ISSN 0971-1554

JOURNAL OF QUANTITATIVE ECONOMICS

New Series Volume 10, No.1	***January, 2012***
Old Series Volume 26, No. 1	*January, 2010*

JOURNAL OF QUANTITATIVE ECONOMICS

(Journal of the Indian Econometric Society)

New Series Volume 10, No. 1	***January, 2012***
Old Series Volume 26, No. 1	*January, 2010*

Managing Editor: **D.M. Nachane** and *Joint Managing Editors:* **G. Mythili and Rupayan Pal**

CONTENTS

Journal of Quantitative Economics, Vol. 10, No. (1), January 2012

A NOTE ON ESTIMATION IN SEEMINGLY UNRELATED SEMI-PARAMETRIC REGRESSION MODELS[1]

RADHEY S. SINGH[2]
LICHUN WANG[3]

Abstract

In this paper a system of two seemingly unrelated semi-parametric regression models is considered, in which, following the partial residual procedure, we first show that the weighted least squares estimator (WLSE) of the regression parameter from the system can be expressed as a matrix series. Then this estimator is shown to be the limit of the covariance-adjusted estimator sequence of the regression parameter. Furthermore, based on the matrix series, we prove that the WLSE actually has only one unique simpler form, which exactly equals to the one-step covariance-adjusted estimator of the regression parameter. We also show that when the variance-covariance matrix of disturbances is unknown, the corresponding two-stage WLSE too has exactly one simpler form, and for any finite $k \geq 2$, the k-step covariance-adjusted estimator degenerates to the one-step covariance-adjusted estimator. Finally, we generalize our above conclusions to the system of $m(m \geq 3)$ seemingly unrelated semi-parametric regressions and point out that the conclusions presented in this paper include the system of $m(m \geq 2)$ seemingly unrelated linear regressions as its special case.

Keywords: Seemingly unrelated regressions; Semi-parametric models; Partial residual method; Covariance-adjusted estimator.

JEL Classifications: C13; C30

1. Introduction

The system of two seemingly unrelated (SU) regressions has been widely applied to many fields including economics, biological sciences, geography and nutritional epidemiology, etc. It is introduced to statistics by the pioneering works of Zellner (1962, 1963) and since then it has received considerable attention in the literature. In particular the contributions to the SU linear regression models by Kementa and Gilbert (1968), Revankar(1974), Swamy and Mehta (1975), Mehta and Swamy (1976), Schmidt (1977), Srivastava and Giles (1987), Kurata (1999), Ng (2002), Carroll et al. (2006), among others are notable. Also, recently Welsh and Yee (2006), Xu

[1] Partly supported by NSERC of Canada Grant (No. 400045) and Scientific Foundation of Beijing Jiaotong University (2007XM046).

[2] Department of Mathematics & Statistics University of Guelph, Guelph, Ontario, Canada, Email: rssingh@uoguelph.ca

[3] Department of Mathematics, Beijing Jiaotong University, Beijing 100044, P. R. China.

et al. (2008) and some other authors consider the case of SU nonparametric regression models and establish some interesting results.

In the past decades, there has been increasing interest and activity in the general area of semi-parametric regression model and many methods and techniques have been proposed and studied (see Chen (1988), Härdle et al. (2000) and among others). This is especially because it is much more flexible than the standard linear model since it combines both parametric and nonparametric components. They are often used in situations where the fully nonparametric model may not perform well or when the researcher wants to use a parametric model but the functional form with respect to a subset of the regressors or the density of the errors is not known.

Let us consider the following two SU semi-parametric regressions

$$\begin{cases} y_1 = X_1\beta_1 + g_1(t^{(1)}) + \varepsilon_1 \\ y_2 = X_2\beta_2 + g_2(t^{(2)}) + \varepsilon_2, \end{cases} \quad \text{...(1.1)}$$

Where $y_i (i = 1,2)$ are $n \times 1$ dependent variables, X_i are non-random explanatory variables that enter linearly with $\text{rank}(X_i) = p_i, \beta_i$ are $p_i \times 1$ vectors of unknown parameters, and $t^{(i)}$ are other non-random explanatory variables that enter in a nonlinear way, $g_i(\cdot)$ denote two unknown smooth functions of $t^{(i)}$ in R , and the variance-covariance matrices of ε_i are given by $\text{Cov}(\varepsilon_i, \varepsilon_j) = \sigma_{ij} I_n$, $i, j = 1,2,$ where $\Sigma^* = (\sigma_{ij})$ is a 2×2 non-diagonal positive definite matrix.

Following the partial residual procedure of Denby (1984), which extends the smoothing procedure of Clark (1977) to the case of multiple regression, for each model we define the nonparametric kernel estimate of $g_\iota(t^{(\iota)})$ $(\iota = 1,2)$ as

$$\bar{g}_\iota(t^{(\iota)}, \beta_\iota) = (\bar{g}_\iota(t_1^{(\iota)}, \beta_\iota), \cdots, \bar{g}_\iota(t_n^{(\iota)}, \beta_\iota))' = S_\iota(y_\iota - X_\iota\beta_\iota) \quad \text{...(1.2)}$$

With $S_\iota = S_\iota(\lambda_\iota) = (K_{\iota;j}^{\lambda_\iota}(t_i^{(\iota)}))_{n\times n}$, where $K_{\iota;j}^{\lambda_\iota}(\cdot; t_1^{(\iota)}, \cdots, t_n^{(\iota)})$ denote the kernel weight functions and $\lambda_\iota (\iota = 1,2)$ are the smoothing parameters. As made in Remark 2.2 below, our kernel weight function and the smoothing parameter are as suggested and adopted in Denby (1984), and therefore we do not discuss them here.

Hence, (1.1) combining with (1.2) formally yields the following SU linear regressions

$$\begin{cases} \tilde{y}_1 = \tilde{X}_1\beta_1 + \varepsilon_1 \\ \tilde{y}_2 = \tilde{X}_2\beta_2 + \varepsilon_2, \end{cases} \quad \text{...(1.3)}$$

where $\tilde{y}_i = (I_n - S_i)y_i$, $\tilde{X}_i = (I_n - S_i)X_i$ and $I_n - S_i$ are assumed to be inverse, and $\text{Cov}(\varepsilon_i, \varepsilon_j) = \sigma_{ij} I_n$.

Denote $y = (\tilde{y}_1', \tilde{y}_2')'$, $X = \text{diag}(\tilde{X}_1, \tilde{X}_2)$, $\beta = (\beta_1', \beta_2')'$, $\varepsilon = (\varepsilon_1', \varepsilon_2')'$ and $\Sigma_{ij} = \text{Cov}(\varepsilon_i, \varepsilon_j)$. Then the system (1.3) can be further expressed as follows

$$y = X\beta + \varepsilon, \quad E\varepsilon = 0, \quad \text{Cov}(\varepsilon) = \Sigma, \quad \text{...(1.4)}$$

where $\Sigma = (\Sigma_{ij}) = \Sigma^* \otimes I_n$ and $\otimes$ denotes the Kronecker product of matrices.

Obviously, if Σ is known, then by the generalized Gauss-Markov Theorem, the best WLSE of β from the system would be

$$\tilde{\beta} = \begin{pmatrix} \tilde{\beta}_1 \\ \tilde{\beta}_2 \end{pmatrix} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y$$
$$= \begin{pmatrix} Q^{11}\tilde{X}_1'\Sigma^{11} + Q^{12}\tilde{X}_2'\Sigma^{21} & Q^{11}\tilde{X}_1'\Sigma^{12} + Q^{12}\tilde{X}_2'\Sigma^{22} \\ Q^{21}\tilde{X}_1'\Sigma^{11} + Q^{22}\tilde{X}_2'\Sigma^{21} & Q^{21}\tilde{X}_1'\Sigma^{12} + Q^{22}\tilde{X}_2'\Sigma^{22} \end{pmatrix} \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{pmatrix}, \qquad ...(1.5)$$

where $\Sigma^{-1} \triangleq (\Sigma^{ij})_{2\times 2}$ and $(X'\Sigma^{-1}X)^{-1} \triangleq (Q^{ij})_{2\times 2}$.

In the following, we focus on estimating the regression parameter β_1; β_2 can be estimated accordingly.

First, note that $\Sigma^{-1} = (\Sigma^*)^{-1} \otimes I_n$, hence, we have

$$\Sigma^{11} = \frac{1}{\sigma_{11}} \cdot \frac{1}{1-\rho^2} I_n; \qquad \Sigma^{12} = -\frac{\sigma_{12}}{\sigma_{11}\sigma_{22}} \cdot \frac{1}{1-\rho^2} I_n;$$
$$\Sigma^{21} = -\frac{\sigma_{21}}{\sigma_{11}\sigma_{22}} \cdot \frac{1}{1-\rho^2} I_n; \quad \Sigma^{22} = \frac{1}{\sigma_{22}} \cdot \frac{1}{1-\rho^2} I_n. \qquad ...(1.6)$$

Secondly, by the inverse of partitioned matrix and matrix power series, we have

$$Q^{11} = \left[\tilde{X}_1'\Sigma^{11}\tilde{X}_1 - \tilde{X}_1'\Sigma^{12}\tilde{X}_2(\tilde{X}_2'\Sigma^{22}\tilde{X}_2)^{-1}\tilde{X}_2'\Sigma^{21}\tilde{X}_1\right]^{-1}$$
$$= \left[\frac{1}{\sigma_{11}(1-\rho^2)}\tilde{X}_1'\tilde{X}_1 - \frac{\rho^2}{\sigma_{11}(1-\rho^2)}\tilde{X}_1'\tilde{P}_2\tilde{X}_1\right]^{-1}$$
$$= \sigma_{11}(1-\rho^2)(\tilde{X}_1'\tilde{X}_1)^{-1}[I_{p1} - \rho^2\tilde{X}_1'\tilde{P}_2\tilde{X}_1(\tilde{X}_1'\tilde{X}_1)^{-1}]^{-1}$$
$$= \sigma_{11}(1-\rho^2)(\tilde{X}_1'\tilde{X}_1)^{-1}\sum_{i=0}^{\infty}[\rho^2\tilde{X}_1'\tilde{P}_2\tilde{X}_1(\tilde{X}_1'\tilde{X}_1)^{-1}]^i, \qquad ...(1.7)$$

and

$$Q^{12} = \sigma_{12}(1-\rho^2)(\tilde{X}_1'\tilde{X}_1)^{-1}\sum_{i=0}^{\infty}[\rho^2\tilde{X}_1'\tilde{P}_2\tilde{X}_1(\tilde{X}_1'\tilde{X}_1)^{-1}]^i\tilde{X}_1'\tilde{X}_2(\tilde{X}_2'\tilde{X}_2)^{-1}, \qquad ...(1.8)$$

where $\rho^2 = \frac{\sigma_{12}\sigma_{21}}{\sigma_{11}\sigma_{22}} = \frac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}}$ and $\tilde{P}_2 = \tilde{X}_2(\tilde{X}_2'\tilde{X}_2)^{-1}\tilde{X}_2'$.

Hence, we obtain

$$Q^{11}\tilde{X}_1'\Sigma^{11} + Q^{12}\tilde{X}_2'\Sigma^{21}$$
$$= (\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\left[I_n - \rho^2\sum_{i=0}^{\infty}(\rho^2\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1\right], \qquad ...(1.9)$$

and

$$
\begin{aligned}
&Q^{11}\tilde{X}_1'\Sigma^{12} + Q^{12}\tilde{X}_2'\Sigma^{22} \\
&= -\frac{\sigma_{12}}{\sigma_{22}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\sum_{i=0}^{\infty}(\rho^2\tilde{P}_2\tilde{P}_1)^i\tilde{N}_2 \\
&= (\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\left[I_n - \rho^2\sum_{i=0}^{\infty}(\rho^2\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1\right](-\frac{\sigma_{12}}{\sigma_{22}}\tilde{N}_2).
\end{aligned}
$$

where we use the fact that $\tilde{P}_2\tilde{P}_1\tilde{N}_2 = -\tilde{P}_2\tilde{N}_1\tilde{N}_2$ and $\tilde{P}_1 = \tilde{X}_1(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'$, $\tilde{N}_i = I_n - \tilde{P}_i$

Finally, note that $\tilde{\beta}_1 = (Q^{11}\tilde{X}_1'\Sigma^{11} + Q^{12}\tilde{X}_2'\Sigma^{21})\tilde{y}_1 + (Q^{11}\tilde{X}_1'\Sigma^{12} + Q^{12}\tilde{X}_2'\Sigma^{22})\tilde{y}_2$, we have the following lemma.

Lemma 1.1. In the model (1.3), if Σ is known, then

$$
\tilde{\beta}_1 = (\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\left[I_n - \rho^2\sum_{i=0}^{\infty}(\rho^2\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1\right](\tilde{y}_1 - \frac{\sigma_{12}}{\sigma_{22}}\tilde{N}_2\tilde{y}_2).
$$

Obviously, $\tilde{\beta}_1$ is superior to $\hat{\beta}_1 = (\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{y}_1$ which comes from the first equation of (1.3), and accordingly the partial residual estimator $\tilde{g}_1(t^{(1)},\tilde{\beta}_1)$ can improve the estimator $\hat{g}_1(t^{(1)},\hat{\beta}_1)$, where $\tilde{g}_1(t^{(1)},\tilde{\beta}_1)$ and $\hat{g}_1(t^{(1)},\hat{\beta}_1)$ equals to $\bar{g}_1(t^{(1)},\beta_1)$ with β_1 replaced $\tilde{\beta}$ and $\hat{\beta}_1$, respectively. However, although $\tilde{\beta}_1$ is the best estimator from the system, its form is a little complicated for use in practice, especially in the case when the variance-covariance Σ is not known. Naturally, given this, a practioner would like to know whether there exists a simpler form of $\tilde{\beta}_1$ which is also better than $\hat{\beta}_1$.

In Section 2, we employ a useful lemma of Rao (1967) to generate an iteration estimator sequence for β_1, which is called the covariance-adjusted estimator sequence, and prove that the limit of the sequence is exactly equal to $\tilde{\beta}_1$. Furthermore, we prove that actually $\tilde{\beta}_1$ has only one unique simpler form which exactly equals to the one-step covariance-adjusted estimator. Hence we conclude when the variance-covariance matrix of disturbances is unknown the corresponding two-stage WLSE of β_1 also has exactly one simpler form; and for any finite $k \geq 2$, the k-step covariance-adjusted estimator degenerates to the one-step covariance-adjusted estimator. In Section 3 we generalize the above results to the system of $m(m \geq 3)$ SU semi-parametric regressions. Finally, we conclude the paper with a few remarks in Section 4.

2. The Covariance-Adjusted Sequence of β_1

First, we state the following covariance-adjusted lemma.

Lemma 2.1. Let T_1 and T_2 be $k_1 \times 1$ and $k_2 \times 1$ statistics with $E[T_1] = \theta$ and $E[T_2] = 0$, where θ is an unknown parameter vector. Denote

$$\mathrm{Cov}\begin{pmatrix} T_1 \\ T_2 \end{pmatrix} = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}.$$

If $V_{12} \neq 0$, then there exists an unbiased estimator $\hat{\theta} = T_1 - V_{12}V_{22}^{-}T_2$ over a class of estimators $D = \{A_1T_1 + A_2T_2 \mid A_1, A_2$ are non-random matrices}, and

$$\mathrm{Cov}(\hat{\theta}) = V_{11} - V_{12}V_{22}^{-}V_{21} \leq V_{11} = \mathrm{Cov}(T_1),$$

where V_{22}^{-} is a generalized inverse of V_{22} and $A \geq B$, i.e., $A - B$ is a real positive semi-definite matrix.

Proof. It is a straightforward result of Rao (1967).

From Lemma 2.1, we note that, as two unbiased estimators of θ, $\hat{\theta}$ is more efficient than T_1 in the sense that $\hat{\theta}$ has less covariance.

To obtain the most efficient estimator of β_1 in the system (1.3), by Lemma 2.1 and noting that $E[\hat{\beta}_1] = \beta_1$ and $E[\tilde{N}_2\tilde{y}_2] = 0$, first we use $\tilde{N}_2\tilde{y}_2$ to improve $\hat{\beta}_1$ and obtain a more efficient estimator $\hat{\beta}_1(1)$ which has less covariance than that of $\hat{\beta}_1$. Then we adjust $\hat{\beta}_1(1)$ by $\tilde{N}_1\tilde{y}_1$ and get $\hat{\beta}_1(2)$ which has less covariance than that of $\hat{\beta}_1(1)$, where $E[\tilde{N}_1\tilde{y}_1] = 0$. Repeating this process, we obtain the following covariance-adjusted estimator sequence for β_1:

$$\begin{aligned} \hat{\beta}_1(2k-1) &= \hat{\beta}_1(2k-2) - \mathrm{Cov}(\hat{\beta}_1(2k-2), \tilde{N}_2\tilde{y}_2)[\mathrm{Cov}(\tilde{N}_2\tilde{y}_2)]^{-}\tilde{N}_2\tilde{y}_2, \\ \hat{\beta}_1(2k) &= \hat{\beta}_1(2k-1) - \mathrm{Cov}(\hat{\beta}_1(2k-1), \tilde{N}_1\tilde{y}_1)[\mathrm{Cov}(\tilde{N}_1\tilde{y}_1)]^{-}\tilde{N}_1\tilde{y}_1, \\ & k=1,2,..., \end{aligned} \qquad ...(2.1)$$

where $\hat{\beta}_1(0) = \hat{\beta}_1$.

Step by step, we have for $k \geq 1$

$$\hat{\beta}_1(2k-1) = (\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'(\tilde{P}_1 + \rho^2\tilde{N}_2\tilde{N}_1)^{k-1}(\tilde{y}_1 - \frac{\sigma_{12}}{\sigma_{22}}\tilde{N}_2\tilde{y}_2).$$

$$\begin{aligned} \hat{\beta}_1(2k) &= (\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'(\tilde{P}_1 + \rho^2\tilde{N}_2\tilde{N}_1)^{k}\tilde{y}_1 \\ &\quad - \frac{\sigma_{12}}{\sigma_{22}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'(\tilde{P}_1 + \rho^2\tilde{N}_2\tilde{N}_1)^{k-1}\tilde{N}_2\tilde{y}_2. \end{aligned} \qquad ...(2.2)$$

Thus, we have the following theorem.

Theorem 2.1. The limit of the adjusted-covariance sequence $\{\hat{\beta}_1(k), k \geq 1\}$ is $\tilde{\beta}_1$.

Proof. Note that any $k \geq 1$,

$$(\tilde{P}_1+\rho^2\tilde{N}_2\tilde{N}_1)^k=\sum_{\iota=0}^{k}\tilde{P}_1^{k-\iota}(\rho^2\tilde{N}_2\tilde{N}_1)^{\iota}=\sum_{\iota=0}^{k}\tilde{P}_1(\rho^2\tilde{N}_2\tilde{N}_1)^{\iota}. \qquad ...(2.3)$$

Hence, we have,

$$\begin{aligned}
&(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'(\tilde{P}_1+\rho^2\tilde{N}_2\tilde{N}_1)^k\\
=\;&(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'[I_n+\rho^2\tilde{N}_2\tilde{N}_1+\cdots+(\rho^2\tilde{N}_2\tilde{N}_1)^k]\\
=\;&(\tilde{X}_1'\tilde{X}_1)^{-1}[\tilde{X}_1'-\tilde{X}_1'\cdot\rho^2\tilde{P}_2\tilde{N}_1-\cdots-\tilde{X}_1'\cdot(\rho^2\tilde{P}_2\tilde{P}_1)^{k-1}\cdot\rho^2\tilde{P}_2\tilde{N}_1]\\
=\;&(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'[I_n-\rho^2\sum_{\iota=0}^{k-1}(\rho^2\tilde{P}_2\tilde{P}_1)^{\iota}\tilde{P}_2\tilde{N}_1].
\end{aligned} \qquad ...(2.4)$$

Together with (2.2) we conclude that $\lim_{k\to\infty}\hat{\beta}\,(k)=\tilde{\beta}_1$. Thus the proof of the theorem is complete.

In the following theorem we obtain a simpler form of $\tilde{\beta}_1$ which is unique.

Theorem 2.2. The WLSE $\tilde{\beta}_1$ has only one simpler form, which is given by

$$\tilde{\beta}_{1,s}=(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{y}_1-\frac{\sigma_{12}}{\sigma_{22}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_2\tilde{y}_2. \qquad ...(2.5)$$

Proof. Following from Lemma 1.1, we know

$$\begin{aligned}
\tilde{\beta}_1 \;=\;& (\tilde{X}_1'\tilde{X}_1)^{-1}\left[\tilde{X}_1'-\sum_{i=0}^{\infty}(\rho^2)^{i+1}\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1\right]\tilde{y}_1\\
&-\frac{\sigma_{12}}{\sigma_{22}}(\tilde{X}_1'\tilde{X}_1)^{-1}\left[\tilde{X}_1'\tilde{N}_2-\sum_{i=0}^{\infty}(\rho^2)^{i+1}\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1\tilde{N}_2\right]\tilde{y}_2
\end{aligned} \qquad ...(2.6)$$

In the following, we point out that if one term in $\sum_{i=0}^{\infty}(\rho^2)^{i+1}\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1$ or $\sum_{i=0}^{\infty}(\rho^2)^{i+1}\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1\tilde{N}_2$ is zero, both infinite sums are actually zero.

(1) Case I: If $\tilde{X}_1'\tilde{P}_2\tilde{N}_1=0$, then obviously $\sum_{i=0}^{\infty}(\rho^2)^{i+1}\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1=0$. Using the fact $\tilde{X}_1'\tilde{P}_2\tilde{N}_1=0\Leftrightarrow\tilde{X}_1'\tilde{P}_2\tilde{N}_1\tilde{N}_2=0$ also $\sum_{i=0}^{\infty}(\rho^2)^{i+1}\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1\tilde{N}_2=0$.

(2) Case II: Firstly, note that for any a fixed $i\geq 1$ If $\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1=0$ then $\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^{i+1}\tilde{P}_2\tilde{N}_1=0$, hence, we have

$$\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^{\iota}\tilde{P}_2\tilde{N}_1=0,\quad \iota\geq i+1. \qquad ...(2.7)$$

Secondly, note that for any a fixed $i\geq 1$, we have

$$\begin{aligned}
&\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1 = 0 \\
\Rightarrow\;& \tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{N}_2\tilde{N}_1 = 0 \\
\Rightarrow\;& \tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^{i-1}\tilde{P}_2\tilde{N}_1\tilde{N}_2\tilde{N}_1 = 0 \\
\Rightarrow\;& \tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^{i-1}\tilde{P}_2\tilde{N}_1\tilde{N}_2 = 0 \\
\Rightarrow\;& \tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^{i-1}\tilde{N}_2\tilde{N}_1\tilde{N}_2 = 0 \\
\Rightarrow\;& \tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^{i-1}\tilde{N}_2\tilde{N}_1 = 0 \\
\Rightarrow\;& \tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^{i-1}\tilde{P}_2\tilde{N}_1 = 0,
\end{aligned} \qquad ...(2.8)$$

where we use the facts that $\tilde{N}_2^2 = \tilde{N}_2$ and $\tilde{N}_1^2 = \tilde{N}_1$. Hence, we obtain

$$\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^\iota\tilde{P}_2\tilde{N}_1 = 0, \quad \iota = 0,1,\cdots,i-1. \qquad ...(2.9)$$

Thus, we conclude $\sum_{i=0}^{\infty}(\rho^2)^{i+1}\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1 = 0.$

Further, combining $\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1 = 0$ with (2.7) and 2.9) and the following fact

$$\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1 = 0 \Leftrightarrow \tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1\tilde{N}_2 = 0, \quad i = 0,1,2,\cdots,$$

we obtain

$$\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^\iota\tilde{P}_2\tilde{N}_1\tilde{N}_2 = 0, \quad \iota = 0,1,2,\cdots, \qquad ...(2.10)$$

Which implies $\sum_{i=0}^{\infty}(\rho^2)^{i+1}\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1\tilde{N}_2 = 0.$

Together, Cases I and II show that for any a fixed $i \geq 0$ if $\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1 = 0$ or $\tilde{X}_1'(\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1\tilde{N}_2 = 0$ then the two infinite sums are zero. Hence, $\tilde{\beta}_1$ only has unique simpler form $\tilde{\beta}_{1,s}$.

The proof of Theorem 2.2 is now complete.

In the situation that Σ is unknown, based on $\tilde{y}_i (i = 1,2)$, we can construct an estimator $\hat{\Sigma} = \hat{\Sigma}^* \otimes I_n$ for Σ, where $\hat{\Sigma}^* = (\hat{\sigma}_{ij})$ with

$$\hat{\sigma}_{ij} = \frac{\tilde{y}_i'\tilde{N}\tilde{y}_j}{n - \text{rank}(\tilde{X})}$$

where $\tilde{N} = I_n - \tilde{X}(\tilde{X}'\tilde{X})^-\tilde{X}'$ and $\tilde{X} = (\tilde{X}_1, \tilde{X}_2)$. Also, by the law of large numbers of Kolmogorov, it is easy to show $\hat{\sigma}_{ij}$ converges to σ_{ij} with probability one.

Accordingly, the two-stage WLSE of β_1 may be defined as

$$\tilde{\beta}_1(\hat{\Sigma}) = (\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\left[I_n - \hat{\rho}^2\sum_{i=0}^{\infty}(\hat{\rho}^2\tilde{P}_2\tilde{P}_1)^i\tilde{P}_2\tilde{N}_1\right](\tilde{y}_1 - \frac{\hat{\sigma}_{12}}{\hat{\sigma}_{22}}\tilde{N}_2\tilde{y}_2), \qquad ...(2.11)$$

where $\hat{\rho}^2 = \hat{\sigma}_{12}\hat{\sigma}_{21}/(\hat{\sigma}_{11}\hat{\sigma}_{22})$ denotes the estimator of ρ^2.

Thus, similar to the proof of Theorem 2.2, we have the following theorem.

Theorem 2.3. The two-stage WLSE $\tilde{\beta}_1(\hat{\Sigma})$ has exactly one simpler form

$$\tilde{\beta}_{1,s}(\hat{\Sigma}) = (\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{y}_1 - \frac{\hat{\sigma}_{12}}{\hat{\sigma}_{22}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_2\tilde{y}_2. \quad ...(2.12)$$

We further note that, by the iteration process (2.1) and using the idempotence of $\tilde{N}_1$,

$$\begin{aligned} \hat{\beta}_1(1) &= \hat{\beta}_1 - \text{Cov}(\hat{\beta}_1, \tilde{N}_2\tilde{y}_2)[\text{Cov}(\tilde{N}_2\tilde{y}_2)]^{-}\tilde{N}_2\tilde{y}_2 \\ &= (\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'y_1 - \sigma_{12}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_2(\sigma_{22}\tilde{N}_2)^{-}\tilde{N}_2\tilde{y}_2 \\ &= (\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'y_1 - \frac{\sigma_{12}}{\sigma_{22}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_2\tilde{y}_2 \quad ...(2.13) \\ &= \tilde{\beta}_{1,s}. \end{aligned}$$

Hence, we come to the following one conclusion.

Theorem 2.4. In the system (1.3), the one-step covariance-adjusted estimator of β_1 equals to the unique simpler form of $\tilde{\beta}_1$ given by $\tilde{\beta}_{1,s}$.

We also have the following interesting result.

Theorem 2.5. In the system (1.3), for any finite $k \geq 2$, the k-step covariance-adjusted estimator $\hat{\beta}_1(k) = \tilde{\beta}_{1,s}$.

Proof. In fact, without loss of generality, let $k = 2$, then by (2.2), we have

$$\hat{\beta}_1(2) = (\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{y}_1 + \rho^2(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_2\tilde{N}_1\tilde{y}_1 - \frac{\sigma_{12}}{\sigma_{22}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_2\tilde{y}_2, \quad ...(2.14)$$

which means $\tilde{X}_1'\tilde{N}_2\tilde{N}_1\tilde{N}_2 = 0$. However, since $\tilde{X}_1'\tilde{N}_2\tilde{N}_1\tilde{N}_2 = 0$ we have $X_1'\tilde{N}_2\tilde{N}_1 = 0$, which makes $\hat{\beta}_1(2)$ equal to $\tilde{\beta}_{1,s}$. Hence, there does not really exist $\hat{\beta}_1(2)$ different from $\tilde{\beta}_{1,s}$. In fact, by Theorem 2.2, for any finite $k \geq 2$, $\hat{\beta}_1(k)$ degenerates to $\tilde{\beta}_{1,s}$.

Remark 2.1. In the system (1.1), if $g_i(\cdot) = 0$ $(i = 1, 2)$, then it becomes the system of two SU linear regressions. Obviously, the above Theorems 2.1-2.5 still hold for this case, which show that in the system of two SU linear models:

(1) Essentially the Gauss-Markov estimator (GME) of regression parameter can be expressed as a matrix series.

(2) The GME has only one unique simpler form which exactly equals to the one-step covariance-adjusted estimator of the regression parameter, and the corresponding two-stage GME also has exactly one simpler form.

(3) For any finite $k \geq 2$, *k*-step covariance-adjusted estimator degenerates to one-step covariance-adjusted estimator.

Remark 2.2. (1) For further clarification of the estimation of g(t) in (1.2), it may be noted that we employ the partial residual procedure of Denby (1984) to formally present the estimators of $g_\iota(t)$ $(\iota = 1,2)$ in (1.2), where $\beta_\iota (\iota = 1,2)$ need to be estimated before $\bar{g}_\iota(t^{(\iota)}, \beta_\iota)$ $(\iota = 1,2)$ being used. Second, we point out that if we estimate $\beta_\iota (\iota = 1,2)$ by $\tilde{\beta}_\iota (\iota = 1,2)$ then the expressions $\bar{g}_\iota(t^{(\iota)}, \tilde{\beta}_\iota)$ $(\iota = 1,2)$ can be simplified into $\bar{g}_\iota(t^{(\iota)}, \tilde{\beta}_{\iota,s})$ $(\iota = 1,2)$, which allow one to study their properties in more detail. Here, we do not pursue the study of the properties of the estimators of $g_\iota(t)$ $(\iota = 1,2)$ as this is not within the objective of this paper.

(2) The choices of the kernel weight functions $K^{\lambda_\iota}_{\iota;j}(\cdot; t^{(\iota)}_1, \cdots, t^{(\iota)}_n)$ and the smoothing parameters λ_ι $(\iota = 1,2)$ are not considered in the present paper for the reasons that they are as suggested and adopted in Denby (1984).

(3) It is not within the scope of this paper to discuss the asymptotic properties of $\tilde{\beta}_1(\hat{\Sigma})$. However, under some regularity conditions, it can be shown that $Cov(\tilde{\beta}_1(\hat{\Sigma}))$ and $Cov(\tilde{\beta}_1)$ are asymptotic equivalent, i.e., $\lim_{n \to \infty} Cov(\tilde{\beta}_1(\hat{\Sigma})) = Cov(\tilde{\beta}_1)$.

(4) We first estimate β_ι in (1.2) by $\tilde{\beta}_\iota$ and then we note that $\bar{g}_\iota(t^{(\iota)}, \tilde{\beta}_\iota)$ $(\iota = 1,2)$ incorporate the correlation in two equations by $\tilde{\beta}_\iota (\iota = 1,2)$ since they are the best estimators of $\beta_\iota (\iota = 1,2)$, which use the information of the whole system.

3. The system of m(≥3) SU semi-parametric models

Let us consider the system of m SU semi-parametric regressions

$$y_i = X_i\beta_i + g_i(t^{(i)}) + \varepsilon_i, \quad i = 1,2,\cdots,m,$$
$$E[\varepsilon_i] = 0, \ E[\varepsilon_i\varepsilon_j] = \sigma_{ij}I_n, \quad i,j = 1,2,\cdots,m, \qquad ...(3.1)$$

where y_i $(i = 1,2,\cdots,m)$ are $n \times 1$ observation vectors, X_i are $n \times p_i$ full column rank matrices, β_i are $p_i \times 1$ regression parameters, $t^{(i)}$ are other explanatory variables, $g_i(\cdot)$ denote unknown smooth functions of $t^{(i)}$ R, and ε_i are random errors.

With the help of the partial residual procedure, we come to

$$\tilde{y}_i = \tilde{X}_i\beta_i + \varepsilon_i, \quad i = 1,2,\cdots,m,$$
$$E[\varepsilon_i] = 0, \ E[\varepsilon_i\varepsilon_j] = \sigma_{ij}I_n, \quad i,j = 1,2,\cdots,m, \qquad ...(3.2)$$

where $\tilde{y}_i = (I_n - S_i)y_i$, $\tilde{X}_i = (I_n - S_i)X_i$ and $I_n - S_i$ are inverse matrices, S_i denote the matrices generated by the corresponding kernel functions and the smoothing parameter. Hence, formally the system (3.1) has been converted into the system of m SU linear regressions.

Actually, many authors have considered and studied the system of m SU linear regressions. Wang (1989) employs the adjusted-covariance method to obtain one-step covariance-adjusted estimator for regression coefficients. Gao and Huang (2000) establish some finite sample properties of a Zellner-type estimator in the m SU linear regressions, whereas, from the viewpoint of computation, Foschi and Kontoghiorghes (2004) propose an efficient method to solve this kind of system.

Without loss of generality, we still investigate the WLSE of β_1 in the following. Denote $\bar{y}_2 = (\tilde{y}_2', \tilde{y}_3', \cdots, \tilde{y}_m')'$, $\bar{X}_2 \text{diag}(\tilde{X}_2, \tilde{X}_3, \cdots, \tilde{X}_m)$, $\bar{\beta}_2 = (\beta_2', \beta_3', \cdots, \beta_m')'$, $\bar{\varepsilon}_2 = (\varepsilon_2', \varepsilon_3', \cdots, \varepsilon_m')'$, we represent (3.2) as a system of two SU linear regressions

$$\begin{cases} \tilde{y}_1 = \tilde{X}_1\beta_1 + \varepsilon_1 \\ \bar{y}_2 = \tilde{X}_2\bar{\beta}_2 + \bar{\varepsilon}_2 \end{cases}, \qquad ...(3.3)$$

where $E[\varepsilon_1] = 0, E[\bar{\varepsilon}_2] = 0$, and denote $\text{Cov}(\varepsilon_1, \varepsilon_1) \triangleq \bar{\Sigma}_{11}$, $\text{Cov}(\varepsilon_1, \bar{\varepsilon}_2) \triangleq \bar{\Sigma}_{12}$, $\text{Cov}(\bar{\varepsilon}_2, \varepsilon_1) \triangleq \bar{\Sigma}_{21}$, $\text{Cov}(\bar{\varepsilon}_2, \bar{\varepsilon}_2) \triangleq \bar{\Sigma}_{22}$ and obviously $\bar{\Sigma} = (\bar{\Sigma}_{ij}) > 0$ is a non-diagonal partitioned matrix.

Note that in the system (3.3) $\bar{X}_2$ is a $(mn - n) \times (\Sigma_{i=2}^m p_i)$ full column rank matrix, if we partition it as $\bar{X}_2 = [\bar{X}_{21}' \vdots \bar{X}_{22}']'$, where $\bar{X}_{21}$ and $\bar{X}_{22}$ are $n \times (\Sigma_{i=2}^m p_i)$ and $(mn - 2n) \times (\Sigma_{i=2}^m p_i)$ matrices, respectively, then we have

$$\mu(\bar{X}_{21}') \cap \mu(\bar{X}_{22}') = \{0\}, \qquad ...(3.4)$$

where $\mu(A)$ denotes the space generated by the column vectors of matrix A.

Thus, we have the following lemma.

Lemma 3.1. In the system (3.3), denote $\bar{X}_2 = [\bar{X}_{21}' \vdots \bar{X}_{22}']'$, then

(i) $\bar{X}_{22} = (\bar{X}_2'\bar{X}_2]^{-1}\bar{X}_{21}' = 0;$

(ii) $\bar{X}_{21}'\bar{X}_{21}(\bar{X}_2'\bar{X}_2)^{-1}\bar{X}_{21}' = \bar{X}_{21}'.$

Proof. (i) Set $W = \bar{X}_{22}'\bar{X}_{22}(\bar{X}_2'\bar{X}_2)^{-1}\bar{X}_{21}'\bar{X}_{21}$, then obviously $\mu(W) \subset \mu(\bar{X}_{22}')$ and $\mu(W') \subset \mu(\bar{X}_{21}')$. Further, note that $W = (\bar{X}_2'\bar{X}_2 - \bar{X}_{21}'\bar{X}_{21})(\bar{X}_2'\bar{X}_2)^{-1}\bar{X}_{21}'\bar{X}_{21} = \bar{X}_{21}'\bar{X}_{21} - \bar{X}_{21}'\bar{X}_{21}(\bar{X}_2'\bar{X}_2)^{-1}\bar{X}_{21}'\bar{X}_{21}$ is a symmetric matrix, and, since $\mu(\bar{X}_{21}') \cap \mu(\bar{X}_{22}') = \{0\}$, we have $W = 0$, which implies $\bar{X}_{22}(\bar{X}_2'\bar{X}_2)^{-1}\bar{X}_{21}' = 0$.

(ii) It follows from $W = \bar{X}_{21}'\bar{X}_{21} - \bar{X}_{21}'\bar{X}_{21}(\bar{X}_2'\bar{X}_2)^{-1}\bar{X}_{21}'\bar{X}_{21} = 0$.

Lemma 3.1 is proved.

Thus, based on Lemma 3.1 and Lemma 1.1, we obtain the following results.

Theorem 3.1. In the system (3.3), the WLSE of β_1 is given by

$$\tilde{\beta}_1 = (\tilde{X}_1'\bar{\Sigma}^{11}\tilde{X}_1)^{-1}\tilde{X}_1'\sum_{i=0}^{\infty}[\bar{P}_{2,\bar{\Sigma}^{-1}}\tilde{P}_{1,\bar{\Sigma}^{-1}}]^i(\bar{\Sigma}^{11}-\bar{P}_{2,\bar{\Sigma}^{-1}})\tilde{y}_1$$
$$+(\tilde{X}_1'\bar{\Sigma}^{11}\tilde{X}_1)^{-1}\tilde{X}_1'\sum_{i=0}^{\infty}[\bar{P}_{2,\bar{\Sigma}^{-1}}\tilde{P}_{1,\bar{\Sigma}^{-1}}]^i(\bar{\Sigma}^{12}-\bar{\Sigma}^{12}\mathrm{diag}(\bar{P}_{21,\bar{\Sigma}^{-1}},\bar{P}_{22,\bar{\Sigma}^{-1}})\bar{\Sigma}^{22})\bar{y}_2,$$

where $\bar{\Sigma}^{-1}=(\bar{\Sigma}^{ij})$ $(i,j=1,2)$ denotes the inverse of $\bar{\Sigma},\tilde{P}_{1,\bar{\Sigma}^{-1}}=\tilde{X}_1(\tilde{X}_1'\bar{\Sigma}^{11}\tilde{X}_1)^{-1}\tilde{X}_1'$ and $\bar{P}_{2,\bar{\Sigma}^{-1}}=\bar{\Sigma}^{12}\mathrm{diag}(\bar{P}_{21,\bar{\Sigma}^{-1}},\bar{P}_{22,\bar{\Sigma}^{-1}})\bar{\Sigma}^{21}$ with $\bar{P}_{2i,\bar{\Sigma}^{-1}}=\bar{X}_{2i}(\bar{X}_2'\bar{\Sigma}^{22}\bar{X}_2)^{-1}\bar{X}_{2i}'$ $(i=1,2)$. Accordingly, the unique simpler form of $\tilde{\beta}_1\tilde{}$ in this case is

$$\tilde{\beta}_{1,s}=\hat{\beta}_1+(\tilde{X}_1'\bar{\Sigma}^{11}\tilde{X}_1)^{-1}\tilde{X}_1'[\bar{\Sigma}^{12}-\bar{\Sigma}^{12}\mathrm{diag}(\bar{P}_{21,\bar{\Sigma}^{-1}},\bar{P}_{22,\bar{\Sigma}^{-1}})\bar{\Sigma}^{22}]\bar{y}_2$$

where $\hat{\beta}_1=(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{y}_1$.

For instance, in the case that $m=4$, $\tilde{\beta}_{1,s}$ would have the form

$$\begin{aligned}\tilde{\beta}_{1,s(4)} = {} & \tilde{\beta}_1-\frac{\sigma_{12}}{\sigma_{22}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_2\tilde{y}_2-\frac{\sigma_{13}}{\sigma_{33}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_3\tilde{y}_3\\ & -\frac{\sigma_{14}}{\sigma_{44}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_4\tilde{y}_4+\frac{\sigma_{12}\sigma_{23}}{\sigma_{22}\sigma_{33}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_3\tilde{y}_3\\ & +\frac{\sigma_{12}\sigma_{24}}{\sigma_{22}\sigma_{44}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_2\tilde{N}_4\tilde{y}_4+\frac{\sigma_{13}\sigma_{34}}{\sigma_{33}\sigma_{44}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_3\tilde{N}_4\tilde{y}_4\\ & -\frac{\sigma_{12}\sigma_{23}\sigma_{34}}{\sigma_{22}\sigma_{33}\sigma_{44}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_2\tilde{N}_3\tilde{N}_4\tilde{y}_4,\end{aligned} \qquad ...(3.5)$$

where $\tilde{N}_i=I_n-\tilde{X}_i(\tilde{X}_i'\tilde{X}_i)^{-1}\tilde{X}_i'(i=2,3,4)$.

Further, some induction computations show the following result.

Theorem 3.2. In the system (3.2), the unique simpler WLSE of β_1 is given by

$$\begin{aligned}\tilde{\beta}_{1,s(m)} = {} & \hat{\beta}_1-\sum_{i=2}^{m}\frac{\sigma_{1i}}{\sigma_{ii}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_i\tilde{y}_i\\ & +\sum_{2\le i<j\le m}\frac{\sigma_{1i}\sigma_{ij}}{\sigma_{ii}\sigma_{jj}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_i\tilde{N}_j\tilde{y}_i\\ & +\sum_{\iota=3}^{m-2}(-1)^{\iota}\sum_{2\le i_1<i_2<\cdots<i_\iota\le m}\frac{\sigma_{1i_1}\sigma_{i_1i_2}\cdots\sigma_{i_{\iota-1}i_\iota}}{\sigma_{i_1i_1}\sigma_{i_2i_2}\cdots\sigma_{i_\iota i_\iota}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_{i_1}\tilde{N}_{i_2}\cdots\tilde{N}_{i_\iota}\tilde{y}_{i_\iota}\\ & +(-1)^{m-1}\frac{\sigma_{12}\sigma_{23}\cdots\sigma_{m-1m}}{\sigma_{22}\sigma_{33}\cdots\sigma_{mm}}(\tilde{X}_1'\tilde{X}_1)^{-1}\tilde{X}_1'\tilde{N}_2\tilde{N}_3\cdots\tilde{N}_m\tilde{y}_m,\end{aligned}$$

which equals to the one-step covariance-adjusted estimator of β_1 from the system, where $\tilde{N}_i = I_n - \tilde{X}_i(\tilde{X}_i'\tilde{X}_i)^{-1}\tilde{X}_i' (i = 2, \cdots, m)$. Further, for any finite $k \geq 2$, the k-step covariance-adjusted estimator of β_1 degenerates to $\tilde{\beta}_{1,s(m)}$.

Remark 3.1. In the system (3.1), if all $g(\cdot)(i = 1, \cdots, m)$ equal to zero, then it degenerates to the system of $m(m \geq 3)$ SU linear regressions, and accordingly the simpler WLSE $\tilde{\beta}_{1,s(m)}$ becomes the simpler GME of regression parameter β_1 from the system. The above results also show that in the system of $m(m \geq 3)$ SU semi-parametric regressions, the two-stage WLSE of regression parameter has only one unique simpler form.

4. Conclusions and Remarks

The system consisting of $m(m \geq 2)$ SU linear regressions has many practical applications related to econometrics, industry, biological science and geography, etc. The SU regressions system has drawn considerable attention of researcher in recent years since the pioneer works of Zellner (1962, 1963). Although some authors propose and study their respective simpler form estimators (see Zellner (1963), Revankar (1974), Liu (2000) and Liu (2002)), a very few have ever examined the uniqueness of their simpler form estimators. Further, although the covariance-adjusted method can be used to construct the estimator sequence of regression parameter, hardly any have studied the relationship between the one-step covariance-adjusted estimator and the simpler form. In the present paper, we consider the system of m SU semi-parametric regressions, which includes the system of linear regressions as its special case. Hence, our conclusions not only answer the above problems but further show that in the system of $m(\geq 2)$ SU semi-parametric regressions:

(1) Essentially the WLSE of regression parameter can be expressed as a matrix series which only has one unique simpler form.

(2) The covariance-adjusted method is efficient and feasible, which can be integrated into the system to generate an efficient estimator sequence in the sense of smaller covariance, and the limit of the sequence exactly equals to the WLSE of the parameter. Moreover, the one-step covariance-adjusted estimator is just the simpler form of the WLSE and for any finite $k \geq 2$, the k-step covariance-adjusted estimator degenerates to the corresponding one-step covariance-adjusted estimator.

(3) The simpler form of the two-stage WLSE is unique, which allows us to study its performance in more details in the case that the variance-covariance of disturbances is not known.

(4) The partial residual estimators of $g_i(\cdot)$ have the corresponding simpler forms whether the variance-covariance of disturbances is known or unknown.

(5) Discussions on the choices of the kernel weight function and the smoothing parameter are not within the objective of this paper, nor is the discussion on asymptotic properties of $\tilde{\beta}_1(\hat{\Sigma})$ or of $\tilde{\beta}_1$. These could be the subject matter of further research.

In conclusion, the present paper generalizes and enriches a number of results in the existing literature involving SU regressions system.

References

Carroll, R.J., Doug M., Larry F. and Victor K. (2006), "Seemingly Unrelated Measurement Error Models, with Application to Nutritional Epidemiology", *Biometrics*, 62(3), 75-84.

Chen, H. (1988), "Convergence Rates for Parametric Components in a Partly Linear Model", *Annals of Statistics*, 16(2), 136-146.

Clark, R. M. (1977), "Non-parametric Estimation of a Smooth Regression Function", *Journal of the Royal Statistical Society*, Ser. B, 39(5), 107-113.

Denby, L. (1984), "Smooth Regression Functions", PHD. Thesis, Dept. of Statist., University of Michigan.

Foschi, P. and Kontoghiorghes, E. J. (2004), "A Computationally Efficient Method for Solving SUR Models with Orthogonal Regressors", *Linear Algebra and its Applications*, 388(7), 193-200.

Gao, D. D. and Huang, R. B. (2000), "Some Finite Sample Properties of Zellner Estimator in the Context of m Seemingly Unrelated Regression Equations", *Journal of Statistical Planning and Inference*, 88(6), 267-283.

Hardle, W., Gao, J. and Liang, H. (2000), *Partially Linear Models*, Springer, New York.

Kementa, J. and Gilbert, R. F. (1968), "Small Sample Properties of Alternative Estimators of Seemingly Unrelated Regressions", *Journal of American Statistical Association*, 63(324), 1180-1200.

Kurata, H. (1999), "On the Efficiencies of Several Generalized Least Squares Estimators in a Seemingly Unrelated Regression Model and a Heteroscedastic Model", *Journal of Multivariate Analysis,* 70(5), 86-94.

Liu, A. Y. (2002), "Efficient Estimation of Two Seemingly Unrelated Regression Equations", *Journal of Multivariate Analysis*, 82(6), 445-456.

Liu, J. S. (2000), "MSEM Dominance of Estimators in Two Seemingly Unrelated Regressions", *Journal of Statistical Planning and Inference*, 88(11), 255-266.

Mehta, J. S. and Swamy P. A. V. B. (1976), "Further Evidence on the Relative Efficiencies of Zellner's Seemingly Unrelated Regressions Estimator", *Journal of American Statistical Association,* 71(355), 634-639.

Ng, V. M. (2002), "Robust Bayesian Inference for Seemingly Unrelated Regressions with Elliptical Errors", *Journal of Multivariate Analysis*, 83(4), 409-414.

Rao, C. R. (1967), "Least Squares Theory Using an Estimated Dispersion Matrix and Its Applications to Measurement of Signals", In *Proceeding of the fifth Berkerley Symposium on Math. Stat. & Prob.,* 2, 355-372.

Revankar, N. S. (1974), "Some Finite Sample Results in the Context of Two Seemingly Unrelated Regression Equations", *Journal of American Statistical Association*, 69(345), 187-190.

Schmidt, P. (1977), "Estimation of Seemingly Unrelated Regressions with Unequal Numbers of Observations", *Journal of Econometrics*, 5(3), 365-377.

Srivastava, V. K. and Giles, E. A. (1987), "Seemingly Unrelated Regression Equations Models", *Estimation and Inference. Statistics: Textbooks and Monographs*, vol. 80. Marcel Dekker, Inc., New York.

Swamy, P. A. V. B. and Mehta, J. S. (1975), "On Bayesian Estimation of Seemingly Unrelated Regressions when Some Observations are Missing", *Journal of Econometrics*, 3(2), 157-169.

Wang, S. G. (1989) "A New Estimate of Regression Coefficients in Seemingly Unrelated Regression System", *Science in China, Ser. A*, 19(7), 808-816.

Welsh, A. H. and Yee, T. W. (2006), "Local Regression for Vector Responses", *Journal of Statistical Planning and Inference*, 135(6), 307-331.

Xu, Q.F., You, J. H. and Zhou, B. (2008), "Seemingly Unrelated Nonparametric Models with Positive Correlation and Constrained Error Variances", *Economics Letters*, 99(3), 223-227.

Zellner, A. (1962), "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias", *Journal of American Statistical Association*, 57(298), 348-368.

Zellner, A. (1963), "Estimators of Seemingly Unrelated Regressions: Some Exact Finite Sample Results", *Journal of American Statistical Association*, 58(304), 977-992.

♦♦♦