Statistics and Its Interface Volume 2 (2009) 83–90

Empirical likelihood based inference for additive partial linear measurement error models

Hua Liang*, Haiyan Su, Sally W. Thurston,
John D. Meeker and Russ Hauser

This paper considers statistical inference for additive partial linear models when the linear covariate is measured with error. To improve the accuracy of the normal approximation based confidence intervals, we develop an empirical likelihood based statistic, which is shown to be asymptotically chi-square distributed. We emphasize the finite-sample performance of the proposed method by conducting simulation experiments. The method is used to analyze the relationship between semen quality and phthalate exposure from an environment study.

AMS 2000 subject classifications: Primary 60G08, 62G10; secondary 62G20.
Keywords and phrases: Backfitting, Correction-for-attenuation, Coverage probability, Error-prone, Local linear regression, Semiparamatric estimation, Undersmoothing.

1. INTRODUCTION

Linear models and nonparametric models are often used to explore the relationship between a response variable and its covariates. In practice, however, these models may not be valid, either because of model misspecification or because the small sample size does not allow one to use nonparametric models. Partial linear models [1] allow the response variable to depend nonlinearly on one covariate and linearly on the remaining variables. Although the partially linear models play an important role in data analysis, they may not be sufficient when more than one covariate is nonlinearly related to the response. Additive partial linear models (APLM) generalize partial linear models to allow more than one covariate to have a nonlinear relationship to the response. APLM have proved to be very useful as they combine the flexibility of additive models [2, 3] and the interpretation of linear models. One of the attractive features of the APLM is that one can derive estimators of the linear parameters, which are root-n consistent and asymptotically normal. Furthermore, in APLM the estimators of the nonparametric components have the same desirable optimal convergence rates as in the traditional nonparametric regression. Recently a variety of algorithms such as the backfitting algorithm [4, 5] and marginal integration [6] have been proposed for the APLM and developed in the commonly used software packages, Splus/R and Matlab. This convenience makes the APLM more and more appealing in practice. See Stone [2] for a more detailed discussion on APLM.

Liang et al. [7] studied estimation of the parameters in an additive partial linear model when the linear covariate is measured with the additive error; i.e.,

$$Y = X^{\mathrm{T}}\beta + \sum_{d=1}^{D} f_d(Z_d) + \varepsilon \quad \text{and} \quad W = X + U, \tag{1}$$

where $X = (x_1, \ldots, x_p)^{\mathrm{T}}$ and $Z = (Z_1, \ldots, Z_D)^{\mathrm{T}}$ are the linear and nonparametric components, $f_1, \ldots, f_D$ are unknown smooth functions, U is the measurement error, independent of (X, Z, Y), and has covariance matrix Σ_{uu}, $\beta = (\beta_1, \ldots, \beta_p)$ is a vector of unknown parameters, and the model error ε has mean zero given (X, Z). The authors proposed attenuation-to-correction and SIMEX estimators of the parameter β, showed that the first resulting estimator is asymptotically normal, and pointed out that no undersmoothing is necessary as was the case for the previously published approaches for the APLM. The motivation, as pointed out by these authors, was to study the relationship between chemical exposures and semen quality, from an environmental study conducted at Massachusetts General Hospital (MGH) because nonlinear relationships between semen quality and several covariates are expected, especially for abstinence time [8], and age. Based on the asymptotic results developed, Liang et al. [7] calculated the estimated values of the parameters and the corresponding standard errors, from which confidence intervals can be derived easily. However, we actually encountered several challenges. For instance, (i) the finite-sample performance of the proposed method may be underestimated because of the complexity of the covariance matrix and the need to plug in several estimated terms (see Theorems 1 and 2 of [7]); (ii) the confidence region derived by this procedure is based on a normal approximation, which may not be realistic; (iii) there is a very large computation burden for the proposed SIMEX approach. The contribution of this paper is to develop a new approach to address this concern by using the empirical likelihood principle [9], where the focus is on constructing a nonparametric likelihood for parameters of interest in a

*Corresponding author.

parametric or semiparametric setting, with nice properties typical for the parametric likelihood such as Wilk's Theorem.

This paper is organized as follows. In Section 2, we briefly mention the correction-for-attenuation estimation procedure proposed in [7]. In Section 3, we propose the empirical likelihood based statistic and show that it has an asymptotic chi-squared distribution. Section 4 reports the results of a simulation experiment, and Section 5 presents the results applied to the same MGH semen study that was analyzed in [7]. All technical derivations are given in the Appendix.

2. CORRECTION-FOR-ATTENUATION ESTIMATION

For notational simplicity, we first consider $D = 2$ in (1). To ensure identifiability of the nonparametric functions, we assume that $E\{f_1(Z_1)\} = E\{f_2(Z_2)\} = 0$. Without loss of generality, we also assume that X and Y are centered.

Let $(X_1, Z_{11}, Z_{12}, W_1, Y_1), \ldots, (X_n, Z_{n1}, Z_{n2}, W_n, Y_n)$ be an iid sample of size n from model (1). Let $\mathbf{X} = (X_1^{\mathrm{T}}, \ldots, X_n^{\mathrm{T}})^{\mathrm{T}}$, $\mathbf{W} = (W_1^{\mathrm{T}}, \ldots, W_n^{\mathrm{T}})^{\mathrm{T}}$, $\mathbf{Y} = (Y_1, \ldots, Y_n)^{\mathrm{T}}$ and similarly for $\mathbf{Z}_1$ and $\mathbf{Z}_2$. Also let $\mathbf{Z}_i = (Z_{i1}, Z_{i2})$. We write the vectors of additive functions at the observations as $\mathbf{f}_1 = \{f_1(Z_{11}), \ldots, f_1(Z_{n1})\}^{\mathrm{T}}$ and $\mathbf{f}_2 = \{f_1(Z_{12}), \ldots, f_1(Z_{n2})\}^{\mathrm{T}}$. In matrix notation, (1) can be expressed as

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{f}_1 + \mathbf{f}_2 + \boldsymbol{\varepsilon}. \tag{2}$$

We first briefly review the estimators proposed by [7]. Using their notation we denote $A^{\otimes 2} = AA^{\mathrm{T}}$ and $A^{\star 2} = A^{\mathrm{T}}A$, $\widetilde{X}_i = X_i - E(X_i|Z_{i1}) - E(X_i|Z_{i2})$, $\widetilde{W}_i = W_i - E(W_i|Z_{i1}) - E(W_i|Z_{i2})$, $\widetilde{Y}_i = Y_i - E(Y_i|Z_{i1}) - E(Y_i|Z_{i2})$, and $\Gamma_{\mathrm{x|z}} = \mathrm{cov}(\widetilde{X})$.

For $j = 1, 2$, let s_{j,z_j}^{T} be the equivalent kernels for the local linear regression at z_j; $s_{j,z_j}^{\mathrm{T}} = \mathbf{e}_1^{\mathrm{T}}(\mathcal{Z}_j^{\mathrm{T}}\boldsymbol{\Omega}_j\mathcal{Z}_j)^{-1}\mathcal{Z}_j^{\mathrm{T}}\boldsymbol{\Omega}_j$, where $\mathbf{e}_1 = (1, 0)^{\mathrm{T}}$, $\boldsymbol{\Omega}_j = \mathrm{diag}\{K_{h_j}(Z_{1j} - z_j), \ldots, K_{h_j}(Z_{nj} - z_j)\}$ for a kernel function $K(\cdot)$ with bandwidth h_j, where $K_h(t) = K(t/h)/h$, and $\mathcal{Z}_j$ is a $n \times 2$ design matrix, whose ith row is $(1, Z_{ij} - z_j)$. Let $\mathbf{S}_1$ and $\mathbf{S}_2$ be the smoother matrices whose rows are the equivalent kernels at the observations $(Z_{11}, \ldots, Z_{n1})^{\mathrm{T}}$ and $(Z_{12}, \ldots, Z_{n2})^{\mathrm{T}}$; i.e.,

$$\mathbf{S}_1 = [s_{1Z_{11}}^{\mathrm{T}}, \ldots, s_{1Z_{n1}}^{\mathrm{T}}]^{\mathrm{T}}, \qquad \mathbf{S}_2 = [s_{2Z_{12}}^{\mathrm{T}}, \ldots, s_{2Z_{n2}}^{\mathrm{T}}]^{\mathrm{T}}.$$

$\mathbf{S}_1^c = (\mathbf{I} - \mathbf{J}/n)\mathbf{S}_1$ is the centered smoother matrix corresponding to $\mathbf{S}_1$, and similarly for $\mathbf{S}_2^c$, where $\mathbf{J}$ is an $n \times n$ matrix of 1's. Let $\mathbf{S}_{12} = \{\mathbf{I} - (\mathbf{I} - \mathbf{S}_1^c\mathbf{S}_2^c)^{-1}(\mathbf{I} - \mathbf{S}_1^c)\} + \{\mathbf{I} - (\mathbf{I} - \mathbf{S}_2^c\mathbf{S}_1^c)^{-1}(\mathbf{I} - \mathbf{S}_2^c)\}$.

Applying the least-squares principle, Liang et al. [7] proposed the profile-based estimator of β as follows:

$$\widehat{\beta}_{n,\mathrm{ac}} = \{\mathbf{W}^{\mathrm{T}}(\mathbf{I} - \mathbf{S}_{12})^{\star 2}\mathbf{W} - n\Sigma_{uu}\}^{-1}\mathbf{W}^{\mathrm{T}}(\mathbf{I} - \mathbf{S}_{12})^{\star 2}\mathbf{Y}. \tag{3}$$

Under a set of assumptions similar to A1–A6 given in the Appendix A.1, the authors derived the asymptotic distribution of the proposed estimators. For $\Sigma_{u,u}$ unknown, the asymptotic variance of $\widehat{\beta}_{n,\mathrm{ac}}$ is $\Gamma_{\mathrm{x|z}}^{-1}\Sigma_{\mathrm{ac}}\Gamma_{\mathrm{x|z}}^{-1}$ with $\Sigma_{\mathrm{ac}} = E[\{(\varepsilon - U^{\mathrm{T}}\beta)\widetilde{X}\}^{\otimes 2}] + E\{(UU^{\mathrm{T}} - \Sigma_{uu})\beta\}^{\otimes 2} + E(UU^{\mathrm{T}}\varepsilon^2)$. These quantities can further be estimated by $\widehat{\Gamma}_{\mathrm{x|z}} = \mathbf{W}^{\mathrm{T}}(\mathbf{I} - \mathbf{S}_{12})^{\star 2}\mathbf{W}/n - \Sigma_{uu}$ and

$$\widehat{\Sigma}_{\mathrm{ac}} = \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{W}_i\widehat{\varepsilon}_i + \Sigma_{uu}\widehat{\beta}_{n,\mathrm{ac}}\right)^{\otimes 2},$$

where $\widehat{W}_i$ is the ith row of $(\mathbf{I} - \mathbf{S}_{12})\mathbf{W}$, $\widehat{Y}_i$ is the ith element of $(\mathbf{I} - \mathbf{S}_{12})\mathbf{Y}$, and $\widehat{\varepsilon}_i$ is the ith element of $(\mathbf{I} - \mathbf{S}_{12})(\mathbf{Y} - \mathbf{W}\widehat{\beta}_{n,\mathrm{ac}})$. Also Σ_{uu} can be estimated using external data or internal replicates [7, 10]. It is easily seen that $\widehat{\Sigma}_{\beta,\mathrm{ac}} = \widehat{\Gamma}_{\mathrm{x|z}}^{-1}\widehat{\Sigma}_{\mathrm{ac}}\Gamma_{\mathrm{x|z}}^{-1}$ is a consistent estimator of $\Sigma_{\beta,\mathrm{ac}}$.

3. INFERENCE BASED ON EMPIRICAL LIKELIHOOD

Based on the estimators of the covariance matrix or its bootstrap version, one can obtain a confidence region for β. However, its finite-sample behavior may be affected by the need to plug in several estimated terms, and the reasons we described earlier. An alternative method is to use the empirical likelihood principle, which uses the likelihood function incorporating auxiliary information such as known constraints on the parameters, adjustments for biased sampling schemes, and does not involve specifying the shape of the confidence region [11]. The most appealing features of the empirical likelihood method include avoiding estimation of the covariance of the estimators, increasing coverage accuracy because it includes auxiliary information, and convenience of implementation. See Owen [9] for a comprehensive survey on empirical likelihood, and [9, 12, 13] for a more detailed discussion on advantages of the empirical likelihood methods over the conventional methods. The methods have been applied in a variety of topics, for example, linear models [11, 14, 15], general estimating equations [13], and partially linear models [16–20].

In the remainder of this section, we assume that the ε_i are independent and identically distributed and independent of (W_i, Z_i). We will study the empirical-likelihood-based confidence interval for β. Let F be the distribution function which assigns probability p_i at points $(W_i, Y_i, \mathbf{Z}_i)$. Then $\sum_{i=1}^n p_i = 1$ and $p_i \geq 0$ for each i. We now motivate and define our semiparametric empirical likelihood ratio estimator.

Recalling the estimator given in (3) and the definition of $\widehat{Y}_i$, we notice that $\widehat{\beta}_{n,\mathrm{ac}}$ is actually the solution of the equation:

$$\mathbf{W}^{\mathrm{T}}(\mathbf{I} - \mathbf{S}_{12})^{\star 2}\mathbf{Y} - \{\mathbf{W}^{\mathrm{T}}(\mathbf{I} - \mathbf{S}_{12})^{\star 2}\mathbf{W} - n\Sigma_{uu}\}\beta = 0,$$

which is equivalent to

$$\sum_{i=1}^{n}\Big\{\widehat{W}_i\widehat{Y}_i-(\widehat{W}_i\widehat{W}_i^{\mathrm{T}}-\Sigma_{uu})\beta\Big\}=0.$$

Consequently, our empirical likelihood ratio function for β is defined as

$$\mathcal{R}_n(\beta)=\max\Bigg\{\prod_{i=1}^{n}np_i:\sum_{i=1}^{n}p_i\{\widehat{W}_i(\widehat{Y}_i-\widehat{W}_i^{\mathrm{T}}\beta)+\Sigma_{uu}\beta\}=0,p_i\geq 0,\sum_{i=1}^{n}p_i=1\Bigg\}. \tag{4}$$

Theorem 1. *Under Assumptions A1–A6 given in the Appendix, $-2\log\{\mathcal{R}_n(\beta)\}$ converges in distribution to a chi-squared distribution with p degrees of freedom.*

Based on this result, a confidence region for β can be given by $\{\beta:-2\log\{\mathcal{R}_n(\beta)\}\leq c_\alpha\}$, where c_α denotes the α quantile of the chi-squared distribution. When Σ_{uu} is unknown, we need replicate data in the usual way. In the special case of $m_i\equiv 2$, we can then replace W_i by $\overline{W}_i$ and $\Sigma_{uu}\beta$ by $\widehat{\Sigma}_{uu}\beta/2$. The resulting statistic still has the property given in Theorem 1. A justification of this last assertion can be easily obtained by using the fact that

$$E[\{\overline{W}_{ir}(\widetilde{Y}_i-\overline{W}_{ir}^{\mathrm{T}}\beta)+\widehat{\Sigma}_{uu}\beta/2\}]=0,$$

where $\overline{W}_{ir}$ is analogous to $\widehat{W}_i$ except replacing $\mathbf{W}$ in $(\mathbf{I}-\mathbf{S}_{12})\mathbf{W}$ by $\overline{\mathbf{W}}$.

Remark 1. Comparing to the attenuation correction method, which requires estimation of Σ_{ac} and $\Gamma_{\mathrm{x|z}}$ for inference, the EL method only requires $\widehat{W}$, $\widehat{Y}$, and an unbiased estimator $\widehat{\Sigma}_{uu}$ if Σ_{uu} is unknown. The implementation of the new method is a simple optimization, which can be achieved in well-developed software packages such as R/Splus or Matlab.

Remark 2. The bandwidth h_1 and h_2 are of order $n^{-1/5}$. Thus, any bandwidths with this order give the same limiting distribution of $\widehat{\beta}_{n,\mathrm{ac}}$, and the same limiting distribution of $-2\log\{\mathcal{R}_n(\beta)\}$. In our implementation, we used an equal-spaced sequence of length 20 in the interval $[.75n^{-1/5},1.25n^{-1/5}]$. The optimal bandwidth is selected as the minimizer of the criterion: $\sum_{i=1}^{n}\{Y_i-\widehat{Y}_i^{(-i)}\}^2$, where $\widehat{Y}_i^{(-i)}$ stands for the fitted value based on the data with the ith observation excluded.

4. A SIMULATION STUDY

To evaluate the performance of the proposed methods, we conducted a small scale simulation experiment. We generated $n=200$ and $n=400$ observations from model (1) with $D=2$ and X, Z_1, and Z_2 independently generated from uniform$(0,1)$. In our simulations, $\beta=4$, $f_1(z_1)=\exp(2z_1)-3.75$ and $f_2(z_2)=0.2z_2^{11}\{10*(1-z_2)\}^6+10^4z_2^3(1-z_2)^{10}-1.4$. We also assume that the measurement error follows $W=X+U$, where $U\sim\mathrm{normal}(0,0.2^2)$.

We considered two different error structures for ϵ, (a) ε follows normal$\{0,\sigma_\varepsilon^2(x,z,\sigma_0^2)\}$, where $\sigma_\varepsilon^2(x,z,\sigma_0^2)=\{\sigma_0\sin(2\pi x^3)+0.5z_1+0.5z_2+0.3\}^2$, and (b) ε follows $\sigma_0^2(\mathcal{X}_2^2-2)$, where $\mathcal{X}_2^2$ is a chi-squared variable with 2 degrees of freedom. We consider this case to see the effect of asymmetric error on the confidence intervals. For each error structure we used one of 4 possible values of σ_0^2: $\sigma_0^2=0.1^2,.25^2,.5^2,1$.

In our nonparametric estimation procedure, we used the kernel function $K(u)=(3/4)(1-u)^2I_{(|u|\leq 1)}$. We selected bandwidth as the minimizer of the criterion: $\sum_{i=1}^{n}\{Y_i-\widehat{Y}_i^{(-i)}\}^2$, where $\widehat{Y}_i^{(-i)}$ stands for the fitted value based on the data with the ith observation excluded. We generated 500 data sets for each case. To estimate the variance of U, we generated duplicate samples of W. The results are presented in Table 1. The lengths of the confidence intervals for the EL method were always as small or smaller than the AC method. Although the coverage probabilities for the AC method were generally somewhat larger than the EL method, the coverage probabilities for both methods were close to 95%. Overall, our simulation results suggest the EL method generally performs better than the AC method.

5. MGH SEMEN STUDY

We now present the results of the empirical likelihood based confidence interval as applied to the same MGH semen study that was examined in [7]. An earlier version of this dataset was also examined in [8, 21]. Our analysis uses complete data on 455 subjects, and our interest focuses on the covariate-adjusted and bias-corrected relationship between sperm concentration and monobutyl phthalate (MBP), the metabolite of di-n-butyl phthalate. In order to account for differences in urinary dilution, MBP was specific gravity-adjusted, consistent with [7, 8, 21]. Both sperm concentration (millions of sperm per ml), and MBP (ng monoester per ml urine) were log-transformed to better satisfy regression assumptions.

MBP is subject to substantial measurement error, not because of the measurement process itself, but because of the very short half-life of MBP. The short half-life of MBP, which is less than a day long, means that short-term changes in di-n-butyl phthalate exposure can give rise to large temporal variations in MBP concentration. Thus a single measurement of MBP is not an accurate representation of the average exposure over a 3-month time window, which is the time interval thought to be important for sperm development. Replicate measurements on 78 subjects [22] allow us to estimate the measurement error variance. On the logarithmic scale, the sample correlation between replicate MBP measurements was only 0.101 [7], indicating an extremely large measurement error.

Table 1. The 95% confidence intervals based on the empirical likelihood (EL) and attenuation correction (AC) methods, the length of CI, and the associated coverage probability (CP) for the simulated data

case	n	σ^2	Est	CI		length		CP	
				AC	EL	AC	EL	AC	EL
1	200	0.1	4.02	(3.75, 4.30)	(3.76, 4.30)	0.55	0.54	95.8	94.5
		0.25	4.03	(3.74, 4.31)	(3.78, 4.28)	0.57	0.50	97.0	94.0
		0.5	4.02	(3.73, 4.31)	(3.73, 4.31)	0.58	0.58	95.6	94.5
		1	4.02	(3.69, 4.35)	(3.71, 4.34)	0.66	0.63	96.2	96.0
	400	0.1	4.01	(3.82, 4.20)	(3.84, 4.19)	0.38	0.35	95.8	96.0
		0.25	4.02	(3.82, 4.21)	(3.83, 4.20)	0.39	0.37	96.0	94.8
		0.5	4.01	(3.81, 4.21)	(3.82, 4.19)	0.40	0.37	96.4	93.5
		1	4.01	(3.78, 4.24)	(3.80, 4.22)	0.46	0.42	95.8	93.5
2	200	0.1	4.02	(3.79, 4.24)	(3.81, 4.26)	0.45	0.45	97.2	94.5
		0.25	4.03	(3.80, 4.25)	(3.80, 4.25)	0.45	0.45	97.8	96.0
		0.5	4.01	(3.79, 4.24)	(3.83, 4.20)	0.45	0.37	98.6	94.8
		1	4.02	(3.79, 4.24)	(3.80, 4.25)	0.45	0.45	97.8	98.5
	400	0.1	4.01	(3.86, 4.17)	(3.88, 4.15)	0.31	0.27	96.2	93.5
		0.25	4.02	(3.86, 4.17)	(3.88, 4.15)	0.31	0.27	96.8	93.0
		0.5	4.01	(3.86, 4.16)	(3.88, 4.15)	0.30	0.27	97.4	93.5
		1	4.01	(3.86, 4.17)	(3.87, 4.15)	0.31	0.28	97.4	94.4

Table 2. The 95% confidence intervals of the parameters from the MGH semen study from two methods: AC (attenuation correction) and EL (empirical likelihood)

	EL	AC
Abstinence time	$(0.0238, 0.1244)$	$(0.0133, 0.1387)$
Race (white)	$(-0.0970, 0.6468)$	$(-0.1056, 0.6156)$
Log(MBP)	$(-0.2980, 0.0437)$	$(-0.3300, 0.0660)$

Liang et al. [7] argued that log(sperm concentration) may nonlinearly depend on age and BMI, but linearly depend on log(MBP) and abstinence time. Age, race, and abstinence time have been justified to be important predictors by [8, 21]. We now apply the model and method described in Section 3 to explore this data set and address the concerns mentioned in Section 1. We present results using empirical likelihood (EL), and compare them to results using attenuation correction (AC): see Table 2. The confidence intervals for the bias-corrected log(MBP) for the EL method were noticeable shorter than for the AC. The length of the confidence intervals for the other covariates were similar for the two methods.

6. DISCUSSION

To simplify statistical inference for the additive partial linear model with an error-prone linear component, we developed an empirical likelihood-based approach to construct a confidence interval for the linear parameter of this model. The proposed approach is simpler than its competitor normal approximation, can easily be implemented in standard software, and computation of the estimators is efficient. The finite-sample performance of the proposed statistics shows promise.

In this article, we only studied continuous response variables. A natural extension is the generalized additive partial linear measurement error models (GAPLMeM) which takes the form

$$E(Y_i|X_i, \mathbf{Z}_i) = \mu\left\{X_i^{\mathrm{T}}\boldsymbol{\beta} + \sum_{k=1}^{K} f_k(Z_{k,i})\right\}, \quad W_i = X_i + U_i$$

where $\mu(\cdot)$ is a link function, and $\mathbf{Z}_i = (Z_{1,i}, \ldots, Z_{K,i})^{\mathrm{T}}$ is a K-dimensional vector. Making statistical inference on the parameters for these GAPLMeM is extremely difficult, if not impossible, using the asymptotic properties established in the literature. However, defining an empirical likelihood ratio and developing an empirical likelihood based method for statistical inference are also not trivial and in fact may be very challenging. There are several reasons why inference for the GAPLMeM model is much more difficult than for the APLMeM model. First, the backfitting algorithm is needed in order to estimate the model parameters, and this involves iterative computations. Second, the involvement of measurement error increases the difficulties because correction for the effect of an error-prone covariates is never trivial. Lastly, implementation of the procedure into standard software may be difficult. Development of an appropriate procedure to solve these problems is an important area of future research.

ACKNOWLEDGEMENT

Liang's research was partially supported by NIH/NIAID grants AI62247 and AI059773, and NSF grant DMS 0806097. Su's research was supported by a Merck Quantitative Sciences Fellowship grant provided by the Merck Company Foundation. Meeker's research was supported by National Institute of Environmental Health Sciences (NIEHS)

grants ES009718 and ES00002. Hauser's research was supported by National Institute of Environmental Health Sciences (NIEHS) grants ES009718 and ES00002. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors thank a referee for her/his comments which improved this article.

APPENDIX

A.1 Assumptions

Assumption A1. The matrix $\Gamma_{x|z}$ is positive-definite, $E(\varepsilon|X,Z)=0$, and $E(|\varepsilon|^3|X,Z)<\infty$.

Assumption A2. The bandwidths h_d for $d=1,2$ are of order $n^{-1/5}$.

Assumption A3. $E|Y|^3<\infty$ and $\sup_{z_1,z_2}E(|Y|^3|Z_1=z_1,Z_2=z_2)<\infty$;

Assumption A4. The kernel function $K(\cdot)$ satisfies the following conditions:

(i) $K(u)=0$ for $u\notin(0,1)$ and bounded by a constant C_k in $(0,1)$;
(ii) $\int|u|K(u)du<\infty$.
(iii) $K(u)$ is differentiable. For some constants L and C_1, $|K'(u)|\le C_1$ for $|u|<L$, and $|K'(u)|\le C_1|u|^{-\gamma}$ for $|u|\ge L$ and some $\gamma>1$.

Assumption A5. The density functions of Z_1 and Z_2 are bounded away from zero and have bounded continuous second partial derivatives.

Assumption A6. The random variable U satisfies $E(\|U\|^3)<\infty$.

A.2 Uniform rate of convergence of local regression

To finish the proof of the main result, we first derive the uniform rate of convergence of local linear regression under the mild assumptions. This result is independently interesting.

Suppose that we model data by $Y=m(Z)+\varepsilon$ and that we have an iid sample $(Z_1,Y_1),\dots,(Z_n,Y_n)$ from the model. Let $\widehat{m}(z,1,h)$ and $\widehat{m}(z,0,h)$ be the local linear and local constant regression estimators of $m(z)$ based on samples. A direct calculation gives the expression of $\widehat{m}(z,1,h)$ as follows.

$$\widehat{m}(z,1,h)=\frac{1}{n}\sum_{i=1}^{n}\frac{\{\widehat{s}_2(z,h)-\widehat{s}_1(z,h)(Z_i-z)\}K_h(Z_i-z)Y_i}{\widehat{s}_2(z,h)\widehat{s}_0(z,h)-\widehat{s}_1^2(z,h)},$$

where

$$\widehat{s}_r(z,h)=\frac{1}{n}\sum_{i=1}^{n}(Z_i-z)^rK_h(Z_i-z),\quad\text{for } r=0,1,2.$$

Bernstein's Inequality [23]: *Let $V_1,\dots,V_n$ be independent random variables with zero means and bounded ranges: $|V_i|\le M$. Then for each $\eta>0$,*

$$P\left(\left|\sum_{i=1}^{n}V_i\right|>\eta\right)\le 2\exp\left[-\frac{\eta^2}{2\{\sum_{i=1}^{n}var(V_i)+M\eta\}}\right].$$

Set $\nu=\int u^2K(u)du$, $a_n=n^{-1/3}\log n$, $a_n^*=h^2+a_n$, $\tau_n=a_n^{-1/2}$.

Lemma 1. *Suppose that $\sup_z E(|Y|^3|Z=z)<\infty$, $E(|Y|^3)<\infty$, and Assumption A4 holds. Let $h=n^{-1/5}$, then*

$$\sup_{z\in[0,1]}|\widehat{m}(z,1,h)-m(z)|=O(n^{-2/5}\log n). \tag{A.1}$$

Proof. We first derive a uniform convergence rate for the local linear constant estimators, and then express the local linear regression as a function of the local constant estimators and complete the proof. This idea was used by Hansen [24].

Write

$$\widehat{g}(z)=\frac{1}{n}\sum_{i=1}^{n}K_h(Z_i-z)Y_i,$$
$$\widehat{f}(z)=\frac{1}{n}\sum_{i=1}^{n}K_h(Z_i-z),$$
$$Q_n(z)=\frac{1}{n}\sum_{i=1}^{n}K_h(Z_i-z)Y_iI_{(|Y_i|>\tau_n)}.$$

Then the local constant kernel estimators $\widehat{m}(z,0,h)$ equals $\widehat{g}(z)/\widehat{f}(z)$. It follows that

$$\begin{aligned}|E\{Q_n(z)\}|&\le\int|K_h(z-u)|E\{YI_{(|Y|>\tau_n)}|Z=u\}f(u)du\\&\le\int|K(u)|E\{|Y|^3\tau_n^{-2}I_{(|Y|>\tau_n)}|Z=z-hu\}f(z-hu)du\\&\le\tau_n^{-2}\int|K(u)|E(|Y|^3|Z=z-hu)f(z-hu)du\\&\le\tau_n^{-2}C=Ca_n.\end{aligned}$$

Then we have $|Q_n(z)-EQ_n(z)|=O_p(a_n)$ by Markov's inequality. It is readily seen that $|\widehat{g}(z)-E\{\widehat{g}(z)\}|=O_p(a_n)$. In the remainder of the proof of Lemma 1, we simply assume that $|Y_i|<\tau_n$.

Cover the interval $[0,1]$ by N equal-length sub-intervals, A_j, centered at z_j such that $N\le a_n^{-1}h^{-1}$. Let $K^*(u)=C_1\{I_{(|u|\le 2L)}+|u-L|^{-\gamma}I_{(|u|\ge L)}\}$. Then if $|u_1-u_2|\le\delta\le L$, $|K(u_1)-K(u_2)|\le\delta K^*(u_1)$ for some $\delta>0$, and $K^*(u)$ still satisfies Assumption A4(i) and (ii).

Write $\widehat{g}^*(z)=\sum_{i=1}^{n}K_h^*(Z_i-z)Y_i/n$. For any $z\in A_j$,

then $|z - z_j| \le a_n h$ and, for large enough M,

$$\begin{aligned}
&\sup_{z\in A_j} |\widehat{g}(z) - E\widehat{g}(z)| \\
&\quad \le |\widehat{g}(z_j) - E\widehat{g}(z_j)| + a_n\{|\widehat{g}^*(z_j) + E\widehat{g}^*(z_j)|\} \\
&\quad \le |\widehat{g}(z_j) - E\widehat{g}(z_j)| + a_n\{|\widehat{g}^*(z_j) - E\widehat{g}^*(z_j)|\} \\
&\qquad + 2a_n E|\widehat{g}^*(z_j)| \\
&\quad \le |\widehat{g}(z_j) - E\widehat{g}(z_j)| + \{|\widehat{g}^*(z_j) - E\widehat{g}^*(z_j)|\} + 2a_n M.
\end{aligned}$$

It follows that

$$\begin{aligned}
&P\{\sup_z |\widehat{g}(z) - E\widehat{g}(z)| > 3Ma_n\} \\
&\quad \le N \max_j P\{\sup_{z\in A_j} |\widehat{g}(z) - E\widehat{g}(z)| > 3Ma_n\} \\
&\quad \le N \max_{1\le j\le N} P\{|\widehat{g}(z_j) - E\widehat{g}(z_j)| > M\} \\
&\qquad + N \max_{1\le j\le N} P\{|\widehat{g}^*(z_j) - E\widehat{g}^*(z_j)| > M\}.
\end{aligned}$$

Denote

$$\zeta_{ni}(z) = K\left(\frac{Z_i - z}{h}\right) Y_i - E\left\{K\left(\frac{Z_i - z}{h}\right) Y_i\right\}.$$

Note that $|Y_i| \le \tau_n$ and $\left|K\left(\frac{Z_i - z}{h}\right)\right| < C_k$ It follows that $|\zeta_{ni}(z)| \le 2C_k\tau_n$.

By using the Bernstein's inequality, we obtain that, for any given z,

$$\begin{aligned}
&P\{|\widehat{g}(z) - E\widehat{g}(z)| > Ma_n\} \\
&\quad = P\left\{\left|\sum_{i=1}^n \zeta_{ni}(z)\right| > Ma_n nh\right\} \\
&\quad \le 2\exp\left\{-\frac{M^2 a_n^2 n^2 h^2}{2(nh + 2C_k\tau_n Ma_n nh)}\right\}.
\end{aligned}$$

For a large enough n, this term is less than Cn^{-2}. By the Borel–Cantelli inequality, we know that $\sup_z |\widehat{g}(z) - E\widehat{g}(z)| = O_p(a_n)$. On the other hand, it is easy to see $E\widehat{g}(z) = g(z) + O(h^2)$. Summarizing these arguments, we conclude that

$$P\{\sup_z |\widehat{g}(z) - E\widehat{g}(z)| > 3Ma_n\} = o(1), \tag{A.2}$$

and (A.1) follows. In the same way as for (A.1), we obtain that

$$\sup_z |\widehat{f}(z) - f(z)| = o(a_n^*). \tag{A.3}$$

A direct manipulation yields that

$$\sup_z |\widehat{m}(z, 0, h) - m(z)| = o(a_n^*). \tag{A.4}$$

We now consider $\widehat{m}(z, 1, h)$, which can be rewritten as

$$\frac{\widehat{g}(z) - \widehat{s}_1(z,h)\widehat{s}_2^{-1}(z,h)\widehat{N}(z,h)}{\widehat{f}(z) - \widehat{s}_1^2(z,h)\widehat{s}_2^{-1}(z,h)},$$

where

$$\widehat{N}(z,h) = \frac{1}{n}\sum_{i=1}^n \frac{Z_i - z}{h} K_h(Z_i - z) Y_i.$$

Using the arguments similar to the proof for (A.2), we can show that, uniformly on z,

$$\begin{aligned}
\widehat{s}_1(z,h) &= h\nu f^{(1)}(z) + o_p(a_n^*), \\
\widehat{s}_2(z,h) &= \nu f(z) + o_p(a_n^*), \\
\widehat{N}(z,h) &= h\nu g^{(1)}(z) + o_p(a_n^*).
\end{aligned}$$

Because $f^{(1)}(z)$ and $g^{(1)}(z)$ are bounded, we know, uniformly on z,

$$\begin{aligned}
f^{(-1)}(z)\widehat{s}_1(z,h) &= o_p(h + a_n^*), \\
f^{(-1)}(z)\widehat{s}_2(z,h) &= \nu + o_p(a_n^*), \\
f^{(-1)}(z)\widehat{N}(z,h) &= O_p(h + a_n^*).
\end{aligned}$$

It follows that $f^{(-1)}(z)\widehat{s}_1(z,h)\widehat{s}_2(z,h) = O_p(a_n^*)$, $f^{(-1)}(z)\widehat{s}_1^2(z,h)\widehat{s}_2^{-1}(z,h) = O_p(a_n^*)$. As a result, $\widehat{m}(z,1,h) = \widehat{m}(z,0,h) + O_p(a_n^*)$ uniformly on z, while $\widehat{m}(z,0,h) = m(z) + O_p(a_n^*)$ by (A.4). We therefore complete the proof of Lemma 1.

We will use the following lemma, whose proof can be found in [25].

Lemma 2. *Assume that random variables a_i and b_i satisfy $Ea_i = 0$ and $\|b_i\| = o_p(n^{-1/4})$. Then*

$$\sum_{i=1}^n a_i b_i \xi_i = o_p(n^{1/2}),$$

where ξ_i are independent variables with zero conditional mean and finite variance.

A.3 Proof of Theorem 1

Let

$$\widehat{\Omega}_i = \widehat{W}_i(\widehat{Y}_i - \widehat{W}_i^{\mathrm{T}}\beta) + \Sigma_{uu}\beta.$$

A standard simplification as in [9] (pp61) yields that

$$p_i = \frac{1}{n(1 + a^{\mathrm{T}}\widehat{\Omega}_i)} \quad \text{for } i = 1, \ldots, n, \tag{A.5}$$

where $a = (a_1, \ldots, a_p)^{\mathrm{T}}$ is the solution of the equation

$$n^{-1}\sum_{i=1}^n \frac{\widehat{\Omega}_i}{1 + a^{\mathrm{T}}\widehat{\Omega}_i} = 0. \tag{A.6}$$

Mimicking the proof of Theorem 3.2 in [9], we have

$$\|a\| = O_p(n^{-1/2}). \tag{A.7}$$

On the other hand, based on the assumptions, Theorem 1 of [7] and the strong law of large numbers, we have

$$\max_{1\le i\le n}\|\widehat{\Omega}_i\| = o_p(n^{1/2}). \tag{A.8}$$

Note that

$$n^{-1}\sum_{i=1}^n \frac{\widehat{\Omega}_i}{1+a^{\mathrm{T}}\widehat{\Omega}_i} = n^{-1}\sum_{i=1}^n \widehat{\Omega}_i(1-a^{\mathrm{T}}\widehat{\Omega}_i) + n^{-1}\sum_{i=1}^n \frac{(a^{\mathrm{T}}\widehat{\Omega}_i)^2\widehat{\Omega}_i}{1+a^{\mathrm{T}}\widehat{\Omega}_i}.$$

The second term is $o_p(n^{-1/2})$ since $|a^{\mathrm{T}}\widehat{\Omega}_i| = o_p(1)$ and

$$\sum_{i=1}^n (a^{\mathrm{T}}\widehat{\Omega}_i)^2\widehat{\Omega}_i \le \|a\|\max_{1\le i\le n}|a^{\mathrm{T}}\widehat{\Omega}_i|\sum_{i=1}^n\|\widehat{\Omega}_i\|^2 = O_p(n^{-1/2})o_p(1)O_p(n) = o_p(1).$$

It then follows from (A.6) that

$$a = \left(\sum_{i=1}^n \widehat{\Omega}_i\widehat{\Omega}_i^{\mathrm{T}}\right)^{-1}\sum_{i=1}^n\widehat{\Omega}_i + o_p(n^{-1/2}). \tag{A.9}$$

Noting $\sum_{i=1}^n p_i = 1$ and using an argument similar to that for (A.9), we have that

$$0 = n^{-1}\sum_{i=1}^n \frac{a^{\mathrm{T}}\widehat{\Omega}_i}{1+a^{\mathrm{T}}\widehat{\Omega}_i} = n^{-1}\sum_{i=1}^n a^{\mathrm{T}}\widehat{\Omega}_i - n^{-1}\sum_{i=1}^n (a^{\mathrm{T}}\widehat{\Omega}_i)^2 + o_p(n^{-1}).$$

Therefore, we have

$$\sum_{i=1}^n a^{\mathrm{T}}\widehat{\Omega}_i = \sum_{i=1}^n (a^{\mathrm{T}}\widehat{\Omega}_i)^2 + o_p(1). \tag{A.10}$$

Consider $\mathcal{R}_n(\beta)$. Using a Taylor expansion of $\log(1+x)$ on x, we have

$$-\log\{\mathcal{R}_n(\beta)\} = \sum_{i=1}^n \log(1+a^{\mathrm{T}}\widehat{\Omega}_i) = \sum_{i=1}^n\left\{a^{\mathrm{T}}\widehat{\Omega}_i - (1/2)(a^{\mathrm{T}}\widehat{\Omega}_i)^2\right\} + Q_n.$$

The remainder term Q_n is bounded by $\|a\|^2\max_{1\le i\le n}|a^{\mathrm{T}}\widehat{\Omega}_i|\sum_{i=1}^n\|\widehat{\Omega}_i\|^2 = O_p(n^{-1})o_p(1)O_p(n) = o_p(1)$. Using (A.9) and (A.10), we have

$$-2\log\{\mathcal{R}_n(\beta)\} = \left(n^{-1/2}\sum_{i=1}^n\widehat{\Omega}_i^{\mathrm{T}}\right)\left(n^{-1}\sum_{i=1}^n\widehat{\Omega}_i\widehat{\Omega}_i^{\mathrm{T}}\right)^{-1} \times\left(n^{-1/2}\sum_{i=1}^n\widehat{\Omega}_i\right) + o_p(1).$$

Recall $\widehat{W}_i = \mathbf{e}_i(\mathbf{I}-\mathbf{S}_{12})\mathbf{W}$, where $\mathbf{e}_i$ is a $n\times 1$ vector of zero except the ith entry being 1, and $\mathbf{S}_{12} = \{\mathbf{I}-(\mathbf{I}-\mathbf{S}_1^c\mathbf{S}_2^c)^{-1}(\mathbf{I}-\mathbf{S}_1^c)\} + \{\mathbf{I}-(\mathbf{I}-\mathbf{S}_2^c\mathbf{S}_1^c)^{-1}(\mathbf{I}-\mathbf{S}_2^c)\}$. Using an argument similar to the proof of Theorem 1 of [7], and recalling the definition of $\widetilde{W}_i$ given in Section 2, we obtain that

$$\widehat{W}_i - \widetilde{W}_i = \mathbf{e}_i(\mathbf{S}_1^c+\mathbf{S}_2^c)\mathbf{W} - E(W_i|Z_{i1}) - E(W_i|Z_{i2}) + o_p(1/n).$$

Lemma 1 implies that

$$\left.\begin{array}{l}\max_i|\widehat{W}_i-\widetilde{W}_i| = o(n^{-1/4}),\\ \max_i|\widehat{Y}_i-\widetilde{Y}_i| = o(n^{-1/4}).\end{array}\right\} \tag{A.11}$$

Write $\widetilde{\Omega}_i = \widetilde{W}_i(\widetilde{Y}_i - \widetilde{W}_i^{\mathrm{T}}\beta) + \Sigma_{uu}\beta$. Then $\widetilde{\Omega}_i - \Omega_i$ can be expressed as

$$\begin{aligned}&(\widehat{W}_i-\widetilde{W}_i)\{(\widehat{Y}_i-\widehat{W}_i^{\mathrm{T}}\beta)-(\widetilde{Y}_i-\widetilde{W}_i^{\mathrm{T}}\beta)\}\\ &\quad+(\widehat{W}_i-\widetilde{W}_i)(\widetilde{Y}_i-\widetilde{W}_i^{\mathrm{T}}\beta)\\ &\quad-\widetilde{W}_i\{(\widehat{Y}_i-\widehat{W}_i^{\mathrm{T}}\beta)-(\widetilde{Y}_i-\widetilde{W}_i^{\mathrm{T}}\beta)\}.\end{aligned}$$

It follows from (A.11) that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n(\widehat{W}_i-\widetilde{W}_i)\{(\widehat{Y}_i-\widehat{W}_i^{\mathrm{T}}\beta)-(\widetilde{Y}_i-\widetilde{W}_i^{\mathrm{T}}\beta)\} = o_p(1).$$

On the other hand, Lemma 2 implies that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n(\widehat{W}_i-\widetilde{W}_i)(\widetilde{Y}_i-\widetilde{W}_i^{\mathrm{T}}\beta) = o_p(1)$$

because $E(\widetilde{Y}_i-\widetilde{W}_i^{\mathrm{T}}\beta) = 0$, and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n\widetilde{W}_i\{(\widehat{Y}_i-\widehat{W}_i^{\mathrm{T}}\beta)-(\widetilde{Y}_i-\widetilde{W}_i^{\mathrm{T}}\beta)\} = o_p(1)$$

because $E(\widetilde{W}_i) = 0$. These arguments imply that $n^{-1/2}\sum_{i=1}^n\widehat{\Omega}_i$ and $n^{-1/2}\sum_{i=1}^n\widetilde{\Omega}_i$ asymptotically have the same normal distribution, and $n^{-1}\sum_{i=1}^n\widehat{\Omega}_i\widehat{\Omega}_i^{\mathrm{T}}$ and $n^{-1}\sum_{i=1}^n\widetilde{\Omega}_i\widetilde{\Omega}_i^{\mathrm{T}}$ have the same limiting value. The proof is thus complete.

Received 25 August 2008

REFERENCES

[1] Härdle, W., Liang, H., and Gao, J. (2000). *Partially Linear Models*. Springer Physica-Verlag, Heidelberg. MR1787637

[2] Stone, C. J. (1985). Additive regression and other nonparametric models. *Ann. Statist.* **13**, 689–705. MR0790566

[3] Opsomer, J. D. and Ruppert, D. (1997). Fitting a bivariate additive model by local polynomial regression. *Ann. Statist.* **25**, 186–211. MR1429922

[4] Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*, Vol. 43, London: Chapman and Hall. MR1082147

[5] Opsomer, J. D. and Ruppert, D. (1999). A root-n consistent backfitting estimator for semiparametric additive modeling. *J. Comp. Graph. Statist.* **8**, 715–732.

[6] Linton, O. B. and Nielsen, J. P. (1995). A kernel method of estimating structured nonparametric regression based on marginal integration. *Biometrika* **82**, 93–101. MR1332841
[7] Liang, H., Thurston, S., Ruppert, D., Apanasovich, T., and Hauser, R. (2008). Additive partial linear models with measurement errors. *Biometrika* **95**, 667–678.
[8] Hauser, R., Meeker, J. D., Duty, S., Silva, M. J., and Calafat, A. M. (2006). Altered semen quality in relation to urinary concentrations of phthalate monoester and oxidative metabolites. *Epidemiology* **17**, 682–691.
[9] Owen, A. B. (2001). *Empirical Likelihood*. London: Chapman and Hall/CRC.
[10] Carroll, R. J., Ruppert, D., Stefanski, L. A., and Crainiceanu, C. M. (2006). *Measurement Error in Nonlinear Models, 2nd ed.* Chapman and Hall, New York. MR2243417
[11] Owen, A. B. (1991). Empirical likelihood for linear models. *Ann. Statist.* **19**, 1725-1747. MR1135146
[12] Qin, J. (1994). Semi-empirical likelihood ratio confidence intervals for the difference of two sample means. *Ann. Statist.* **46**, 117–126. MR1272752
[13] Qin, J. and Lawless, J. (1994). Empirical likelihood and general estimating equations. *Ann. Statist.* **22**, 300–325. MR1272085
[14] Chen, S. X. (1993). On the accuracy of empirical likelihood confidence regions for linear regression model. *Ann. Inst. Statist. Math.* **45**, 621–637. MR1252944
[15] Chen, S. X. (1994). Empirical likelihood confidence intervals for linear regression coefficients. *J. Mult. Anal.* **49**, 24–40. MR1275041
[16] Shi, J. and Lau, T. S. (2000). Empirical likelihood for partially linear models. *J. Mult. Anal.* **72**, 132–148. MR1747427
[17] Qin, G. S. and Jing, B. Y. (2001). Censored partial linear models and empirical likelihood. *J. Mult. Anal.* **78**, 37–61. MR1856265
[18] Wang, Q. H. and Jing, B. Y. (2003). Empirical likelihood for partial linear models. *Ann. Inst. Statist. Math.* **55**, 585–595. MR2007800
[19] Zheng, M. and Li, S. H. (2005). Empirical likelihood in partial linear error-in-covariable model with censored data. *Commun. Statist.* **34**, 389–404. MR2163361
[20] Liang, H., Wang, S., and Carroll, R. J. (2007). Partially linear models with missing response variables and error-prone covariates. *Biometrika* **94**, 185–198. MR2307903
[21] Duty, S. M., Silva, M. J., Barr, D. B., Brock, J. W., Ryan, L., Chen, Z. Y., Herrick, R., Christiani, D. C., and Hauser, R. (2003). Phthalate exposure and human semen parameters. *Epidemiology* **14**, 269–277.
[22] Hauser, R., Meeker, J. D., Park, S., Silva, M. J., and Calafat, A. M. (2004). Temporal variability of urinary phthalate metabolite levels in men of reproductive age. *Environmental Health Perspectives* **112**, 1734–1739.
[23] Petrov, V. V. (1975). *Sums of Independent Random Variables*. Berlin: Springer-Verlag. MR0388499
[24] Hansen, B. (2008). Uniform convergence rates for kernel estimation with dependent data. *Economet. Theory* **24**, 726–748. MR2409261
[25] Liang, H., Härdle, W., and Carroll, R. J. (1999). Estimation in a semiparametric partially linear errors-in-variables model. *Ann. Statist.* **27**, 1519–1535. MR1742498

Hua Liang
Department of Biostatistics and Computational Biology
University of Rochester Medical Center
Rochester, New York 14642, U.S.A.
E-mail address: hliang@bst.rochester.edu

Haiyan Su
Department of Biostatistics and Computational Biology
University of Rochester Medical Center
Rochester, New York 14642, U.S.A.
E-mail address: Haiyan_Su@urmc.rochester.edu

Sally W. Thurston
Department of Biostatistics and Computational Biology
University of Rochester Medical Center
Rochester, New York 14642, U.S.A.
E-mail address: thurston@bst.rochester.edu

John D. Meeker
Department of Environmental Health Sciences
University of Michigan
Ann Arbor, Michigan 48109, U.S.A.
E-mail address: meekerj@umich.edu

Russ Hauser
Department of Environmental Health
Harvard School of Public Health
Boston, Massachusetts 02115, U.S.A.
E-mail address: rhauser@hohp.harvard.edu