Asymptotic Expansions of the Distributions of Semi-Parametric Estimators in a Linear Simultaneous Equations System *

Naoto Kunitomo †
and
Yukitoshi Matsushita

October 2005

Abstract

Asymptotic expansions are made of the distributions of a class of semi-parametric estimators including the Maximum Empirical Likelihood (MEL) method and the Generalized Method of Moments (GMM) for the coefficients of a single structural equation in the linear simultaneous equations system. The expansions in terms of the sample size, when the non-centrality parameters increase proportionally, are carried out to the order of $O(n^{-2})$. Comparisons of the distribuitions of the MEL and GMM estimators are made. Also we relate the asymptotic expansions of the distributions of the MEL and GMM estimators to the corresponding expansions for the Limited Information Maximum Likelihood (LIML), and the Two-Stage Least Squares (TSLS) estimators. We give useful information on the properties of the distributions in alternative estimation methods.

Key Words

Asymptotic Expansions, Maximum Empirical Likelihood (MEL), Generalized Method of Moments (GMM), Limited Information Maximum Likelihood (LIML), Two-Stage Least Squares (TSLS), Linear Simultaneous Equations System

---

*This paper is a revision of Discussion Paper CIRJE-F-237, August 2003, Graduate School of Economics, University of Tokyo.

†Graduate School of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, JAPAN, E-mail: kunitomo@e.u-tokyo.ac.jp.

1. Introduction

The study of estimating a single structural equation in econometric models has led to develop several estimation methods as the alternatives to the least squares estimation method. The classical examples in the econometric literatures are the limited information maximum likelihood (LIML) method and the instrumental variables (IV) method including the two-stage least squares (TSLS) method. See Anderson, Kunitomo, and Sawa (1982) and Anderson, Kunitomo, and Morimune (1986) for their finite sample properties, for instance. In addition to these classical methods the maximum empirical likelihood (MEL) method has been proposed and has gotten some attention recently in the statistical and econometric literatures. It is probably because the MEL method gives asymptotically efficient estimator in the semi-parametric sense and also improves the serious bias problem known in the generalized method of moments (GMM) method when the number of instruments is large in econometric models. See Owen (2001), Qin and Lawless (1994), Kitamura and Stutzer (1997), and Kitamura, Tripathi, and Ahn (2004) on the details of the MEL method.

In the econometric literatures the generalized method of moments (GMM) estimation method has been quite popular in the past two decades. The GMM method was originally proposed by Hansen (1982) in the econometric literature and it is essentially the same as the estimating equation (EE) method proposed by Godambe (1960) which has been used in statistical applications. It has been known that both the MEL estimator and the GMM estimator are asymptotically normally distributed and efficient when the sample size is large. Because we have two semi-parametric estimation methods for econometric models and they are asymptotically equivalent, it is interesting to make comparison of the finite sample properties of alternatives estimation methods.

The main purpose of this study is to develop the method of deriving the asymptotic expansions of the distributions of a class of semi-parametric estimators for the coefficients of a single structural equation in the linear simultaneous equations system. The estimation methods under the present study include both the MEL estimator and the GMM estimator as special cases. Since it is quite difficult to investigate the exact distributions of these estimators in the general case, their asymptotic expansions give some information on their finite sample properties. In this paper the asymptotic expansions shall be carried out in terms of the sample size which is proportional to the non-centrality parameters. Comparisons of the distributions of the MEL and GMM methods can be made. We shall illustrate the usefulness of the asymptotic expansion method developed in this paper by giving some information on the distribution functions of the MEL and GMM estimators. Also we shall relate our results to the earlier studies on the asymptotic expansions of the distributions of the limited information maximum likelihood (LIML) and the two-stage least squares (TSLS) estimators. It will give some insights on the statistical properties of the alternative estimation methods for a single structural equation.

In order to compare alternative estimators, it is much more easier to investigate the asymptotic expansions of their mean and mean squared errors (MSE) than their exact distribution functions. Since the exact distributions of estimators can be quite different from the normal distribution, it should be certainly better to investigate the asymptotic expansions of their exact distributions directly. Also the asymptotic expansions of the mean and the MSE of estimators are not necessarily the same as the mean and the

MSE of the asymptotic expansions of the distributions of estimators. In fact it has been known in econometrics that the LIML estimator does not possess any moments of positive integer order under a set of reasonable assumptions. This paper may be the first attempt to develop the asymptotic expansions of the distribution functions of semi-parametric estimators and to find their explicit form in the estimating equation model and the simultaneous equation models in econometrics.

Our formulation and method of this paper are intentionally parallel to the classical studies on the single equation estimation methods in the linear simultaneous equations by Fujikoshi et. al. (1982) and Anderson et. al. (1986). It is mainly because useful interpretation can be drawn in the light of past studies on the finite sample properties of alternative estimators in the classical parametric framework as well as the semi-parametric framework. The problem dealt with our paper would be related to the studies of higher order asymptotic efficiency of estimation by Takeuchi and Morimune (1985), and Newey and Smith (2004).

In Section 2 we define the single equation econometric model and their estimation methods. Then in Section 3 we give the stochastic expansions of a class of estimators. In Section 4, we give the results of the asymptotic expansions of the distribution functions of estimators under a set of assumptions on the disturbances. Then we shall compare the asymptotic expansions of alternative estimators including their asymptotic biases and the mean-squared errors in a simple case and a more general case. Some discussion and concluding remarks are given in Section 5. The proofs of Lemmas and Theorems and some useful formulas for our results will be given in Appendices.

2. Estimating a Single Structural Equation by the Maximum Empirical Likelihood Method

Let a linear structural equation in the econometric model be given by

$$y_{1i} = (\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix} + u_i \quad (i = 1, \cdots, n), \tag{2.1}$$

where y_{1i} and $\mathbf{y}_{2i}$ are 1×1 and $G_1 \times 1$ (vector of) endogenous variables, $\mathbf{z}_{1i}$ is a $K_1 \times 1$ vector of exogenous variables, $\boldsymbol{\theta}^{'} = (\boldsymbol{\beta}^{'}, \boldsymbol{\gamma}^{'})$ is a $1 \times p$ $(p = K_1 + G_1)$ vector of unknown coefficients, and $\{u_i\}$ are mutually independent disturbance terms with $\mathbf{E}(u_i) = 0$ $(i = 1, \cdots, n)$. We assume that (2.1) is the first equation in a system of $(G_1 + 1)$ structural equations relating the vector of $G_1 + 1$ endogenous variables $\mathbf{y}_i^{'} = (y_{1i}, \mathbf{y}_{2i}^{'})$ and the vector of K $(= K_1 + K_2)$ exogenous variables $\{\mathbf{z}_i\}$ which includes $\{\mathbf{z}_{1i}\}$. The set of exogenous variables $\{\mathbf{z}_i\}$ are often called the instrumental variables and we have the orthogonal condition

$$\mathbf{E}(u_i\, \mathbf{z}_i) = \mathbf{0} \quad (i = 1, \cdots, n)\ . \tag{2.2}$$

Because we do not specify the equations except (2.1) and we only have the limited information on the set of instrumental variables (or instruments), we only consider the limited information estimation methods.

When all structural equations in the econometric model are linear, the reduced form equations for $\mathbf{y}_i^{'} = (y_{1i}, \mathbf{y}_{2i}^{'})$ can be represented as

$$\mathbf{y}_i = \boldsymbol{\Pi}^{'}\mathbf{z}_i + \mathbf{v}_i \quad (i = 1, \cdots, n)\ , \tag{2.3}$$

where $\mathbf{v}_i^{'} = (v_{1i}, \mathbf{v}_{2i}^{'})$ is a $1 \times (1+G_1)$ disturbance terms with $\mathbf{E}[\mathbf{v}_i] = \mathbf{0}$ and $\mathbf{\Pi}^{'} = (\boldsymbol{\pi}_1, \mathbf{\Pi}_2)^{'}$ is a $(1+G_1) \times K$ partitioned matrix of the linear reduced form coefficients. By multiplying $(1, -\beta^{'})$ to (2.3) from the left-hand side, we have the linear restriction $(1, -\boldsymbol{\beta}^{'})\mathbf{\Pi}^{'} = (\boldsymbol{\gamma}^{'}, \mathbf{0}^{'})$ and $u_i = v_{1i} - \boldsymbol{\beta}^{'}\mathbf{v}_{2i}$ $(i = 1, \cdots, n)$.

The maximum empirical likelihood (MEL) estimator for the vector of unknown parameters $\boldsymbol{\theta}$ in (2.1) is defined by maximizing the Lagrange form

$$
(2.4) \qquad L_n^*(\boldsymbol{\lambda}, \boldsymbol{\theta}) = \sum_{i=1}^{n} \log p_i - \mu(\sum_{i=1}^{n} p_i - 1) - n\boldsymbol{\lambda}^{'} \sum_{i=1}^{n} p_i \, \mathbf{z}_i[y_{1i} - (\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\boldsymbol{\theta})] \ ,
$$

where μ and $\boldsymbol{\lambda}$ are a scalar and a $K \times 1$ vector of Lagrange multipliers, and p_i $(i = 1, \cdots, n)$ are the weighted probability functions to be chosen. It has been known (see Qin and Lawless (1994) or Owen (2001)) that the above maximization problem is the same as to maximize

$$
(2.5) \qquad L_n(\boldsymbol{\lambda}, \boldsymbol{\theta}) = -\sum_{i=1}^{n} \log\{1 + \boldsymbol{\lambda}^{'}\mathbf{z}_i \, [y_{1i} - (\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\boldsymbol{\theta})]\} \ ,
$$

where we used the conditions $\hat{\mu} = n$ and $[n\hat{p}_i]^{-1} = 1 + \lambda^{'}\mathbf{z}_i[y_{1i} - (\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\boldsymbol{\theta})]$.
By differentiating (2.5) with respect to λ and combining the resulting equation with the restriction $\sum_{i=1}^{n} p_i = 1$, we have the relation

$$
(2.6) \qquad \sum_{i=1}^{n} \hat{p}_i \mathbf{z}_i \, [y_{1i} - (\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\hat{\boldsymbol{\theta}})] = \mathbf{0}
$$

and

$$
(2.7) \qquad \hat{\boldsymbol{\lambda}} = [\sum_{i=1}^{n} \hat{p}_i u_i^2(\hat{\theta})\mathbf{z}_i\mathbf{z}_i^{'}]^{-1}[\frac{1}{n}\sum_{i=1}^{n} u_i(\hat{\theta})\mathbf{z}_i] \ ,
$$

where $u_i(\hat{\theta}) = y_{1i} - (\mathbf{y}_{2i}, \mathbf{z}_{1i})\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\theta}}$ is the maximum empirical likelihood (MEL) estimator for the vector of unknown parameters $\boldsymbol{\theta}$. From (2.6) and (2.7) the MEL estimator of $\boldsymbol{\theta}$ is the solution of the set of p equations

$$
\begin{aligned}
(2.8) \qquad & [\sum_{i=1}^{n} \hat{p}_i \begin{pmatrix} \mathbf{y}_{2i} \\ \mathbf{z}_{1i} \end{pmatrix} \mathbf{z}_i^{'}][\sum_{i=1}^{n} \hat{p}_i u_i(\hat{\theta})^2 \mathbf{z}_i\mathbf{z}_i^{'}]^{-1}[\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i \, y_{1i}] \\
= \ & [\sum_{i=1}^{n} \hat{p}_i \begin{pmatrix} \mathbf{y}_{2i} \\ \mathbf{z}_{1i} \end{pmatrix} \mathbf{z}_i^{'}][\sum_{i=1}^{n} \hat{p}_i u_i(\hat{\theta})^2 \mathbf{z}_i\mathbf{z}_i^{'}]^{-1}[\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i(\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})] \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\gamma}} \end{pmatrix} .
\end{aligned}
$$

If we substitute $1/n$ for $\hat{p}_i$ $(i = 1, \cdots, n)$ in (2.8) and set an initial (consistent) estimator $\tilde{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$, then we have the generalized method of moments (GMM) estimator for the vector of coefficient parameters $\boldsymbol{\theta}^{'} = (\boldsymbol{\beta}^{'}, \boldsymbol{\gamma}^{'})$, which can be written as the solution of

$$
\begin{aligned}
(2.9) \qquad & [\frac{1}{n}\sum_{i=1}^{n} \begin{pmatrix} \mathbf{y}_{2i} \\ \mathbf{z}_{1i} \end{pmatrix} \mathbf{z}_i^{'}][\frac{1}{n}\sum_{i=1}^{n} u_i(\tilde{\theta})^2 \mathbf{z}_i\mathbf{z}_i^{'}]^{-1}[\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i \, y_{1i}] \\
= \ & [\frac{1}{n}\sum_{i=1}^{n} \begin{pmatrix} \mathbf{y}_{2i} \\ \mathbf{z}_{1i} \end{pmatrix} \mathbf{z}_i^{'}][\frac{1}{n}\sum_{i=1}^{n} u_i(\tilde{\theta})^2 \mathbf{z}_i\mathbf{z}_i^{'}]^{-1}[\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i(\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})] \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\gamma}} \end{pmatrix} .
\end{aligned}
$$

(See Hayashi (2000) on the details of the GMM method in econometrics, for instance.)

By generalizing the weight probabilities p_i $(i = 1, \cdots, n)$ in (2.8), we can introduce a class of estimators. Let

$$\hat{p}_i^* = \frac{1}{n[1 + \delta\, \boldsymbol{\lambda}^{'} \mathbf{z}_i u_i(\hat{\theta})]} \ , \tag{2.10}$$

where δ is a positive constant $(0 \leq \delta \leq 1)$ and $\hat{\theta}$ is the MEL estimator of $\boldsymbol{\theta}$. Then we define the modification of the MEL estimator (the MMEL estimator) by substituting $\hat{p}_i$ $(i = 1, \cdots, n)$ into (2.6)-(2.8). We shall denote the resulting Lagrange multiplier and the modified estimator as $\hat{\lambda}$ and $\hat{\theta}$ whenever we can avoid any confusion.

In the rest of the present paper, we shall consider the standardized error of estimators as

$$\hat{\mathbf{e}} = \sqrt{n} \begin{pmatrix} \hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \\ \hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma} \end{pmatrix} \ , \tag{2.11}$$

where $\hat{\boldsymbol{\theta}}^{'} = (\hat{\boldsymbol{\beta}}^{'}, \hat{\boldsymbol{\gamma}}^{'})$. We sometimes denote $\hat{\mathbf{e}}$ for the MEL estimator and its modification whenever we do not make any confusion. Under a set of regularity conditions, the asymptotic variance-covariance matrix of the asymptotically efficient estimators in the semi-parametric framework is given by

$$\mathbf{Q}^{-1} = \mathbf{D}^{'} \mathbf{M} \mathbf{C}^{-1} \mathbf{M} \mathbf{D} \ , \tag{2.12}$$

where $K \times K$ matrices $\mathbf{M}_n$ and $\mathbf{C}_n$ are defined by

$$\mathbf{M}_n = \frac{1}{n} \sum_{i=1}^{n} \mathbf{z}_i \mathbf{z}_i^{'} \xrightarrow{p} \mathbf{M} \ , \ \ \mathbf{C}_n = \frac{1}{n} \sum_{i=1}^{n} \mathbf{z}_i \mathbf{z}_i^{'} u_i^2 \xrightarrow{p} \mathbf{C} \ , \tag{2.13}$$

and

$$\mathbf{D} = [\boldsymbol{\Pi}_2 \ , \begin{pmatrix} \mathbf{I}_{K_1} \\ \mathbf{O} \end{pmatrix}] \ . \tag{2.14}$$

We assume that the (constant) matrices $\mathbf{M}$ and $\mathbf{C}$ are positive definite, and the rank condition

$$\mathbf{rank}(\mathbf{D}) = p \ (= G_1 + K_1) \ . \tag{2.15}$$

These conditions assure that the limiting variance-covariance matrix $\mathbf{Q}$ is non-degenerate. The above rank condition implies the order condition $L = K - p \geq 0$, which is the degree of over-identification. When the disturbance terms are (conditionally or unconditionally) homoscedastic random variables such as an i.i.d. sequence, then we have $\mathbf{C} = \sigma^2 \mathbf{M}$ and $\mathbf{E}(u_i^2) = \sigma^2$.

In order to compare alternative efficient estimation methods in the finite sample sense, we need to derive the asymptotic expansions of the density functions of the standardized estimators (2.16) in the form of

$$f(\boldsymbol{\xi}) = \phi_{\mathbf{Q}}(\boldsymbol{\xi}) [\ 1 + \frac{1}{\sqrt{n}} H_1(\boldsymbol{\xi}) + \frac{1}{n} H_2(\boldsymbol{\xi})\ ] + o(\frac{1}{n}) \ , \tag{2.16}$$

where $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_p)^{'}$, $\phi_{\mathbf{Q}}(\boldsymbol{\xi})$ is the multivariate normal density function with mean $\mathbf{0}$ and the variance-covariance matrix $\mathbf{Q}$, and $H_i(\boldsymbol{\xi})$ $(i = 1, 2)$ are some polynomial functions of elements of $\boldsymbol{\xi}$. In order to derive the asymptotic expansions of the distributions of estimators in a simple manner, however, we need a set of regularity conditions.

Assumption I :
(i) The sequences $\{\mathbf{v}_i\}$ (and hence $\{u_i\}$) are mutually independent random vectors which have the positive density with respect to Lebesgue measure with $\mathbf{E}[\|\mathbf{v}_i\|^{7+\epsilon}] < \infty$ for some $\epsilon > 0$.
(ii) The (constant) matrices $\mathbf{M}$ and $\mathbf{C}$ are positive definite and the rank condition is satisfied with

$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'} = \mathbf{M} + o_p(\frac{1}{n}) \ . \tag{2.17}$$

(iii) The sequence of $\mathbf{z}_i = (z_i^{(j)})$ $(i = 1, \cdots, n; j = 1, \cdots, K)$ are independent random variables (or non-random) vectors, which are mutually independent of $\mathbf{v}_i$ $(i = 1, \cdots, n)$. The vectors $\{\mathbf{z}_i\}$ are bounded or $\mathbf{E}[\|\mathbf{z}_i\|^{4+\epsilon}]$ for some $\epsilon > 0$. There exist finite $M_3(j_1, j_2, j_3)$ such that $(1/n)\sum_{i=1}^{n} u_i^3 z_i^{(j_1)} z_i^{(j_2)} z_i^{(j_3)} = M_3(j_1, j_2, j_3) + o_p(\frac{1}{\sqrt{n}})$.

We need some moment conditions on disturbance terms to derive higher order stochastic expansions of the associated random variables up to $O(n^{-1})$. The conditions in (ii) and (iii) of **Assumption I** are rather strong and it is possible to weaken them. Then the resulting formulas and their derivations become more complicated than those reported in this paper while the essential method of derivations will not to be changed. We can treat both cases when $\{\mathbf{z}_i\}$ are stochastic and deterministic, and also it is possible to replace the independence assumption with $\{u_i\}$ by using a martingale assumption on the random vector sequence of $\sum_{i=1}^{n}\mathbf{z}_i u_i$. In order to avoid the inessential arguments, however, we mostly treat $\{\mathbf{z}_i\}$ as if they were deterministic variables.

We shall use the mean operator $AM_n(\hat{\mathbf{e}})$, which is defined as the mean of $\hat{\mathbf{e}}$ with respect to the asymptotic expansion of its density function of the standardized estimator up to $O(n^{-1})$ in the form of (2.16). We write the asymptotic bias and the asymptotic MSE of the standardized estimator by

$$ABIAS_n(\hat{\mathbf{e}}) = AM_n(\hat{\mathbf{e}}) \ , \tag{2.18}$$

and

$$AMSE_n(\hat{\mathbf{e}}) = AM_n(\hat{\mathbf{e}}\,\hat{\mathbf{e}}^{'}) \ . \tag{2.19}$$

These quantities are useful because the asymptotic expansion of the distribution of estimators are quite complicated in the general case. However, it should be noted that they are not necessarily the same as the asymptotic expansions of the exact moments and some care should be taken in this respect [1] .

3. Stochastic Expansions of Estimators

3.1 Stochastic Expansions

First, we apply the similar arguments used in Owen (1990) and Qin and Lawless (1994) on the probability limits and the consistency of the MEL estimator. Then we have $n\hat{p}_i \xrightarrow{p} 1$, $\hat{\theta}_{EL} \xrightarrow{p} \theta_0$, ($\theta_0$ is the true value of θ) and $\sqrt{n}\hat{\lambda}$ converges to a random

[1] It has been well-known in econometrics that the LIML estimator does not have any positive integer moments in our setting.

vector as $n \to \infty$. By substituting (2.1) into (2.8), we have the representation of the standardized estimator $\hat{\mathbf{e}}$ as

$$\text{(3.1)} \qquad [\sum_{i=1}^{n} \hat{p}_i \begin{pmatrix} \mathbf{y}_{2i} \\ \mathbf{z}_{1i} \end{pmatrix} \mathbf{z}_i^{'}][\sum_{i=1}^{n} \hat{p}_i u_i(\hat{\theta})^2 \mathbf{z}_i \mathbf{z}_i^{'}]^{-1}[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{z}_i u_i]$$

$$= [\sum_{i=1}^{n} \hat{p}_i \begin{pmatrix} \mathbf{y}_{2i} \\ \mathbf{z}_{1i} \end{pmatrix} \mathbf{z}_i^{'}][\sum_{i=1}^{n} \hat{p}_i u_i(\hat{\theta})^2 \mathbf{z}_i \mathbf{z}_i^{'}]^{-1}[\frac{1}{n} \sum_{i=1}^{n} \mathbf{z}_i (\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})]\hat{\mathbf{e}} \ ,$$

where we use the notation $\hat{\boldsymbol{\theta}}$ for $\hat{\boldsymbol{\theta}}_{EL}$ without any subscript whenever we do not have any confusion. As $n \to \infty$, we write the first order term of $\hat{\mathbf{e}}$ as $\tilde{\mathbf{e}}_0$, which is given by

$$\text{(3.2)} \qquad \tilde{\mathbf{e}}_0 = (\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{D})^{-1}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{z}_i u_i) \ .$$

In the following derivation we shall utilize the asymptotic equivalence under **Assumption I** that $\tilde{\mathbf{e}}_0 - \mathbf{e}_0 = o_p(1)$ with

$$\text{(3.3)} \qquad \mathbf{e}_0 = (\mathbf{D}^{'}\mathbf{M}_n\mathbf{C}_n^{-1}\mathbf{M}_n\mathbf{D})^{-1}\mathbf{D}^{'}\mathbf{M}_n\mathbf{C}_n^{-1}(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{z}_i u_i) \ .$$

By using the central limit theorem (CLT) to the last term, we have the weak convergence of the random vector

$$\text{(3.4)} \qquad \mathbf{X}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{z}_i u_i \xrightarrow{w} N_p(\mathbf{0}, \mathbf{C})$$

and then $\tilde{\mathbf{e}}_0 \xrightarrow{w} N_p(\mathbf{0}, \mathbf{Q})$, where a $p \times p$ matrix $\mathbf{Q}$ has been defined by (2.12) and $\xrightarrow{w}$ means the convergence of distribution as $n \to \infty$.
Also we notice that

$$\text{(3.5)} \quad \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{z}_i u_i(\hat{\theta}) = \mathbf{X}_n + \frac{1}{n}[-\sum_{i=1}^{n} \mathbf{z}_i(\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\hat{\mathbf{e}}] = \mathbf{X}_n - \mathbf{M}_n\mathbf{D}\hat{\mathbf{e}} + O_p(\frac{1}{\sqrt{n}}) \ .$$

By utilizing the representation of (2.7) for $\hat{\boldsymbol{\lambda}}$, we have $\sqrt{n}\boldsymbol{\lambda} - \boldsymbol{\lambda}_0 \xrightarrow{p} 0$ and

$$\text{(3.6)} \ \boldsymbol{\lambda}_0 = \mathbf{C}_n^{-1/2}[\mathbf{I}_K - \mathbf{C}_n^{-1/2}\mathbf{M}_n\mathbf{D}(\mathbf{D}^{'}\mathbf{M}_n\mathbf{C}_n^{-1}\mathbf{M}_n\mathbf{D})^{-1}\mathbf{D}^{'}\mathbf{M}_n\mathbf{C}_n^{-1/2}][\mathbf{C}_n^{-1/2}\mathbf{X}_n] \ .$$

Since the limiting distribution of $\mathbf{B}_n = \mathbf{C}^{-1/2}\mathbf{X}_n$ is given by $N_K(\mathbf{0}, \mathbf{I}_K)$, we have the convergence $\mathbf{C}_n^{1/2}\sqrt{n}\hat{\boldsymbol{\lambda}} \xrightarrow{w} N_K(\mathbf{0}, \bar{\mathbf{P}}_{D^*})$, where the projection operator is defined by $\bar{\mathbf{P}}_{D^*} = \mathbf{I}_K - \mathbf{D}^*(\mathbf{D}^{*'}\mathbf{D}^*)^{-1}\mathbf{D}^{*'}$ which is constructed by a $K \times p$ matrix $\mathbf{D}^* = \mathbf{C}^{-1/2}\mathbf{M}\mathbf{D}$. The variance-covariance matrix of the limiting distribution of random vector $\boldsymbol{\lambda}_0$ is given by

$$\text{(3.7)} \qquad \mathbf{A} = \mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1} \ ,$$

which plays an important role in our analysis. The sample analogue of $\mathbf{D}^*$ is given by $\mathbf{D}_n^* = \mathbf{C}_n^{-1/2}\mathbf{M}_n\mathbf{D} \xrightarrow{p} \mathbf{D}^*$ as $n \longrightarrow +\infty$.

The method we shall use to derive the asymptotic expansion of the density function of the standardized estimator $\hat{\mathbf{e}}$ is similar to the one used in Fujikoshi et. al. (1982) and Anderson et. al. (1986). We shall expand $\hat{\mathbf{e}}$ by the perturbation method in terms of the components of random matrices $\mathbf{X}_n = (X_n^{(j)}), \mathbf{Y}_n = (Y_n^{(j,k)}), \mathbf{Z}_n = (Z_n^{(j,k)})$ and $\mathbf{U}_n = (U_n^{(jk)})$ which are defined by

$$(3.8)\ \ \mathbf{Y}_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(u_i^2 - \mathbf{E}(u_i^2)), \mathbf{Z}_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{v}_{2i}^{'}, \mathbf{0}^{'}), \mathbf{U}_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{w}_i\mathbf{z}_i^{'}\ ,$$

where we have $p \times 1$ random vectors $\mathbf{w}_i = (\mathbf{v}_{2i}^{'}, \mathbf{0}^{'})^{'} - \mathbf{q}_i u_i + O_p(n^{-1})$ and $\mathbf{q}_i^{'} = \mathbf{E}[\begin{pmatrix} \mathbf{v}_{2i} \\ \mathbf{0} \end{pmatrix}^{'} u_i]/\mathbf{E}[u_i^2]$ $(i = 1, \cdots, n)$. In order to facilitate our analysis, we assume that there exists a vector $\mathbf{q}$ such that

$$(3.9)\qquad \mathbf{q}_i = \mathbf{q} + O(\frac{1}{n})\ \ (i = 1, \cdots, n)$$

in the following analysis which can be relaxed.
Then if $\mathbf{E}[|u_i|^s] < \infty$ for $s \geq 3$, we can take a positive (bounded) constant $c_n(1, n)$ depending on n which satisfies

$$(3.10)\qquad \mathbf{P}(\|\mathbf{X}_n\| > [(s-1)\Lambda_n(1)\log n]^{1/2}) \leq c_n(1, s)\frac{(1/\sqrt{n})^{s-2}}{(\log n)^{s/2}}\ ,$$

where $\Lambda_n(1)$ as the maximum of the characteristic roots of $\mathbf{C}_n$. Also for the random variables $\mathbf{Y}_n, \mathbf{Z}_n$ and $\mathbf{U}_n$ we can take positive constants $c_n(i, s)$ $(i = 2, 3, 4)$ and similar inequalities for $s \geq 3$ under **Assumption I** with Condition (3.9). These arguments on the validity of the asymptotic expansions of random variables have been given by Bhattacharya and Rao (1976), and Bhattacharya and Ghosh (1978) for the i.i.d. random vector sequences. They can be easily extended to the present situation while their derivations and resulting explanations become quite lengthy. The validity of the asymptotic expansions based on the inversion of the characteristic functions, which will be utilized in Section 4 of this paper, was also briefly discussed by Fujikoshi et. al. (1982). In the econometric literatures the asymptotic expansion method has been previously discussed by Sargan and Mikhail (1971), and Phillips (1983), for instance.

By expanding (2.7) and (2.8) with respect to $\mathbf{e}_0$, formally we can write

$$(3.11)\qquad \hat{\mathbf{e}} = \tilde{\mathbf{e}}_0 + [\mathbf{e}_0 - \tilde{\mathbf{e}}_0] + \frac{1}{\sqrt{n}}\mathbf{e}_1 + \frac{1}{n}\mathbf{e}_2 + o_p(\frac{1}{n})\ ,$$

and

$$(3.12)\qquad \sqrt{n}\hat{\lambda} = \lambda_0 + \frac{1}{\sqrt{n}}\lambda_1 + \frac{1}{n}\lambda_2 + o_p(\frac{1}{n})\ .$$

By substituting these expansions into p_i $(i = 1, \cdots, n)$, we can also expand the estimated probability function as

$$(3.13)\qquad n\ \hat{p}_i = 1 + \frac{1}{\sqrt{n}}p_i^{(1)} + \frac{1}{n}p_i^{(2)} + o_p(\frac{1}{n})\ ,$$

where $p_i^{(1)} = -\lambda_0^{'}\mathbf{z}_i u_i$ and

$$p_i^{(2)} = -\lambda_1^{'}\mathbf{z}_i[u_i - \frac{1}{\sqrt{n}}(\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\mathbf{e}_0] + \lambda_0^{'}\mathbf{z}_i[(\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})(\mathbf{e}_0 + \frac{1}{\sqrt{n}}\mathbf{e}_1)] + (\lambda_0^{'}\mathbf{z}_i)^2[u_i - \frac{1}{\sqrt{n}}(\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\mathbf{e}_0]^2.$$

In these representations it is possible to show that $\max_{1\le i\le n}|\hat{p}_i - 1/n| = o_p(1/n)$ since $(n\hat{p}_i)^{-1} = 1 + \boldsymbol{\lambda}'\mathbf{z}_i u_i$ (see Owen (1990)). By using the similar arguments, we have $\max_{1\le i\le n}|\hat{p}_i - 1/n - p_i^{(1)}/(n\sqrt{n}) - p_i^{(2)}/n^2| = o_p(1/n^2)$. By using the representation $(\mathbf{y}_{2i}', \mathbf{z}_{1i}') = \mathbf{z}_i'\mathbf{D} + \mathbf{w}_i' + \mathbf{q}'u_i$ $(i = 1, \cdots, n)$ and recursive substitutions, we expand

$$(3.14) \qquad \hat{\mathbf{C}}_n = \sum_{i=1}^n \hat{p}_i u_i^2(\hat{\theta})\mathbf{z}_i\mathbf{z}_i' = \mathbf{C}_n + \frac{1}{\sqrt{n}}\mathbf{C}_n^{(1)} + \frac{1}{n}\mathbf{C}_n^{(2)} + o_p(\frac{1}{n}) ,$$

$$(3.15) \qquad \hat{\mathbf{E}}_n = \sum_{i=1}^n \hat{p}_i \begin{pmatrix} \mathbf{y}_{2i} \\ \mathbf{z}_{1i} \end{pmatrix} \mathbf{z}_i' = \mathbf{E}_n^{(0)} + \frac{1}{\sqrt{n}}\mathbf{E}_n^{(1)} + \frac{1}{n}\mathbf{E}_n^{(2)} + o_p(\frac{1}{n}) ,$$

where we have recursively defined the random matrices as $\mathbf{C}_n^{(0)} = \mathbf{C}_n$, $\mathbf{E}_n^{(0)} = \mathbf{D}'\mathbf{M}_n$, and

$$\mathbf{C}_n^{(1)} = \frac{1}{n}\sum_{i=1}^n \mathbf{z}_i\mathbf{z}_i'[p_i^{(1)}u_i^2 - 2u_i(\mathbf{y}_{2i}', \mathbf{z}_{1i}')\mathbf{e}_0] ,$$

$$\mathbf{C}_n^{(2)} = \frac{1}{n}\sum_{i=1}^n \mathbf{z}_i\mathbf{z}_i'[\{(\mathbf{y}_{2i}', \mathbf{z}_{1i}')\mathbf{e}_0\}^2 - 2u_i(\mathbf{y}_{2i}', \mathbf{z}_{1i}')\mathbf{e}_1 - 2u_i p_i^{(1)}(\mathbf{y}_{2i}', \mathbf{z}_{1i}')\mathbf{e}_0 + u_i^2 p_i^{(2)}],$$

$$\mathbf{E}_n^{(1)} = \mathbf{Z}_n' + \mathbf{D}'\frac{1}{n}\sum_{i=1}^n p_i^{(1)}\mathbf{z}_i\mathbf{z}_i' + \frac{1}{n}\sum_{i=1}^n p_i^{(1)} \begin{pmatrix} \mathbf{v}_{2i} \\ \mathbf{0} \end{pmatrix} \mathbf{z}_i' ,$$

$$\mathbf{E}_n^{(2)} = \mathbf{D}'\frac{1}{n}\sum_{i=1}^n p_i^{(2)}\mathbf{z}_i\mathbf{z}_i' + \frac{1}{n}\sum_{i=1}^n p_i^{(2)} \begin{pmatrix} \mathbf{v}_{2i} \\ \mathbf{0} \end{pmatrix} \mathbf{z}_i' .$$

By substituting the above expressions into (3.1) for $\hat{\mathbf{e}}$, $\hat{\lambda}$, and $\hat{p}_i$ $(i = 1, \cdots, n)$, we can determine each terms of the stochastic expansions of $\hat{\mathbf{e}}$ in the recursive way. The leading two terms are given by

$$(3.16) \qquad \mathbf{e}_1 = -\mathbf{Q}_n\mathbf{D}'\mathbf{M}_n\mathbf{C}_n^{-1}\mathbf{Z}_n\mathbf{e}_0 + \mathbf{Q}_n[\mathbf{A}_1][\mathbf{X}_n - \mathbf{M}\mathbf{D}'\mathbf{e}_0],$$

$$(3.17) \qquad \mathbf{e}_2 = \mathbf{Q}_n[\mathbf{A}_2][\mathbf{X}_n - \mathbf{M}_n\mathbf{D}'\mathbf{e}_0] - \mathbf{Q}_n[\mathbf{A}_1][\mathbf{M}_n\mathbf{D}\mathbf{e}_1 + \mathbf{Z}_n\mathbf{e}_0] - \mathbf{Q}_n\mathbf{D}'\mathbf{M}_n\mathbf{C}_n^{-1}\mathbf{Z}_n\mathbf{e}_1,$$

where we have used the corresponding random matrices $\mathbf{Q}_n^{-1} = \mathbf{D}'\mathbf{M}_n\mathbf{C}_n^{-1}\mathbf{M}_n\mathbf{D}$,

$$\mathbf{A}_1 = -\mathbf{D}'\mathbf{M}_n\mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1} + \mathbf{E}_n^{(1)}\mathbf{C}_n^{-1} ,$$

$$\mathbf{A}_2 = \mathbf{D}'\mathbf{M}_n[-\mathbf{C}_n^{-1}\mathbf{C}_n^{(2)}\mathbf{C}_n^{-1} + \mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1}] - \mathbf{E}_n^{(1)}\mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1} + \mathbf{E}_n^{(2)}\mathbf{C}_n^{-1}.$$

3.2 Effects of $\mathbf{C}_n$

We need to investigate the effects of estimating the variance-covariance matrix $\mathbf{C}$ by $\hat{\mathbf{C}}_n$ in the semi-parametric estimation methods. We notice that under a set of regularity conditions, each components of $\mathbf{Y}_n$ have the asymptotic normality as $n \to +\infty$. The covariance of the (j,k)-th elements of $\mathbf{Y}_n$ and the l-th element of $\mathbf{X}_n$ is given by

$$(3.18) \quad Cov(\frac{1}{\sqrt{n}}\sum_{i=1}^n z_i^{(j)}z_i^{(k)}(u_i^2 - \sigma^2),\ \frac{1}{\sqrt{n}}\sum_{i=1}^n z_i^{(l)}u_i) = \frac{1}{n}\sum_{i=1}^n \mathbf{E}(u_i^3)z_i^{(j)}z_i^{(k)}z_i^{(l)} .$$

From this relation we see that $\mathbf{X}_n$ and $\mathbf{Y}_n$ are asymptotically independent when $\mathbf{E}(u_i^3) = 0$ $(i = 1, \cdots, n)$ and then our analyses can be simplified considerably in this situation. It is the case when the disturbances are normally distributed, for instance.

Next, we expand the inverse of the asymptotic variance-covariance matrix of efficient estimators as

$$\begin{aligned}\mathbf{Q}_n^{-1} &= \mathbf{D}^{'}\mathbf{M}_n[\mathbf{C}^{-1}+\mathbf{C}_n^{-1}(\mathbf{C}-\mathbf{C}_n)\mathbf{C}]\mathbf{M}_n\mathbf{D} \\ &= \mathbf{Q}^{-1}-\frac{1}{\sqrt{n}}[\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}]+\frac{1}{n}[\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}]+O_p(\frac{1}{n^{3/2}})\end{aligned}$$

and

$$\mathbf{D}^{'}\mathbf{M}_n\mathbf{C}_n^{-1}=\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}-\frac{1}{\sqrt{n}}[\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}]+\frac{1}{n}[\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}]+O_p(\frac{1}{n^{3/2}}),$$

(3.19)

where $\mathbf{Q}=(\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{D})^{-1}$.
Then by using the inversion of the matrix $\mathbf{Q}_n$, we can represent

$$\begin{aligned}\mathbf{Q}_n &= \mathbf{Q}+\mathbf{Q}_n(\mathbf{Q}^{-1}-\mathbf{Q}_n^{-1})\mathbf{Q} \\ &= \mathbf{Q}+\frac{1}{\sqrt{n}}[\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}] \\ &\quad +\frac{1}{n}[-\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q} \\ &\quad +\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}]+O_p(\frac{1}{n^{3/2}}),\end{aligned}$$

and

$$\begin{aligned}&\mathbf{Q}_n\mathbf{D}^{'}\mathbf{M}_n\mathbf{C}_n^{-1} \\ =\ &\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}+\frac{1}{\sqrt{n}}[-\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}]+\frac{1}{n}[\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{Y}_n\mathbf{A}]+O_p(\frac{1}{n^{3/2}}),\end{aligned}$$

where we have $\mathbf{A}=\mathbf{C}^{-1/2}\bar{\mathbf{P}}_{D^*}\mathbf{C}^{-1/2}=\mathbf{C}^{-1}-\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}$ and $\mathbf{A}\mathbf{C}\mathbf{A}=\mathbf{A}$.
By using the above relations, the first term of the stochastic expansion of $\hat{\mathbf{e}}$ can be represented as

(3.20) $$\mathbf{e}_0=\tilde{\mathbf{e}}_0+\frac{1}{\sqrt{n}}\mathbf{e}_0^{(1)}+\frac{1}{n}\mathbf{e}_0^{(2)}+O_p(\frac{1}{n^{3/2}})\ ,$$

where $\tilde{\mathbf{e}}_0=\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{X}_n$, $\mathbf{e}_0^{(1)}=-\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n$ and $\mathbf{e}_0^{(2)}=\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n$.
Also by using the stochastic expansions of $\mathbf{C}_n$ and $\mathbf{Q}_n$, we can derive a representation of $\boldsymbol{\lambda}_0$ $(=[\mathbf{C}_n^{-1}-\mathbf{C}_n^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}_n\mathbf{D}^{'}\mathbf{M}\mathbf{C}_n^{-1}]\mathbf{X}_n)$ as

(3.21) $$\boldsymbol{\lambda}_0=\mathbf{A}\mathbf{X}_n+\frac{1}{\sqrt{n}}[-\mathbf{A}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n]+O_p(\frac{1}{n})\ .$$

The above formulas help simplifying our derivations considerably.

3.3 Terms involving $\mathbf{e}_1$

We shall investigate each terms involving $\mathbf{e}_1$. For this purpose we decompose $\mathbf{e}_1$ as $\mathbf{e}_1=\mathbf{e}_{1.1}+\mathbf{e}_{1.2}+\mathbf{e}_{1.3}$, where $\mathbf{e}_{1.1}=\mathbf{Q}_n[\mathbf{A}_1][\mathbf{X}_n-\mathbf{M}\mathbf{D}\mathbf{e}_0]$, $\mathbf{e}_{1.2}=-\mathbf{e}_0(\mathbf{q}^{'}\mathbf{e}_0)$ and $\mathbf{e}_{1.3}=-\mathbf{Q}_n\mathbf{D}^{'}\mathbf{M}\mathbf{C}_n^{-1}\mathbf{U}_n^{'}\mathbf{e}_0$. The last two terms can be investigated rather easily and we treat these terms first. By using the representations in Section 3.2, we first rewrite

$$\mathbf{e}_{1.2}$$

$$
\begin{aligned}
&= -[\tilde{\mathbf{e}}_0 - \frac{1}{\sqrt{n}}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n]\mathbf{q}^{'}[\tilde{\mathbf{e}}_0 - \frac{1}{\sqrt{n}}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n] + O_p(\frac{1}{n}) \\
&= -\tilde{\mathbf{e}}_0(\mathbf{q}^{'}\tilde{\mathbf{e}}_0) + \frac{1}{\sqrt{n}}[\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n(\mathbf{q}^{'}\tilde{\mathbf{e}}_0) + \tilde{\mathbf{e}}_0\mathbf{q}^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n] + O_p(\frac{1}{n}) \\
&= \mathbf{e}_{1.2}^{(0)} + \frac{1}{\sqrt{n}}\mathbf{e}_{1.2}^{(1)} + O_p(\frac{1}{n}) \ ,
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbf{e}_{1.3} \\
&= -[\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1} - \frac{1}{\sqrt{n}}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}]\mathbf{U}_n^{'}[\tilde{\mathbf{e}}_0 - \frac{1}{\sqrt{n}}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n] + O_p(\frac{1}{n}) \\
&= -\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n^{'}\tilde{\mathbf{e}}_0 \\
&+ \frac{1}{\sqrt{n}}[\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n + \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{U}_n^{'}\tilde{\mathbf{e}}_0] + O_p(\frac{1}{n}) \\
&= \mathbf{e}_{1.3}^{(0)} + \frac{1}{\sqrt{n}}\mathbf{e}_{1.3}^{(1)} + O_p(\frac{1}{n}) \ .
\end{aligned}
$$

In the above expressions we have defined $\mathbf{e}_{1.2}^{(1)}$ and $\mathbf{e}_{1.3}^{(1)}$ implicitly. The analysis of $\mathbf{e}_{1.1}$ becomes substantially more complicated because there are many terms involved in the random matrices $\mathbf{C}_n^{(1)}$ and $\mathbf{E}_n^{(1)}$. We rewrite

$$
\begin{aligned}
&\mathbf{C}_n^{(1)} \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\{-2u_i(\mathbf{z}_i^{'}\mathbf{D} + \mathbf{w}_i^{'} + \mathbf{q}^{\prime}u_i)[\tilde{\mathbf{e}}_0 - \frac{1}{\sqrt{n}}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n] + u_i^3(-\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)\} \\
&= \Big\{ -2(\mathbf{q}^{\prime}\tilde{\mathbf{e}}_0)\mathbf{C} - \frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\lambda_{\mathbf{0}})\Big\} \\
&+ \frac{1}{\sqrt{n}}\Big\{ -2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{Y}_n - 2\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{\prime}u_i(\mathbf{z}_i^{'}\mathbf{D} + \mathbf{w}_i^{'})\tilde{\mathbf{e}}_0 + 2\mathbf{C}_n\mathbf{q}^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n \\
&\qquad - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)(u_i^3 - \mathbf{E}(u_i^3))\Big\} + O_p(\frac{1}{n}).
\end{aligned}
$$

Then we have

$$
\begin{aligned}
&-\mathbf{Q}_n\mathbf{D}^{'}\mathbf{M}_n\mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1} \\
&= \Big\{2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1} + \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)\mathbf{C}^{-1}\Big\} \\
&+ \frac{1}{\sqrt{n}}\Big\{ -2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A} - \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)\mathbf{C}^{-1}\mathbf{Y}_n \\
&\quad -\mathbf{Y}_n\mathbf{A}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)]\mathbf{C}^{-1} \\
&\quad -\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[-\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)(u_i^3 - \mathbf{E}(u_i^3)) - 2\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{\prime}u_i(\mathbf{z}_i^{'}\mathbf{D} + \mathbf{w}_i^{'})\tilde{\mathbf{e}}_0 \\
&\qquad +2\mathbf{C}\mathbf{q}^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n]\mathbf{C}^{-1}\Big\} + O_p(\frac{1}{n}) \ .
\end{aligned}
$$

On the other hand, we represent

$$\mathbf{E}_n^{(1)} = \mathbf{E}_n^{(1.0)} + \frac{1}{\sqrt{n}}\mathbf{E}_n^{(1.1)} + O_p(\frac{1}{n}) \ , \tag{3.22}$$

where $\mathbf{E}_n^{(1.0)} = \mathbf{U}_n + \mathbf{q}(\mathbf{X}_n^{'} - \boldsymbol{\lambda}_0^{'}\mathbf{C}_n)$ and

$$\mathbf{E}_n^{(1.1)} = -\mathbf{D'}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)u_i - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{w}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)u_i \ .$$

Then we have

$$\begin{aligned}
& \mathbf{Q}_n\mathbf{E}_n^{(1)}\mathbf{C}_n^{-1} \\
= \ & \Big\{\mathbf{Q}\mathbf{U}_n\mathbf{C}^{-1} + \mathbf{Q}\mathbf{q}(\mathbf{X}_n^{'} - \boldsymbol{\lambda}_0^{'}\mathbf{C})\mathbf{C}^{-1}\Big\} \\
+ \ & \frac{1}{\sqrt{n}}\Big\{\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}[\mathbf{U}_n + \mathbf{q}(\mathbf{X}_n^{'} - \boldsymbol{\lambda}_0^{'}\mathbf{C}_n)]\mathbf{C}^{-1} \\
& -\mathbf{Q}[\mathbf{U}_n + \mathbf{q}(\mathbf{X}_n^{'} - \boldsymbol{\lambda}_0^{'}\mathbf{C}_n)]\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1} \\
& +\mathbf{Q}[-\mathbf{D'}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\lambda_0)u_i - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{w}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\lambda_0)u_i]\mathbf{C}^{-1} - \mathbf{Q}\mathbf{q}\boldsymbol{\lambda}_0^{'}\mathbf{Y}_n\mathbf{C}^{-1}\Big\} + O_p(\frac{1}{n}).
\end{aligned}$$

By using the algebraic relation $\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1} = \mathbf{C}^{-1} - \mathbf{A}$ and $\mathbf{X}_n - \mathbf{M}_n\mathbf{D}\mathbf{e}_0 = \mathbf{C}\mathbf{A}\mathbf{X}_n + (1/\sqrt{n})[\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n] + O_p(\frac{1}{n})$, we rewrite

$$\begin{aligned}
& \mathbf{e}_{1.1} \\
= \ & \Big\{[2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1} + \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)\mathbf{C}^{-1} \\
& \quad +\mathbf{Q}\mathbf{U}_n\mathbf{C}^{-1} + \mathbf{Q}\mathbf{q}(\mathbf{X}_n^{'} - \lambda_0^{'}\mathbf{C})\mathbf{C}^{-1}][\mathbf{X}_n - \mathbf{M}_n\mathbf{D}\mathbf{e}_0]\Big\} \\
+ \ & \frac{1}{\sqrt{n}}\Big\{\Big(-2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A} \\
& \quad -\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)\mathbf{C}\mathbf{Y}_n + \mathbf{Y}_n\mathbf{A}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)]\mathbf{C}^{-1} \\
& \quad -\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[-2\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{\prime}u_i(\mathbf{z}_i^{'}\mathbf{D} + \mathbf{w}_i^{'})\tilde{\mathbf{e}}_0 - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{\prime}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)(u_i^3 - \mathbf{E}(u_i^3)) \\
& \quad +2\mathbf{C}\mathbf{q}^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n]\mathbf{C}^{-1} \\
& +\mathbf{Q}[-\mathbf{D'}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\lambda_0)u_i - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{w}_i\mathbf{z}_i^{'}(\lambda_0^{'}\mathbf{z}_i)u_i]\mathbf{C}^{-1} - \mathbf{Q}\mathbf{q}\boldsymbol{\lambda}_0^{'}\mathbf{Y}_n\mathbf{C}^{-1} \\
& +\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}[\mathbf{U}_n + \mathbf{q}(\mathbf{X}_n^{'} - \boldsymbol{\lambda}_0^{'}\mathbf{C}_n)]\mathbf{C}^{-1} \\
& -\mathbf{Q}[\mathbf{U}_n + \mathbf{q}(\mathbf{X}_n^{'} - \boldsymbol{\lambda}_0^{'}\mathbf{C}_n)]\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\Big)[\mathbf{X}_n - \mathbf{M}_n\mathbf{D}\tilde{\mathbf{e}}_0]\Big\} \ + \ O_p(\frac{1}{n}).
\end{aligned}$$

In order to derive the asymptotic expansions of the distributions of the class of the modified estimators, we need to use $\hat{p}_i^*$ instead of $\hat{p}_i$ $(i = 1, 2)$. For an arbitrary (fixed) δ $(0 \le \delta \le 1)$, we substitute $\delta\boldsymbol{\lambda}_0$ (and $\delta\boldsymbol{\lambda}_1$) into $\boldsymbol{\lambda}_0$ (and $\boldsymbol{\lambda}_1$) in the stochastic expansion of the MEL estimator. Thus we can further express $\mathbf{e}_{1.1}$ as

$$\mathbf{e}_{1.1} = \mathbf{e}_{1.1}^{(0)} + \frac{1}{\sqrt{n}}\mathbf{e}_{1.1}^{(1)} \ + \ O_p(\frac{1}{n}) \ , \tag{3.23}$$

where

$$\mathbf{e}_{1.1}^{(0)} = \mathbf{QU}_n\mathbf{AX}_n + (1-\delta)\mathbf{Qq}(\mathbf{X}_n^{'}\mathbf{AX}_n) + \delta\mathbf{QD}^{'}\mathbf{MC}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{AX}_n)^2$$

and

$$\begin{aligned}
&\mathbf{e}_{1.1}^{(1)} \\
= &\Big(\delta\mathbf{QD}^{'}\mathbf{MC}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\mathbf{AX}_n)(\mathbf{C}^{-1}-\mathbf{A})\mathbf{Y}_n\mathbf{AX}_n \\
&+\mathbf{QU}_n(\mathbf{C}^{-1}-\mathbf{A})\mathbf{Y}_n\mathbf{AX}_n + \mathbf{Qq}\tilde{\mathbf{e}}_0^{'}\mathbf{D}^{'}\mathbf{MC}^{-1}\mathbf{Y}_n\mathbf{AX}_n\Big) \\
+ &\Big(-\delta\mathbf{QD}^{'}\mathbf{MC}^{-1}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)\mathbf{C}_n\mathbf{Y}_n\mathbf{AX}_n + \mathbf{Y}_n\mathbf{A}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)\mathbf{AX}_n] \\
&-\mathbf{QD}^{'}\mathbf{MC}^{-1}[-2\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{AX}_n u_i(\mathbf{z}_i^{'}\mathbf{D}+\mathbf{w}_i^{'})\tilde{\mathbf{e}}_0 - \delta\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{AX}_n)^2(u_i^3-\mathbf{E}(u_i^3))] \\
&+\mathbf{Q}[-\delta\mathbf{D}'\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{AX}_n)^2u_i - \delta\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{w}_i(\mathbf{z}_i^{'}\mathbf{AX}_n)^2u_i] \\
&-\mathbf{Q}[\mathbf{U}_n + \mathbf{q}(\mathbf{X}_n^{'}-\delta\lambda_0^{'}\mathbf{C}_n)]\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{AX}_n \\
&+\mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{MDQ}[\mathbf{U}_n+\mathbf{q}(\mathbf{X}_n^{'}-\delta\lambda_0^{'}\mathbf{C}_n)]\mathbf{AX}_n\Big) \\
+ &\Big(2\mathbf{q}^{'}\mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{Y}_n\mathbf{AX}_n\mathbf{AX}_n - \delta\mathbf{QD}^{'}\mathbf{MC}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i^{'}\mathbf{AY}_n\mathbf{AX}_n\mathbf{z}_i^{'}\mathbf{z}_i^{'}\mathbf{AX}_n\Big) + O_p(\frac{1}{n}),
\end{aligned}$$

By collecting each terms of $\mathbf{e}_1$, we summarize them as

$$\mathbf{e}_1 = \mathbf{e}_1^{(0)} + \frac{1}{\sqrt{n}}\mathbf{e}_1^{(1)} + O_p(\frac{1}{n}), \tag{3.24}$$

where $\mathbf{e}_1^{(0)} = \mathbf{e}_{1.1}^{(0)} + \mathbf{e}_{1.2}^{(0)} + \mathbf{e}_{1.3}^{(0)}$ and $\mathbf{e}_1^{(1)} = \mathbf{e}_{1.1}^{(1)} + \mathbf{e}_{1.2}^{(1)} + \mathbf{e}_{1.3}^{(1)}$.
Then the leading terms of $\mathbf{e}_1$ can be represented as

$$\begin{aligned}
\mathbf{e}_1^{(0)} \;=\; &\mathbf{QU}_n\mathbf{AX}_n + (1-\delta)\mathbf{QqX}_n^{'}\mathbf{AX}_n + \delta\mathbf{QD}^{'}\mathbf{MC}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{AX}_n)^2 \\
&-\tilde{\mathbf{e}}_0(\mathbf{q}^{'}\tilde{\mathbf{e}}_0) - \mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{U}_n\tilde{\mathbf{e}}_0\ .
\end{aligned}$$

The random vector $\tilde{\mathbf{e_0}}$ is asymptotically normal and it is asymptotically uncorrelated with the random vector $\mathbf{AX}_n$ by using CLT. Then by using Lemma 4.3 in Section 4, given $\tilde{\mathbf{e}}_0 = \mathbf{x}$ the conditional expectation of $\mathbf{e}_1^{(0)}$ is given by [2]

$$\mathbf{E}[\mathbf{e}_1^{(0)}|\tilde{\mathbf{e}}_0 = \mathbf{x}] = (1-\delta)L\mathbf{Qq} - \mathbf{xx}^{'}\mathbf{q} + \delta\mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{m}_3 \;+\; O_p(\frac{1}{\sqrt{n}})\ , \tag{3.25}$$

where

$$\mathbf{m}_3 = \mathrm{plim}_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{Az}_i)\ . \tag{3.26}$$

[2] We need to evaluate the terms of $O_p(n^{-1})$ in order to derive the asymptotic expansions of distributions which shall be done in Section 4. Two terms in $\tilde{\mathbf{e}}_{1.1}^{(1)}$ have some important roles to the final results.

Also the conditional second order moments of $\mathbf{e}_1^{(0)}$ can be calculated as

$$
\begin{aligned}
& \mathbf{E}[\mathbf{e}_1^{(0)}\mathbf{e}_1^{(0)'}|\tilde{\mathbf{e}}_0=\mathbf{x}] \\
= \; & \delta^2\mathbf{E}\{\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i'\mathbf{A}\mathbf{X}_n)^2\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i'(\mathbf{z}_i'\mathbf{A}\mathbf{X}_n)^2\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}|\mathbf{x}\} \\
+ \; & \delta\mathbf{E}\{[\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i'\mathbf{A}\mathbf{X}_n)^2] \\
& \times[\mathbf{Q}\mathbf{U}_n\mathbf{A}\mathbf{X}_n+(1-\delta)\mathbf{Q}\mathbf{q}\mathbf{X}_n'\mathbf{A}\mathbf{X}_n-\tilde{\mathbf{e}}_0(\mathbf{q}'\tilde{\mathbf{e}}_0)-\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n'\tilde{\mathbf{e}}_0]'|\mathbf{x}\} \\
+ \; & \delta\mathbf{E}\{[\mathbf{Q}\mathbf{U}_n\mathbf{A}\mathbf{X}_n+(1-\delta)\mathbf{Q}\mathbf{q}\mathbf{X}_n'\mathbf{A}\mathbf{X}_n-\tilde{\mathbf{e}}_0(\mathbf{q}'\tilde{\mathbf{e}}_0)-\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n\mathbf{e}_0] \\
& \times[\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i'(\mathbf{z}_i'\mathbf{A}\mathbf{X}_n)^2\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}]|\mathbf{x}\} \\
+ \; & \mathbf{E}\{[\mathbf{Q}\mathbf{U}_n\mathbf{A}\mathbf{X}_n+(1-\delta)\mathbf{Q}\mathbf{q}\mathbf{X}_n'\mathbf{A}\mathbf{X}_n-\tilde{\mathbf{e}}_0(\mathbf{q}'\tilde{\mathbf{e}}_0)-\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n'\tilde{\mathbf{e}}_0] \\
& \times[\mathbf{Q}\mathbf{U}_n\mathbf{A}\mathbf{X}_n+(1-\delta)\mathbf{Q}\mathbf{q}\mathbf{X}_n'\mathbf{A}\mathbf{X}_n-\tilde{\mathbf{e}}_0(\mathbf{q}'\tilde{\mathbf{e}}_0)-\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n\tilde{\mathbf{e}}_0]|\mathbf{x}\} \\
= \; & \delta^2\Big\{\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3\mathbf{m}_3'\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}+2\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}(\frac{1}{n})^2\sum_{i,j}(\mathbf{E}(u_i^3))^2\mathbf{z}_i\mathbf{z}_j'(\mathbf{z}_i'\mathbf{A}\mathbf{z}_j)^2\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\Big\} \\
+ \; & \delta\{\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3[(1-\delta)(L+2)\mathbf{Q}\mathbf{q}-\mathbf{x}\mathbf{x}'\mathbf{q}]'+[(1-\delta)(L+2)\mathbf{Q}\mathbf{q}-\mathbf{x}\mathbf{x}'\mathbf{q}]\mathbf{m}_3'\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\Big\} \\
+ \; & \Big\{(\mathbf{x}'\mathbf{C}_1^*\mathbf{x}'\mathbf{x}\mathbf{x}'+\mathbf{Q}\mathbf{Q}^*\mathbf{Q}\mathbf{x}'(\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(\mathbf{w}_i\mathbf{w}_i'))\mathbf{x}+\mathbf{Q}'(\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(\mathbf{w}_i\mathbf{w}_i'))\mathbf{Q}\mathbf{tr}(\mathbf{A}\mathbf{M}) \\
& \qquad +(1-\delta)^2L(L+2)\mathbf{Q}\mathbf{C}_1^*\mathbf{Q}-(1-\delta)L[\mathbf{Q}\mathbf{C}_1^*\mathbf{x}\mathbf{x}'+\mathbf{x}\mathbf{x}'\mathbf{C}_1^*\mathbf{Q}]\Big\}+O_p(\frac{1}{\sqrt{n}}),
\end{aligned}
$$

where we use the notations $\mathbf{C}_1^*=\mathbf{q}\mathbf{q}'$ and $\mathbf{Q}^*=\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}$. In the above calculations we have used some relations on moments by applying Lemma 4.2 in Section 4 as $\mathbf{E}[(\mathbf{X}_n'\mathbf{A}\mathbf{X}_n)^2]=L(L+2)+O(n^{-1/2})$ and $\mathbf{E}[(\mathbf{z}_i'\mathbf{A}\mathbf{X}_n)^2(\mathbf{X}_n'\mathbf{A}\mathbf{X}_n)]=(L+2)\mathbf{z}_i'\mathbf{A}\mathbf{z}_i+O(n^{-1/2})$. It is because we use the fact that the limiting random vectors $\tilde{\mathbf{e}}_0$ and $\mathbf{A}\mathbf{X}_n$ are asymptotically normal and uncorrelated.

3.4 Terms involving $\mathbf{e}_2$

We shall investigate the terms associated with $\mathbf{e}_2$. For this purpose we decompose $\mathbf{e}_2=\mathbf{e}_{2.1}+\mathbf{e}_{2.2}+\mathbf{e}_{2.3}$ and $\mathbf{e}_{2.i}$ $(i=1,2,3)$ correspond to each terms of (3.17). Because we can estimate $\mathbf{Q}$ and $\mathbf{C}$ consistently by using $\mathbf{Q}_n$ and $\mathbf{C}_n$, their estimates do not affect the terms much involving $\mathbf{e}_2$. We first consider $\mathbf{e}_{2.3}$ and by using the stochastic expansion of $\mathbf{e}_1$, we have

$$
\mathbf{e}_{2.3}=-\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}[\mathbf{U}_n'+\mathbf{X}_n\mathbf{q}']\mathbf{e}_1^{(0)}+O_p(\frac{1}{\sqrt{n}})\ . \tag{3.27}
$$

Because $\tilde{\mathbf{e}}_0$ and $\mathbf{A}\mathbf{X}_n$ are asymptotically orthogonal, the conditional expectation of the first term, $\mathbf{e}_{2.3.1}=-\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n\mathbf{q}'\mathbf{e}_1^{(0)}$, given $\tilde{\mathbf{e}}_0=\mathbf{x}$ can be calculated as

$$
-\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}(\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i')\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{E}(\mathbf{w}_i\mathbf{w}_i')\mathbf{x}+O_p(\frac{1}{\sqrt{n}})\ .
$$

The second term of $\mathbf{e}_{2.3}$ can be also expressed as

$\mathbf{e}_{2.3.2}$

$$= -(\tilde{\mathbf{e}}_0\mathbf{q}^{'})[\mathbf{Q}\mathbf{U}_n\mathbf{A}\mathbf{X}_n + (1-\delta)\mathbf{Q}\mathbf{q}\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n + \delta\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n)^2 - \tilde{\mathbf{e}}_0(\mathbf{q}^{'}\tilde{\mathbf{e}}_0) - \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n\tilde{\mathbf{e}}_0] .$$

Hence its conditional expectation given $\tilde{\mathbf{e}}_0 = \mathbf{x}$ can be expressed as

$$-(1-\delta)L\mathbf{q}^{'}\mathbf{Q}\mathbf{q}\mathbf{x} + \mathbf{x}(\mathbf{q}^{'}\mathbf{x})^2 - \delta\mathbf{x}\mathbf{q}^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i) + O_p(\frac{1}{\sqrt{n}}).$$

By combining these terms, we have

$$(3.28)\quad \mathbf{E}[\mathbf{e}_{2.3}|\tilde{\mathbf{e}}_0 = \mathbf{x}] = \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}(\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'})\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{E}(\mathbf{w}_i\mathbf{w}_i^{'})\mathbf{x} + \mathbf{x}\mathbf{x}^{'}\mathbf{C}_1^{*}\mathbf{x} - (1-\delta)L\mathbf{tr}(\mathbf{C}_1^{*}\mathbf{Q})\mathbf{x} - \delta\mathbf{x}\mathbf{q}^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3 + O_p(\frac{1}{\sqrt{n}}) .$$

Secondly, we shall evaluate the terms involving $\mathbf{e}_{2.2}$. For this purpose we notice that we only need to investigate two terms as $\mathbf{e}_{2.2.1} = -\mathbf{Q}[\mathbf{A}_1]\mathbf{M}\mathbf{D}\mathbf{e}_1^{(0)}$ and $\mathbf{e}_{2.2.2} = -\mathbf{Q}[\mathbf{A}_1][\mathbf{U}_n^{'} + \mathbf{X}_n\mathbf{q}^{'}]\tilde{\mathbf{e}}_0$ because it is straightforward to show that the rest of terms $\mathbf{e}_{2.2} - \mathbf{e}_{2.2.1} - \mathbf{e}_{2.2.2} = O_p(n^{-1/2})$. The second term can be further rewritten as

$$\mathbf{e}_{2.2.2} = -\{2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1} + \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\lambda_0)\mathbf{C}^{-1} + \mathbf{Q}_n\mathbf{U}_n + \mathbf{Q}\mathbf{q}(\mathbf{X}_n^{'} - \lambda_0^{'}\mathbf{C}_n)\mathbf{C}^{-1}\}[\mathbf{U}_n^{'} + \mathbf{X}_n\mathbf{q}^{'}]\tilde{\mathbf{e}}_0 + O_p(\frac{1}{\sqrt{n}}) .$$

Then the conditional expectation is given by

$$\mathbf{E}[\mathbf{e}_{2.2.2}|\tilde{\mathbf{e}}_0 = \mathbf{x}] = -2\mathbf{q}^{'}\mathbf{x}\mathbf{x}(\mathbf{q}^{'}\mathbf{x}) - \mathbf{Q}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(\mathbf{w}_i\mathbf{w}_i^{'})\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{z}_i]\mathbf{x} - \mathbf{Q}\mathbf{q}\mathbf{E}[\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{X}_n|\mathbf{x}] - \delta\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)](\mathbf{q}^{'}\mathbf{x}) + \delta\mathbf{Q}\mathbf{q}\mathbf{E}[\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n|\mathbf{x}]\mathbf{q}^{'}\tilde{\mathbf{e}}_0 + O_p(\frac{1}{\sqrt{n}}) .$$

Now we use the decomposition $\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{X}_n = \mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n + \tilde{\mathbf{e}}_0^{'}\mathbf{Q}^{-1}\tilde{\mathbf{e}}_0$, and also we notice the relation

$$\mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\lambda_0)\mathbf{C}^{-1}\mathbf{X}_n|\mathbf{x}] = \delta\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i) + O_p(\frac{1}{\sqrt{n}}) .$$

Then the conditional expectation can be rewritten as

$$(3.29)\quad \mathbf{E}[\mathbf{e}_{2.2.2}|\mathbf{x}] = -2\mathbf{x}\mathbf{x}^{'}\mathbf{C}_1^{*}\mathbf{x} - [\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(\mathbf{w}_i\mathbf{w}_i^{'})\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{z}_i]\mathbf{x} - \mathbf{Q}\mathbf{C}_1^{*}\mathbf{x}[L\mathbf{I}_K + \mathbf{x}^{'}\mathbf{Q}^{-1}\mathbf{x}] + \delta L\mathbf{Q}\mathbf{C}_1^{*}\mathbf{x} - \delta\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3(\mathbf{q}^{'}\mathbf{x}) + O_p(\frac{1}{\sqrt{n}}) .$$

On the other hand, the first term of $\mathbf{e}_{2.2}$ can be expressed as

$$\begin{aligned}
\mathbf{e}_{2.2.1} &= -\Big\{2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1} + \delta\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\lambda_0)\mathbf{C}^{-1} \\
&\qquad +\mathbf{Q}_n\mathbf{U}_n\mathbf{C}^{-1} + \mathbf{Q}\mathbf{q}(\mathbf{X}_n^{'} - \delta\boldsymbol{\lambda}_0^{'}\mathbf{C}_n)\mathbf{C}^{-1}\Big\}\mathbf{M}\mathbf{D} \\
&\times \Big\{\mathbf{Q}\mathbf{U}_n\mathbf{A}\mathbf{X}_n + (1-\delta)\mathbf{Q}\mathbf{q}\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n + \delta\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2 \\
&\qquad -\tilde{\mathbf{e}}_0(\mathbf{q}^{'}\tilde{\mathbf{e}}_0) - \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n^{'}\tilde{\mathbf{e}}_0\Big\} + O_p(\frac{1}{\sqrt{n}})\ .
\end{aligned}$$

Given $\tilde{\mathbf{e}}_0 = \mathbf{x}$, the conditional expectation is given by

$$\begin{aligned}
& \mathbf{E}[\mathbf{e}_{2.2.1}|\tilde{\mathbf{e}}_0 = \mathbf{x}] \\
= \ & [-2(1-\delta)L(\mathbf{q}^{'}\mathbf{x})\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{q} - 2\delta(\mathbf{q}^{'}\mathbf{x})\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3] \\
& -\delta\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{E}[(\mathbf{z}_i^{'}\lambda_0)\mathbf{C}^{-1}\mathbf{M}\mathbf{D}|\mathbf{x}] \\
& \quad \times[(1-\delta)L\mathbf{Q}\mathbf{q} + \delta\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)] \\
& -\mathbf{Q}\mathbf{q}\mathbf{E}[\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}((1-\delta)\mathbf{Q}\mathbf{q}L + \delta\mathbf{Q}\mathbf{D}^{'}\mathbf{m}_3)|\mathbf{x}] \\
& +2(\mathbf{q}^{'}\mathbf{x})^2\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{x} + \mathbf{Q}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}[\mathbf{w}_i\mathbf{w}_i^{'}]\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{z}_i\mathbf{x} \\
& +\mathbf{Q}\mathbf{q}(\mathbf{q}^{'}\mathbf{e}_0)\mathbf{E}[\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\tilde{\mathbf{e}}_0 - \delta\mathbf{X}_n^{'}\mathbf{A}\mathbf{M}\mathbf{D}\tilde{\mathbf{e}}_0|\mathbf{x}]\ .
\end{aligned}$$

In order to evaluate each terms in the above expression, we use the relations $\mathbf{A}\mathbf{C}\mathbf{M}\mathbf{D} = \mathbf{O}$, $\mathbf{E}[\mathbf{X}_n^{'}\mathbf{A}\tilde{\mathbf{e}}_0|\tilde{\mathbf{e}}_0 = \mathbf{x}] = \mathbf{O}_p(n^{-1/2})$, and

$$\begin{aligned}
\mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\lambda_0)\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}|\tilde{\mathbf{e}}_0 = \mathbf{x}] &= \mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{X}_n^{'}\mathbf{A}\mathbf{z}_i|\mathbf{x}] = O_p(\frac{1}{\sqrt{n}}), \\
\mathbf{E}[\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\tilde{\mathbf{e}}_0|\tilde{\mathbf{e}}_0 = \mathbf{x}] &= \mathbf{E}[\mathbf{X}_n^{'}\mathbf{A}\mathbf{M}\mathbf{D}\tilde{\mathbf{e}}_0 + \mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\tilde{\mathbf{e}}_0|\mathbf{x}] \\
&= \tilde{\mathbf{e}}_0^{'}\mathbf{Q}^{-1}\tilde{\mathbf{e}}_0\ .
\end{aligned}$$

Then we have the conditional expectation

(3.30)
$$\begin{aligned}
& \mathbf{E}[\mathbf{e}_{2.2.1}|\tilde{\mathbf{e}}_0 = \mathbf{x}] \\
= \ & 2(\mathbf{q}^{'}\mathbf{x})^2\mathbf{x} + \mathbf{Q}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}[\mathbf{w}_i\mathbf{w}_i^{'}]\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{z}_i\mathbf{x} \\
+ \ & \mathbf{Q}\mathbf{q}\mathbf{q}^{'}\mathbf{x}\mathbf{x}^{'}\mathbf{Q}^{-1}\mathbf{x} - 3(1-\delta)L\mathbf{Q}\mathbf{C}_1^{*}\mathbf{x} \\
- \ & 2\delta(\mathbf{q}^{'}\mathbf{x})\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3 - \delta\mathbf{Q}\mathbf{q}\mathbf{x}^{'}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3 + O_p(\frac{1}{\sqrt{n}}).
\end{aligned}$$

Hence we obtain the explicit expression of the conditional expectation $\mathbf{E}[\mathbf{e}_{2.2}|\tilde{\mathbf{e}}_0 = \mathbf{x}] = \mathbf{E}[\mathbf{e}_{2.2.1}|\mathbf{x}] + \mathbf{E}[\mathbf{e}_{2.2.2}|\mathbf{x}]$ up to $O_p(n^{-1/2})$.

Thirdly, we shall evaluate all terms involving $\mathbf{e}_{2.1}$ which is the first term of (3.17) and we need more complicated evaluations. We notice that it is asymptotically equivalent to

(3.31)
$$\mathbf{e}_{2.1}^{*} = \mathbf{e}_{2.1}(A) + \mathbf{e}_{2.1}(B) + \mathbf{e}_{2.1}(C) + \mathbf{e}_{2.1}(D)\ ,$$

where we use the notations $\mathbf{e}_{2.1}(A) = -\mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{C}_n^{(2)}\mathbf{AX}_n$, $\mathbf{e}_{2.1}(B) = \mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{C}_n^{(1)}\mathbf{C}^{-1}\mathbf{C}_n^{(1)}\mathbf{AX}_n$, $\mathbf{e}_{2.1}(C) = -\mathbf{QE}_n^{(1)}\mathbf{C}^{-1}\mathbf{C}_n^{(1)}\mathbf{AX}_n$ and $\mathbf{e}_{2.1}(D) = \mathbf{QE}_n^{(2)}\mathbf{AX}_n$. Because the above terms contain $p_i^{(2)}$ $(i = 1, \cdots, n)$, we need to use the explicit expression of $\boldsymbol{\lambda}_1$, which is the solution of the equation

$$\begin{aligned}
&\boldsymbol{\lambda}_0 + \frac{1}{\sqrt{n}}\boldsymbol{\lambda}_1 + O_p(n^{-1}) \\
= \;& \{\mathbf{C}_n^{-1} + \frac{1}{\sqrt{n}}[-\mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1}]\}\{\mathbf{X}_n - \mathbf{M}_n\mathbf{De}_0] + \frac{1}{\sqrt{n}}[-\mathbf{M}_n\mathbf{De}_1 - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{v}_{2i}^{'}, \mathbf{0})\mathbf{e}_0]\}.
\end{aligned}$$

In the above representation we have used the representation of residuals as $u_i(\hat{\theta}) = y_{1i} - (\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\hat{\boldsymbol{\theta}} = u_i - (\frac{1}{\sqrt{n}})[\mathbf{z}_i^{'}\mathbf{D} + (\mathbf{v}_{2i}^{'}, \mathbf{0}^{'})]\hat{\mathbf{e}}$. Then by using the stochastic expansion of $\mathbf{C}_n^{(1)}$, we have the representation as

$$\begin{aligned}
\boldsymbol{\lambda}_1 &= -\mathbf{C}^{-1}\mathbf{MDe}_1^{(0)} - \mathbf{C}^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{v}_{2i}^{'}, \mathbf{0}^{'})\mathbf{e}_0 - \mathbf{C}^{-1}\mathbf{C}_n^{(1)}\mathbf{AX}_n + O_p(\frac{1}{\sqrt{n}}) \\
&= -\mathbf{C}^{-1}\mathbf{U}_n^{'}\tilde{\mathbf{e}}_0 - \mathbf{C}^{-1}\mathbf{MDQU}_n\mathbf{AX}_n + \mathbf{C}^{-1}\mathbf{MDQD}^{'}\mathbf{MC}^{-1}\mathbf{U}_n^{'}\tilde{\mathbf{e}}_0 \\
&\quad +\mathbf{C}^{-1}\mathbf{MD}\tilde{\mathbf{e}}_0(\mathbf{q}^{'}\tilde{\mathbf{e}}_0) + (\mathbf{q}^{'}\tilde{\mathbf{e}}_0)[2\mathbf{AX}_n - \mathbf{C}^{-1}\mathbf{X}_n] - (1-\delta)\mathbf{C}^{-1}\mathbf{MDQqX}_n^{'}\mathbf{AX}_n \\
&\quad -\delta\mathbf{C}^{-1}\mathbf{MDQD}^{'}\mathbf{MC}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{AX}_n)^2 + \delta\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{AX}_n)^2 \\
&\quad + O_p(\frac{1}{\sqrt{n}}).
\end{aligned}$$

By using the relations $\mathbf{C}^{-1}\mathbf{X}_n = \mathbf{AX}_n + \mathbf{C}^{-1}\mathbf{MDQD}^{'}\mathbf{MC}^{-1}\mathbf{X}_n$ and $2\mathbf{AX}_n - \mathbf{C}^{-1}\mathbf{X}_n = \mathbf{AX}_n - \mathbf{C}^{-1}\mathbf{MD}\tilde{\mathbf{e}}_0$, we have

$$\begin{aligned}
\boldsymbol{\lambda}_1 &= -\mathbf{AU}_n^{'}\tilde{\mathbf{e}}_0 - \mathbf{C}^{-1}\mathbf{MDQU}_n\mathbf{AX}_n + (\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{AX}_n \\
&\quad -(1-\delta)\mathbf{C}^{-1}\mathbf{MDQqX}_n^{'}\mathbf{AX}_n + \delta\mathbf{A}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{AX}_n)^2 \\
&\quad +O_p(\frac{1}{\sqrt{n}}).
\end{aligned}$$

Then we need to evaluate each terms by using the expression for $\mathbf{e}_1$ and $\boldsymbol{\lambda}_1$. By using the explicit representation of $p_i^{(1)}$ $(i = 1, \cdots, n)$, it can be written as

$$\begin{aligned}
&\mathbf{e}_{2.1}(A) \\
= \;& -\mathbf{QD}^{'}\mathbf{MC}^{-1}\Big\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}[(\mathbf{z}_i^{'}\mathbf{D}\tilde{\mathbf{e}}_0 + \mathbf{w}_i^{'}\tilde{\mathbf{e}}_0 + u_i\mathbf{q}^{'}\tilde{\mathbf{e}}_0)^2 \\
&\quad -2(u_i\mathbf{z}_i^{'}\mathbf{De}_1 + u_i\mathbf{w}_i^{'}\mathbf{e}_1 + u_i^2\mathbf{q}^{'}\mathbf{e}_1) + 2u_i^2(\mathbf{z}_i^{'}\lambda_0)(\mathbf{z}_i^{'}\mathbf{D}\tilde{\mathbf{e}}_0 + \mathbf{w}_i^{'}\tilde{\mathbf{e}}_0 + u_i\mathbf{q}^{'}\tilde{\mathbf{e}}_0) + u_i^2p_i^{(2)}]\Big\}\mathbf{AX}_n \\
= \;& -\mathbf{QD}^{'}\mathbf{MC}^{-1}\Big\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}[(\mathbf{z}_i^{'}\mathbf{D}\tilde{\mathbf{e}}_0)^2 + (\mathbf{w}_i^{'}\tilde{\mathbf{e}}_0)^2 + u_i^2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)^2 \\
&\quad -2u_i^2\mathbf{q}^{'}\mathbf{e}_1 + 2u_i^2(\mathbf{z}_i^{'}\lambda_0)(\mathbf{z}_i^{'}\mathbf{D}\tilde{\mathbf{e}}_0 + \mathbf{w}_i^{'}\tilde{\mathbf{e}}_0 + u_i\mathbf{q}^{'}\tilde{\mathbf{e}}_0) + u_i^2p_i^{(2)}]\Big\}\mathbf{AX}_n + O_p(\frac{1}{\sqrt{n}})\;.
\end{aligned}$$

Since the random vector $\mathbf{AX}_n$ is asymptotically uncorrelated with $\tilde{\mathbf{e}}_0$, we have the convergence in probability as $2\mathbf{QD}^{'}\mathbf{MC}^{-1}(\frac{1}{n}\sum_{i=1}^{n}u_i^2\mathbf{z}_i\mathbf{z}_i^{'})\mathbf{AX}_n(\mathbf{q}^{'}\mathbf{e}_1) \xrightarrow{p} \mathbf{O}$, we use the relation $\mathbf{E}\Big\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}[(\mathbf{z}_i^{'}\mathbf{D}\tilde{\mathbf{e}}_0)^2+(\mathbf{w}_i^{'}\tilde{\mathbf{e}}_0)^2+u_i^2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)^2]\mathbf{A}\mathbf{X}_n|\tilde{\mathbf{e}}_0=\mathbf{x}\Big\}=O_p(n^{-1/2})$. Hence for $\mathbf{e}_{2.1}(A)$ we only need to evaluate the conditional expectation of the last four terms as $-\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\Big\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}[-2\delta(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)(u_i^2\mathbf{z}_i^{'}\mathbf{D}\tilde{\mathbf{e}}_0+u_i^2\mathbf{w}_i^{'}\tilde{\mathbf{e}}_0+u_i^3\mathbf{q}^{'}\tilde{\mathbf{e}}_0)+u_i^2p_i^{(2)}]\Big\}\mathbf{A}\mathbf{X}_n$ up to the order of $O_p(n^{-1/2})$. For the last term involving $p_i^{(2)}$, we use the stochastic expansion of $\boldsymbol{\lambda}_1$ and it becomes $-\mathbf{Q}\mathbf{M}\mathbf{C}^{-1}$ times

$$
\begin{aligned}
&\Big\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)u_i^2[\mathbf{z}_i^{'}\delta\lambda_0(\mathbf{z}_i^{'}\mathbf{D}\tilde{\mathbf{e}}_0+\mathbf{w}_i^{'}\tilde{\mathbf{e}}_0+u_i\mathbf{q}^{'}\tilde{\mathbf{e}}_0)-\mathbf{z}_i^{'}u_i\delta\boldsymbol{\lambda}_1+u_i^2(\mathbf{z}_i^{'}\delta\boldsymbol{\lambda}_0)^2]\Big\}\\
=\;&\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)u_i^2(\mathbf{z}_i^{'}\delta\boldsymbol{\lambda}_0)(\mathbf{z}_i^{'}\mathbf{D}\tilde{\mathbf{e}}_0+\mathbf{w}_i^{'}\tilde{\mathbf{e}}_0+\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\\
&-[\delta\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)u_i^3][-\mathbf{A}\mathbf{U}_n^{'}\mathbf{e}_0-\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{U}_n\mathbf{A}\mathbf{X}_n\\
&\qquad+\mathbf{q}^{'}\mathbf{e}_0\mathbf{A}\mathbf{X}_n-(1-\delta)\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n+\delta\mathbf{A}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2]\\
&+\delta^2\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^3u_i^4\ .
\end{aligned}
$$

Here we illustrate the typical argument of order evaluations. For the final term, we write

$$
\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^3u_i^4=\frac{1}{n}\sum_{i=1}^{n}E(u_i^4)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^3+\frac{1}{n}\sum_{i=1}^{n}[u_i^4-E(u_i^4)]\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^3
$$

and the second term is of order $O_p(n^{-1/2})$. Then we take the conditional expectations of each terms. By applying Lemma 4.3, we find that the first term in the above equation is of order $O_p(n^{-1/2})$. Also we use the relation

$$
\begin{aligned}
&\mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_iu_i^2(\mathbf{z}_i^{'}\mathbf{D}\tilde{\mathbf{e}}_0)(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2|\tilde{\mathbf{e}}_0=\mathbf{x}]\\
=\;&\mathbf{E}\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i[\mathbf{E}(u_i^2)+(u_i^2-\mathbf{E}(u_i^2))](\mathbf{z}_i^{'}\mathbf{D}\tilde{\mathbf{e}}_0)(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2|\tilde{\mathbf{e}}_0=\mathbf{x}\}\\
=\;&\mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{D}\mathbf{x})]+\ O_p(\frac{1}{\sqrt{n}})\ .
\end{aligned}
$$

Then we can show that $\mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_iu_i^2\mathbf{w}_i^{'}\tilde{\mathbf{e}}_0(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2|\mathbf{x}]=\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{E}[u_i^2\mathbf{w}_i^{'}]\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i\mathbf{x}$ and $\mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_iu_i^3\mathbf{q}^{'}\tilde{\mathbf{e}}_0(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2|\mathbf{x}]=\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{q}^{'}\mathbf{x}\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i$, where we have ignored the terms of $O_p(n^{-1/2})$. After tedious calculations of conditional expectations, we then can obtain

$$
\mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_iu_i^2p_i^{(2)}|\tilde{\mathbf{e}}_0=\mathbf{x}]=\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i[E(u_i^2\mathbf{w}_i^{'})\mathbf{x}+\mathbf{E}(u_i^2)\mathbf{z}_i^{'}\mathbf{D}\mathbf{x}]+O_p(\frac{1}{\sqrt{n}})\ .
$$

Then together with the conditional expectations of other terms we can derive the expression

$$
\mathbf{E}[\mathbf{e}_{2.1}(A)|\tilde{\mathbf{e}}_0=\mathbf{x}]\ \ =\ \ -3\delta\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{E}(u_i^2)(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}]\mathbf{D}\mathbf{x}
$$

(3.32)
$$-3\delta \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)]\mathbf{E}[u_i^2\mathbf{w}_i^{'}]\mathbf{x}$$
$$-2\delta \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)]\mathbf{q}^{'}\mathbf{x}+O_p(\frac{1}{\sqrt{n}}).$$

Similarly, the second term of $\mathbf{e}_{2.1}$ is expressed as

(3.33)
$$\mathbf{e}_{2.1}(B) \;=\; \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}[-2u_i(\mathbf{y}_{2i}^{'},\mathbf{z}_{1i}^{'})\mathbf{e}_0+u_ip_i^{(1)}]\}$$
$$\times\mathbf{C}^{-1}\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}[-2u_i(\mathbf{y}_{2i}^{'},\mathbf{z}_{1i}^{'})\mathbf{e}_0+u_ip_i^{(1)}]\}\mathbf{A}\mathbf{X}_n\ .$$

Then by using the asymptotic uncorrelatedness of $\mathbf{A}\mathbf{X}_n$ and $\tilde{\mathbf{e}}_0$, the conditional expectation can be reduced to

(3.34)
$$\mathbf{E}[\mathbf{e}_{2.1}(B)|\tilde{\mathbf{e}}_0=\mathbf{x}]=4\delta(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3\ +\ O_p(n^{-1/2})\ .$$

For the third term, we write $\mathbf{e}_{2.1}(C)=-\mathbf{Q}\mathbf{E}_n^{(1)}\mathbf{C}^{-1}\mathbf{C}_n^{(1)}\mathbf{A}\mathbf{X}_n$, and we use the fact that $\mathbf{Q}\mathbf{E}_n^{(1)}\mathbf{C}^{-1}=[\mathbf{Q}\mathbf{U}_n+\mathbf{Q}\mathbf{q}(\mathbf{X}_n^{'}-\lambda_0^{'}\mathbf{C}_n)]\mathbf{C}^{-1}+O_p(\frac{1}{\sqrt{n}})$ and $\mathbf{C}_n^{(1)}\mathbf{A}\mathbf{X}_n=[-2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{C}_n-\delta(1/n)\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)]\mathbf{A}\mathbf{X}_n+O_p(\frac{1}{\sqrt{n}})$. Then the conditional expectation of $\mathbf{e}_{2.1}(C)$ is given by

(3.35)
$$\begin{aligned}
&\mathbf{E}[\mathbf{e}_{2.1}(C)|\tilde{\mathbf{e}}_0=\mathbf{x}]\\
=\;&2\mathbf{Q}\mathbf{q}(\mathbf{q}^{'}\mathbf{x})(1-\delta)\mathbf{E}[\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{C}\mathbf{A}\mathbf{X}_n|\mathbf{x}]\\
&+\mathbf{Q}\mathbf{q}(\mathbf{X}_n^{'}-\delta\lambda_0^{'}\mathbf{C})\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2+O_p(\frac{1}{\sqrt{n}})\ ,\\
=\;&2(1-\delta)L\mathbf{Q}\mathbf{q}\mathbf{q}^{'}\mathbf{x}+\delta\mathbf{Q}\mathbf{q}\mathbf{e}_0^{'}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3+O_p(\frac{1}{\sqrt{n}})\ .
\end{aligned}$$

The fourth term of $\mathbf{e}_{2.1}$ is expressed as

(3.36)
$$\mathbf{e}_{2.1}(D)=\mathbf{Q}\{\mathbf{D}^{'}\frac{1}{n}\sum_{i=1}^{n}p_i^{(2)}\mathbf{z}_i\mathbf{z}_i^{'}+\frac{1}{n}\sum_{i=1}^{n}p_i^{(2)}\begin{pmatrix}\mathbf{v}_{2i}\\ \mathbf{0}\end{pmatrix}\mathbf{z}_i^{'}\}\mathbf{A}\mathbf{X}_n$$

Since the first term of $\mathbf{e}_{2.1}(D)$ is similar to the last term of $\mathbf{e}_{2.1}(A)$ and we can utilize the expression of λ_1 in (3.32), it is easy to find that the conditional expectation, which is given by

$$\begin{aligned}
&\mathbf{E}[\mathbf{Q}\mathbf{D}^{'}\frac{1}{n}\sum_{i=1}^{n}p_i^{(2)}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n|\tilde{\mathbf{e}}_0=\mathbf{x}]\\
=\;&\mathbf{Q}\mathbf{D}^{'}\mathbf{E}\Bigg(\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)(\delta\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)(\mathbf{z}_i^{'}\mathbf{D}\mathbf{e}_0+\mathbf{w}_i^{'}\mathbf{e}_0+u_i\mathbf{q}^{'}\mathbf{e}_0)\\
&-\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)[-\tilde{\mathbf{e}}_0^{'}\mathbf{U}_n\mathbf{A}-\mathbf{X}_n^{'}\mathbf{A}\mathbf{U}_n^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}+\mathbf{q}^{'}\tilde{\mathbf{e}}_0\mathbf{X}_n^{'}\mathbf{A}\\
&-(1-\delta)\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n\mathbf{q}^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}+\delta\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}(u_j^3)\mathbf{z}_j(\mathbf{z}_j^{'}\mathbf{A}\mathbf{X}_n)^2\mathbf{A}]\mathbf{z}_iu_i
\end{aligned}$$

$$+\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)^2u_i^2|\mathbf{x}\Big)$$

$$= \delta\mathbf{E}\Big\{\mathbf{Q}\mathbf{D}^{'}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}]\mathbf{D}\mathbf{x}+\mathbf{Q}\mathbf{D}^{'}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2\mathbf{A}\mathbf{X}_n|\mathbf{x}\Big\}+O_p(\frac{1}{\sqrt{n}})$$

$$= \delta\mathbf{Q}\mathbf{D}^{'}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}]\mathbf{D}\mathbf{x}\ +\ O_p(\frac{1}{\sqrt{n}})\ .$$

For the second term of $\mathbf{e}_{2.1}(D)$, we rewrite

$$\mathbf{Q}\frac{1}{n}\sum_{i=1}^{n}p_i^{(2)}\begin{pmatrix}\mathbf{v}_{2i}\\ \mathbf{0}\end{pmatrix}\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n$$

$$= \mathbf{Q}\frac{1}{n}\sum_{i=1}^{n}(\mathbf{w}_i+\mathbf{q}u_i)\{\boldsymbol{\lambda}_0^{'}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{D}+\mathbf{w}_i^{'}+\mathbf{q}u_i^{'})\tilde{\mathbf{e}}_0-\boldsymbol{\lambda}_1^{'}\mathbf{z}_iu_i+(\boldsymbol{\lambda}_0^{'}\mathbf{z}_i)u_i^2\}\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n$$

$$= \mathbf{Q}\frac{1}{n}\sum_{i=1}^{n}(\mathbf{w}_i+\mathbf{q}u_i)(\mathbf{z}_i^{'}\mathbf{D}+\mathbf{w}_i^{'}+\mathbf{q}^{'}u_i)\tilde{\mathbf{e}}_0(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)$$

$$-\mathbf{Q}\frac{1}{n}\sum_{i=1}^{n}(\mathbf{w}_i+\mathbf{q}u_i)(\mathbf{z}_i^{'}u_i)\boldsymbol{\lambda}_1(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)+\mathbf{Q}\frac{1}{n}\sum_{i=1}^{n}(\mathbf{w}_i+\mathbf{q}u_i)(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)^2u_i^2(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)+O_p(\frac{1}{\sqrt{n}}).$$

For the sake of exposition, we denote each term of the above expression as $\mathbf{e}_{2.1.1}(D)$, $\mathbf{e}_{2.1.2}(D)$, $\mathbf{e}_{2.1.3}(D)$, respectively. Then

$$\mathbf{E}[\mathbf{e}_{2.1.1}(D)|\tilde{\mathbf{e}}_0=\mathbf{x}] = \mathbf{E}[\mathbf{Q}\frac{1}{n}\sum_{i=1}^{n}(\mathbf{w}_i\mathbf{w}_i^{'}+\mathbf{q}\mathbf{q}^{'}u_i^2)\tilde{\mathbf{e}}_0\boldsymbol{\lambda}_0^{'}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n|\mathbf{x}]$$

$$= \delta\mathbf{Q}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(\mathbf{w}_i\mathbf{w}_i^{'})\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i]\mathbf{x}+\delta L\mathbf{Q}\mathbf{C}_1^{*}\mathbf{x}+O_p(\frac{1}{\sqrt{n}}).$$

Also we find that

$$\mathbf{E}[\mathbf{e}_{2.1.3}(D)|\tilde{\mathbf{e}}_0=\mathbf{x}] = \mathbf{E}[\mathbf{Q}\frac{1}{n}\sum_{i=1}^{n}\mathbf{q}u_i^3(\mathbf{z}_i^{'}\delta\boldsymbol{\lambda}_0)^2\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n|\tilde{\mathbf{e}}_0=\mathbf{x}]$$

$$= \delta^2\mathbf{Q}\mathbf{q}E[\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^3|\tilde{\mathbf{e}}_0=\mathbf{x}]+O_p(\frac{1}{\sqrt{n}})\ .$$

But the random vector $\mathbf{A}\mathbf{X}_n$ is asymptotically normal, and its limiting distribution is uncorrelated with that of $\tilde{\mathbf{e}}_0$. Hence we find $\mathbf{E}[\mathbf{e}_{2.1.3}(D)|\tilde{\mathbf{e}}_0=\mathbf{x}]=O_p(n^{-1/2})$. Then we evaluate the conditional expectation of $\mathbf{e}_{2.1.2}(D)$. Because the pairs of random vectors $(\mathbf{w}_i^{'},u_i)$ are uncorrelated, we have the convergence in probability $(1/n)\sum_{i=1}^{n}\mathbf{w}_iu_iz_i^{(j)}z_i^{(k)}\xrightarrow{p}\mathbf{0}$. Therefore, as for the remaining conditional expectation terms, we use the explicit expression for $\boldsymbol{\lambda}_1$, we find

$$\mathbf{E}[\mathbf{e}_{2.1.2}(D)|\tilde{\mathbf{e}}_0=\mathbf{x}] = -\delta\mathbf{Q}\mathbf{q}\mathbf{E}[\boldsymbol{\lambda}_1^{'}(\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}u_i^2)\mathbf{A}\mathbf{X}_n|\mathbf{x}]+O_p(\frac{1}{\sqrt{n}})$$

$$= -\delta L\mathbf{Q}\mathbf{q}\mathbf{q}^{'}\mathbf{x}+O_p(\frac{1}{\sqrt{n}})\ .$$

Hence we summarize the conditional expectation

$$\text{(3.37)}\qquad \mathbf{E}[\mathbf{e}_{2.1}(D)|\tilde{\mathbf{e}}_0 = \mathbf{x}] = \delta \mathbf{Q}\mathbf{D}^{'}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}]\mathbf{D}\mathbf{x}$$

$$+\delta\mathbf{Q}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(\mathbf{w}_i\mathbf{w}_i^{'})\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i]\mathbf{x} + \delta L\mathbf{Q}\mathbf{C}_1^{*}\mathbf{x} - \delta L\mathbf{Q}\mathbf{q}\mathbf{q}^{'}\mathbf{x} + O_p(n^{-1/2}) .$$

Finally, we can derive the conditional expectation of $\mathbf{E}[\mathbf{e}_{2.1}|\tilde{\mathbf{e}}_0 = \mathbf{x}]$ by collecting the conditional expectations of $\mathbf{E}[\mathbf{e}_{2.1}(A)|\mathbf{x}]$, $\mathbf{E}[\mathbf{e}_{2.1}(B)|\mathbf{x}]$, $\mathbf{E}[\mathbf{e}_{2.1}(C)|\mathbf{x}]$ and $\mathbf{E}[\mathbf{e}_{2.1}(D)|\mathbf{x}]$. At the first glance of these many terms, it looks formidable to calculate them. However, the resulting formulas become relatively simple if we can ignore the third order moments because it has turned out that many terms have disappeared in the conditional expectation formulas eventually.

4. Asymptotic Expansions of Distributions

4.1 Conditional Expectation Formulas

In order to derive the asymptotic expansions of the distribution functions, we prepare some useful formulas on the conditional expectations and their proofs are given in Appendix A. They are the result of straightforward calculations, but we shall give the derivations of Lemma 4.1 and Lemma 4.3 in Appendix A for the expository purpose.

Lemma 4.1 : Let the random vectors $\tilde{\mathbf{e}}_0, \mathbf{X}_n$, and $\mathbf{Y}_n$ be defined as in Section 3. Then

$$\text{(4.1)}\qquad \mathbf{E}[\mathbf{Y}_n\mathbf{A}\mathbf{X}_n|\tilde{e}_0 = \mathbf{x}] = \mathbf{m}_3 + O_p(\frac{1}{\sqrt{n}}) ,$$

where $\mathbf{m}_3$ is given by (3.26).

Lemma 4.2 : Let a set of random vectors $\mathbf{X} = (X_i)$ and $\mathbf{Z} = (Z_i)$ be normally distributed and the conditional expectation of X_i given $\mathbf{Z}$ be $\mathbf{E}(X_i|\mathbf{Z})$. Then

$$\text{(4.2)}\quad \mathbf{E}[X_iX_jX_k|\mathbf{Z}] = Cov(X_i, X_j|\mathbf{Z})\mathbf{E}(X_k|\mathbf{Z}) + Cov(X_j, X_k|\mathbf{Z})\mathbf{E}(X_i|\mathbf{Z}) + Cov(X_k, X_i|\mathbf{Z})\mathbf{E}(X_j|\mathbf{Z}) + \mathbf{E}(X_i|\mathbf{Z})\mathbf{E}(X_j|\mathbf{Z})\mathbf{E}(X_k|\mathbf{Z})$$

and

$$\text{(4.3)}\qquad \begin{aligned} & \mathbf{E}[X_iX_jX_kX_l|\mathbf{Z}] \\ = & \; Cov(X_i, X_j|\mathbf{Z})Cov(X_k, X_l|\mathbf{Z}) + Cov(X_i, X_k|\mathbf{Z})Cov(X_j, X_l|\mathbf{Z}) \\ & + Cov(X_i, X_l|\mathbf{Z})Cov(X_j, X_k|\mathbf{Z}) \\ + & \; Cov(X_i, X_j|\mathbf{Z})\mathbf{E}(X_k|\mathbf{Z})\mathbf{E}(X_l|\mathbf{Z}) + Cov(X_i, X_k|\mathbf{Z})\mathbf{E}(X_j|\mathbf{Z})\mathbf{E}(X_l|\mathbf{Z}) \\ & + Cov(X_i, X_l|\mathbf{Z})\mathbf{E}(X_j|\mathbf{Z})\mathbf{E}(X_k|\mathbf{Z}) + Cov(X_j, X_k|\mathbf{Z})\mathbf{E}(X_i|\mathbf{Z})\mathbf{E}(X_l|\mathbf{Z}) \\ & + Cov(X_j, X_l|\mathbf{Z})\mathbf{E}(X_i|\mathbf{Z})\mathbf{E}(X_k|\mathbf{Z}) + Cov(X_k, X_l|\mathbf{Z})\mathbf{E}(X_i|\mathbf{Z})\mathbf{E}(X_j|\mathbf{Z}) \\ + & \; \mathbf{E}(X_i|\mathbf{Z})\mathbf{E}(X_j|\mathbf{Z})\mathbf{E}(X_k|\mathbf{Z})\mathbf{E}(X_l|\mathbf{Z}) . \end{aligned}$$

The above formulas are used repeatedly in our derivations by setting $\mathbf{Z} = \tilde{e}_0$, which is the leading term of our stochastic expansions of estimators. In order to evaluate the conditional expectation operations appeared in the stochastic expansions of estimators, we also need the next formula.

Lemma 4.3 : Let $\mathbf{u}_n = (u_i)$ and v_n be $p \times 1$ random vector and a random variable with $\mathbf{E}(u_i) = 0, \mathbf{E}(v_n) = 0, \mathbf{E}(u_i u_j) = \delta(i,j), \mathbf{E}(v_n^2) = 1$ and they have finite fourth order moments. Assume that they are sum of i.i.d. (non-lattice) random vectors and asymptotically normally distributed and admit the asymptotic expansion of their distribution function up to $O_p(n^{-1})$. Then

(4.4) $\mathbf{E}[v_n|\mathbf{u}_n]$

$$= \boldsymbol{\rho}^{'}\mathbf{u}_n + \frac{1}{6\sqrt{n}}\{3\sum_{l,l'=1}^{p}\beta_{l,l',v}h_2(u_l,u_{l'}) - 3\sum_{l',l''=1}^{p}[\sum_{l=1}^{p}\beta_{l,l',l''}\rho_l]h_2(u_{l'},u_{l''})\} + O_p(\frac{1}{n}) ,$$

where $\beta_{l,l',v} = \mathbf{E}(u_l u_{l'} v_n)$, $\beta_{l.l',l''} = \mathbf{E}(u_l u_{l'} u_{l''})$, $h_2(u_l,u_{l'}) = u_l u_{l'} - \delta(l,l')$ ($\delta(l,l') = 1$ if $l = l'$ and $\delta(l,l') = 0$ if $l \neq l'$), and $\boldsymbol{\rho} = Cov(v,\mathbf{u}_n)$.
In particular, if $\mathbf{E}(u_i u_j u_k) = 0$ $(i \neq j \neq k)$, then $\beta_{l.l',l''} = 0$ and the second term in the order of $O_p(n^{-1/2})$ in (4.4) vanishes.

Although we have obtained several conditional expectation formulas associated with the terms $\mathbf{e}_1^{(0)}$ and $\mathbf{e}_2$ in Section 3, we need to evaluate several additional terms for deriving the asymptotic expansions of density functions of estimators up to $O(n^{-1})$. By applying Lemma 4.3, the conditional expectation of $\mathbf{e}_0^{(1)}$ in (3.20) given $\mathbf{X}_n$ can be calculated as

$$\begin{aligned}
&\mathbf{E}[\mathbf{e}_0^{(1)}|\mathbf{X}_n] \\
= &-\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{z}_i] \\
&+\frac{1}{6\sqrt{n}}\{-3\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{n}\sum_{i=1}^{n}(\mathbf{E}(u_i^4) - (\mathbf{E}(u_i^2))^2)\sum_{l,l'}z_i^l z_i^{l'}h_2(x_l,x_{l'})]\mathbf{A}\mathbf{X}_n \\
&\qquad +3\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{n}\sum_{i=1}^{n}(\mathbf{E}(u_i^3))^2\mathbf{z}_i\mathbf{z}_i^{'}\sum_{l,l'}z_i^l z_i^{l'}h_2(x_l,x_{l'})]\mathbf{A}\mathbf{X}_n\} + O_p(n^{-1/2}).
\end{aligned}$$

By decomposing $\mathbf{A}\mathbf{X}_n\mathbf{X}_n^{'}\mathbf{C}^{-1} = \mathbf{A}\mathbf{X}_n\mathbf{X}_n^{'}\mathbf{A} + \mathbf{A}\mathbf{X}_n\tilde{\mathbf{e}}_0^{'}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}$ and taking the conditional expectations, we have

(4.5) $\mathbf{E}[\mathbf{e}_0^{(1)}|\tilde{\mathbf{e}}_0 = \mathbf{x}]$

$$\begin{aligned}
= &-\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3 \\
&+\frac{1}{6\sqrt{n}}\Big\{-3\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{n}\sum_{i=1}^{n}(\mathbf{E}(u_i^4) - (\mathbf{E}(u_i^2))^2) \times 2\mathbf{z}_i\mathbf{z}_i{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\tilde{\mathbf{e}}_0(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)] \\
&\qquad +3\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{E}(u_i^3))^2(2\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\tilde{\mathbf{e}}_0\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)]\Big\} + O_p(\frac{1}{\sqrt{n}}) .
\end{aligned}$$

By using the conditional expectation formulas we have obtained, we can summarize

(4.6) $$\mathbf{E}[\mathbf{e}_1^{(0)} + \mathbf{e}_0^{(1)}|\tilde{\mathbf{e}}_0 = \mathbf{x}] = (1-\delta)L\mathbf{Q}\mathbf{q} - \mathbf{x}\mathbf{x}^{'}\mathbf{q} - (1-\delta)\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3 + O_p(\frac{1}{\sqrt{n}}).$$

Then it is immediate to obtain the (unconditional) asymptotic bias of the class of standardized estimators up to the order $o(1/\sqrt{n})$.

It is possible to continue our procedure to derive the asymptotic expansions of the density function of estimators in the general case, but the resulting expressions become more messy due to the terms associated with $\mathbf{e}_1^{(1)}$ and $\mathbf{e}_0^{(2)}$. It is because the conditional expectations of some random variables of order $O_p(n^{-1/2})$ have the terms of $O_p(n^{-1/2})$ and *also* some additional terms of $O_p(n^{-1})$, which make our derivation some complications.

When we can ignore the effects of the third order moments of the disturbance, the asymptotic expansions of estimators for an arbitrary δ $(0 \leq \delta \leq 1)$ can be simplified greatly in evaluating many terms of order $O_p(n^{-1})$. Then we shall make the following assumption and show how our derivations can be simplified greatly.

Assumption II :
(i) For the second and fourth moments of disturbances in (2.1) and (2.3) we assume $\mathbf{E}(u_i^2) = \sigma^2 > 0$, $\mathbf{C}_2^* = \mathbf{E}(\mathbf{w}_i \mathbf{w}_i^{'})$, the condition given by (3.9) and $\kappa = \mathbf{E}(u_i^4)/\sigma^4 - 3$ for $i = 1, \cdots, n$.
(ii) For the third order moments of disturbances in (2.1) and (2.3) we assume $\mathbf{E}[u_i^3] = 0$ and $\mathbf{E}[u_i^2 \mathbf{w}_i] = \mathbf{0}$ for $i = 1, \cdots, n$.

It is immediate that the conditions in Assumption II can be relaxed as

(4.7)
$$\frac{1}{n} \sum_{i=1}^{n} z_i^{(j)} z_i^{(k)} z_i^{(l)} u_i^3 = o_p(\frac{1}{\sqrt{n}})$$

and the similar conditions on the third order moments on $\{u_i^2 \mathbf{w}_i\}$. In this section we shall derive the asymptotic distribution functions of estimators under **Assumption I** and **Assumption II**.

Now we shall evaluate the conditional expectations of $\mathbf{e}_1^{(1)} = \mathbf{e}_{1.1}^{(1)} + \mathbf{e}_{1.2}^{(1)} + \mathbf{e}_{1.3}^{(1)}$ given $\tilde{\mathbf{e}}_0 = \mathbf{x}$. This term plays an important role and makes some complications of our analyses. We first note that the conditional expectations of $\mathbf{e}_{1.2}^{(1)}$ and $\mathbf{e}_{1.3}^{(1)}$ given $\tilde{\mathbf{e}}_0 = \mathbf{x}$ can be evaluated easily. By using Lemma 4.3, their conditional expectations are given by

(4.8)
$$\begin{aligned} & \mathbf{E}[\mathbf{e}_{1.2}^{(1)} | \tilde{\mathbf{e}}_0 = \mathbf{x}] \\ &= \mathbf{E}[\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n^{'}\mathbf{A}\mathbf{X}_n(\mathbf{q}^{'}\tilde{\mathbf{e}}_0) + \tilde{\mathbf{e}}_0\mathbf{q}^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n^{'}\mathbf{A}\mathbf{X}_n | \tilde{\mathbf{e}}_0 = \mathbf{x}] \\ &= (\mathbf{q}^{'}\mathbf{x})\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3 + \mathbf{x}\mathbf{q}^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3 + O_p(\frac{1}{\sqrt{n}}) \end{aligned}$$

and

(4.9)
$$\begin{aligned} & \mathbf{E}[\mathbf{e}_{1.3}^{(1)} | \tilde{\mathbf{e}}_0 = \mathbf{x}] \\ &= \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{E}[\mathbf{U}_n^{'} | \tilde{\mathbf{e}}_0 = \mathbf{x}]\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{m}_3 + \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i\mathbf{E}(u_i^2\mathbf{w}_i^{'})\mathbf{x} , \end{aligned}$$

which are both the terms of $O(1/n)$ under **Assumption II**.
Next, we evaluate the conditional expectation of $\mathbf{e}_{1.1}^{(1)}$ in (3.23) which has many terms.

We notice that in $\mathbf{e}_{1.1}^{(1)}$ two terms associated with the random matrices $\mathbf{C}_n^{(1)}$ and $\mathbf{E}_n^{(1)}$ have been cancelled out. Then we try to evaluate each remaining terms of order $O_p(n^{-1/2})$ and the terms in the first two lines are given by

$$\begin{aligned} & \mathbf{E}\{\delta \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n\mathbf{z}_i^{'}(\mathbf{C}^{-1}-\mathbf{A})\mathbf{E}(\mathbf{Y}_n|\mathbf{X}_n)\mathbf{A}\mathbf{X}_n|\tilde{\mathbf{e}}_0=\mathbf{x}\} \\ + \ & \mathbf{E}\{\mathbf{Q}\mathbf{U}_n(\mathbf{C}^{-1}-\mathbf{A})\mathbf{E}(\mathbf{Y}_n|\mathbf{X}_n)\mathbf{A}\mathbf{X}_n|\tilde{\mathbf{e}}_0=\mathbf{x}\} \\ + \ & \mathbf{E}\{\mathbf{Q}\mathbf{q}\tilde{\mathbf{e}}_0^{'}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{E}(\mathbf{Y}_n|\mathbf{X}_n)\mathbf{A}\mathbf{X}_n|\tilde{\mathbf{e}}_0=\mathbf{x}\}\ . \end{aligned}$$

Similarly, the terms in the third line are given by

$$\begin{aligned} & \mathbf{E}\{-\delta \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{E}(\mathbf{Y}_n|\mathbf{X}_n)\mathbf{A}\mathbf{X}_n|\tilde{\mathbf{e}}_0=\mathbf{x}\} \\ + \ & \mathbf{E}\{-\delta \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{E}(\mathbf{Y}_n|\mathbf{X}_n)\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i|\tilde{\mathbf{e}}_0=\mathbf{x}\}\ , \end{aligned}$$

which are of order $O_p(n^{-1/2})$ under **Assumption II**. The most important terms in $O_p(n^{-1/2})$ are the next two terms in the fourth and fifth lines because they are dependent on the fourth order moments of $\{u_i\}$, which are given by

$$\begin{aligned} & 2\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}u_i(\mathbf{z}_i^{'}\mathbf{D}+\mathbf{w}_i^{'})\tilde{e}_0]\mathbf{A}\mathbf{X}_n \\ & +\delta \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)(u_i^3-\mathbf{E}(u_i^3)-u_i\mathbf{E}(u_i^2))]\mathbf{A}\mathbf{X}_n \end{aligned}$$

up to $O_p(n^{-1/2})$. When we have homoscedastic disturbances, we can use the fourth order cumulant given by $\kappa=[E(u_i^4)-3\sigma^4)]/\sigma^4$. The conditional expectations of the first term and the second term are given by

$$\begin{aligned} (4.10)\qquad & 2\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{E}[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}u_i(\mathbf{z}_i^{'}\mathbf{D}+\mathbf{w}_i^{'})\tilde{e}_0]\mathbf{A}\mathbf{X}_n|\tilde{\mathbf{e}}_0=\mathbf{x}] \\ & = 2\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_iu_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\mathbf{D}\mathbf{x})\mathbf{A}\mathbf{z}_i+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{E}[u_i\mathbf{w}_i^{'}]\mathbf{x}\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n\} \\ & \quad +O_p(\frac{1}{\sqrt{n}})\ , \end{aligned}$$

and

$$\begin{aligned} (4.11)\qquad & \delta \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{E}[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2(u_i^3-\mathbf{E}(u_i^3)-\sigma^2u_i)|\tilde{\mathbf{e}}_0=\mathbf{x}] \\ & = \delta \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{E}[u_i^4-\sigma^2u_i^2]\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{x}+O_p(\frac{1}{\sqrt{n}}) \\ & = \delta(2+\kappa)\mathbf{Q}\mathbf{D}^{'}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)]\mathbf{D}\mathbf{x}+O_p(\frac{1}{\sqrt{n}}). \end{aligned}$$

For the expression of $\mathbf{e}_{1.1}^{(1)}$, there are many remaining terms in the last three lines, which are given by

$$\begin{aligned}
& \mathbf{E}\{-\mathbf{Q}[\mathbf{U}_n\mathbf{C}^{-1} + \mathbf{q}(\mathbf{X}_n^{'}\mathbf{C}^{-1} - \delta\mathbf{X}_n^{'}\mathbf{A})]\mathbf{E}(\mathbf{Y}_n|\mathbf{X}_n)\mathbf{A}\mathbf{X}_n|\tilde{\mathbf{e}}_0 = \mathbf{x}\} \\
+ \;& \mathbf{E}\{\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}[\mathbf{U}_n\mathbf{A}\mathbf{X}_n + (1-\delta)\mathbf{q}\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n] \\
+ \;& \mathbf{E}\{2\mathbf{q}^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n\mathbf{A}\mathbf{X}_n - \delta\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i^{'}\mathbf{A}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n\mathbf{z}_i^{'}\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n|\tilde{\mathbf{e}}_0 = \mathbf{x}\},
\end{aligned}$$

which are of the order $O_p(n^{-1/2})$.

Finally, after many terms involving the third order moments disappear due to **Assumption II**, we have only two terms involving the fourth order moments terms in (4.10) and (4.11). Hence under **Assumption II** in this section the conditional expectation of $\mathbf{e}_1^{(1)}$ is given by

$$\mathbf{E}[\mathbf{e}_1^{(1)}|\tilde{\mathbf{e}}_0 = \mathbf{x}] = [2 + \delta(2+\kappa)]\mathbf{Q}\mathbf{D}^{'}\mathbf{F}\mathbf{D}\mathbf{x} + O_p(\frac{1}{\sqrt{n}}), \tag{4.12}$$

where we define

$$\mathbf{F} = \text{plim}_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}\ . \tag{4.13}$$

We note that (4.12) is one of key aspects in our analysis of the semi-parametric estimation of a single structural equation because we do not have the corresponding terms $O_p(n^{-1})$ in the asymptotic expansions of the classical estimation methods.

Similarly, under **Assumption II**,

$$\mathbf{E}[\mathbf{e}_0^{(1)}|\tilde{\mathbf{e}}_0 = \mathbf{x}] = \frac{1}{\sqrt{n}}\{-(2+\kappa)\mathbf{Q}\mathbf{D}^{'}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}]\mathbf{D}\mathbf{x}\} + O_p(\frac{1}{n})\ . \tag{4.14}$$

In order to evaluate $\mathbf{E}[\mathbf{e}_0^{(2)}|\tilde{\mathbf{e}}_0 = \mathbf{x}]$, we need the relation that for a constant matrix $\mathbf{A}$ $(= (A_{jk}))$

$$\begin{aligned}
\mathbf{E}[\mathbf{Y}_n\mathbf{A}\mathbf{Y}_n|\mathbf{X}_n] &= \mathbf{E}[(u_i^2 - \mathbf{E}(u_i^2))^2][\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}] \\
&+ [\mathbf{E}(u_i^3)]^2\sum_{j,k=1}^{p}A_{jk}(\frac{1}{n})\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}z_i^{(j)}\mathbf{C}^{-1}(\mathbf{X}_n\mathbf{X}_n^{'} - \mathbf{C})\mathbf{C}^{-1}(\frac{1}{n}\sum_{i=1}^{n}z_i^{(k)}\mathbf{z}_i^{'}) \\
&= \mathbf{E}[(u_i^2 - \mathbf{E}(u_i^2))^2][\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}] + O_p(\frac{1}{\sqrt{n}})
\end{aligned}$$

under **Assumption II**. Hence we have

$$\mathbf{E}[\mathbf{e}_0^{(2)}|\tilde{\mathbf{e}}_0 = \mathbf{x}] = \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{E}[\mathbf{Y}_n\mathbf{A}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n|\mathbf{x}] = O_p(\frac{1}{\sqrt{n}}) \tag{4.15}$$

because the random vector $\mathbf{A}\mathbf{X}_n$ is asymptotically normally distributed and uncorrelated with $\tilde{\mathbf{e}}_0$. By using direct calculations we have

$$\begin{aligned}
& \mathbf{E}[\mathbf{e}_0^{(1)}\mathbf{e}_0^{(1)'}|\tilde{\mathbf{e}}_0 = \mathbf{x}] \\
= \;& \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{E}[\mathbf{Y}_n\mathbf{A}\mathbf{X}_n\mathbf{X}_n^{'}\mathbf{A}\mathbf{Y}_n|\mathbf{x}]\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}
\end{aligned}$$

$$= \quad (2+\kappa)\mathbf{Q}\mathbf{D}^{'}\mathbf{F}\mathbf{D}\mathbf{Q} + \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\mathbf{m}_3\mathbf{m}_3^{'} + (\frac{1}{n})^2 \sum_{i,j=1}^{n} (\mathbf{E}(u_i^3))^2 \mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_j)^2\mathbf{z}_j^{'}]\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}$$

$$+O_p(\frac{1}{\sqrt{n}})$$

in the general case.
For the random matrix $\mathbf{U}_n = (u_{jk}) = (1/\sqrt{n})\sum_{i=1}^{n} \mathbf{w}_i\mathbf{z}_i^{'}$, we apply Lemma 4.3 and use the fact that $Cov(u_{jk}, \tilde{\mathbf{e}}_0) = \mathbf{0}$. Then we have

$$\begin{aligned}
\mathbf{E}[u_{jk}|\mathbf{X}_n] &= \frac{1}{2\sqrt{n}} \sum_{l,l'=1}^{K} \{\frac{1}{n}\sum_{i=1}^{n}(\mathbf{C}^{-1/2}\mathbf{z}_i)_l(\mathbf{C}^{-1/2}\mathbf{z}_i)_l^{'} z_i^{(k)}\mathbf{E}(u_i w_i^{(j)}) \\
&\qquad \times[(\mathbf{C}^{-1/2}\mathbf{X}_n)_l(\mathbf{C}^{-1/2}\mathbf{X}_n)_{l'} - \delta(l,l^{'})]\} + O_p(\frac{1}{n}) \\
&= \frac{1}{2\sqrt{n}}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^2 w_i^{(j)}) z_i^{(k)}[\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{X}_n\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{z}_i] + O_p(\frac{1}{\sqrt{n}}) \ .
\end{aligned}$$

Then by decomposing $\mathbf{C}^{-1} = \mathbf{A} + \mathbf{C}^{-1}\mathbf{M}\mathbf{D}^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}$ and $\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{X}_n\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{z}_i = 2\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n\tilde{\mathbf{e}}_0^{'}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{z}_i + \mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\tilde{\mathbf{e}}_0\tilde{\mathbf{e}}_0^{'}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{z}_i$, and taking the conditional expectation, we have

$$\mathbf{E}[u_{jk}|\tilde{\mathbf{e}}_0 = \mathbf{x}] = \frac{1}{2\sqrt{n}}\frac{1}{n}\sum_{i=1}^{n} E(u_i^2\mathbf{w}_i^{(j)})\mathbf{z}_i^{(k)'}[\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}(\mathbf{x}\mathbf{x}^{'} - \mathbf{Q})\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{z}_i] + O_p(\frac{1}{\sqrt{n}}).$$

By using the above type of the conditional expectations formulas and lengthy calculation under **Assumption II**, we find that

$$\begin{aligned}
(4.16)\ \ & \mathbf{E}[\mathbf{e}_0^{(1)}\mathbf{e}_1^{(0)'}|\tilde{\mathbf{e}}_0 = \mathbf{x}] \\
= \quad & -\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i\mathbf{E}[u_i^2\mathbf{w}_i^{'}]\mathbf{Q} + \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\tilde{\mathbf{e}}_0^{'} \\
+ \quad & (1-\delta)\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}(L+2)\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(u_i^3)\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i\mathbf{q}^{'}\mathbf{Q} \\
- \quad & \delta[\mathbf{E}(u_i^3)]^2\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\left((\frac{1}{n})^2\sum_{i,j}^{n}\mathbf{z}_i\mathbf{z}_j\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i\mathbf{z}_j^{'}\mathbf{A}\mathbf{z}_j + 2(\frac{1}{n})^2\sum_{i,j}^{n}\mathbf{z}_i\mathbf{z}_j(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_j)^2]\right) \\
& \qquad \times\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q},
\end{aligned}$$

which is of $O_p(n^{-1/2})$.

4.2 Asymptotic Expansions of Density Functions

Since there are many terms appeared in the stochastic expansion of $\hat{\mathbf{e}}$, at first it looks formidable to evaluate all terms and to derive the asymptotic expansions of the density functions of alternative estimators. Under **Assumption I** and **Assumption II**, however, it is possible to derive them in a compact form which gives useful information on the exact distributions of estimators.

In order derive the asymptotic expansions of density functions, we define $\mathbf{e}_0^* = \tilde{\mathbf{e}}_0$ as the leading term. As the second term, we define $\mathbf{e}_1^*(\mathbf{x})$ as the sum of constant order terms of the conditional expectations $\mathbf{E}[\mathbf{e}_1^{(0)}+\mathbf{e}_0^{(1)}|\tilde{\mathbf{e}}_0=\mathbf{x}]$. As the third order term, we define $\mathbf{e}_2^*(\mathbf{x})$ as the sum of $O_p(n^{-1/2})$ terms of the conditional expectations $\mathbf{E}[\mathbf{e}_1^{(0)}+\mathbf{e}_0^{(1)}|\tilde{\mathbf{e}}_0=\mathbf{x}]$ plus the constant order terms of the conditional expectations $\mathbf{E}[\mathbf{e}_0^{(2)}+\mathbf{e}_1^{(1)}+\mathbf{e}_2|\tilde{\mathbf{e}}_0=\mathbf{x}]$. As the cross order term, we define $\mathbf{e}_{11}^*(\mathbf{x})$ as the sum of constant order terms of the conditional expectations $\mathbf{E}[(\mathbf{e}_1^{(0)}+\mathbf{e}_0^{(1)})(\mathbf{e}_1^{(0)}+\mathbf{e}_0^{(1)})'|\tilde{\mathbf{e}}_0=\mathbf{x}]$.

By using the conditional expectation formulae we have obtained under **Assumption II** in the previous sections, we summarize

$$\mathbf{e}_1^*(\mathbf{x}) = (1-\delta)L\mathbf{Q}\mathbf{q}-\mathbf{x}\mathbf{x}'\mathbf{q}\,, \tag{4.17}$$

$$\begin{aligned}\mathbf{e}_2^*(\mathbf{x}) &= -(2+\kappa)\mathbf{Q}\mathbf{D}'\mathbf{F}\mathbf{D}\mathbf{x}+[2+\delta(2+\kappa)]\mathbf{Q}\mathbf{D}'\mathbf{F}\mathbf{D}\mathbf{x}+\mathbf{x}\mathbf{x}'\mathbf{C}_1^*\mathbf{x} \\ &\quad +\mathbf{Q}\mathbf{Q}^*\mathbf{Q}\mathbf{C}_2^*\mathbf{x} \\ &\quad -(1-\delta)L[\mathbf{x}\,\mathbf{tr}(\mathbf{C}_1^*\mathbf{Q})+2\mathbf{Q}\mathbf{C}_1^*\mathbf{x}]-(1-\delta)\mathbf{Q}\mathbf{C}_2^*\mathbf{x}\,\mathbf{tr}(\mathbf{M}\mathbf{A}) \\ &\quad +[-3\delta+\delta]\mathbf{Q}\mathbf{D}'\mathbf{F}\mathbf{D}\mathbf{x} \\ &= (\delta-1)\kappa\mathbf{Q}\mathbf{D}'\mathbf{F}\mathbf{D}\mathbf{x}+\mathbf{x}\mathbf{x}'\mathbf{C}_1^*\mathbf{x}+\mathbf{Q}\mathbf{Q}^*\mathbf{Q}\mathbf{C}_2^*\mathbf{x} \\ &\quad -(1-\delta)L[\mathbf{x}\,\mathbf{tr}(\mathbf{C}_1^*\mathbf{Q})+2\mathbf{Q}\mathbf{C}_1^*\mathbf{x}]-(1-\delta)\mathbf{Q}\mathbf{C}_2^*\mathbf{x}\,\mathbf{tr}(\mathbf{M}\mathbf{A}).\end{aligned} \tag{4.18}$$

Also the second order conditional moments of $\mathbf{e}_{11}^*$ under **Assumption II** can be represented as

$$\begin{aligned}\mathbf{e}_{11}^*(\mathbf{x}) &= (2+\kappa)\mathbf{Q}\mathbf{D}'\mathbf{F}\mathbf{D}\mathbf{Q}+\mathbf{x}'\mathbf{C}_1^*\mathbf{x}\mathbf{x}\mathbf{x}'+\mathbf{Q}\mathbf{Q}^*\mathbf{x}'\mathbf{C}_2^*\mathbf{x}+\mathbf{Q}\mathbf{C}_2^*\mathbf{Q}\mathbf{tr}(\mathbf{M}\mathbf{A}) \\ &\quad +(1-\delta)^2L(L+2)\mathbf{Q}\mathbf{C}_1^*\mathbf{Q}-(1-\delta)L[\mathbf{Q}\mathbf{C}_1^*\mathbf{x}\mathbf{x}'+\mathbf{x}\mathbf{x}'\mathbf{C}_1^*\mathbf{Q})].\end{aligned} \tag{4.19}$$

Next, we consider the characteristic function of the standardized estimator $\hat{\mathbf{e}}$ in order to derive the asymptotic expansion of its distribution function. We shall calculate

$$\begin{aligned}C(t) &= \mathbf{E}[\exp(i\mathbf{t}'\mathbf{x})] \\ &\quad +\frac{1}{\sqrt{n}}\mathbf{E}[i\mathbf{t}'\mathbf{e}_1^*(\mathbf{x})\exp(i\mathbf{t}'\mathbf{x})] \\ &\quad +\frac{1}{2n}\mathbf{E}\{2i\mathbf{t}'\mathbf{e}_2^*(\mathbf{x})\exp(i\mathbf{t}'\mathbf{x})+i^2\mathbf{t}'\mathbf{e}_{11}^*(\mathbf{x})\mathbf{t}\exp(i\mathbf{t}'\mathbf{x})\}+O(\frac{1}{n\sqrt{n}})\,,\end{aligned} \tag{4.20}$$

where $\mathbf{x}=\tilde{\mathbf{e}}_0$, $\mathbf{t}=(t_i)$ is a $p\times 1$ vector of real variables and $i^2=-1$. By modifying the Fourier Inversion Formulae developed by *Appendix B* of Fujikoshi et. al. (1982), we can invert the characteristic function in (4.20). Although the intermediate computations are quite tedious but straightforward. We first consider the asymptotic expansion of the density function of $\tilde{\mathbf{e}}_0$ and we know that its limiting distribution as $n\to+\infty$ is normal. By expanding its characteristic function $\mathbf{E}[exp(i\mathbf{t}'\tilde{\mathbf{e}}_0)]$ and inverting it as the standard practice in the asymptotic theory, we have

$$\begin{aligned}&\phi_{\mathbf{Q}}^*(\boldsymbol{\xi}) \\ &= \phi_{\mathbf{Q}}(\xi)\Big\{1+\frac{1}{6\sqrt{n}}\sum_{l,l',l''=1}^{p}\beta_{l,l',l''}h_3(\xi_l,\xi_{l'},\xi_{l''}) \\ &\qquad +\frac{1}{24n}\sum_{l,l',l'',l'''=1}^{p}\beta_{l,l',l'',l'''}h_4(\xi_l,\xi_{l'},\xi_{l''},\xi_{l'''}) \\ &\qquad +\frac{1}{72n}\sum_{l,l',l'',m,m',m''=1}^{p}\beta_{l,l',l''}\beta_{m,m',m''}h_6(\xi_l,\xi_{l'},\xi_{l''},\xi_m,\xi_{m'},\xi_{m''})\Big\} \\ &\quad +O(\frac{1}{n^{3/2}}),\end{aligned} \tag{4.21}$$

where $\phi_{\mathbf{Q}}(\boldsymbol{\xi})$ is the p-dimensional normal density function with means $\mathbf{0}$ and the variance-covariance matrix $\mathbf{Q}$,

$$\beta_{l,l',l''} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}(u_i^3) z_{il}^* z_{il'}^* z_{il''}^* \ ,$$

$$\beta_{l,l',l'',l'''} = (\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}(u_i^4) z_{il}^* z_{il'}^* z_{il''}^* z_{il'''}^*)$$

$$- \sum_{l,l',l'',l'''} [\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}(u_i^2) z_{il}^* z_{il'}^*][\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}(u_i^2) z_{il''}^* z_{il'''}^*],$$

where $\mathbf{z}_i^* = (z_i^{(l)*}) = \mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{z}_i$ $(i = 1, \cdots, n)$, and $\sum_{l,l',l'',l'''}$ means the combinations of two pairs such as $(l, l^{'})$ and $(l^{''}, l^{'''})$ (it is 3 when $l = l^{'} = l^{''} = l^{'''}$, for instance). Also $h_3(x_l, x_{l'}, x_{l''})$ and $h_4(x_l, x_{l'}, x_{l''}, x_{l'''})$ are the Hermitian polynomials as

$$h_3(x_l, x_{l'}, x_{l''}) = (-1)^3 \frac{1}{\phi_{\mathbf{Q}}(x)} \frac{\partial^3 \phi_{\mathbf{Q}}(x)}{\partial x_l \partial x_{l'} \partial x_{l''}}, \tag{4.22}$$

$$h_4(x_l, x_{l'}, x_{l''}, x_{l'''}) = (-1)^4 \frac{1}{\phi_{\mathbf{Q}}(x)} \frac{\partial^4 \phi_{\mathbf{Q}}(x)}{\partial x_l \partial x_{l'} \partial x_{l''} \partial x_{l'''}}. \tag{4.23}$$

We notice that (4.21) is common for all efficient estimators and then it does not make any effects on the comparisons of alternative efficient estimators. When the third order moments of disturbances are zeros, the terms of $O_p(n^{-1/2})$ on the right-hand side vanish (i.e. $\beta_{l,l',l''} = 0$) and we only have extra terms in the order of n^{-1}. In these cases we can directly use the Fourier inversion formulae reported in Appendix B because only terms on the effects of non-normality of disturbance terms appear as $\mathbf{QD}^{'}\mathbf{FDx}$ in the order of $O_p(n^{-1})$ and the resulting expressions become considerably simplified. Also we notice that when the disturbance terms are conditionally homoscedastic as in **Assumption II**, we have $\mathbf{C} = \sigma^2\mathbf{M}$, $\mathbf{Q} = \sigma^2(\mathbf{D}^{'}\mathbf{MD})^{-1}$, $\mathbf{Q}^* = \mathbf{D}^{'}\mathbf{MC}^{-1}\mathbf{MC}^{-1}\mathbf{MD} = \sigma^{-2}\mathbf{Q}^{-1}$, and $\mathbf{tr}(\mathbf{MA}) = \sigma^{-2}L$.
Then we can state our main result after lengthy but straightforward computations by using the formulas given in Appendix B.

Theorem 4.1 : Under *Assumption I* and *Assumption II*, the asymptotic expansion of the joint density function of $\hat{\mathbf{e}}$ for the class of modified MEL estimators as $n \to \infty$ is given by

$$\begin{aligned}
&f(\boldsymbol{\xi}) = \phi_{\mathbf{Q}}^*(\boldsymbol{\xi}) \\
&+\frac{1}{\sqrt{n}}\phi_{\mathbf{Q}}(\boldsymbol{\xi})(\mathbf{q}^{'}\boldsymbol{\xi})[p+1+(1-\delta)L-\boldsymbol{\xi}^{'}\mathbf{Q}^{-1}\boldsymbol{\xi}] \\
&+\frac{1}{2n}\phi_{\mathbf{Q}}(\boldsymbol{\xi})\Big(\boldsymbol{\xi}^{'}\mathbf{C}_1\boldsymbol{\xi}\{[p+1+(1-\delta)L-\boldsymbol{\xi}^{'}\mathbf{Q}^{-1}\boldsymbol{\xi}]^2+p+1-3\boldsymbol{\xi}^{'}\mathbf{Q}^{-1}\boldsymbol{\xi}+2(1-\delta)^2L\} \\
&\qquad +\mathbf{tr}(\mathbf{C}_1\mathbf{Q})[(1-\delta)L][2-(1-\delta)(L+2)] \\
&\qquad +\boldsymbol{\xi}^{'}\mathbf{C}_2\boldsymbol{\xi}\{L[1-2(1-\delta)]-p-2+\boldsymbol{\xi}^{'}\mathbf{Q}^{-1}\boldsymbol{\xi}\}+\mathbf{tr}(\mathbf{C}_2\mathbf{Q})\{L[2(1-\delta)-1]\} \\
&\qquad +[2+(2\delta-1)\kappa][\boldsymbol{\xi}^{'}\mathbf{D}^{'}\mathbf{FD}\boldsymbol{\xi}-\mathbf{tr}(\mathbf{D}^{'}\mathbf{FDQ})]\Big) \\
&+o(\frac{1}{n}) \ ,
\end{aligned}$$

where $\boldsymbol{\xi}$ is a $p \times 1$ ($p = G_1 + K_1$) vector, $\phi^*_{\mathbf{Q}}(\xi)$ and $\mathbf{F}$ are given by (4.21) and (4.13), respectively, and $\mathbf{C}_1 = \mathbf{C}_1^* \ (= \mathbf{q}\mathbf{q}^{'})$, $\mathbf{C}_2 = \sigma^{-2}\mathbf{C}_2^* \ (= \sigma^{-2}E[\mathbf{w}_i\mathbf{w}_i^{'}])$, and $\kappa = [E(u_i^4) - 3\sigma^4)/\sigma^4$.

The first term $\phi^*_{\mathbf{Q}}(\boldsymbol{\xi})$ are common among all asymptotically efficient estimators and we need to make comparison on the terms of the second term of $O(n^{-1/2})$ and the third term of $O(n^{-1})$. When the disturbance terms are normally distributed all terms except the leading term vanish in (4.21) and $\phi^*_{\mathbf{Q}}(\mathbf{x}) = \phi_{\mathbf{Q}}(\mathbf{x})$. There is an interesting observation in Theorem 4.1 that if we further drop the last term

$$[2 + (2\delta - 1)\kappa][\xi^{'}\mathbf{D}^{'}\mathbf{F}\mathbf{D}\xi - \mathbf{tr}(\mathbf{D}^{'}\mathbf{F}\mathbf{D}\mathbf{Q})]$$

and the disturbance terms are normally distributed, the resulting formulas are identical to those for the limited information maximum likelihood (LIML) estimator and the two stage least squares (TSLS) estimator, which have been reported by Fujikoshi et. al. (1982).
By using the asymptotic expansion of the density function, we can evaluate the asymptotic mean and the asymptotic mean squared errors of the modified MEL estimator. We summarize the resulting formulas.

Theorem 4.2 : Under the assumptions of *Theorem 4.1*, the asymptotic bias and the asymptotic mean squared errors of $\hat{\mathbf{e}}$ for the modified (MEL) estimators based on the asymptotic expansion of the density function as $n \to \infty$ up to $O(n^{-1})$ are given by

$$ABIAS_n(\hat{\mathbf{e}}) = \frac{1}{\sqrt{n}}[(1-\delta)L - 1]\mathbf{Q}\mathbf{q} + o(\frac{1}{\sqrt{n}}) \ , \tag{4.24}$$

and

$$\begin{aligned} & AMSE_n(\hat{\mathbf{e}}) \\ = \ & \mathbf{Q} + \frac{1}{n}\Big\{\mathbf{Q}\mathbf{C}_1\mathbf{Q}[6 - 6(1-\delta)L + (1-\delta)^2L(L+2)] \\ & +\mathbf{Q}\mathbf{tr}(\mathbf{C}_1\mathbf{Q})[3 - 2(1-\delta)L] + \mathbf{Q}\mathbf{tr}(\mathbf{C}_2\mathbf{Q}) + [L + 2 - 2L(1-\delta)]\mathbf{Q}\mathbf{C}_2\mathbf{Q} \\ & +[2 + (2\delta-1)\kappa]\mathbf{Q}\mathbf{D}^{'}\mathbf{F}\mathbf{D}\mathbf{Q}\Big\} + o(\frac{1}{n}) \ , \end{aligned} \tag{4.25}$$

where we use the notations in Theorem 4.1.

4.3 A Simple Case

We notice that the exact density functions of estimators and their asymptotic expansions are quite complicated in the general case. Hence it is interesting to derive the asymptotic expansions of the distribution functions of estimators in the simplest case when $G_1 = 1$. We take the estimator on the coefficient of an endogenous variable in the right hand side and standardize

$$\mathbf{P}(\frac{\sqrt{\mathbf{\Pi}_{22}^{'}\mathbf{A}_{22.1}\mathbf{\Pi}_{22}}}{\sigma}(\hat{\beta} - \beta) \le x) \tag{4.26}$$

since its limiting distribution is the standard normal.

In the univariate case we use the notation

$$Q_{11} = \sigma^2 \left(\mathbf{\Pi}_{22}^{'}\mathbf{M}_{22.1}\mathbf{\Pi}_{22}\right)^{-1}$$

as the $(1,1)-$element of $\mathbf{Q}$ and we partition a $[1+(p-1)] \times [1+(p-1)]$ matrix as

$$\mathbf{Q} = \begin{pmatrix} Q_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix} .$$

The right-hand side of $\phi^*(x)$ for the standardized estimator in (4.26) can be simplified and it is in the form of

$$(4.27)\phi(x)\{1+\frac{1}{\sqrt{n}}[\beta_3(x^3-3x)]+\frac{1}{n}[\frac{\beta_4}{24}(x^4-6x^2+3)+\frac{\beta_3^2}{72}(x^6-15x^4+45x^2-15)]\},$$

where β_3 and β_4 are the third and fourth order cumulants in (4.21) by replacing z_i^{**} $(= Q_{11}^{-1/2} z_i^{(1)*})$ for $\mathbf{z}_i^*$ $(i=1,\cdots,n)$ and $\phi(x)$ is the density function of the standard normal distribution.
For any $p-$dimensional normal density $\phi_{\mathbf{Q}}(\xi)$, we partition any $[1+(p-1)]\times[1+(p-1)]$ matrix $\mathbf{B}$ and p $(=[1+(p-1)])$ vector $\xi^{'} = (\xi_1, \xi_2^{'})^{'}$, we can write

$$(4.28) \qquad \int_{\mathbf{R}^{p-1}} [\xi^{'}\mathbf{B}\xi - \mathbf{tr}(\mathbf{BQ})]\phi_{\mathbf{Q}}(\xi)d\xi_2 = E_{\xi_2|\xi_1}\{\mathbf{tr}[\mathbf{B}(\xi\xi^{'}-\mathbf{Q})]\phi_{Q_{11}}(\xi_1)\} .$$

By using the fact that $\xi_2|\xi_1$ follows the $(p-1)-$dimensional (conditional) normal distribution $N_{p-1}[\mathbf{Q}_{21}Q_{11}^{-1}\xi_1, \mathbf{Q}_{22.1}]$ and $\mathbf{Q}_{22.1} = \mathbf{Q}_{22} - \mathbf{Q}_{21}Q_{11}^{-1}\mathbf{Q}_{21}$, we can evaluate (4.28) explicitly and it is rewritten as

$$(4.29) \quad \mathbf{tr}\left[\mathbf{B}\begin{pmatrix} \xi_1^2 - Q_{11} & \xi_1^2 Q_{11}^{-1}\mathbf{Q}_{12} - \mathbf{Q}_{12} \\ \mathbf{Q}_{21}Q_{11}^{-1}\xi_1^2 - \mathbf{Q}_{21} & \mathbf{Q}_{21}Q_{11}^{-1}\xi_1^2 Q_{11}^{-1}\mathbf{Q}_{12} + \mathbf{Q}_{22.1} - \mathbf{Q}_{22} \end{pmatrix}\right]\phi_{Q_{11}}(\xi_1)$$
$$= \mathbf{tr}[Q_{11}^{-1}(1,\mathbf{0}^{'})\mathbf{QCQ}\begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}Q_{11}^{-1}(\xi_1^2 - Q_{11})]\phi_{Q_{11}}(\xi_1).$$

When the third order moments of disturbances are zeros, the asymptotic expansion of the density function of the standardized estimator can be simplified. We set the standardized form as (4.26) and notice that two matrices $\mathbf{C}_1$ and $\mathbf{C}_2$ appeared in the asymptotic expansions have non-zero elements only in the upper-left parts. By evaluating the integrations with respect to the $(p-1)$ last elements of ξ in Theorem 4.1, the asymptotic expansion of the density becomes

$$(4.30) \qquad \phi^*(x) + \frac{1}{\sqrt{n}}\phi(x)(-\alpha_* x)[2+(1-\delta)L - x^2]$$
$$+\frac{1}{2n}\phi(x)\{\alpha_*^2 x^2[(2+(1-\delta)L-x^2)^2 + 2 - 3x^2 + 2(1-\delta)^2 L]$$
$$+\alpha_*^2[(1-\delta)L][2-(1-\delta)(L+2)]$$
$$+\eta_*\left(x^2[L(1-2(1-\delta)) - 3 + x^2] + L[2(1-\delta)-1]\right)$$
$$+\tau_*[2+(2\delta-1)\kappa](x^2-1)\}$$
$$+O_p(\frac{1}{n^{3/2}}) ,$$

where $\phi^*(\cdot)$ is given by (4.28) with $\beta_3 = 0$, $\alpha_* = -Cov(\mathbf{v}_{2i}, u_i)/|\mathbf{\Omega}|^{1/2}$, $\eta_* = |\Omega|/\sigma^2$, and

$$\tau_* = (1, \mathbf{0}) Q_{11}^{-1} \mathbf{Q} \mathbf{D}' \mathbf{F} \mathbf{D} \mathbf{Q} \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} .$$

If we further assume the normal disturbances, then the formula can be further simplified. By setting $\delta = 1$ for the MEL estimator, we have

$$\begin{aligned} \text{(4.31)} \qquad & \mathbf{P}\Big(\frac{\sqrt{\mathbf{\Pi}_{22}' \mathbf{A}_{22.1} \mathbf{\Pi}_{22}}}{\sigma}(\hat{\beta}_{MEL} - \beta) \le x)\Big) \\ = \quad & \Phi(x) + \{-\frac{\alpha}{\mu}x^2 - \frac{1}{2\mu^2}[(\tau + L)x + (1 - 2\alpha^2)x^3 + \alpha^2 x^5]\}\phi(x) \\ & +O(\mu^{-3}), \end{aligned}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the cdf and the density function of the standard normal distribution, respectively, and $\alpha = \omega_{22}[\beta - \frac{\omega_{12}}{\omega_{22}})]/|\Omega|^{1/2}$,

$$\text{(4.32)} \qquad \mu^2 = (1 + \alpha^2)\frac{\mathbf{\Pi}_{22}' \mathbf{A}_{22.1} \mathbf{\Pi}_{22}}{\omega_{22}} ,$$

$$\text{(4.33)} \qquad \tau = 2\frac{1 + \alpha^2}{\omega_{22}}(1, \mathbf{0}) Q_{11}^{-1} \mathbf{Q} \mathbf{D}' \mathbf{F} \mathbf{D} \mathbf{Q} Q_{11}^{-1} \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} .$$

Also under the normal disturbances and we set $\delta = 0$ for the GMM estimator, the resulting asymptotic expansion of the distribution function becomes

$$\begin{aligned} \text{(4.34)} \quad & \mathbf{P}\Big(\frac{\sqrt{\mathbf{\Pi}_{22}' \mathbf{A}_{22.1} \mathbf{\Pi}_{22}}}{\sigma}(\hat{\beta}_{GMM} - \beta) \le x)\Big) \\ = \quad & \Phi(x) + \{-\frac{\alpha}{\mu}[x^2 - L] - \frac{1}{2\mu^2}[(\tau + L^2\alpha^2)x + (1 - 2(L+1)\alpha^2)x^3 + \alpha^2 x^5]\}\phi(x) \\ & +O(\mu^{-3}) . \end{aligned}$$

Furthermore if we set $\tau = 0$ in the above expressions, the resulting formulas in (4.31) and (4.34) are identical to those for the limited information maximum likelihood (LIML) estimator and the two stage least squares (TSLS) estimator obtained by Anderson (1974), and Anderson and Sawa (1973), respectively.

4.4 A More General Case

When we remove **Assumption II**, it is straightforward to obtain the asymptotic expansion of the density function of the MEL and GMM estimators up to the order $O(n^{-1/2})$. We find that the explicit form of the asymptotic expansion of the density function of the MEL estimator given in *Theorem 4.1* is unchanged up to the order $O(n^{-1/2})$ with $\phi_{\mathbf{Q}}^*(\boldsymbol{\xi})$. We also see that the factor $\phi_{\mathbf{Q}}(\boldsymbol{\xi})(\mathbf{q}'\boldsymbol{\xi})[p + 1 - \boldsymbol{\xi}'\mathbf{Q}^{-1}\boldsymbol{\xi}]$ in the term $O(1/\sqrt{n})$ is symmetric around zeros. Then if we write $\hat{e}_i$ is the i-th component of $\hat{\mathbf{e}}$ for the MEL estimator,

$$\text{(4.35)} \qquad \mathbf{P}(\hat{e}_i \ge 0) = \frac{1}{2} + o(\frac{1}{\sqrt{n}})$$

when the third order moments of disturbances are zeros. On the other hand, the asymptotic expansion of the density function of the GMM estimator has an additional term involving the third order moments and the term of $O(n^{-1/2})$ is proportional to $L(1/\sqrt{n})$, where $L = K_2 - G_1$. Hence if K_2 (the number of excluded instruments) is large, the probability bias of the GMM (or the TSLS) estimator becomes sustantial while the MEL (or the LIML) estimator concentrates its probability around the true parameter values (see Tables in Anderson et. al. (2005) when $G_1 = 1$).

Although it is possible to derive the explicit expressions of the asymptotic expansions of the terms of $O(n^{-1})$, they become quite tedious because there are many additional terms as we have illustrated in the derivation of the asymptotic expansions in Sections 3 and 4. However, it is rather straightforward to derive the asymptotic bias of the estimator in the general case. By using (4.6) we have

$$(4.36)\ ABIAS_n(\hat{\mathbf{e}}) = \frac{1}{\sqrt{n}}\Big\{[(1-\delta)L-1]\mathbf{Qq} - (1-\delta)\mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{m}_3\Big\} + o(\frac{1}{\sqrt{n}}).$$

The above formula agrees with the recent result by Newey and Smith (2004) on the asymptotic bias of the MEL and GMM estimators in the more general nonlinear setting of estimating equation model.

Although it is straightforward to proceed our step to the mean-squared errors of alternative estimators, it is quite tedious to obtain the explicit formula of the additional terms in the asymptotic MSE of the modified estimators in the general case. In principle it can be calculated by

$$\begin{aligned} AM_n(\hat{\mathbf{e}}\hat{\mathbf{e}}^{'}) &= \mathbf{E}\Big\{[\tilde{\mathbf{e}}_0 + \frac{1}{\sqrt{n}}(\mathbf{e}_0^{(1)} + \mathbf{e}_1^{(0)}) + \frac{1}{n}(\mathbf{e}_0^{(2)} + \mathbf{e}_1^{(1)} + \mathbf{e}_2)] \\ &\quad \times[\tilde{\mathbf{e}}_0 + \frac{1}{\sqrt{n}}(\mathbf{e}_0^{(1)} + \mathbf{e}_1^{(0)}) + \frac{1}{n}(\mathbf{e}_0^{(2)} + \mathbf{e}_1^{(1)} + \mathbf{e}_2)]^{'}\Big\} + o(\frac{1}{n})\ . \end{aligned}$$

In additon to $AMSE^*(\hat{\mathbf{e}})$ of $AM(\hat{\mathbf{e}}\hat{\mathbf{e}}^{'})$ in *Theorem 4.2*, there are several additional terms for an arbitrary δ $(0 \leq \delta \leq 1)$ when we cannot ignore the effects of third order moments of disturbance terms. For the MEL estimator case, however, there are only a few additional terms. Although it is straightforward to write down those terms, we have omitted to report the details since they are complicated and may not be useful at the present stage of our investigation.

5. Concluding Remarks

In this paper we have developed the asymptotic expansions of the density functions for the class of semi-parametric estimators including the MEL estimator and the GMM estimator. Although the general forms of the asymptotic expansions are quite complicated at the first glance, it is possible to obtain some interesting observations which make possible to compare alternative estimation methods. It would be desirable to investigate the finite sample distributions of estimators and Anderson et. al. (2005), for instance, have investigated the finite sample properties of the distribution functions of the MEL and GMM estimators and have given extensive tables when $G_1 = 1$ in a systematic way. In the more general case, however, it would not be possible to investigate the finite sample properties directly and the asymptotic expansion method we have developed should be useful for comparing different estimators. The explicit formulas we have derived in this paper give some useful information on the exact ditributions of alternative estimators in more general cases.

The issue of comparing the finite sample distributions of alternative estimators are closely related to the problem of higher order asymptotic efficiency of estimation. Takeuchi and Morimune (1985) gave the classic result on the simultaneous equations system in the parametric framework and shown that the LIML estimator is third order asymptotically efficient when the dusturbances are normally distributed. Recently, Newey and Smith (2004) concluded that the MEL estimator is third order asymptotically efficient when the disturbances are the multinomial (lattice) distribution in the more general nonlinear estimating equation framework. In general, the asymptotic expansions of distributions of estimators in the lattice case are different from those in the non-lattice case due to the lack of so-called *the Cramér condition* (see Chapter 5 of Bhattacharya and Rao (1976), for instance). However, it is important to note that the differences between the distributions of the LIML and MEL estimators are very small as we have discussed in Section 4 at least when the disturbances are i.i.d. non-lattice random variables with zero third moments. It may be interesting to see if these differences would be substantial for the practical purposes.

References

[1] Anderson, T.W. (1974), "An Asymptotic Expansion of the Distribution of the Limited Information Maximum Likelihood Estimate of a Coefficient in a Simultaneous Equation System," *Journal of the American Statistical Asociation*, Vol. 69, 565-573.

[2] Anderson, T.W. and T. Sawa (1973), "Distributions of Estimates of Coefficients of a Single Equation in a Simultaneous System and Their Asymptotic Expansion," *Econometrica*, Vol. 41, 683-714.

[3] Anderson, T.W., N. Kunitomo, and T. Sawa (1982), "Evaluation of the Distribution Function of the Limited Information Maximum Likelihood Estimator," *Econometrica*, Vol. 50, 1009-1027.

[4] Anderson, T.W., N. Kunitomo, and K. Morimune (1986), "Comparing Single Equation Estimators in a Simultaneous Equation System," *Econometric Theory*, Vol. 2, 1-32.

[5] Anderson, T.W., N. Kunitomo, and Y. Matsushita (2005), "A New Light from Old Wisdoms : Alternative Estimation Method of Simultaneous Equations and Microeconometric Models," Discussion Paper CIRJE-F-321, Graduate School of Economics, University of Tokyo (http://www.e.u-tokyo.ac.jp/cirje/research/dp/2005).

[6] Bhattacharya, R.N. and Ghosh, J.K. (1978), "On the Validity of the Formal Edgeworth Expansion," *The Annals of Statistics*, Vol. 6, 434-451.

[7] Bhattacharya, R.N. and R. Rao (1976) *Normal Approximations and Asymptotic Expansions*, Jhon-Wiley.

[8] Fujikoshi, Y., K. Morimune, N. Kunitomo, and M. Taniguchi (1982), "Asymptotic Expansions of the Distributions of the Estimates of Coefficients in a Simultaneous Equation System," *Journal of Econometrics*, Vol. 18, 2, 191-205.

[9] Hansen, L. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, Vol. 50, 1029-1054.

[10] Hayashi, F. (2000) *Econometrics*, Princeton University Press.

[11] Godambe, V.P. (1960), "An Optimum Property of Regular Maximum Likelihood Equation," *Annals of Mathematical Statistics*, Vol. 31, 1208-1211.

[12] Kitamura, Y. and M. Stutzer (1997), "An Informational-Theoretic Alternative to Genaralized method of Moments," *Econometrica*, 65, 861-874.

[13] Kitamura, Y. G. Tripathi, and H. Ahn (2004), "Empirical Likelihood-Based Inference in Conditional Moment Restriction Models," *Econometrica*, 72, 1667-1714.

[14] Newey, W. K. and R. Smith (2004), "Higher Order Properties of GMM and Generalized Empirical Likelihood Estimator," *Econometrica*, 72, 219-255.

[15] Owen, A. B. (1990), "Empirical Likelihood Ratio Confidence Regions," *The Annals of Statistics*, Vol. 22, 300-325.

[16] Owen, A. B. (2001), *Empirical Likelihood*, Chapman and Hall.

[17] Phillips, P.C.B. (1983), "Exact Small Sample Theory in the Simultaneous Equations Model," *Handbook of Econometrics*, Vol. 1, 449-516, North-Holland.

[18] Qin, J. and Lawless, J. (1994), "Empirical Likelihood and General Estimating Equations," *The Annals of Statistics*, Vol. 22, 300-325.

[19] Sargan, J. D. and Mikhail, W.M. (1971), "A General Approximation to the Distribution of Instrumental Variables Estimates," *Econometrica*, Vol. 39, 131-169.

[20] Takeuchi, K. and Morimune, K. (1985), "Third-Order Efficiency of the Extended Maximum Likelihood Estimators in a Simultaneous Equations System," *Econometrica*, Vol. 53, 177-200.

Appendices

In Appendix A, we give the proofs of two lemmas used in Section 4. In Appendix B we gather some useful formulae on the Fourier inversion because it seems they are not readily available.

Appendix A : Proof of Lemmas

[A-1] : Proof of Lemma 4.1

Let $X_1 = (\mathbf{Y}_n)_{ij}$, $X_2 = (\mathbf{A}\mathbf{X}_n)_k$ and $X_3 = (\tilde{\mathbf{e}}_0)_l$. Since the limiting distribution of random vector $(X_1, X_2, X_3)^{'}$ is normal, the conditional distribution of $(X_1, X_2)^{'}$ given X_3 is also asymptotically normal. Then we have

$$\mathbf{E}[X_1X_2|X_3] \cong \mathbf{E}[X_1|X_3]E[X_2|X_3] + [Cov(X_1, X_2) - \frac{Cov(X_1, X_3)Cov(X_2, X_3)}{Var(X_3)}].$$

Because X_2 and X_3 are asymptotically orthogonal, $E[X_2|X_3] \cong 0$ and $Cov(X_2, X_3) \cong 0$. Also by using the notation $\mathbf{z}_{\alpha}^{(j)}$ and noting that

$$\text{(A.1)} \qquad Cov(X_1, X_2) \cong \frac{1}{n}\sum_{\alpha=1}^{n} z_{\alpha}^{(i)} z_{\alpha}^{(j)} (\mathbf{A}\mathbf{z}_{\alpha})^{(k)} \mathbf{E}[u_{\alpha}^3],$$

we have the result. (Q.E.D)

[A-2] : Proof of Lemma 4.3

Let $\mathbf{z}_n = (\mathbf{u}_n^{'}, v_n)^{'}$ be a $(p+1) \times 1$ random vector which is a sum of i.i.d. random vectors $\mathbf{z}_j^{(n)}$ $(j = 1, \cdots, n)$:

$$\text{(A.2)} \qquad \mathbf{z}_n = \frac{1}{\sqrt{n}}\sum_{j=1}^{n} \mathbf{z}_j^{(n)}$$

and $\mathbf{E}[\mathbf{z}_j^{(n)}] = \mathbf{0}$, $\mathbf{E}[\mathbf{z}_j^{(n)}\mathbf{z}_j^{(n)'}] = \mathbf{\Sigma} > \mathbf{0}$. Then under a set of regularity conditions (Bhattacharya and Rao (1976) for instance) the characteristic function of $\mathbf{z}_n$ can be expressed as

$$\begin{aligned} \varphi(\mathbf{t}) &= \prod_{j=1}^{n} \mathbf{E}[e^{i\sum_{k=1}^{p+1} t_j z_{jk}^{(n)}}] \\ &= e^{-\frac{1}{2}\mathbf{t}^{'}\mathbf{\Sigma}\mathbf{t}}\{1 + \frac{1}{6\sqrt{n}} \sum_{l,l^{'},l^{''}=1}^{p+1} \beta_{l,l^{'},l^{''}}(it_l)(it_{l^{'}})(it_{l^{''}})\} + O(\frac{1}{n}) , \end{aligned}$$

where $\beta_{l,l^{'},l^{''}}$ are the third order moments of $\mathbf{z}_j^{(n)}$. Then the density function of $\mathbf{z}_n$ has a representation

$$\text{(A.3)} \qquad f_n(\mathbf{z}) = \phi_{\Sigma}(\mathbf{z})\Big\{1 + \frac{1}{6\sqrt{n}} \sum_{l,l^{'},l^{''}=1}^{p+1} \beta_{l,l^{'},l^{''}} h_3(z_l, z_{l^{'}}, z_{l^{''}})\Big\} + O(\frac{1}{n}) ,$$

where $h_3(z_l, z_{l^{'}}, z_{l^{''}})$ are the third-order Hermitian polynomials and we set a $(p+1) \times (p+1)$ variance-covariance matrix of $\mathbf{z}_n$ as

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{I}_p & \boldsymbol{\rho} \\ \boldsymbol{\rho}^{'} & 1 \end{pmatrix}$$

for the mathematical convenience. Let $f_n(\mathbf{u}_n)$ be the marginal density and $f_n(v_n|\mathbf{u}_n)$ be the conditional density, which can be represented as

$$
\begin{aligned}
& f_n(v_n|\mathbf{u}_n) = \phi(v|\boldsymbol{\rho}^{'}\mathbf{u}_n, 1-\boldsymbol{\rho}^{'}\boldsymbol{\rho}) \\
& \times\Big\{1+\frac{1}{6\sqrt{n}}[\sum_{l,l^{'},l^{''}=1}^{p}\beta_{l,l^{'},l^{''}}h_{3,\cdot}(u_l,u_{l^{'}},u_{l^{''}})+3\sum_{l,l^{'}}^{p}\beta_{l,l^{'},p+1}h_{3,\cdot}(u_l,u_{l^{'}},v_n) \\
& \quad +3\sum_{l=1}^{p}\beta_{l,p+1,p+1}h_{3,\cdot}(u_l,v_n,v_n)+\beta_{p+1,p+1,p+1}h_{3,\cdot}(v_n,v_n,v_n) \\
& \quad -\sum_{l,l^{'},l^{''}=1}^{p}\beta_{l,l^{'},l^{''}}h_3(u_l,u_{l^{'}},u_{l^{''}})]\Big\}+O(\frac{1}{n})\ ,
\end{aligned}
$$

where $\phi(v|\boldsymbol{\rho}^{'}\mathbf{u}_n, 1-\boldsymbol{\rho}^{'}\boldsymbol{\rho})$ is the conditional density function, and $h_{3,\cdot}(\cdot)$ are the third order Hermitian polynomials for the $(p+1)-$dimensional random vector $(\mathbf{u}_n, v)$ and $h_3(\cdot)$ are the third order Hermitian polynomials for the $p-$dimensional random vector $\mathbf{u}_n$. Then the conditional expectation is given by

$$
\begin{aligned}
& \mathbf{E}[v_n|\mathbf{u}_n] \\
= \ & \boldsymbol{\rho}^{'}\mathbf{u}_n+\frac{1}{6\sqrt{n}}\Big\{\sum_{l,l^{'},l^{''}=1}^{p}\beta_{l,l^{'},l^{''}}\int v(-1)^3\frac{\partial^3 f_n(\mathbf{u}_n,v)}{\partial u_l\partial u_{l^{'}}\partial u_{l^{''}}}\frac{1}{f_n(\mathbf{u}_n)}dv \\
& +3\sum_{l,l^{'}=1}^{p}\beta_{l,l^{'},p+1}(-1)^3\frac{\partial^2}{\partial u_l\partial u_{l^{'}}}\int v\frac{\partial f_n(\mathbf{u}_n,v)}{\partial v}\frac{1}{f_n(\mathbf{u}_n)}dv \\
& +3\sum_{l=1}^{p}\beta_{l,p+1,p+1}(-1)^3\frac{\partial}{\partial u_l}\int v\frac{\partial^2 f_n(\mathbf{u}_n,v)}{\partial v^2}\frac{1}{f_n(\mathbf{u}_n)}dv \\
& +\beta_{p+1,p+1,p+1}(-1)^3\int v\frac{\partial^3 f_n(\mathbf{u}_n,v)}{\partial v^3}\frac{1}{f_n(\mathbf{u}_n)}dv \\
& -(\boldsymbol{\rho}^{'}\mathbf{u}_n)\sum_{l,l^{'},l^{''}=1}^{p}\beta_{l,l^{'},l^{''}}h_3(u_l,u_{l^{'}},u_{l^{''}})\Big\}+O(\frac{1}{n}).
\end{aligned}
$$

By using the integral-by-parts calculations, the third term and the fourth term of the right-hand side of $O(n^{-1/2})$ are zeros. Hence

$$
\begin{aligned}
& \mathbf{E}[v_n|\mathbf{u}_n] \\
& = \rho^{'}\mathbf{u}_n \\
& \quad +\frac{1}{6\sqrt{n}}\Big\{(-1)\sum_{l,l^{'},l^{''}=1}^{p}\beta_{l,l^{'},l^{''}}[\frac{\partial^3 f_n(\cdot)}{\partial u_l\partial u_{l^{'}}\partial u_{l^{''}}}(\boldsymbol{\rho}^{'}\mathbf{u}_n f_n(\mathbf{u}_n))]/f_n(\mathbf{u}_n) \\
& \quad +3\sum_{l,l^{'}=1}^{p}\beta_{l,l^{'},p+1}[\frac{\partial^2}{\partial u_l\partial u_{l^{'}}}f_n(\mathbf{u}_n)]/f_n(\mathbf{u}_n)-\boldsymbol{\rho}^{'}\mathbf{u}_n\sum_{l,l^{'},l^{''}=1}^{p}\beta_{l,l^{'},l^{''}}h_3(u_l,u_{l^{'}},u_{l^{''}})\Big\}+O(\frac{1}{n}).
\end{aligned}
$$

$$= \rho' \mathbf{u}_n$$

$$+ \frac{1}{6\sqrt{n}} \Big\{ 3 \sum_{l,l'=1}^{p} \beta_{l,l',p} h_2(u_l, u_{l'}) - \sum_{l,l',l''=1}^{p} \beta_{l,l',l''} [\boldsymbol{\rho}' \mathbf{u}_n h_3(u_l, u_{l'}, u_{l''})]$$

$$+ \sum_{l,l',l''=1}^{p} \beta_{l,l',l''} [\rho' \mathbf{u}_n h_3(u_l, u_{l'}, u_{l''}) - \rho_l h_2(u_{l'}, u_{l''}) - \rho_{l'} h_2(u_{l'}, u_{l''}) - \rho_{l''} h_2(u_{l'}, u_{l''})] \Big\}$$

$$+ O(\frac{1}{n}) ,$$

where $h_2(u_l, u_{l'})$ are the second order Hermite polynomials of $p-$dimensional vector $\mathbf{u}_n$. Since two terms in the above expressions on the right-hand side are cancelled out, we have the desired result. (Q.E.D.)

Appendix B : Useful Formulas

This appendix gives the useful formulas, which correspond to the inversion of the characteristic function from the conditional expectations given $\mathbf{z}$ and $\mathbf{z}$ follows the p-dimensional normal distribution $N_p(\mathbf{0}, \mathbf{Q})$. All inversion results we have needed in Sections 3 and 4 can be reduced the Fourier Inversion formulas for the density function as

(A.4) $$\mathcal{F}^{-1}\{h(-i\mathbf{t})\mathcal{E}[g(\mathbf{z})\exp(i\mathbf{t}'\mathbf{z})]\} = h(\frac{\partial}{\partial \boldsymbol{\xi}})g(\boldsymbol{\xi})\phi_{\mathbf{Q}}(\boldsymbol{\xi})$$

for any polynomials $h(\,\cdot\,)$ and $g(\,\cdot\,)$, where $\mathbf{t} = (t_i)$ is a $p \times 1$ vector, $i^2 = -1$, and the differentiation vector

$$\frac{\partial}{\partial \boldsymbol{\xi}'} = (\frac{\partial}{\partial \xi_1}, \cdots, \frac{\partial}{\partial \xi_p}) .$$

The method adopted here was developed by Fujikoshi et. al. (1982) and they were given by Anderson et. al. (1986). We present some useful results.

Lemma B.1 : Let $\boldsymbol{\alpha}' = (\alpha_1, \cdots, \alpha_p)$ be a $1 \times p$ constant vector, $\mathbf{A}$ be a symmetric constant matrix, and $\mathbf{tr}\frac{\partial^2}{\partial \boldsymbol{\xi} \partial \boldsymbol{\xi}'}[\,\cdot\,]$ stands for $\sum_i \sum_j \partial^2 / \partial \xi_i \partial \xi_j [\,\cdot\,]_{ij}$. Then

(i) $\frac{\partial}{\partial \boldsymbol{\xi}'}[\boldsymbol{\alpha}\phi_{\mathbf{Q}}(\boldsymbol{\xi})] = -\boldsymbol{\alpha}\mathbf{Q}^{-1}\boldsymbol{\xi}\phi_{\mathbf{Q}}(\boldsymbol{\xi})$,

(ii) $\frac{\partial}{\partial \boldsymbol{\xi}'}[\boldsymbol{\xi}(\boldsymbol{\alpha}'\boldsymbol{\xi})\phi_{\mathbf{Q}}(\boldsymbol{\xi})] = (\boldsymbol{\alpha}'\boldsymbol{\xi})(p+1-\boldsymbol{\xi}'\mathbf{Q}^{-1}\boldsymbol{\xi})\phi_{\mathbf{Q}}(\boldsymbol{\xi})$,

(iii) $\frac{\partial}{\partial \boldsymbol{\xi}'}[\mathbf{QA}\boldsymbol{\xi}\phi_{\mathbf{Q}}(\boldsymbol{\xi})] = [\mathbf{tr}(\mathbf{AQ}) - \boldsymbol{\xi}'\mathbf{A}\boldsymbol{\xi})]\phi_{\mathbf{Q}}(\boldsymbol{\xi})$,

(iv) $\frac{\partial}{\partial \boldsymbol{\xi}'}[\boldsymbol{\xi}\boldsymbol{\xi}'\mathbf{A}\phi_{\mathbf{Q}}(\boldsymbol{\xi})] = (\boldsymbol{\xi}'\mathbf{A}\boldsymbol{\xi})(\boldsymbol{\alpha}'\boldsymbol{\xi})(p+2-\boldsymbol{\xi}'\mathbf{Q}^{-1}\boldsymbol{\xi})\phi_{\mathbf{Q}}(\boldsymbol{\xi})$,

(v) $\mathbf{tr}\frac{\partial^2}{\partial \boldsymbol{\xi}\partial \boldsymbol{\xi}'}[\mathbf{QAQ}\phi_{\mathbf{Q}}(\boldsymbol{\xi})], = [\boldsymbol{\xi}'\mathbf{A}\boldsymbol{\xi} - \mathbf{tr}(\mathbf{AQ})]\phi_{\mathbf{Q}}(\boldsymbol{\xi})$,

(vi) $\mathbf{tr}\frac{\partial^2}{\partial \boldsymbol{\xi}\partial \boldsymbol{\xi}'}[\mathbf{Q}\boldsymbol{\xi}'\mathbf{A}\boldsymbol{\xi}\phi_{\mathbf{Q}}(\boldsymbol{\xi})] = [2\mathbf{tr}(\mathbf{AQ}) - (p+4-\boldsymbol{\xi}'\mathbf{Q}^{-1}\boldsymbol{\xi})\boldsymbol{\xi}'\mathbf{A}\boldsymbol{\xi}]\phi_{\mathbf{Q}}(\boldsymbol{\xi})$,

(vii) $\mathbf{tr}\frac{\partial^2}{\partial \boldsymbol{\xi}\partial \boldsymbol{\xi}'}[\mathbf{QA}\boldsymbol{\xi}\boldsymbol{\xi}'\phi_{\mathbf{Q}}(\boldsymbol{\xi})] = [(p+1-\boldsymbol{\xi}'\mathbf{Q}^{-1}\boldsymbol{\xi})(\mathbf{tr}(\mathbf{AQ}) - \boldsymbol{\xi}'\mathbf{A}\boldsymbol{\xi})) - 2\boldsymbol{\xi}'\mathbf{A}\boldsymbol{\xi}]\phi_{\mathbf{Q}}(\boldsymbol{\xi})$,

(viii) $\mathbf{tr}\frac{\partial^2}{\partial \boldsymbol{\xi}\partial \boldsymbol{\xi}'}[\boldsymbol{\xi}\boldsymbol{\xi}'\boldsymbol{\xi}'\mathbf{A}\boldsymbol{\xi}\phi_{\mathbf{Q}}(\boldsymbol{\xi})] = \boldsymbol{\xi}'\mathbf{A}\boldsymbol{\xi}[(p+1-\boldsymbol{\xi}'\mathbf{Q}^{-1}\boldsymbol{\xi})^2+3(p+1)+2-5\boldsymbol{\xi}'\mathbf{Q}^{-1}\boldsymbol{\xi})]\phi_{\mathbf{Q}}(\boldsymbol{\xi})$.