Supplemental Digital Content II

Title: Per- and Polyfluoroalkyl Substances in Drinking Water and Birthweight in the US: A County-level Study

Authors: Yachen Zhu[1], Scott M. Bartell[1,2]

Author affiliations:
1. Program in Public Health, University of California, Irvine, CA 92697-3957, USA
2. Department of Statistics, University of California, Irvine, CA 92697-1250, USA

Corresponding Author:
Yachen Zhu
Program in Public Health
Anteater Instruction and Research Building
Irvine, CA 92697-3957
USA
Email: yachenz1@uci.edu

Using county-level multiple-stratified average birthweights (grouped data) in weighted regression models produces equivalent results to those that would be obtained from using individual-level data in unweighted regression models.

Proof. Suppose there are n individuals, each with birthweight $Y_i,\ i = 1, ..., n$. There are p $(p \geq 1)$ variables in the regression model. For individual $i, X_{i1}, X_{i2}, ...,$ and X_{ip} are the values of the p explanatory variables. Suppose the n individuals are divided into m groups $(m < n)$. For group j $(j = 1, 2, ..., m)$, there are n_j members, $\sum_{j=1}^{m} n_j = n$. In each group, the individuals share the same vector of explanatory variables $\boldsymbol{X_i^T} = (1, X_{i1}, \cdots, X_{ip})$. To simplify the notation, let $\boldsymbol{Z_j^T} = (1, Z_{j1}, \cdots, Z_{jp})$ be the vector for group j's explanatory variables. Let $\boldsymbol{U^T} = (U_1, U_2, \cdots, U_m)$ denote the mean of the response variable for the m groups. $U_j = \frac{1}{n_j} \sum_{i=1}^{n_j} Y_i$ is the average birthweight for group j.

$$\mathbf{X} = \begin{pmatrix} 1 & X_{1,1} & \cdots & X_{1,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n_1,1} & \cdots & X_{n_1,p} \\ 1 & X_{n_1+1,1} & \cdots & X_{n_1+1,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n_1+n_2,1} & \cdots & X_{n_1+n_2,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1+\sum_{j=1}^{m-1} n_j,1} & \cdots & X_{1+\sum_{j=1}^{m-1} n_j,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n,1} & \cdots & X_{n,p} \end{pmatrix}_{n\times(1+p)} \triangleq \begin{pmatrix} 1 & Z_{1,1} & \cdots & Z_{1,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Z_{1,1} & \cdots & Z_{1,p} \\ 1 & Z_{2,1} & \cdots & Z_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Z_{2,1} & \cdots & Z_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Z_{m,1} & \cdots & Z_{m,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Z_{m,1} & \cdots & Z_{m,p} \end{pmatrix}_{n\times(1+p)}$$

Using individual-level data, we have

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon} = \begin{pmatrix} \boldsymbol{X_1^T} \\ \boldsymbol{X_2^T} \\ \vdots \\ \boldsymbol{X_n^T} \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

$$= \begin{pmatrix} 1 & X_{11} & \cdots & X_{1p} \\ 1 & X_{21} & \cdots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \cdots & X_{np} \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} \beta_0 + \beta_1 X_{11} + \cdots + \beta_p X_{1p} + \varepsilon_1 \\ \beta_0 + \beta_1 X_{21} + \cdots + \beta_p X_{2p} + \varepsilon_2 \\ \vdots \\ \beta_0 + \beta_1 X_{n1} + \cdots + \beta_p X_{np} + \varepsilon_n \end{pmatrix}$$

Using grouped data, we have

$$\mathbf{U} = \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_m \end{pmatrix} = \mathbf{Z} \cdot \tilde{\boldsymbol{\beta}} + \tilde{\boldsymbol{\varepsilon}} = \begin{pmatrix} \boldsymbol{Z_1^T} \\ \boldsymbol{Z_2^T} \\ \vdots \\ \boldsymbol{Z_m^T} \end{pmatrix} \cdot \begin{pmatrix} \tilde{\beta}_0 \\ \tilde{\beta}_1 \\ \vdots \\ \tilde{\beta}_p \end{pmatrix} + \begin{pmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \\ \vdots \\ \tilde{\varepsilon}_m \end{pmatrix}$$

$$= \begin{pmatrix} 1 & Z_{11} & \cdots & Z_{1p} \\ 1 & Z_{21} & \cdots & Z_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Z_{m1} & \cdots & Z_{mp} \end{pmatrix} \cdot \begin{pmatrix} \tilde{\beta}_0 \\ \tilde{\beta}_1 \\ \vdots \\ \tilde{\beta}_p \end{pmatrix} + \begin{pmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \\ \vdots \\ \tilde{\varepsilon}_m \end{pmatrix} = \begin{pmatrix} \tilde{\beta}_0 + \tilde{\beta}_1 Z_{11} + \cdots + \tilde{\beta}_p Z_{1p} + \tilde{\varepsilon}_1 \\ \tilde{\beta}_0 + \tilde{\beta}_1 Z_{21} + \cdots + \tilde{\beta}_p Z_{2p} + \tilde{\varepsilon}_2 \\ \vdots \\ \tilde{\beta}_0 + \tilde{\beta}_1 Z_{m1} + \cdots + \tilde{\beta}_p Z_{mp} + \tilde{\varepsilon}_m \end{pmatrix}$$

$\mathbf{Z} = \begin{pmatrix} 1 & Z_{11} & \cdots & Z_{1p} \\ 1 & Z_{21} & \cdots & Z_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Z_{n1} & \cdots & Z_{mp} \end{pmatrix}$ only includes the unique vectors in $\mathbf{X}$.

(1) Assuming constant variance:
Using individual-level data and according to the ordinary least squares

$$
\begin{aligned}
\hat{\boldsymbol{\beta}}_1 =& (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} \\
=& \left[\begin{pmatrix} 1 & 1 & \cdots & 1 \\ X_{11} & X_{21} & \cdots & X_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1p} & X_{2p} & \cdots & X_{np} \end{pmatrix} \cdot \begin{pmatrix} 1 & X_{11} & \cdots & X_{1p} \\ 1 & X_{21} & \cdots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \cdots & X_{np} \end{pmatrix}\right]^{-1} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ X_{11} & X_{21} & \cdots & X_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1p} & X_{2p} & \cdots & X_{np} \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \\
=& \begin{pmatrix} n & \sum_{i=1}^{n} X_{i1} & \cdots & \sum_{i=1}^{n} X_{ip} \\ \sum_{i=1}^{n} X_{i1} & \sum_{i=1}^{n} X_{i1}^2 & \cdots & \sum_{i=1}^{n} X_{i1}X_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} X_{ip} & \sum_{i=1}^{n} X_{ip}X_{i1} & \cdots & \sum_{i=1}^{n} X_{ip}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_{i1}Y_i \\ \vdots \\ \sum_{i=1}^{n} X_{ip}Y_i \end{pmatrix} \\
=& \begin{pmatrix} n & \sum_{j=1}^{m} n_jZ_{j1} & \cdots & \sum_{j=1}^{m} n_jZ_{jp} \\ \sum_{j=1}^{m} n_jZ_{j1} & \sum_{j=1}^{m} n_jZ_{j1}^2 & \cdots & \sum_{j=1}^{m} n_jZ_{j1}Z_{jp} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^{m} n_jZ_{jp} & \sum_{j=1}^{m} n_jZ_{jp}Z_{j1} & \cdots & \sum_{j=1}^{m} n_jZ_{jp}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{j=1}^{m} n_jU_j \\ \sum_{j=1}^{m} n_jZ_{j1}U_j \\ \vdots \\ \sum_{j=1}^{m} n_jZ_{jp}U_j \end{pmatrix}
\end{aligned}
$$

Using grouped data and according to the weighted least squares using the number of births in each stratum for the weights

$$
\begin{aligned}
\hat{\boldsymbol{\beta}}_2 =& (\mathbf{Z}^T\mathbf{W}_1\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{W}_1\mathbf{U} \\
=& \left[\begin{pmatrix} 1 & 1 & \cdots & 1 \\ Z_{11} & Z_{21} & \cdots & Z_{m1} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{1p} & Z_{2p} & \cdots & Z_{mp} \end{pmatrix} \cdot \begin{pmatrix} n_1 & & & \\ & n_2 & & \\ & & \ddots & \\ & & & n_m \end{pmatrix} \cdot \begin{pmatrix} 1 & Z_{11} & \cdots & Z_{1p} \\ 1 & Z_{21} & \cdots & Z_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Z_{n1} & \cdots & Z_{mp} \end{pmatrix}\right]^{-1} \\
& \cdot \begin{pmatrix} 1 & 1 & \cdots & 1 \\ Z_{11} & Z_{21} & \cdots & Z_{m1} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{1p} & Z_{2p} & \cdots & Z_{mp} \end{pmatrix} \begin{pmatrix} n_1 & & & \\ & n_2 & & \\ & & \ddots & \\ & & & n_m \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_m \end{pmatrix} \\
=& \begin{pmatrix} n & \sum_{j=1}^{m} n_jZ_{j1} & \cdots & \sum_{j=1}^{m} n_jZ_{jp} \\ \sum_{j=1}^{m} n_jZ_{j1} & \sum_{j=1}^{m} n_jZ_{j1}^2 & \cdots & \sum_{j=1}^{m} n_jZ_{j1}Z_{jp} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^{m} n_jZ_{jp} & \sum_{j=1}^{m} n_jZ_{jp}Z_{j1} & \cdots & \sum_{j=1}^{m} n_jZ_{jp}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{j=1}^{m} n_jU_j \\ \sum_{j=1}^{m} n_jZ_{j1}U_j \\ \vdots \\ \sum_{j=1}^{m} n_jZ_{jp}U_j \end{pmatrix}
\end{aligned}
$$

We have

$$\hat{\boldsymbol{\beta}}_1 = \hat{\boldsymbol{\beta}}_2$$

Therefore, using the number of births in each stratum for the weights produces equivalent results to those that would be obtained from using unweighted multiple linear regression with the individual-level data. In particular, $E(\hat{\boldsymbol{\beta}}_1) = E(\hat{\boldsymbol{\beta}}_2)$ and $Var(\hat{\boldsymbol{\beta}}_1) = Var(\hat{\boldsymbol{\beta}}_2)$, meaning that using county-level multiple-stratified average birthweights in weighted regression models produces estimates of parameters with the same mean and variance as the regression models using individual-level data.

(2) Assuming non-constant variance:

Using individual-level data in inverse-variance weighted regression (here we use the standard deviation of birthweight in the group to represent the standard deviation of birthweight for the individual). According to the weighted least squares (WLS), we have

$$
\begin{aligned}
\hat{\boldsymbol{\beta}}_3 =& (\mathbf{X}^T\mathbf{W}_2\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}_2\mathbf{Y} \\
=& \left[\begin{pmatrix} 1 & \cdots & 1 & \cdots & 1 & \cdots & 1 \\ Z_{11} & \cdots & Z_{11} & \cdots & Z_{m1} & \cdots & Z_{m1} \\ Z_{12} & \cdots & Z_{12} & \cdots & Z_{m2} & \cdots & Z_{m2} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ Z_{1p} & \cdots & Z_{1p} & \cdots & Z_{mp} & \cdots & Z_{mp} \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma_1^2} & & & & & & \\ & \ddots & & & & & \\ & & \frac{1}{\sigma_1^2} & & & & \\ & & & \ddots & & & \\ & & & & \frac{1}{\sigma_m^2} & & \\ & & & & & \ddots & \\ & & & & & & \frac{1}{\sigma_m^2} \end{pmatrix} \cdot \begin{pmatrix} 1 & Z_{1,1} & \cdots & Z_{1,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Z_{1,1} & \cdots & Z_{1,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Z_{m,1} & \cdots & Z_{m,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Z_{m,1} & \cdots & Z_{m,p} \end{pmatrix} \right]^{-1} \\
& \cdot \begin{pmatrix} 1 & \cdots & 1 & \cdots & 1 & \cdots & 1 \\ Z_{11} & \cdots & Z_{11} & \cdots & Z_{m1} & \cdots & Z_{m1} \\ Z_{12} & \cdots & Z_{12} & \cdots & Z_{m2} & \cdots & Z_{m2} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ Z_{1p} & \cdots & Z_{1p} & \cdots & Z_{mp} & \cdots & Z_{mp} \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma_1^2} & & & & & & \\ & \ddots & & & & & \\ & & \frac{1}{\sigma_1^2} & & & & \\ & & & \ddots & & & \\ & & & & \frac{1}{\sigma_m^2} & & \\ & & & & & \ddots & \\ & & & & & & \frac{1}{\sigma_m^2} \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \\
=& \begin{pmatrix} \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} & \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{j1} & \cdots & \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{jp} \\ \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{j1} & \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{j1}^2 & \cdots & \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{j1} Z_{jp} \\ \vdots & \vdots & \ddots & \vdots \\ \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{jp} & \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{jp} Z_{j1} & \cdots & \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{jp}^2 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{\sigma_1^2} & \cdots & \frac{1}{\sigma_1^2} & \cdots & \frac{1}{\sigma_m^2} & \cdots & \frac{1}{\sigma_m^2} \\ \frac{Z_{11}}{\sigma_1^2} & \cdots & \frac{Z_{11}}{\sigma_2^2} & \cdots & \frac{Z_{m1}}{\sigma_m^2} & \cdots & \frac{Z_{m1}}{\sigma_m^2} \\ \frac{Z_{12}}{\sigma_1^2} & \cdots & \frac{Z_{12}}{\sigma_2^2} & \cdots & \frac{Z_{m2}}{\sigma_m^2} & \cdots & \frac{Z_{m2}}{\sigma_m^2} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{Z_{1p}}{\sigma_1^2} & \cdots & \frac{Z_{1p}}{\sigma_2^2} & \cdots & \frac{Z_{mp}}{\sigma_m^2} & \cdots & \frac{Z_{mp}}{\sigma_m^2} \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \\
=& \begin{pmatrix} \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} & \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{j1} & \cdots & \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{jp} \\ \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{j1} & \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{j1}^2 & \cdots & \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{j1} Z_{jp} \\ \vdots & \vdots & \ddots & \vdots \\ \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{jp} & \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{jp} Z_{j1} & \cdots & \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{jp}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} U_j \\ \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{j1} U_j \\ \vdots \\ \sum\limits_{j=1}^{m} \frac{n_j}{\sigma_j^2} Z_{jp} U_j \end{pmatrix}
\end{aligned}
$$

Using grouped data in inverse-variance weighted regression,

$$
\tilde{\mathbf{W}}_2 = \begin{pmatrix} \frac{1}{Var(U_1)} & & & \\ & \frac{1}{Var(U_2)} & & \\ & & \ddots & \\ & & & \frac{1}{Var(U_m)} \end{pmatrix} = \begin{pmatrix} \frac{n_1}{\sigma_1^2} & & & \\ & \frac{n_2}{\sigma_2^2} & & \\ & & \ddots & \\ & & & \frac{n_m}{\sigma_m^2} \end{pmatrix} \text{ is the weight matrix,}
$$

because $Var(U_j) = Var(\frac{1}{n_j}\sum_{j=1}^{n_j} Y_j) = \frac{1}{n_j^2}\sum_{j=1}^{n_j} Var(Y_j) = \frac{\sigma_j^2}{n_j}$.

According to the weighted least squares (WLS), we have

$$\begin{aligned}
\hat{\boldsymbol{\beta}}_4 =&(\mathbf{Z}^T\tilde{\mathbf{W}}_2\mathbf{Z})^{-1}\mathbf{Z}^T\tilde{\mathbf{W}}_2\mathbf{U}\\
=&\left[\begin{pmatrix} 1 & 1 & \cdots & 1\\ Z_{11} & Z_{21} & \cdots & Z_{m1}\\ \vdots & \vdots & \ddots & \vdots\\ Z_{1p} & Z_{2p} & \cdots & Z_{mp}\end{pmatrix}\cdot\begin{pmatrix}\frac{n_1}{\sigma_1^2} & & & \\ & \frac{n_2}{\sigma_2^2} & & \\ & & \ddots & \\ & & & \frac{n_m}{\sigma_m^2}\end{pmatrix}\cdot\begin{pmatrix}1 & Z_{11} & \cdots & Z_{1p}\\ 1 & Z_{21} & \cdots & Z_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & Z_{n1} & \cdots & Z_{mp}\end{pmatrix}\right]^{-1}\\
&\cdot\begin{pmatrix} 1 & 1 & \cdots & 1\\ Z_{11} & Z_{21} & \cdots & Z_{m1}\\ \vdots & \vdots & \ddots & \vdots\\ Z_{1p} & Z_{2p} & \cdots & Z_{mp}\end{pmatrix}\begin{pmatrix}\frac{n_1}{\sigma_1^2} & & & \\ & \frac{n_2}{\sigma_2^2} & & \\ & & \ddots & \\ & & & \frac{n_m}{\sigma_m^2}\end{pmatrix}\begin{pmatrix}U_1\\ U_2\\ \vdots\\ U_m\end{pmatrix}\\
=&\begin{pmatrix}\sum\limits_{j=1}^m\frac{n_j}{\sigma_j^2} & \sum\limits_{j=1}^m\frac{n_j}{\sigma_j^2}Z_{j1} & \cdots & \sum\limits_{j=1}^m\frac{n_j}{\sigma_j^2}Z_{jp}\\ \sum\limits_{j=1}^m\frac{n_j}{\sigma_j^2}Z_{j1} & \sum\limits_{j=1}^m\frac{n_j}{\sigma_j^2}Z_{j1}^2 & \cdots & \sum\limits_{j=1}^m\frac{n_j}{\sigma_j^2}Z_{j1}Z_{jp}\\ \vdots & \vdots & \ddots & \vdots\\ \sum\limits_{j=1}^m\frac{n_j}{\sigma_j^2}Z_{jp} & \sum\limits_{j=1}^m\frac{n_j}{\sigma_j^2}Z_{jp}Z_{j1} & \cdots & \sum\limits_{j=1}^m\frac{n_j}{\sigma_j^2}Z_{jp}^2\end{pmatrix}^{-1}\begin{pmatrix}\sum\limits_{j=1}^m\frac{n_j}{\sigma_j^2}U_j\\ \sum\limits_{j=1}^m\frac{n_j}{\sigma_j^2}Z_{j1}U_j\\ \vdots\\ \sum\limits_{j=1}^m\frac{n_j}{\sigma_j^2}Z_{jp}U_j\end{pmatrix}
\end{aligned}$$

We have

$$\hat{\boldsymbol{\beta}}_4 = \hat{\boldsymbol{\beta}}_3$$

Therefore, using grouped data in inverse-variance weighted regression produces equivalent results to those that would be obtained from individual-level weighted regression allowing for heteroscedasticity. In particular, $E(\hat{\boldsymbol{\beta}}_3) = E(\hat{\boldsymbol{\beta}}_4)$ and $Var(\hat{\boldsymbol{\beta}}_3) = Var(\hat{\boldsymbol{\beta}}_4)$.

□