEX}$, we use $B = 100$, $\lambda_m = (m-1)/2$, $m = 1, \ldots, 5$ and the rational linear extrapolant $g_{RL}(\lambda, \Gamma)$ given in Section 3.1. For the variance estimation of $\widehat{\beta}_{SIMEX}$, two methods are considered. First, we use the simulation-extrapolation method by Stefanski and Cook (1995). To be more specific, let $\widehat{\tau}_b(\lambda)$ be the estimated variance of $\widehat{\beta}_b(\lambda)$ by Severini and Staniswalis (1994), $\widehat{\tau}(\lambda) = \sum_{b=1}^{B} \widehat{\tau}_b(\lambda)/B$, and $s_\Delta^2(\lambda)$ the sample variance of $\widehat{\beta}_b(\lambda)$, $b = 1, \ldots, B$. Then $\mathrm{Var}(\widehat{\beta}_{SIMEX})$ is estimated by the extrapolation of the data $\{\lambda_m, \widehat{\tau}(\lambda_m) - s_\Delta^2(\lambda_m)\}_{m=1}^{M}$ to $\lambda = -1$ with quadratic exptrapolant $g_Q(\lambda, \Gamma)$. Here quadratic extrapolant is used to guarantee the nonnegativity of variance estimate. This estimate is denoted by $\widehat{V}_{SIMEX}$. The estimating equation method of Carroll *et al.* (1996) is easily evaluated by simple calculation and we denote $\widehat{V}_{EE}$ as the estimated variance of $\widehat{\beta}_{SIMEX}$ by this method.

Based on the asymptotic normality of $\widehat{\beta}_{naive}$, $\widehat{\beta}_L$, and $\widehat{\beta}_{SIMEX}$, the asymptotic 95% coverage probabilities are calculated after 100 repeats. The results are given

TABLE 1 *The asymptotic 95% coverage probabilities for $\widehat{\beta}_{naive}$, $\widehat{\beta}_L$, $\widehat{\beta}_{SIMEX}$ with $\widehat{V}_{SIMEX}$, and $\widehat{\beta}_{SIMEX}$ with $\widehat{V}_{EE}$*

estimator	*coverage probability*
$\widehat{\beta}_{naive}$	0.01
$\widehat{\beta}_L$	0.94
$\widehat{\beta}_{SIMEX}$ with $\widehat{V}_{SIMEX}$	0.72
$\widehat{\beta}_{SIMEX}$ with $\widehat{V}_{EE}$	0.93

in Table 1. As we expected, the coverage probability of $\widehat{\beta}_{naive}$ is far below the nominal level. Note that $\widehat{\beta}_L$ and $\widehat{\beta}_{SIMEX}$ with estimated variance $\widehat{V}_{EE}$ have almost nominal coverage probabilities. Even though $\widehat{\beta}_{SIMEX}$ is aymptotically equivalent to $\widehat{\beta}_L$, the coverage probability is smaller than that of $\widehat{\beta}_L$ when the variance is estimated by the simulation-extrapolation method. This is not due to the SIMEX estimation scheme of β, but is instead an artifact of using the variance estimation method of Stefanski and Cook (1995) on a quadratic extrapolant.

APPENDIX : PROOF FOR ASYMPTOTIC EQUIVALENCE BETWEEN $\widehat{\boldsymbol{\beta}}_{SIMEX}$ AND $\widehat{\boldsymbol{\beta}}_L$

Let $\widetilde{Y}_i = Y_i - E(Y_i|Z_i)$, $\widetilde{\mathbf{W}}_i = \mathbf{W}_i - E(\mathbf{W}_i|Z_i)$, and $\widetilde{\mathbf{W}}_{b,i}(\lambda) = \mathbf{W}_{b,i}(\lambda) - E(\mathbf{W}_{b,i}(\lambda)|Z_i)$ as $n \to \infty$. Note that $\widetilde{\mathbf{W}}_{b,i}(\lambda) = \widetilde{\mathbf{W}}_i + \sqrt{\lambda}\mathbf{U}_{b,i}$, and $\mathbf{U}_{b,i}$ are *iid* random variables from $N(\mathbf{0}, \Sigma_{uu})$ and independent of $(Y_i, \mathbf{W}_i, Z_i)$.

For fixed λ and b, we have

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\widetilde{\mathbf{W}}_{b,i}(\lambda)\widetilde{\mathbf{W}}_{b,i}^t(\lambda) &= \frac{1}{n}\sum_{i=1}^{n}\widetilde{\mathbf{W}}_i\widetilde{\mathbf{W}}_i^t + \frac{2\sqrt{\lambda}}{n}\sum_{i=1}^{n}\widetilde{\mathbf{W}}_i\mathbf{U}_{b,i}^t + \frac{\lambda}{n}\sum_{i=1}^{n}\mathbf{U}_{b,i}\mathbf{U}_{b,i}^t \\ &= \frac{1}{n}\sum_{i=1}^{n}\widetilde{\mathbf{W}}_i\widetilde{\mathbf{W}}_i^t + \lambda\Sigma_{uu} + O_p(n^{-1/2}) \qquad (A.1)\end{aligned}$$

by using $A = E(A) + O_p\{\text{Var}(A)^{1/2}\}$ for a random variable A. Similarly, we obtain

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\widetilde{\mathbf{W}}_{b,i}(\lambda)\widetilde{Y}_i &= \frac{1}{n}\sum_{i=1}^{n}\widetilde{\mathbf{W}}_i\widetilde{Y}_i + \frac{\sqrt{\lambda}}{n}\sum_{i=1}^{n}\mathbf{U}_{b,i}\widetilde{Y}_i \\ &= \frac{1}{n}\sum_{i=1}^{n}\widetilde{\mathbf{W}}_i\widetilde{Y}_i + O_p(n^{-1/2}). \qquad (A.2)\end{aligned}$$

By (A.1) and (A.2), it follows that

$$\begin{aligned}\widehat{\beta}_b(\lambda) &= \left\{\sum_{i=1}^{n} \widetilde{\mathbf{W}}_{b,i}(\lambda)\widetilde{\mathbf{W}}_{b,i}^t(\lambda)\right\}^{-1} \sum_{i=1}^{n} \widetilde{\mathbf{W}}_{b,i}(\lambda)\widetilde{Y}_i \\ &= \left\{\sum_{i=1}^{n} \widetilde{\mathbf{W}}_i\widetilde{\mathbf{W}}_i^t + n\lambda\Sigma_{uu} + O_p(n^{1/2})\right\}^{-1} \left\{\sum_{i=1}^{n} \widetilde{\mathbf{W}}_i\widetilde{Y}_i + O_p(n^{1/2})\right\} \\ &= \left\{\sum_{i=1}^{n} \widetilde{\mathbf{W}}_i\widetilde{\mathbf{W}}_i^t + n\lambda\Sigma_{uu}\right\}^{-1} \sum_{i=1}^{n} \widetilde{\mathbf{W}}_i\widetilde{Y}_i + O_p(n^{-1/2}).\end{aligned}$$

Therefore, the average $\widehat{\beta}(\lambda) = \sum_{b=1}^{B} \widehat{\beta}_b(\lambda)/B$ obtained in the simulation step is equal to

$$\left\{\sum_{i=1}^{n} \widetilde{\mathbf{W}}_i\widetilde{\mathbf{W}}_i^t + n\lambda\Sigma_{uu}\right\}^{-1} \sum_{i=1}^{n} \widetilde{\mathbf{W}}_i\widetilde{Y}_i \tag{A.3}$$

at the rate $O_p(n^{-1/2})$. Note that (A.3) is also equivalent to $\widehat{\beta}_L$ when $\lambda = -1$. From this fact, it is proved that $\widehat{\beta}_{SIMEX}$ obtained with the rational linear extrapolant $g_{RL}(\lambda, \Gamma)$ in the extrapolation step is asymptotically equivalent to $\widehat{\beta}_L$ in a partially linear measurement error model.

Acknowledgements

Both authors are members of Research Institute of Computer, Information and Communication at the Pusan National University. The authors appreciate to professor Raymond J. Carroll of Texas A&M University for many helpful comments.

References

Carroll, R. J., Küchenhoff, H., Lombard, F. and Stefanski, L. A. (1996). "Asymptotics for the SIMEX estimator in nonlinear measurement error models", *Journal of the American Statistical Association*, **91**, 242–250.

Carroll, R. J., Ruppert, D. and Stefanski, L. A. (1995). *Measurement Error in Nonlinear Models*, Chapman & Hall, New York.

Carroll, R. J. and Stefanski, L. A. (1990). "Approximate quasi-likelihood estimation in models with surrogate predictors", *Journal of the American Statistical Association*, **85**, 652–663.

Chen, H. (1988). "Convergence rates for parametric components in a partially linear model", *The Annals of Statistics*, **16**, 136–146.

Cook, J. and Stefanski, L. A. (1995). "A simulation extrapolation method for parametric measurement error models", *Journal of the American Statistical Association*, **89**, 1314–1328.

FULLER, W. A. (1987). *Measurement Error Models*, John Wiley & Sons, New York.

HECKMAN, N. E. (1986). "Spline smoothing in partly linear models", *Journal of the Royal Statistical Society*, **B48**, 244–248.

KIM, C., HONG, C. AND JEONG, M. (2000). "Simulation-extrapolation via the Bezier curve in measurement error models", *Communications in Statistics-Simulation and Computation*, **29**, 1135–1148.

LIANG, H., HÄRDLE, W. AND CARROLL, R. J. (1999). "Large sample theory in a semiparametric partially linear errors-in-variables model", *The Annals of Statistics*, **27**, 1519–1535.

SEVERINI, T. A. AND STANISWALIS, J. G. (1994). "Quasi-likelihood estimation in semiparametric models", *Journal of the American Statistical Association*, **89**, 501–511.

SPECKMAN, P. (1988). "Kernel smoothing in partial linear models", *Journal of the Royal Statistical Society*, **B50**, 413–436.

STEFANSKI, L. A. AND COOK, J. (1995). "Simulation-extrapolation : The measurement error jackknife", *Journal of the American Statistical Association*, **90**, 1247–1256.