Published in final edited form as:
J Stat Plan Inference. 2012 May 1; 142(5): 1047–1062. doi:10.1016/j.jspi.2011.11.003.

Accounting for Uncertainty in Heteroscedasticity in Nonlinear Regression

Changwon Lim[a], **Pranab K. Sen**[b,c], and **Shyamal D. Peddada**[a,*]

[a]Biostatistics Branch, NIEHS, NIH, 111 T. W. Alexander Dr, RTP, NC 27709

[b]Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, 338 Hanes Hall, CB#3260, Chapel Hill, NC 27599

[c]Department of Biostatistics, University of North Carolina at Chapel Hill, 3101 McGavran-Greenberg, CB#7420, Chapel Hill, NC 27599

Abstract

Toxicologists and pharmacologists often describe toxicity of a chemical using parameters of a nonlinear regression model. Thus estimation of parameters of a nonlinear regression model is an important problem. The estimates of the parameters and their uncertainty estimates depend upon the underlying error variance structure in the model. Typically, a priori the researcher would know if the error variances are homoscedastic (i.e., constant across dose) or if they are heteroscedastic (i.e., the variance is a function of dose). Motivated by this concern, in this article we introduce an estimation procedure based on preliminary test which selects an appropriate estimation procedure accounting for the underlying error variance structure. Since outliers and influential observations are common in toxicological data, the proposed methodology uses M-estimators. The asymptotic properties of the preliminary test estimator are investigated; in particular its asymptotic covariance matrix is derived. The performance of the proposed estimator is compared with several standard estimators using simulation studies. The proposed methodology is also illustrated using a data set obtained from the National Toxicology Program.

Keywords

Asymptotic normality; Dose-response study; Heteroscedasticity; Hill model; M-estimation procedure; Preliminary test estimation; Toxicology

1 Introduction

Often toxicologists are interested in investigating the dose-response relationship when animals are exposed to varying doses of a chemical. Usually a nonlinear regression model such as a Hill model is used to describe the relationship (Gaylor and Aylward, 2004; Sand et al., 2004; Crofton et al., 2007). There may be several problems when fitting nonlinear models. Among them, one important concern is the error variance structure. Depending upon various factors, including the bioassay, dose-spacing and the endpoint of interest etc., the variability in response may not be constant across dose groups (heteroscedasticity). The

*Corresponding author. peddada@niehs.nih.gov.

standard asymptotic confidence intervals and test procedures based on the ordinary least squares (OLS) methodology may not be robust to heteroscedasticity and consequently may produce inaccurate coverage probabilities and Type I error rates. On the other hand, the standard iterated weighted least squares (IWLS) based methodology may not be efficient when the variances are approximately equal across dose groups (homoscedasticity). However, in practice one generally does not know if the data are homoscedastic or heteroscedastic.

The problem of heteroscedasticity has been extensively discussed in the literature in a wide range of contexts involving linear and nonlinear models. For instance, Hoferkamp and Peddada (2002) considered the problem of heteroscedasticity in the context of groups of experiments, as in fertilizer trials, where the error variances are ordered. Cysneiros et al. (2007) derived a joint iterative process for estimating the location and dispersion parameters in heteroscedastic linear models with symmetrical errors. Guo and Koul (2008) developed asymptotic theory for long memory time series based on heteroscedastic linear models. Recently, the problem of heteroscedasticity has also been addressed in the context of semi-parametric partially linear models (Ma et al., 2006; You et al., 2007; Lu, 2009). A Bayesian method for testing for equality of regression parameters in a heteroscedastic linear model has also been considered in the literature (Moreno et al., 2005).

Several authors modeled error variance as a function of dose in dose-response models (cf. Davidian and Carroll, 1987). Wang and Zhou (2007) even developed a nonparametric test for checking the adequacy of a given variance function. Bellio et al. (2000) proposed the use of higher order likelihood based methods for inference in heteroscedastic nonlinear models with application to dose-response models in herbicide bioassays.

Although IWLS based methods perform well under heteroscedasticity, they may lose efficiency relative to other methods when the data are homoscedastic. To illustrate this, consider the simulated data presented in Fig. 1 which is based on homoscedastic errors. The data were generated using the Hill model (Hill, 1910), $y=\theta_0+\theta_1 x^{\theta_2}/(\theta_3^{\theta_2}+x^{\theta_2})+\varepsilon$, which is usually used to study *in vivo* concentration response relationships, where y is the response at dose x, θ_0 is the intercept parameter, θ_1 is the difference between the maximum effect of a drug (E_{max}) and the intercept, θ_2 is the slope parameter that reflects the steepness of the effect-concentration curve, and θ_3 is the sensitivity parameter, the drug concentration producing 50% of E_{max} (ED_{50}).

In Fig. 1, the point estimate (with standard error in parentheses) for the ED_{50} (whose true value is 120) based on the OLS estimator (OLSE) is 149.7 (31.90). The IWLS estimator using the sample variances, denoted as $IWLSE_N$, is 226.4 (65.37), and the IWLS estimator using a variance model, denoted as $IWLSE_M$, is 233.0 (64.77). This example illustrates that a method designed for heteroscedastic data may not perform well when the data are homoscedastic.

Because the performance of a method relies on whether the data are homoscedastic or heteroscedastic, it is important to develop an estimation procedure which is robust to whether the error variance is homoscedastic or heteroscedastic. To make the procedure robust to the structure of the error variance, a preliminary test estimation (PTE) based methodology is developed in this paper. PTE has been well studied in the literature in a variety of contexts (Judge and Bock, 1978). For instance, Sen (1986) studied the asymptotic distributional risks for the preliminary test version of a maximum likelihood estimator. Recently, Ahmed et al. (2007) investigated the asymptotic properties of a pretest semiparametric estimator under quadratic loss and examined its performance using asymptotic analysis of quadratic risk functions in a partially linear model. Hoque et al.

(2009) studied the performance of the PTE of the slope parameter of a simple linear regression model under a linex loss function and derived the risk function and the moment generating function of the PTE.

Outliers and influential observations are common in toxicological data. To make the proposed procedure robust against outliers, we use the principle of M-estimation. Thus in this paper the PTE is either the ordinary M-estimator (OME) or the weighted M-estimator (WME) depending upon the outcome of a preliminary test for heteroscedasticity.

The PTE methodology based on OME and WME is proposed in Section 2. Results of a sample of simulation studies are provided in Section 3 and the proposed methodology is illustrated using a data set from the National Toxicology Program (NTP) in Section 4. Proofs of the main results are provided in Appendix while the theorems needed for proving these main results are provided in the online supplementary material.

2 The proposed methodology

2.1 Weighted M-estimation

Let y_i denote an $n_i \times 1$ response vector corresponding to an $m \times 1$ vector of covariates x_i in the ith sample, $i = 1, 2, \ldots, k$. Let

$$y_i = f(x_i, \theta) + \sigma_i \varepsilon_i, \quad i = 1, \ldots, k,$$

denote the nonlinear regression model, where $f(x_i, \theta)$ is some pre-specified nonlinear function of a $p \times 1$ parameter vector $\theta = (\theta_1, \theta_2, \ldots, \theta_p)^{\mathrm{T}}$ and ε_i are independent $n_i \times 1$ vectors that are identically distributed as $N(0, I)$. The total sample size n is given by $\sum_{i=1}^{k} n_i$. The components of y_i are denoted by y_{ij}, $j = 1, 2, \ldots, n_i$.

It is assumed that $\sigma_i \equiv \sigma(z_i, \tau)$, where $\sigma(\cdot, \cdot)$ is a known function of a known $q \times 1$ covariate vector z_i and an unknown $q \times 1$ parameter vector τ. Such models are commonly used in practice. For example, in some studies it may be reasonable to assume that the error variance σ_i^2 is a linear function of dose. For more examples one may refer to Carroll and Ruppert (1988).

The definition of an M-estimator depends upon the Huber score function which is defined as follows. For a pre-specified positive constant k_0, the Huber score function $h(u)$ is given by:

$$h(u) = \begin{cases} u/\sqrt{2}, & \text{if} |u| < k_0 \\ \{k_0(|u| - k_0/2)\}^{1/2}, & \text{otherwise.} \end{cases}$$

Throughout this paper we took k_0 to be 1.5. Then the ordinary M-estimator for θ is obtained by solving the following minimization problem:

$$S_o(\theta) = \sum_{i,j} h^2(y_{ij} - f(x_i, \theta)).$$

If we take $h(u) = u$ then we obtain the classical OLSE. Note that the OME does not account for heteroscedasticity in the data. The estimating equations for solving the above optimization problem are given by (Sanhueza and Sen, 2001):

$$\sum_{i,j}\lambda_o(x_i, y_{ij}, \tilde{\theta}_n)=0,$$

where

$$\lambda_o(x_i, y_{ij}, \theta)=\psi(y_{ij} - f(x_i, \theta))f_\theta(x_i, \theta),$$

$f_\theta(x_i, \theta) = (\partial/\partial\theta)f(x_i, \theta)$, and $\psi(u) = (\partial/\partial u)h^2(u)$.

To deal with heteroscedasticity, one may define a weighted M-estimation procedure similar in spirit to the popular IWLSE methodology using a variance function. Thus the WME is obtained by solving the following minimization problem:

$$\begin{pmatrix} \widehat{\theta}_n \\ \widehat{\tau}_n \end{pmatrix}=Argmin\left[\sum_{i,j}\left\{h^2\left(\frac{y_{ij} - f(x_i, \theta)}{\sigma(z_i, \tau)}\right)+\log\sigma(z_i, \tau)\right\} :\theta \in \Re^p, \tau \in \Re^q\right],$$

where $\log \sigma(z_i, \tau)$ is added within the sum, which is analogous to maximum likelihood estimation when the errors are normally distributed. The above minimization problem can be solved using the following estimating equations:

$$\sum_{i,j}\lambda(x_i, y_{ij}, \widehat{\theta}_n, \widehat{\tau}_n)=0, \tag{1}$$

where

$$\lambda(x_i, y_{ij}, \theta, \tau)=\begin{pmatrix} k(z_i, \tau)\psi(\varepsilon_{ij})f_\theta(x_i, \theta) \\ k(z_i, \tau)\{\psi(\varepsilon_{ij})\varepsilon_{ij} - 1\}\sigma_\tau(z_i, \tau) \end{pmatrix}, \tag{2}$$

$\sigma_\tau(z_i, \tau) = (\partial/\partial\tau)\sigma(z_i, \tau)$, and $k(z_i, \tau) = 1/\sigma(z_i, \tau)$.

We now derive the asymptotic normality of the WME. It is important to note that all the asymptotic results discussed in this paper are valid as long as n, the total sample size, goes to infinity. Two types of asymptotics often discussed in the literature are (a) the n_i is finite (e.g., $n_i = 1$) for all i and the number of doses becomes large (c.f., Wu, 1986), and (b) n_i goes to infinity for $i = 1, \ldots, k$ (c.f., Peddada and Smith, 1997). Although our asymptotic results are valid under both situations, in the following theorem we provide the proofs for (a). The proof for (b) can be obtained similarly and hence omitted. Essentially, all the asymptotic results obtained in this paper are valid as long as the total sample size goes to infinity.

We require the following sets of regularity conditions concerning (A) the score function ψ, (B) the function f and (C) the function σ.

[A1] Let $\varepsilon = \{y - f(x, \theta)\}/\sigma(z, \tau)$. Then $\gamma_1 = E\{\psi(\varepsilon)\varepsilon\}(\neq 0)$; $E\psi'(\varepsilon) = \gamma_2 (\neq 0)$; $E\{\psi'(\varepsilon)\varepsilon^2\} = \gamma_3 (\neq 0)$; $E\psi^2(\varepsilon)=\sigma^2_{\psi 1}<\infty$;$\text{var}\{\psi(\varepsilon)\varepsilon\}=\sigma^2_{\psi 2}<\infty$.

[B1] **i.** $\lim_{n\to\infty} n^{-1}\Gamma_{1n}(\theta, \tau) = \Gamma_1(\theta, \tau)$, where

$$\Gamma_{1n}(\theta,\tau)=\gamma_2\sum_{i=1}^{n}k^2(z_i,\tau)f_\theta(x_i,\theta)f_\theta^{\mathrm{T}}(x_i,\theta).$$

ii. $\lim_{n\to\infty}n^{-1}\Gamma_{31n}(\theta,\tau)=\Gamma_{31}(\theta,\tau)$, where

$$\Gamma_{31n}(\theta,\tau)=\sigma_{\psi 1}^2\sum_{i=1}^{n}k^2(z_i,\tau)f_\theta(x_i,\theta)f_\theta^{\mathrm{T}}(x_i,\theta),$$

and $\Gamma_{31}(\theta,\tau)$ is a positive definite matrix.

iii. $\max_i\{k^2(z_i,\tau)f_\theta^{\mathrm{T}}(x_i,\theta)\Gamma_{31n}^{-1}(\theta,\tau)f_\theta(x_i,\theta)\}\to 0$, as $n\to\infty$

[C] **i.** $\lim_{n\to\infty}n^{-1}\Gamma_{2n}(\theta,\tau)=\Gamma_2(\theta,\tau)$, where

$$\Gamma_{2n}(\theta,\tau)=\sum_{i=1}^{n}\left\{\frac{2\gamma_1+\gamma_3-1}{\sigma^2(z_i,\tau)}\sigma_\tau(z_i,\tau)\sigma_\tau^{\mathrm{T}}(z_i,\tau)+\frac{1-\gamma_1}{\sigma(z_i,\tau)}\sum\nolimits_\tau(z_i,\tau)\right\},$$

and $\Sigma_\tau(z_i,\tau)=(\partial^2/\partial\tau\partial\tau^{\mathrm{T}})\sigma(z_i,\tau)$.

ii. $\lim_{n\to\infty}n^{-1}\Gamma_{32n}(\theta,\tau)=\Gamma_{32}(\theta,\tau)$, where

$$\Gamma_{32n}(\theta,\tau)=\sigma_{\psi 2}^2\sum_{i=1}^{n}k^2(z_i,\tau)\sigma_\tau(z_i,\tau)\sigma_\tau^{\mathrm{T}}(z_i,\tau),$$

and $\Gamma_{32}(\theta,\tau)$ is a positive definite matrix.

iii. $\max_i\{k^2(z_i,\tau)\sigma_\tau^{\mathrm{T}}(z_i,\tau)\Gamma_{32n}^{-1}(\theta,\tau)\sigma_\tau(z_i,\tau)\}\to 0$, as $n\to\infty$

The asymptotic normality of the WME is established in Theorem 1 under the above regularity conditions.

Theorem 1—Under the conditions [A1], [B1] and [C]; [S1] – [S9] in the supplementary material,

$$\widehat{\Gamma}^{-\frac{1}{2}}\sqrt{n}\begin{pmatrix}\widehat{\theta}_n-\theta\\ \widehat{\tau}_n-\tau-\nu_n(\theta,\tau)\end{pmatrix}\to N_{p+q}(0,I_{p+q}),\tag{3}$$

where

$$\nu_n(\theta,\tau)=\left(\tfrac{1}{n}\Gamma_{2n}(\theta,\tau)\right)^{-1}\tfrac{\gamma_1-1}{n}\sum_{i=1}^{n}k(z_i,\tau)\sigma_\tau(z_i,\tau),$$
$$\widehat{\Gamma}=\left(\tfrac{1}{n}\Gamma_{5n}(\widehat{\theta}_n,\widehat{\tau}_n)\right)^{-1}\left(\tfrac{1}{n}\Gamma_{3n}(\widehat{\theta}_n,\widehat{\tau}_n)\right)\left(\tfrac{1}{n}\Gamma_{5n}(\widehat{\theta}_n,\widehat{\tau}_n)\right)^{-1},$$
$$\Gamma_{3n}(\theta,\tau)=\begin{pmatrix}\Gamma_{31n}(\theta,\tau) & 0\\ 0 & \Gamma_{32n}(\theta,\tau)\end{pmatrix},$$

and

$$\Gamma_{5n}(\theta,\tau)=\begin{pmatrix}\Gamma_{1n}(\theta,\tau) & 0\\ 0 & \Gamma_{2n}(\theta,\tau)\end{pmatrix}.$$

Note that the above theorem also provides the asymptotic covariance matrix of the WME.

2.2 Preliminary test estimation

We now describe the PTE procedure in the context of dose-response studies. Based on our experience with dose-response studies in toxicology, it is reasonable to assume that log-variance is a linear function of dose. Thus we assume $\log \sigma_i = \tau_0 + \tau_1 x_i$. Then $\tau_1 = 0$ corresponds to homoscedasticity, while τ_1 not equal to 0 corresponds to heteroscedasticity. Without loss of generality, in this article we consider assays where the response increases with dose. Accordingly, under heteroscedasticity, we assume that potentially the variance increases with dose (i.e., $\tau_1 > 0$). Thus we determine if the data are heteroscedastic by testing the hypotheses H_0: $\tau_1 = 0$ vs. H_1: $\tau_1 > 0$. Depending upon the researcher's belief, one could test for $\tau_1 \neq 0$ or $\tau_1 < 0$. Let r_{ij} denote the residual based on the OME. Then under suitable conditions $\hat{\tau} = (Z^{\mathrm{T}}Z)^{-1}Z^{\mathrm{T}}u$ is asymptotically normally distributed (Sen et al., 2009), where $Z = (z_{11}, z_{12}, \ldots, z_{1n_1}, \ldots, z_{k,n_k})^{\mathrm{T}}$ is an $n \times 2$ matrix and $u = (u_{11}, u_{12}, \ldots, u_{1n_1}, \ldots, u_{k,n_k})^{\mathrm{T}}$ is an $n \times 1$ vector, with $z_{ij} = (1, x_i)^{\mathrm{T}}$ and $u_{ij} = \log |r_{ij}|$, $i = 1, \ldots, k$, $j = 1, \ldots, n_i$, $\sum_{i=1}^{k} n_i = n$. Hence we may test the above hypotheses using $T_n = \hat{\tau}_{1n}/\sqrt{\mathrm{var}(\hat{\tau}_{1n})}$, where $\hat{\tau}_{1n}$ is the least squares estimator of τ_1. One can use a variety of test statistics for testing for τ_1, but here we use a simple statistic which can be derived directly from residuals based on the OME.

Then, the PTE is defined as

$$\widehat{\theta}_n^{\mathrm{PT}} = \begin{cases} \tilde{\theta}_n & \text{if } T_n \leq t_{\alpha,n-2} \\ \tilde{\theta}_n & \text{if } T_n > t_{\alpha,n-2}, \end{cases} \tag{4}$$

where $t_{\alpha,n-2}$ is the critical value of the t-distribution with $n - 2$ degrees of freedom having probability $1 - \alpha$ and α is the significance level of the preliminary test.

In order to derive asymptotic results regarding PTE, we require the following sets of regularity conditions.

[A2] Let $\varepsilon = \{y - f(x, \theta)\}/\sigma(z, \tau)$. Then, $E\psi'(\sigma(z, \tau)\,\varepsilon) = \gamma_4\ (\neq 0)$, $E\psi^2(\sigma(z,\tau)\varepsilon) = \sigma_{\psi 3}^2 w_1(x) < \infty$ and $E\{\psi(\varepsilon)\psi(\sigma(z,\tau)\varepsilon)\} = \sigma_{\psi 4}^2 w_2(x) < \infty$.

[B2] **i.** $\lim_{n\to\infty} n^{-1}\Gamma_{4n}(\theta) = \Gamma_4(\theta)$, where

$$\Gamma_{4n}(\theta) = \gamma_4 \sum_{i=1}^{k} n_i f_\theta(x_i, \theta) f_\theta^{\mathrm{T}}(x_i, \theta).$$

ii. $\lim_{n\to\infty} n^{-1}\Gamma_{33n}(\theta) = \Gamma_{33}(\theta)$, where

$$\Gamma_{33n}(\theta) = \sigma_{\psi 3}^2 \sum_{i=1}^{k} n_i w_1(x_i) f_\theta(x_i, \theta) f_\theta^{\mathrm{T}}(x_i, \theta),$$

and $\Gamma_{33}(\theta)$ is a positive definite matrix.

iii. $\lim_{n\to\infty} n^{-1}\Gamma_{34n}(\theta, \tau) = \Gamma_{34}(\theta, \tau)$, where

$$\Gamma_{34n}(\theta,\tau)=\sigma_{\psi 4}^2\sum_{i=1}^{k}\frac{n_i w_2(x_i)}{\sigma_i}f_\theta(x_i,\theta)f_\theta^{\mathrm{T}}(x_i,\theta),$$

iv. $\lim_{n\to\infty} n^{-1}G_{2n}(\theta, \tau) = G_2(\theta, \tau)$, where

$$G_{2n}(\theta,\tau)=\begin{pmatrix}\Gamma_{31n}(\theta,\tau) & \Gamma_{34n}(\theta,\tau) & 0\\ \Gamma_{34n}(\theta,\tau) & \Gamma_{33n}(\theta) & 0\\ 0 & 0 & 2n^2\sum_{i=1}^{k}n_i w_{i12}^2\end{pmatrix}, \tag{5}$$

w_{i12} is the second element of $w_{i1} = (Z^{\mathrm{T}}Z)^{-1}z_{i1}$, and $G_2(\theta, \tau)$ is a positive definite matrix.

v. $\max_i ch_1\{G_1(x_i, \theta, \tau)\, G_{2n}(\theta, \tau))^{-1}\} \to 0$, as $n \to \infty$, where

$$G_1(x_i,\theta,\tau)=\begin{pmatrix}\sigma_{\psi 1}^2\sigma_i^{-2}H_i & \sigma_{\psi 4}^2 w_2(x_i)\sigma_i^{-1}H_i & 0\\ \sigma_{\psi 4}^2 w_2(x_i)\sigma_i^{-1}H_i & \sigma_{\psi 3}^2 w_1(x_i)H_i & 0\\ 0 & 0 & 2n^2 w_{i12}^2\end{pmatrix},$$

and $H_i = f_\theta(x_i, \theta)f_\theta^{\mathrm{T}}(x_i, \theta)$.

We begin by proving the asymptotic joint normality of OME and WME which is then used for deriving the asymptotic covariance matrix of the PTE. The proof, provided in the appendix, follows arguments similar to that of Theorem 1.

Theorem 2—Let $\beta = (\theta^{\mathrm{T}}, \theta^{\mathrm{T}}, \tau_1)^{\mathrm{T}}$ and $\widehat{\beta}_n=\left(\widehat{\theta}_n^{\mathrm{T}}, \tilde{\theta}_n^{\mathrm{T}}, \widehat{\tau}_{1n}\right)^{\mathrm{T}}$. Then, under the conditions [A1], [A2], [B1], [B2] and [C]; [S1] – [S9] in the supplementary material,

$$\sqrt{n}\left(\widehat{\beta}_n - \beta\right) \to N_{2p+1}(0, G(\theta,\tau))\ as\ n \to \infty, \tag{6}$$

where

$$G(\theta,\tau)=G_3^{-1}(\theta,\tau)G_2(\theta,\tau)G_3^{-1}(\theta,\tau),$$
$$G_3(\theta,\tau)=\begin{pmatrix}\Gamma_1(\theta,\tau) & 0 & 0\\ 0 & \Gamma_4(\theta) & 0\\ 0 & 0 & 2\end{pmatrix}.$$

From the above theorem we deduce the asymptotic covariance matrix of PTE in the following theorem.

Theorem 3—Under the conditions [A1], [A2], [B1], [B2] and [C]; [S1] – [S9] in the supplementary material,

$$
\begin{aligned}
&E\left[n\left(\widehat{\theta}_n^{\mathrm{PT}}-\theta\right)\left(\widehat{\theta}_n^{\mathrm{PT}}-\theta\right)^{\mathrm{T}}\right]\\
=&F_t\left(t_{\alpha,n-2}-\frac{\tau_1}{\sqrt{\mathrm{var}(\widehat{\tau}_{1n})}}\right)E\left[n\left(\tilde{\theta}_n-\theta\right)\left(\tilde{\theta}_n-\theta\right)^{\mathrm{T}}\right]\\
+&\left\{1-F_t\left(t_{\alpha,n-2}-\frac{\tau_1}{\sqrt{\mathrm{var}(\widehat{\tau}_{1n})}}\right)\right\}E\left[n\left(\widehat{\theta}_n-\theta\right)\left(\widehat{\theta}_n-\theta\right)^{\mathrm{T}}\right]\\
=&F_t\left(t_{\alpha,n-2}-\frac{\tau_1}{\sqrt{\mathrm{var}(\widehat{\tau}_{1n})}}\right)\left(\tfrac{1}{n}\Gamma_{4n}(\theta)\right)^{1}\left(\tfrac{1}{n}\Gamma_{33n}(\theta)\right)\left(\tfrac{1}{n}\Gamma_{4n}(\theta)\right)^{-1}\\
+&\left\{1-F_t\left(t_{\alpha,n-2}-\frac{\tau_1}{\sqrt{\mathrm{var}(\widehat{\tau}_{1n})}}\right)\right\}\left(\tfrac{1}{n}\Gamma_{1n}(\theta)\right)^{-1}\left(\tfrac{1}{n}\Gamma_{31n}(\theta,\tau)\right)\left(\tfrac{1}{n}\Gamma_{1n}(\theta)\right)^{-1},
\end{aligned}
\tag{7}
$$

where F_t is the cdf of the t-distribution with $n-2$ degrees of freedom.

As we see from the above theorem, the asymptotic covariance matrix of PTE is a weighted average of those of OME and WME. The weights are directly related to τ_1. In particular, the weight corresponding to the covariance of OME is monotonically decreasing in τ_1. It equals $1-\alpha$ when $\tau_1 = 0$.

3 Simulation studies

3.1 Study design

We simulated data using the following Hill model under three different error variance structures. In each case the errors are normally distributed with mean 0.

$$
y_{ij}=f(x_i,\theta)+\varepsilon_{ij}=\theta_0+\frac{\theta_1 x_i^{\theta_2}}{x_3^{\theta_2}+x_i^{\theta_2}}+\varepsilon_{ij},\qquad i=1,\ldots,8,\qquad j=1,\ldots,5. \tag{8}
$$

The values of x_i were set to be 0, 1, 3, 10, 30, 100, 400, 600, and $(\theta_0, \theta_1, \theta_2, \theta_3) = (1, 4, 1.\ 5, 120)$. In Data 1 the errors are homoscedastic with variance e^{-3}. In Data 2 the variances were chosen to follow the log-linear model in dose as in the previous section. Thus the generated data are heteroscedastic with variance $e^{-6+0.01x_i}$. Lastly, to evaluate the performance of the proposed method when the variance model is mis-specified, in Data 3 we generated data according to the variance model, $0.\ 01f^2(x_i;\theta)$. We also investigated the performance of the proposed methodology in the presence of outliers. Typically, in toxicological studies outliers are observed in the high dose group where the observed response may drop below the expected response because of deaths due to treatment toxicity. For this reason, we generated data with outliers in the two highest dose groups using a shifted normal error with mean centered at -3 rather than 0.

There are two parts to the simulation study. Firstly, for illustration purposes, we generated one data set of each type (i.e., Data 1, Data 2 and Data 3) and the parameters were estimated using OLSE, IWLSE_N, IWLSE_M, OME, WME, and PTE methods. We used 0.05 as the significance level for the preliminary test in the PTE methodology.

Secondly, using 1,000 simulation runs, we compared the performance of the estimators in terms of three standard criteria: (i) mean squared error (MSE) of individual parameters as well as total MSE, (ii) the coverage probabilities of the 95% confidence intervals (CIs) of individual parameters as well as the simultaneous confidence ellipsoid defined below, and (iii) the length of the 95% CI of individual parameters as well as the volume of 95% confidence ellipsoid for each estimator. The $100(1-\alpha)$% confidence ellipsoid centered at an estimator $\hat{\theta}$ is defined as

$$(\widehat{\theta}-\theta)^{\mathrm{T}}[\widehat{\mathrm{var}}(\widehat{\theta})]^{-1}(\widehat{\theta}-\theta) \leq pF_{p,n-p}(\alpha),$$

where p is the number of parameters and $\widehat{\mathrm{var}}(\widehat{\theta})$ is the appropriate variance estimator. Note that for simplicity we are using the critical values based on WME for the critical values for PTE because the exact critical values for PTE are not available. This is because the asymptotic distribution of PTE is not normal but a mixture of two normals.

3.2 Results

The results of various estimates (and their standard errors) for the three simulated data sets without outliers are summarized in Table 1. As described in the introduction, for homoscedastic data (Data 1), although the fitted curves using IWLSE_N, IWLSE_M, and WME methods may seem reasonable based on the data (Fig. 2(a)), their estimated values of θ_3 (ED_{50}) differed from the true value (120) much more than those of the other methods and also they had very large standard errors. However, PTE automatically selected OME and the standard errors of PTE were much less than those of WME. Similarly, as expected, the converse was true in the case of heteroscedastic data (Data 2 and 3).

Note that if the data are homoscedastic (Data 1), then the "correct" choice of estimator is OME (OLSE when there are no outliers), whereas for heteroscedastic data (Data 2), the "correct" choice is WME (IWLSE_M when there are no outliers). However, in a practical setting, for a given data set one does not know *a priori* whether the data are homoscedastic or heteroscedastic. In all three data sets, PTE automatically chose the "correct" estimation procedure (either OME or WME) while keeping the standard error nearly as small as that of the "correct" estimation procedure.

The results of various estimates (and their standard errors) for the three data sets with outliers are summarized in Table 2. Because there were outliers in the data, θ_1 ($\mathrm{E}_{\max}$) and θ_3 (ED_{50}) were severely underestimated using the least square estimators (true values of θ_1 and θ_3 are 4 and 120, respectively), while OME and WME were closer to the true values. Figure 3 also reflects this result. It is noted that because of the outliers the preliminary test rejected the null hypothesis of homoscedasticity and PTE selected WME, which was closer to the true value than OME, even though the data were homogeneous.

Table 3 shows the results of the 1,000 simulations using data without outliers. When data were generated from homoscedastic model (Data 1), as expected, the estimated mean squared errors of OLSE and OME were smaller than those of the other estimators (except θ_0, the intercept) and the estimated MSEs of PTE were slightly larger than those of OME and much smaller than those of IWLSE_N and IWLSE_M, especially for θ_3 (ED_{50}). Relative to OME, the loss in efficiency (in terms of total MSE) due to PTE was less than 0.1%. However, the PTE gained substantially relative to all the weighted estimators. For example, there was a 44% gain in efficiency relative to IWLSE_N, an 18% gain relative to IWLSE_M and a 22% gain in efficiency relative to WME. Furthermore, the coverage probability of PTE was closer to the nominal level (0.95) than that of IWLSE_N and IWLSE_M for θ_3 with similar length of CI.

In the case of heteroscedastic data (Table 3) we observe that OLSE and OME could potentially perform extremely poorly. This is because when there is a large variation in the higher dose groups the observed data may fail to "plateau" at higher dose groups. Consequently, estimators such as the OLSE and OME would tend to overestimate θ_1 ($\mathrm{E}_{\max}$) and θ_3 (ED_{50}). For this reason, for some random samples, the estimates of ED_{50} became

arbitrarily large. As a consequence the estimated MSEs of OLSE and OME became extremely large! However, as one would expect, estimators such as IWLSE and WME performed well for such data with smaller MSE, and approximately correct coverage probability. The PTE performed as well as IWLSE and WME in terms of MSE and coverage probability. It competed very well in terms of efficiency relative to the weighted estimators (in terms of total MSE). For example, in the case of Data 2, it was 23% more efficient than IWLSE_N, about 4% less efficient than IWLSE_M and was just as efficient as WME. We see similar relative efficiencies in the case of Data 3, but the striking result here is that PTE was almost 270% more efficient than IWLSE_N. The PTE performed well in attaining the true coverage probability (0.95) although it had slightly wider confidence region than the weighted estimators since for some samples it used the unweighted estimator, OME. As expected, PTE performed substantially better than the unweighted estimators such as OLSE and OME in terms of all criteria. The reduction in total MSE was substantial.

Table 4 shows the results of the 1,000 simulations using data with outliers. As expected, the least squares based methods performed very poorly in terms of MSE and coverage probability in the presence of outliers, while the M-estimation based methods performed much better. In some cases the coverage probability of CIs centered at the least squares estimators were substantially smaller than the nominal level. For example, in the case of Data 2, the coverage probability of CIs centered at IWLSE_M for parameter θ_1 was as low as 18%.

In all the cases, the gains in efficiency (in terms of total MSE) of WME and PTE relative to IWLSE_N and IWLSE_M was almost 100% whether the data were homoscedastic or heteroscedastic. Because the PTE was developed using the M-estimators (OME and WME), the PTE also performed much better than the least squares methods. It is also noted that the MSE of PTE was exactly same as that of WME because the preliminary test rejected the null hypothesis of homoscedasticity for all 1,000 data sets. As explained earlier, in the presence of outliers, the test for heteroscedasticity can potentially have a higher Type I error rate. In the present context that is not an undesirable feature.

Our simulation studies made a strong case for the use of the proposed methodology. The gains in terms of MSE were generally substantial.

As commonly understood, the performance of an estimator may be affected much by dose placement. To illustrate this point, we generated a data set from the Hill model with homoscedastic error. The true values of the parameters are $(\theta_0, \theta_1, \theta_2, \theta_3) = (2, 4, 2, 30)$ and the values of the dose are $x = 0, 1, 3, 10, 30, 100, 300$. Because all estimators considered in this paper are affected, the OLSE was used to fit the curve for simplicity. The fitted curve and the estimation result are presented in Fig. 4(a) and Table 5, respectively. Even though the fitted curve is visually reasonable based on the data, the estimate of the parameter θ_2 (slope), its standard error and the standard error of the estimate of θ_3 (ED_{50}) are extremely large. We then chose additional dose arbitrarily, that is, $x = 70$, to generate observations at the dose and estimate the parameters and their standard errors based on the new data set. The fitted curve and the estimation result are presented in Fig. 4(b) and Table 5, respectively. Now the estimate of θ_2, its standard error and the standard error of the estimate of θ_3 are all reasonably small. This illustration suggests the same argument presented in Lim et al. (2011) that “dose-spacing plays a major role when estimating parameters of nonlinear models, especially the ED_{50} and the slope parameters of a Hill model”.

4 Application to hexavalent chromium data

In this section the proposed PTE methodology is applied to a data set from a toxicological study that was designed to examine the relationship between concentrations of Hexavalent Chromium (CrVI), as sodium dichromate dihydrate, in drinking water and accumulation of total chromium in tissue for three species (rats, mice, and guinea pigs) (NTP, 2007).

As commonly done in toxicology, we use the Hill model (8) to describe the dose-response relationship. In our model, x denotes the dose (in mg/L), ranging from 0 to 300, and y denotes the total chromium concentration (in mg/L) in the tissues. The proposed methodology is illustrated using rat kidney and blood data sets where chromium concentration (y) is modeled. There were 7 dose levels and 4 observations at each dose level except $x = 0$ (3 observation). Thus, total sample size is 27. Based on each of the data sets the parameters (and their standard errors) are estimated using OLSE, IWLSE$_N$, IWLSE$_M$, OME, WME, and PTE methods. Estimates of parameters and their standard errors are summarized in Table 6. The corresponding fitted curves are plotted in Fig. 5.

Visually the scatter plot in Fig. 5(a) seems to suggest that there is some amount of heteroscedasticity in the data. The sample variances of kidney chromium for the seven dose groups were 0.0012, 0.0046, 0.0006, 0.054, 0.720, 0.213, and 6.507, respectively. Thus they ranged from about 0.0006 to 6.507, which indicates potential heteroscedasticity in the data. When the log-linear model for the absolute residual based on the OME was fitted against doses, the estimated value of τ_1 was 0.005 with a standard error of 0.002. The slope was significant at the 5% level ($p = 0.009$). Thus the data appears to be heteroscedastic.

The data in Fig. 5(a) appear to be heteroscedastic and hence it is not surprising that the point estimates and their standard errors are quite different between ordinary estimates (OLSE and OME) and weighted estimates (IWLSE$_N$, IWLSE$_M$ and WME) (see Table 6). Since the preliminary test rejected the null hypothesis, PTE is the same as WME and both estimators had similar standard errors.

On the other hand, Fig. 5(b) shows that the data might be homoscedastic. The sample variances of blood chromium were 0.0006, 0.0002, 0.0003, 0.0001, 0.0007, 0.0016, and 0.0013, respectively. Thus they ranged from about 0.0001 to 0.0016. The result of the preliminary test using the absolute residuals based on the OME revealed that the slope (τ_1) was not significantly greater than zero at the 5% level ($p = 0.531$). Thus the data appear homoscedastic.

Visually the fitted curves are almost identical except IWLSE$_N$. However, Table 6 shows that point estimates and their standard errors are not similar, although OLSE and OME are exactly the same. Again, point estimates of θ_3 (ED$_{50}$) and their standard errors using OLSE (OME) and WME are quite different from each other. For this data set, the preliminary test could not reject the null hypothesis, and hence PTE is same as OME, although standard error for OME is larger than that for WME.

5 Concluding remarks

In this paper, PTE based methodology has been developed for analyzing nonlinear models that are possibly subject to heteroscedastic variance structure. The methodology proposed here allows researchers to use estimation procedures that are robust to the error variance structure in nonlinear models. We demonstrated its utility using simulation studies and a real data set obtained by the NTP on chromium VI.

The proposed methodology depends on the model used for describing the heteroscedasticity. In our experience, a log-linear model for variance is plausible for data observed in these toxicological experiments because of the underlying sigmoidal shaped dose-response curve and variance being monotonic with mean response. Also, the log-linear model offers a simple interpretation of variability with fewer parameters. If, however, an experimenter is interested in using a different parametric model to describe the variance, then the proposed methodology can be modified easily. For example, Lim et al. (2011a) have studied WME using a different variance model, where the error standard deviation was assumed to be a nonlinear function of three unknown parameters.

According to their document “Benchmark Dose Technical Guidance Document” (USEPA, 2000), for independent continuous outcome variables, the EPA determines BMD using nonlinear least squares methodology. They typically estimate various parameters of the model, including BMD, using the ordinary least squares estimator (OLSE) for homoscedastic data and the weighted least squares estimator (WLSE) for heteroscedastic data. According to the above document, they determine whether the data are homoscedastic or heteroscedastic by visual judgment using a scatter plot of the data. After making decisions regarding the error variance on the basis of the observed data, the EPA either uses the OLSE or the WLSE. However, since the choice of WLSE and OLSE depends upon the observed data, the standard errors of the estimators of various parameters (including BMD) should account for the uncertainty in decision made regarding the error variance. This is not done by the EPA’s methodology, which is the point of our paper where we account for the uncertainty induced by the preliminary evaluation of data regarding the error variance. Hence we believe that the methodology developed in this paper can be extended to estimate the BMD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported, in part, by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences [Z01 ES101744-04]. We thank Drs Gregg Dinse and Paramita Saha as well as the editors and the referees for many important comments which helped improve the presentation of the manuscript.

References

Ahmed SE, Doksum KA, Hossain S, You J. Shrinkage, pretest and absolute penalty estimators in partially linear models. Aust N Z J Stat. 2007; 49:435–454.

Bellio R, Jensen JE, Seiden P. Applications of likelihood asymptotics for nonlinear regression in herbicide bioassays. Biometrics. 2000; 56:1204–1212. [PubMed: 11129480]

Carroll, RJ.; Ruppert, D. Transformation and Weighting in Regression. Chapman and Hall; New York: 1988.

Crofton KM, Paul KB, DeVito MJ, Hedge JM. Short-term in vivo exposure to the water contaminant triclosan: evidence for disruption of thyroxine. Environ Toxicol Pharmacol. 2007; 24:194–197. [PubMed: 21783810]

Cysneiros FJA, Paula GA, Galea M. Heteroscedastic symmetrical linear models. Statist Prob Lett. 2007; 77:1084–1090.

Davidian M, Carroll RJ. Variance function estimation. J Amer Statist Assoc. 1987; 82:1079–1091.

Gaylor DW, Aylward LL. An evaluation of benchmark dose methodology for non-cancer continuous-data health effects in animals due to exposures to dioxin (TCDD). Regul Toxicol Pharm. 2004; 40:9–17.

Guo H, Koul HL. Asymptotic inference in some heteroscedastic regression models with long memory design and errors. Ann Statist. 2008; 36:458–487.

Hill AV. The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. J Physiol. 1910; 40(Suppl):iv–vii.

Hoferkamp C, Peddada SD. Parameter estimation in linear models with heteroscedastic variances subject to order restrictions. J Multivariate Anal. 2002; 82:65–87.

Hoque Z, Khan S, Wesolowski J. Performance of preliminary test estimator under linex loss function. Commun Stat–Theor M. 2009; 38:252–261.

Judge, GG.; Bock, ME. The Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics. North-Holland, Amsterdam: 1978.

Lim C, Sen PK, Peddada SD. Statistical inference in nonlinear regression under heteroscedasticity. Sankhya B. 2011a; 72:202–218.

Lim C, Sen PK, Peddada SD. Asymptotic linearity of an M-estimator in heteroscedastic nonlinear regression models. 2011b submitted.

Lu X. Empirical likelihood for heteroscedastic partially linear models. J Multivariate Anal. 2009; 100:387–396.

Ma Y, Chiou JM, Wang N. Efficient semiparametric estimator for heteroscedastic partially linear models. Biometrika. 2006; 93:75–84.

Moreno E, Torres F, Casella G. Testing equality of regression coefficients in heteroscedastic normal regression models. J Statist Plann Inference. 2005; 131:117–134.

National Toxicology Program. Toxicity Report Series 72. U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, RTP; North Carolina, U.S.A: 2007. NTP Toxicity Studies of Sodium Dichromate Dihydrate (CAS No. 7789-12-0) Administered in Drinking Water to Male and Female F344/N Rats and B6C3F1 Mice and Male BALB/c and am3-C57BL/6 Mice; p. 1-G4.

Sand S, von Rosen D, Eriksson P, Fredriksson A, Viberg H, Victorin K, Filpsson AF. Dose-response modeling and benchmark calculations from spontaneous behavior data on mice neonatally exposed to 2,2′,4,4′,5-pentabromodiphenyl ether. Toxicol Sci. 2004; 81:491–501. [PubMed: 15254340]

Sanhueza, AI.; Sen, PK. M-methods in generalized nonlinear models. In: Charalambides, ChA; Koutras, MV.; Balakrishnan, N., editors. Probability and Statistical Models with Applications. Chapman & Hall/CRC; New York: 2001. p. 359-375.

Sen PK. On the asymptotic distributional risks of shrinkage and preliminary test versions of maximum likelihood estimators. Sankhya Ser A. 1986; 48:354–371.

Sen, PK.; Singer, JM.; Pedroso de Lima, AC. From Finite Samples to Asymptotic Methods in Statistics. Cambridge Univ. Press; New York: 2009.

Peddada SD, Smith T. Consistency of a class of variance estimators in linear models under heteroscedasticity. Sankhya Ser B. 1997; 59:1–10.

US Environmetal Protection Agency. Benchmark Dose Technical Guidance Document. 2000. EPA/630/R-00/001

Wang L, Zhou XH. Assessing the adequacy of variance function in heteroscedastic regression models. Biometrics. 2007; 63:1218–1225. [PubMed: 17484775]

Wu CFJ. Jackknife, bootstrap and other resampling methods in regression analysis. Ann Stat. 1986; 14:1261–1295.

You J, Chen G, Zhou Y. Statistical inference of partially linear regression models with heteroscedastic errors. J Multivariate Anal. 2007; 98:1539–1557.

Appendix A. Proofs of the main results

Proof of Theorem 1

The proof relies on the asymptotic linearity of the WME (Lim et al., 2011b) and the existence of a solution to (1) which yields a $\sqrt{n}$-consistent estimator of $(\theta^{\mathrm{T}}, \tau^{\mathrm{T}})^{\mathrm{T}}$ (Theorem 4 in the supplementary material).

From Theorem 5 in the supplementary material we have that:

$$\sqrt{n}\begin{pmatrix}\widehat{\theta}_n-\theta\\ \widehat{\tau}_n-\tau\end{pmatrix}=\left(\frac{1}{n}\Gamma_{5n}(\theta,\tau)\right)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\lambda(x_i,y_i,\theta,\tau)+o_p(1).$$

Then from Theorem 6 in the supplementary material and the Slutsky Theorem we have the expression in (3).

Proof of Theorem 2

From the estimating equation (1) for the WME, we let

$$\lambda_w(x_i,y_{ij},\theta)=\frac{1}{\sigma_i}\psi\left(\frac{y_{ij}-f(x_i,\theta)}{\sigma_i}\right)f_\theta(x_i,\theta).$$

Then, from (13) in Theorem 5 in the supplementary material, the following asymptotic representation of $\hat{\theta}_n$ is obtained:

$$\sqrt{n}(\widehat{\theta}_n-\theta)=\left(\frac{1}{n}\Gamma_{1n}(\theta,\tau)\right)^{-1}\frac{1}{\sqrt{n}}\sum_{i,j}\lambda_w(x_i,y_{ij},\theta)+o_p(1). \tag{9}$$

Now, similarly as in Theorem 4 and 5 in the supplementary material, the uniform asymptotic linearity on the OME can be shown as:

$$\sup_{\|t\|\le C}\left\|\frac{1}{\sqrt{n}}\sum_{i,j}\left\{\lambda_o(x_i,y_{ij},\theta+n^{-\frac{1}{2}}t)-\lambda_o(x_i,y_{ij},\theta)\right\}+\frac{1}{n}\Gamma_{4n}(\theta)t\right\|=o_p(1),$$

and hence the following asymptotic representation of $\tilde{\theta}_n$ is obtained:

$$\sqrt{n}(\widehat{\theta}_n-\theta)=\left(\frac{1}{n}\Gamma_{4n}(\theta)\right)^{-1}\frac{1}{\sqrt{n}}\sum_{i,j}\lambda_o(x_i,y_{ij},\theta,)+o_p(1). \tag{10}$$

Then, from (9), (10) and (16) in the supplementary material,

$$\sqrt{n}\left(\widehat{\beta}_n-\beta\right)=G_{3n}^{-1}(\theta,\tau)\frac{1}{\sqrt{n}}\sum_{i,j}\lambda^*(x_i,y_{ij},\theta)+o_p(1),$$

where

$$G_{3n}(\theta,\tau)=\begin{pmatrix}n^{-1}\Gamma_{1n}(\theta,\tau) & 0 & 0\\ 0 & n^{-1}\Gamma_{4n}(\theta) & 0\\ 0 & 0 & 2\end{pmatrix}.$$

Then from Theorem 7 in the supplementary material and the Slutsky Theorem the expression in (5) is shown.

Proof of Theorem 3

From (4), for arbitrary $x \in \Re^p$,

$$\begin{aligned} P\left\{\sqrt{n}\left(\widehat{\theta}_n^{\mathrm{PT}}-\theta\right)\le x\right\} &= P\left\{\sqrt{n}\left(\tilde{\theta}_n-\theta\right)\le x, T_n \le t_{\alpha,n-2}\right\} + P\left\{\sqrt{n}\left(\widehat{\theta}_n-\theta\right)\le x, T_n > t_{\alpha,n-2}\right\} \\ &= P\left\{\sqrt{n}\left(\tilde{\theta}_n-\theta\right)\le x\right\} P\{T_n \le t_{\alpha,n-2}\} + P\left\{\sqrt{n}\left(\widehat{\theta}_n-\theta\right)\le x\right\} P\{T_n > t_{\alpha,n-2}\} \\ &= F_t\left(t_{\alpha,n-2}-\frac{\tau_1}{\sqrt{\mathrm{var}(\widehat{\tau}_{1n})}}\right) P\left\{\sqrt{n}\left(\tilde{\theta}_n-\theta\right)\le x\right\} + \left\{1-F_t\left(t_{\alpha,n-2}-\frac{\tau_1}{\sqrt{\mathrm{var}(\widehat{\tau}_{1n})}}\right)\right\} P\left\{\sqrt{n}\left(\widehat{\theta}_n-\theta\right)\le x\right\}. \end{aligned}$$

The second equality above holds because of the asymptotic independence of $\tilde{\theta}_n$ and T_n as well as the asymptotic independence of $\hat{\theta}_n$ and T_n proved in Theorem 2. Then from (6),

$$\begin{aligned} &E\left[n\left(\widehat{\theta}_n^{\mathrm{PT}}-\theta\right)\left(\widehat{\theta}_n^{\mathrm{PT}}-\widehat{\theta}\right)^{\mathrm{T}}\right] \\ &= F_t\left(t_{\alpha,n-2}-\frac{\tau_1}{\sqrt{\mathrm{var}(\widehat{\tau}_{1n})}}\right) E\left[n\left(\tilde{\theta}_n-\theta\right)\left(\tilde{\theta}_n-\theta\right)^{\mathrm{T}}\right] \\ &+ \left\{1-F_t\left(t_{\alpha,n-2}-\frac{\tau_1}{\sqrt{\mathrm{var}(\widehat{\tau}_{1n})}}\right)\right\} E\left[n\left(\widehat{\theta}_n-\theta\right)\left(\widehat{\theta}_n-\theta\right)^{\mathrm{T}}\right], \end{aligned}$$

and since from (6),

$$E\left[n\left(\tilde{\theta}_n-\theta\right)\left(\tilde{\theta}_n-\theta\right)^{\mathrm{T}}\right] = \left(\frac{1}{n}\Gamma_{4n}(\theta)\right)^{-1}\left(\frac{1}{n}\Gamma_{33n}(\theta)\right)\left(\frac{1}{n}\Gamma_{4n}(\theta)\right)^{-1}$$

and

$$E\left[n\left(\widehat{\theta}_n-\theta\right)\left(\widehat{\theta}_n-\theta\right)^{\mathrm{T}}\right] = \left(\frac{1}{n}\Gamma_{1n}(\theta,\tau)\right)^{-1}\left(\frac{1}{n}\Gamma_{31n}(\theta,\tau)\right)\left(\frac{1}{n}\Gamma_{1n}(\theta,\tau)\right)^{-1},$$

the expression in (7) is proved.

Appendix B. Supplementary material

Supplementary material associated with this article can be found in the online version.

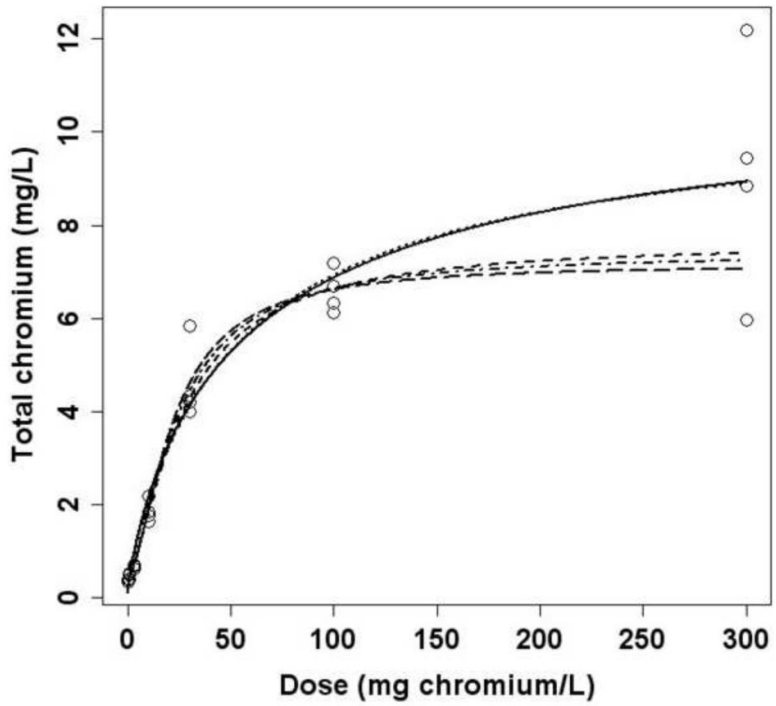
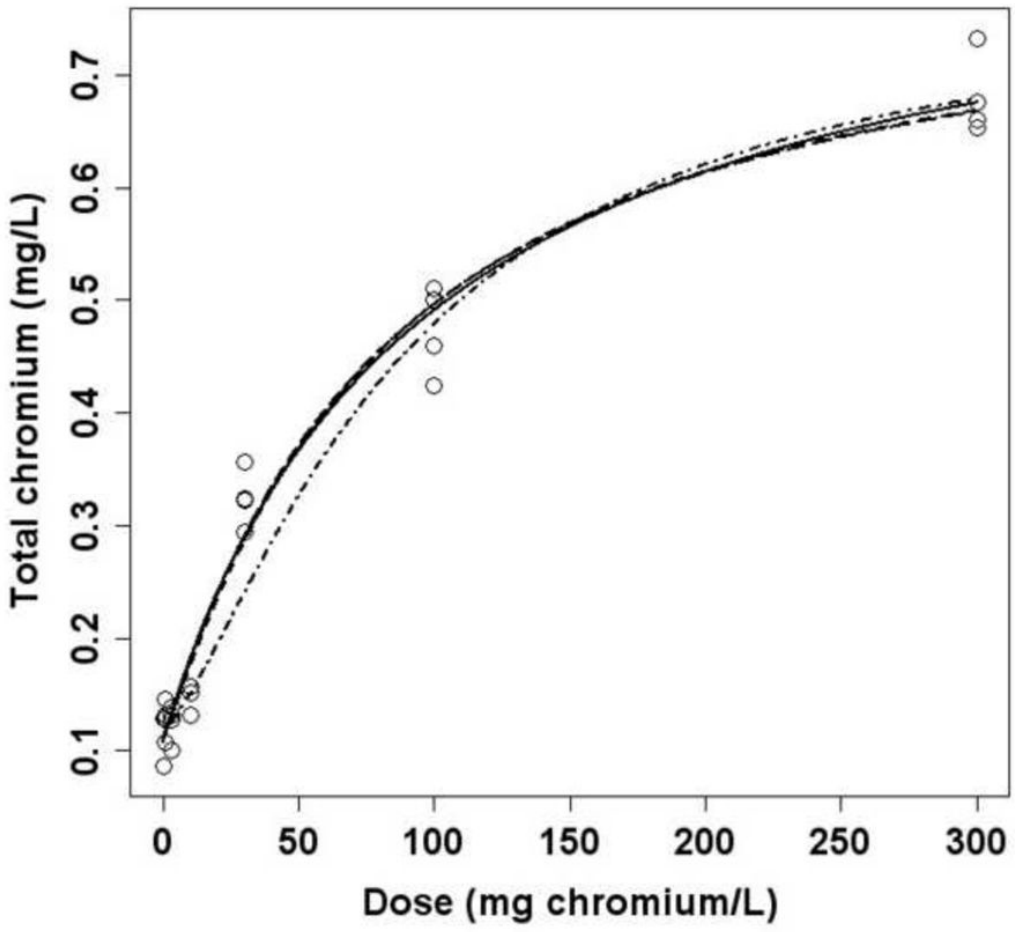
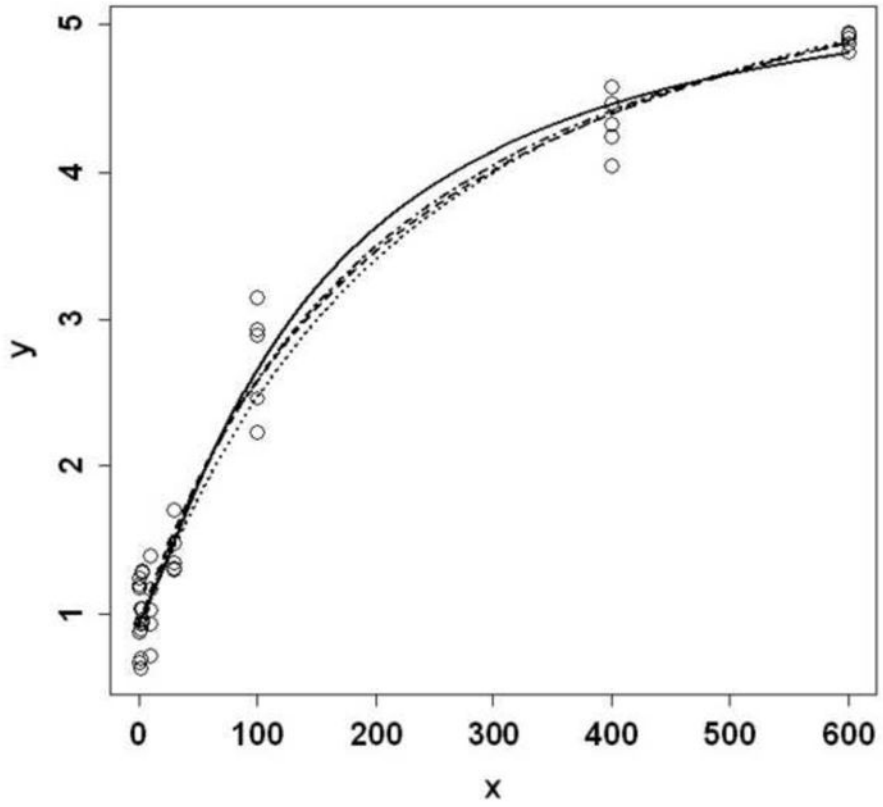


Figure 1.
Example of model predictions by OLSE (solid), IWLSE$_N$ (dot), IWLSE$_M$ (dashes) methods for homoscedastic data.





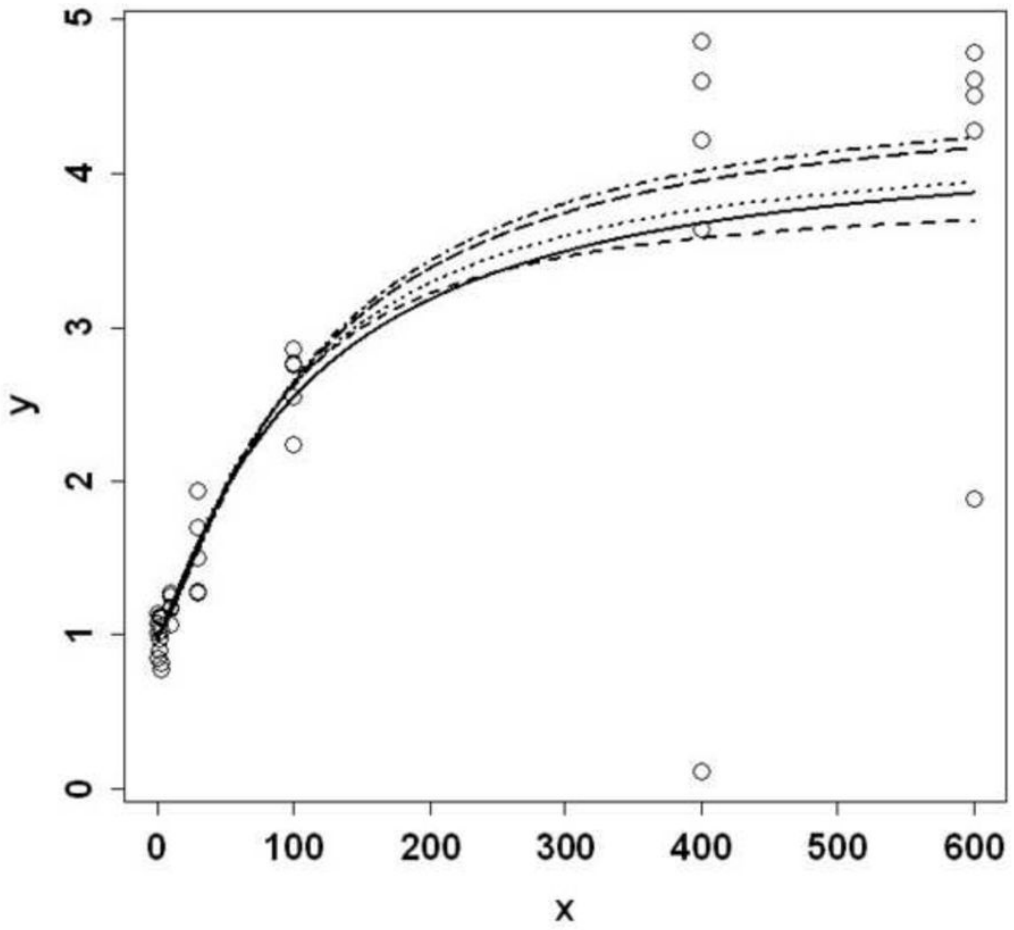
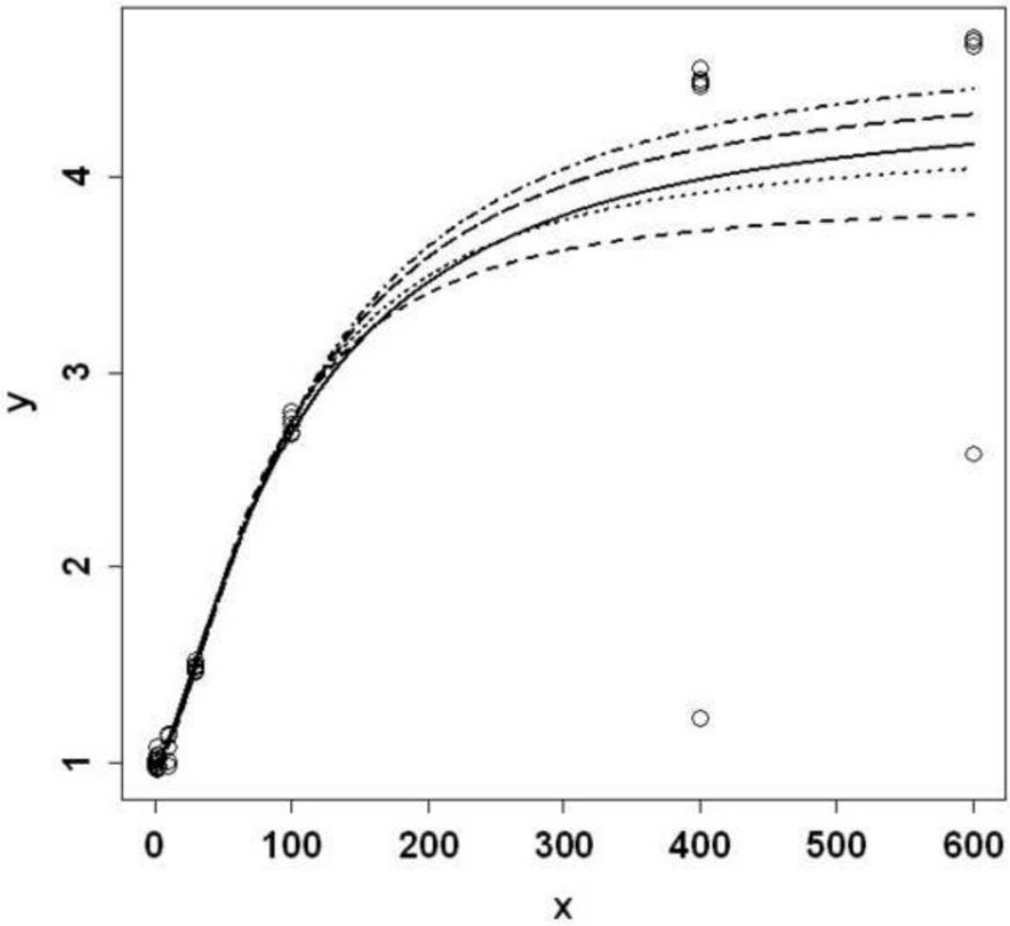
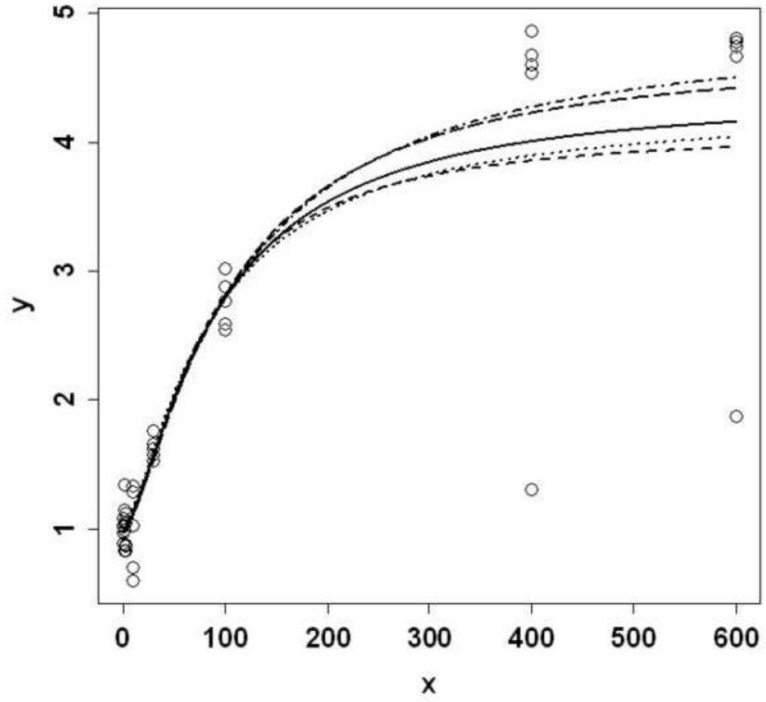


Figure 2.
Model predictions by OLSE (solid), $IWLSE_N$ (dot), $IWLSE_M$ (dashes), OME (long dashes), WME (dot-dash) methods when there are no outliers for: (a) homoscedastic data (Data 1), (b) heteroscedastic data (Data 2), and (c) mismodeled heteroscedastic data (Data 3).





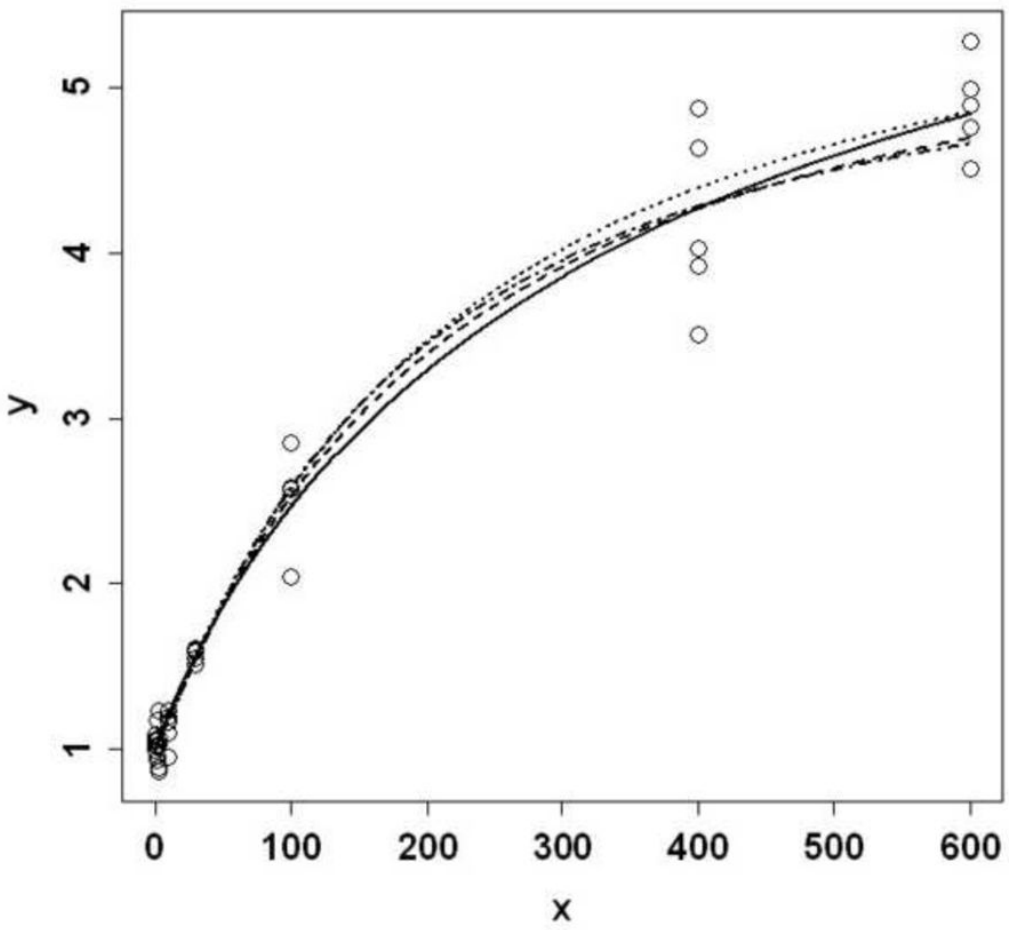


Figure 3.
Model predictions by OLSE (solid), IWLSE$_N$ (dot), IWLSE$_M$ (dashes), OME (long dashes), WME (dot-dash) methods when there are outliers for: (a) homoscedastic data (Data 1), (b) heteroscedastic data (Data 2), and (c) mismodeled heteroscedastic data (Data 3).

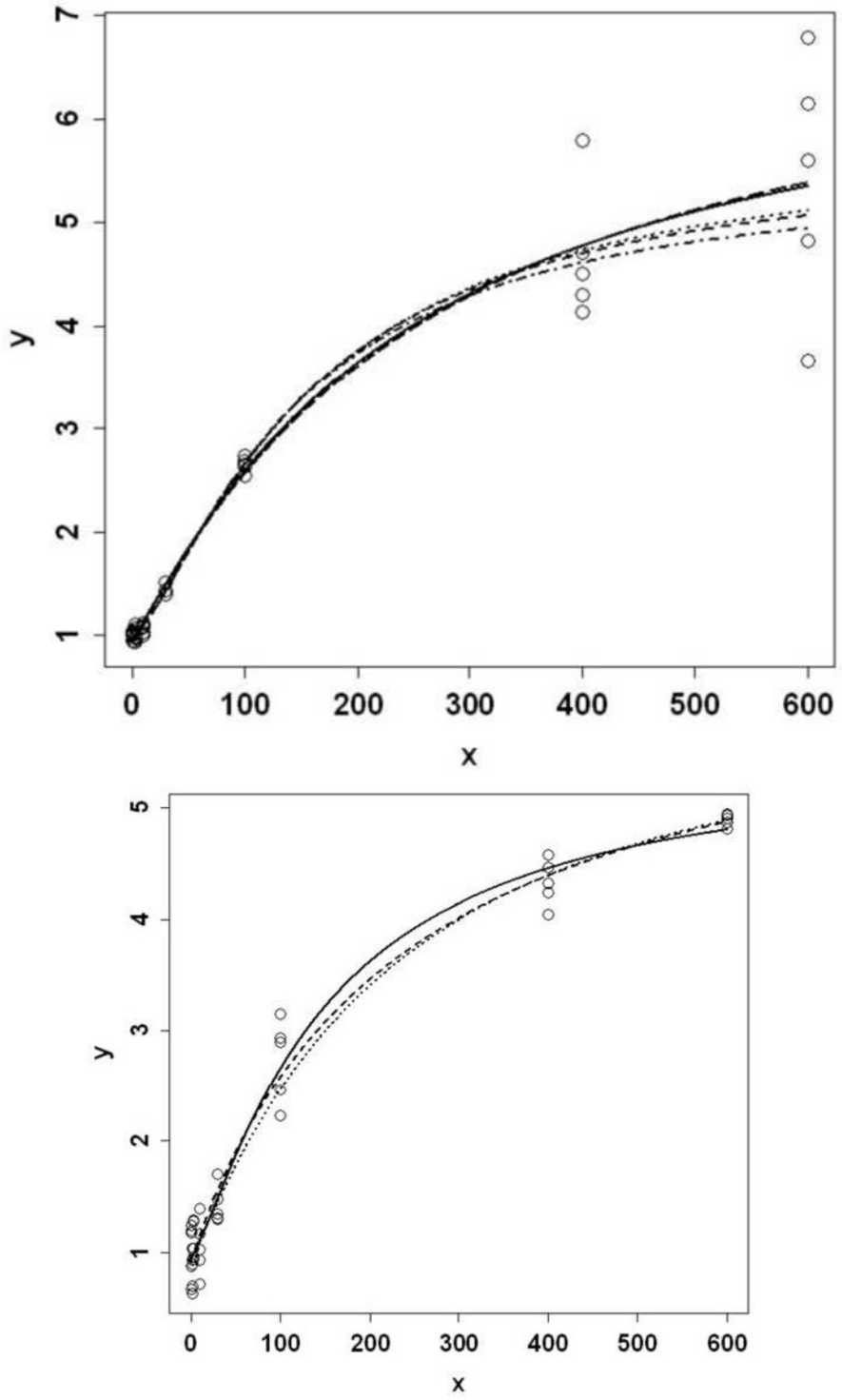
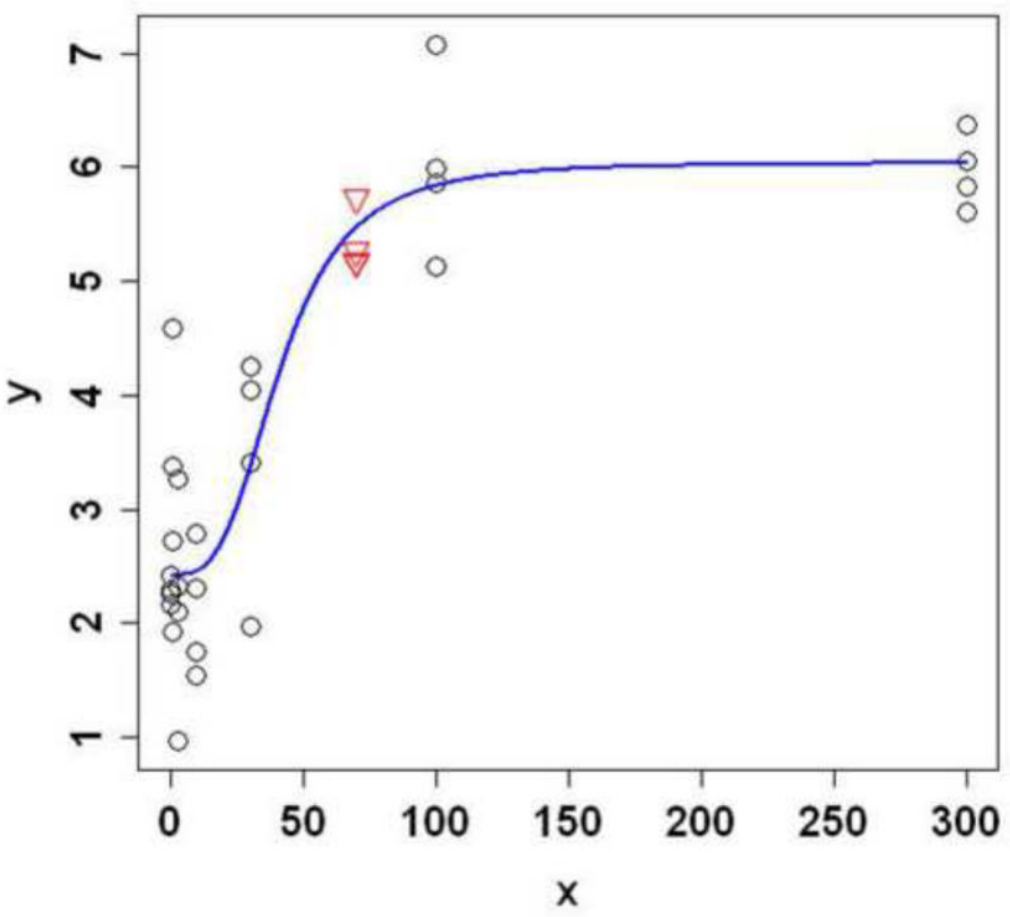


Figure 4.
Data generated from the Hill model and the fitted curve (a) without $x = 70$ and (b) with $x = 70$ using OLSE.



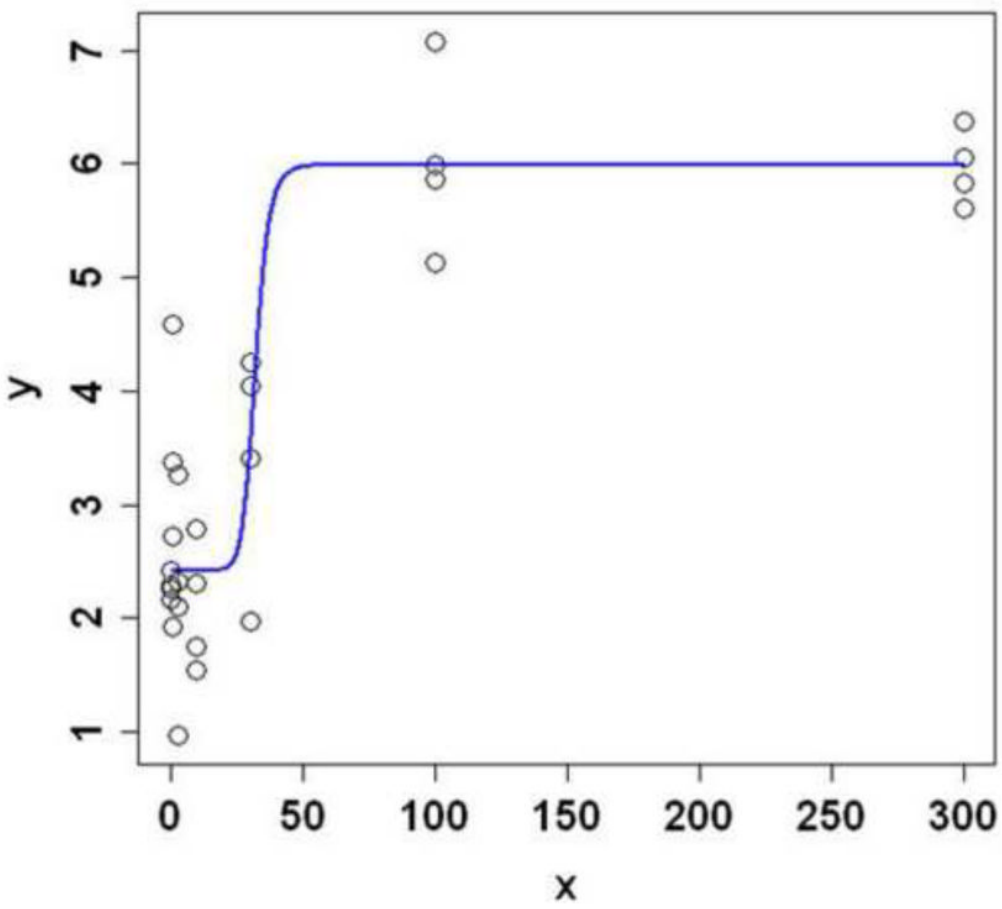


Figure 5.
Chromium concentration in (a) rat kidney and (b) rat blood using OLSE (solid), IWLSE$_N$ (dot-dash), IWLSE$_M$ (long dashes), OME (dot), WME (dashes) methods.

Table 1

Estimate and Standard Error for parameters of the models for Data 1, 2 and 3 without outliers using OLSE, IWLSE$_N$, IWLSE$_M$, OME, WME and PTE methods.

			Data 1		Data 2		Data 3	
PAR[a]	TRUE[b]	Method	EST[c]	SE[d]	EST[c]	SE[d]	EST[c]	SE[d]
θ_0	1	OLSE	0.954	0.062	0.964	0.126	0.975	0.075
		IWLSE$_N$	0.943	0.060	0.994	0.012	1.025	0.016
		IWLSE$_M$	0.919	0.077	0.992	0.013	0.995	0.031
		OME	0.954	0.062	0.965	0.124	0.975	0.075
		WME	0.920	0.079	0.993	0.013	0.997	0.033
		PTE	0.954	0.063	0.993	0.021	0.997	0.033
θ_1	4	OLSE	4.526	0.459	5.948	1.836	6.235	1.805
		IWLSE$_N$	5.291	0.639	4.791	0.530	5.065	0.695
		IWLSE$_M$	5.503	0.691	4.694	0.468	4.829	0.664
		OME	4.526	0.459	6.212	2.069	6.235	1.805
		WME	5.239	0.651	4.496	0.437	4.558	0.634
		PTE	4.526	0.471	4.496	0.513	4.558	0.634
θ_2	1.5	OLSE	1.252	0.207	1.121	0.354	0.922	0.187
		IWLSE$_N$	1.106	0.149	1.370	0.074	1.089	0.124
		IWLSE$_M$	0.991	0.143	1.385	0.082	1.092	0.125
		OME	1.252	0.207	1.093	0.339	0.922	0.187
		WME	1.049	0.163	1.402	0.083	1.143	0.141
		PTE	1.252	0.205	1.402	0.094	1.143	0.141
θ_3	120	OLSE	149.698	31.897	238.756	154.209	353.539	231.724
		IWLSE$_N$	226.409	65.366	159.133	23.177	213.067	74.478
		IWLSE$_M$	233.013	64.773	153.893	21.310	202.618	56.115
		OME	149.698	31.897	262.832	183.123	353.539	231.724
		WME	206.488	56.028	145.463	19.512	174.726	46.950
		PTE	149.698	33.518	145.463	31.118	174.726	47.010

[a] parameter;

[b] true value;

[c] estimate;

[d] standard error.

Table 2

Estimate and Standard Error for parameters of the models for Data 1, 2 and 3 with outliers using OLSE, $IWLSE_N$, $IWLSE_M$, OME, WME and PTE methods.

			Data 1		Data 2		Data 3	
PAR[a]	TRUE[b]	Method	EST[c]	SE[d]	EST[c]	SE[d]	EST[c]	SE[d]
θ_0	1	OLSE	0.977	0.181	0.980	0.156	0.976	0.226
		$IWLSE_N$	0.991	0.028	0.999	0.006	0.995	0.029
		$IWLSE_M$	0.984	0.041	1.000	0.008	0.993	0.039
		OME	0.979	0.121	0.995	0.107	0.978	0.132
		WME	0.984	0.043	0.997	0.008	0.994	0.041
		PTE	0.984	0.043	0.997	0.023	0.994	0.042
θ_1	4	OLSE	3.400	0.623	3.457	0.629	3.302	1.165
		$IWLSE_N$	3.263	0.636	3.189	0.429	3.266	0.705
		$IWLSE_M$	3.107	0.505	2.890	0.276	2.868	0.592
		OME	3.720	0.480	3.557	0.408	3.577	0.659
		WME	3.899	0.590	3.724	0.313	3.609	0.768
		PTE	3.899	0.590	3.724	0.318	3.609	0.768
θ_2	1.5	OLSE	1.409	0.668	1.381	0.600	1.150	0.751
		$IWLSE_N$	1.359	0.203	1.575	0.118	1.240	0.174
		$IWLSE_M$	1.538	0.254	1.695	0.113	1.398	0.254
		OME	1.407	0.444	1.497	0.453	1.250	0.460
		WME	1.311	0.196	1.476	0.068	1.300	0.219
		PTE	1.311	0.196	1.476	0.111	1.300	0.220
θ_3	120	OLSE	90.618	34.984	101.953	38.578	108.630	86.225
		$IWLSE_N$	85.918	26.908	89.029	15.975	99.952	36.816
		$IWLSE_M$	78.892	19.465	78.036	9.880	81.891	26.886
		OME	102.124	27.197	103.416	23.756	112.249	44.537
		WME	111.378	28.420	109.007	12.278	113.553	38.870
		PTE	111.378	28.420	109.007	12.906	113.553	38.889

[a] parameter;

[b] true value;

[c] estimate;

[d] standard error.

Table 3

Simulation results based on 1,000 replications generated from Data 1, 2 and 3 without outliers for OLSE, $IWLSE_N$, $IWLSE_M$, OME, WME and PTE.

		Data 1			Data 2			Data 3		
PAR[a]	Method	MSE[b]	COV[c]	LEN[d]	MSE[b]	COV[c]	LEN[d]	MSE[b]	COV[c]	LEN[d]
θ_0	OLSE	0.003	0.96	0.115	0.001	1.00	0.178	0.001	1.00	0.133
	$IWLSE_N$	0.004	0.85	0.098	<0.001	0.86	0.023	0.001	0.86	0.046
	$IWLSE_M$	0.003	0.95	0.112	<0.001	0.95	0.026	0.001	0.97	0.061
	OME	0.003	0.96	0.115	0.001	1.00	0.173	0.001	1.00	0.133
	WME	0.003	0.97	0.120	<0.001	0.96	0.028	0.001	0.98	0.067
	PTE	0.003	0.96	0.116	<0.001	0.98	0.033	0.001	0.98	0.068
θ_1	OLSE	0.093	0.96	0.594	4.620	0.82	2.013	0.530	0.93	0.868
	$IWLSE_N$	0.124	0.92	0.535	0.106	0.86	0.496	0.236	0.92	0.635
	$IWLSE_M$	0.106	0.94	0.571	0.080	0.93	0.547	0.144	0.93	0.632
	OME	0.093	0.96	0.594	5.110	0.82	2.104	0.530	0.93	0.869
	WME	0.108	0.95	0.603	0.084	0.95	0.583	0.143	0.95	0.687
	PTE	0.094	0.96	0.598	0.084	0.95	0.699	0.144	0.95	0.689
θ_2	OLSE	0.060	0.95	0.507	0.106	0.92	0.787	0.079	0.94	0.584
	$IWLSE_N$	0.090	0.90	0.522	0.007	0.91	0.148	0.038	0.90	0.338
	$IWLSE_M$	0.066	0.94	0.486	0.006	0.95	0.160	0.036	0.91	0.324
	OME	0.060	0.95	0.507	0.105	0.92	0.753	0.079	0.94	0.584
	WME	0.070	0.95	0.524	0.007	0.96	0.173	0.036	0.93	0.353
	PTE	0.061	0.95	0.509	0.007	0.97	0.190	0.036	0.93	0.356
θ_3	OLSE	444.373	0.94	37.924	>100000	0.88	340.976	8682.630	0.91	78.273
	$IWLSE_N$	639.520	0.91	35.310	180.871	0.89	21.195	2595.800	0.91	46.745
	$IWLSE_M$	525.002	0.91	36.538	141.608	0.94	22.932	668.802	0.87	34.180
	OME	444.186	0.94	37.921	>100000	0.88	376.064	8686.730	0.91	78.311
	WME	543.100	0.93	38.781	146.986	0.96	24.508	691.734	0.88	37.388
	PTE	444.536	0.94	38.148	146.986	0.96	60.318	701.087	0.88	37.856
Total	OLSE	444.529	0.89	0.529	>100000	0.73	19.005	8683.240	0.79	1.711
	$IWLSE_N$	639.738	0.67	0.283	180.985	0.73	0.015	2596.080	0.69	0.233

		Data 1			Data 2			Data 3		
PAR[a]	Method	MSE[b]	COV[c]	LEN[d]	MSE[b]	COV[c]	LEN[d]	MSE[b]	COV[c]	LEN[d]
	$IWLSE_M$	525.178	0.91	0.642	141.695	0.95	0.034	668.983	0.85	0.342
	OME	444.342	0.89	0.529	>100000	0.72	21.852	8687.340	0.79	1.712
	WME	543.282	0.92	0.801	147.076	0.97	0.046	691.913	0.89	0.487
	PTE	444.693	0.95	0.788	147.076	0.99	1.028	701.268	0.90	0.539

[a] parameter;

[b] mean squared error;

[c] coverage probability of 95% confidence interval;

[d] length of 95% confidence interval.

J Stat Plan Inference. Author manuscript; available in PMC 2013 May 1.

Table 4

Simulation results based on 1,000 replications generated from Data 1, 2 and 3 with outliers for OSLE, IWLSE$_N$, IWLSE$_M$, OME, WME and PTE.

		Data 1			Data 2			Data 3		
PAR[a]	Method	MSE[b]	COV[c]	LEN[d]	MSE[b]	COV[c]	LEN[d]	MSE[b]	COV[c]	LEN[d]
θ_0	OLSE	0.003	1.00	0.348	0.001	1.00	0.342	0.001	1.00	0.354
	IWLSE$_N$	0.004	0.83	0.098	<0.001	0.84	0.022	0.001	0.83	0.046
	IWLSE$_M$	0.003	0.93	0.107	<0.001	0.91	0.024	0.001	0.96	0.058
	OME	0.003	1.00	0.247	0.001	1.00	0.217	0.001	1.00	0.251
	WME	0.004	0.93	0.113	<0.001	0.91	0.024	0.001	0.96	0.061
	PTE	0.004	0.93	0.113	<0.001	1.00	0.045	0.001	0.97	0.064
θ_1	OLSE	0.674	0.70	1.165	0.799	0.77	1.548	0.716	0.66	1.288
	IWLSE$_N$	0.729	0.82	1.176	0.712	0.85	0.995	0.761	0.77	1.132
	IWLSE$_M$	0.836	0.61	1.078	0.808	0.18	0.704	0.901	0.54	0.988
	OME	0.300	0.81	0.982	0.437	1.00	1.098	0.386	0.76	1.092
	WME	0.275	0.93	1.105	0.412	0.99	0.819	0.349	0.89	1.089
	PTE	0.275	0.93	1.106	0.412	0.99	0.845	0.349	0.89	1.092
θ_2	OLSE	0.097	1.00	1.662	0.063	0.99	1.522	0.089	1.00	1.666
	IWLSE$_N$	0.140	0.94	0.619	0.040	0.96	0.276	0.087	0.95	0.484
	IWLSE$_M$	0.138	0.98	0.707	0.035	0.88	0.248	0.120	0.93	0.574
	OME	0.082	1.00	1.140	0.032	0.95	0.924	0.074	1.00	1.146
	WME	0.100	0.98	0.661	0.009	0.97	0.188	0.057	0.95	0.422
	PTE	0.100	0.98	0.663	0.009	0.98	0.248	0.057	0.95	0.432
θ_3	OLSE	898.739	0.93	71.196	3958.830	1.00	153.474	1648.200	0.86	89.490
	IWLSE$_N$	1018.610	0.73	47.587	1066.280	0.83	36.768	1064.130	0.72	45.639
	IWLSE$_M$	1147.850	0.63	47.397	1053.200	0.23	25.832	1236.080	0.52	38.738
	OME	492.203	0.92	61.199	3389.390	1.00	99.998	1476.520	0.86	76.690
	WME	514.383	0.91	54.042	680.140	0.99	31.998	618.571	0.82	45.283
	PTE	514.383	0.91	54.106	680.140	0.99	41.431	618.571	0.82	46.197
Total	OLSE	899.513	0.55	33.603	3959.700	0.53	49.346	1649.010	0.55	41.796
	IWLSE$_N$	1019.490	0.62	1.512	1067.030	0.54	0.017	1064.970	0.61	0.507

		Data 1			Data 2			Data 3		
PAR[a]	**Method**	**MSE**[b]	**COV**[c]	**LEN**[d]	**MSE**[b]	**COV**[c]	**LEN**[d]	**MSE**[b]	**COV**[c]	**LEN**[d]
	$IWLSE_M$	1148.830	0.90	3.822	1054.040	0.61	0.038	1237.100	0.71	0.726
	OME	492.588	0.71	9.007	3389.860	0.84	7.649	1476.980	0.64	11.117
	WME	514.762	0.96	4.111	680.561	0.96	0.054	618.978	0.87	0.910
	PTE	514.762	0.97	4.252	680.561	0.99	1.050	618.978	0.92	1.103

[a] parameter;

[b] mean squared error;

[c] coverage probability of 95% confidence interval;

[d] length of 95% confidence interval.

J Stat Plan Inference. Author manuscript; available in PMC 2013 May 1.

Table 5

Estimate and Standard Error for parameters of the models for data generated from the Hill model with/without $x = 70$ using OLSE method.

		Without $x = 70$		With $x = 70$	
PAR[a]	TRUE[b]	EST[c]	SE[d]	EST[c]	SE[d]
θ_0	2	2.425	0.212	2.426	0.196
θ_1	4	3.565	0.466	3.613	0.420
θ_2	2	12.406	135165	3.228	1.328
θ_3	30	32.388	27020	41.663	7.760

[a] parameter;

[b] true value;

[c] estimate;

[d] standard error.

Table 6

Estimate and Standard Error for parameters of the models for chromium rat kidney and blood data using OLSE, $IWLSE_N$, $IWLSE_M$, OME, WME and PTE methods.

		Kidney		Blood	
	Method	**EST**[a]	**SE**[b]	**EST**[a]	**SE**[b]
θ_0	OLSE	0.112	0.487	0.108	0.013
	$IWLSE_N$	0.380	0.015	0.120	0.006
	$IWLSE_M$	0.415	0.040	0.109	0.011
	OME	0.187	0.311	0.108	0.013
	WME	0.386	0.056	0.111	0.012
	PTE	0.386	0.058	0.108	0.013
θ_1	OLSE	10.992	2.703	0.752	0.112
	$IWLSE_N$	7.020	0.390	0.661	0.055
	$IWLSE_M$	6.749	0.312	0.699	0.096
	OME	10.402	1.502	0.752	0.112
	WME	7.253	0.454	0.689	0.090
	PTE	7.253	0.461	0.752	0.111
θ_2	OLSE	0.850	0.277	0.987	0.171
	$IWLSE_N$	1.497	0.087	1.393	0.160
	$IWLSE_M$	1.619	0.094	1.055	0.168
	OME	0.931	0.200	0.987	0.171
	WME	1.401	0.108	1.094	0.181
	PTE	1.401	0.108	0.987	0.172
θ_3	OLSE	57.482	35.508	95.972	31.847
	$IWLSE_N$	24.249	3.259	87.957	15.194
	$IWLSE_M$	22.222	1.518	81.371	23.632
	OME	51.665	17.612	95.972	31.847
	WME	27.115	2.756	80.302	21.925
	PTE	27.115	2.921	95.972	31.425

[a] estimate;

[b] standard error.