CIRJE-F-611

Asymptotic Expansions and Higher Order Properties of Semi-Parametric Estimators in a System of Simultaneous Equations

Naoto Kunitomo
University of Tokyo

Yukitoshi Matsushita
JSPS and Graduate School of Economics, University of Tokyo

February 2009

Asymptotic Expansions and Higher Order Properties of Semi-Parametric Estimators in a System of Simultaneous Equations *

Naoto Kunitomo †
and
Yukitoshi Matsushita

December 26, 2008
(Final Version)

Abstract

Asymptotic expansions are made for the distributions of the Maximum Empirical Likelihood (MEL) estimator and the Estimating Equation (EE) estimator (or the Generalized Method of Moments (GMM) in econometrics) for the coefficients of a single structural equation in a system of linear simultaneous equations, which corresponds to a reduced rank regression model. The expansions in terms of the sample size, when the non-centrality parameters increase proportionally, are carried out to $O(n^{-1})$. Comparisons of the distributions 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 higher order properties of alternative estimators including the semi-parametric inefficiency factor under the homoscedasticity assumption.

Key Words

Asymptotic Expansions, Empirical Likelihood, Estimating Equation, Generalized Method of Moments, Limited Information Maximum Likelihood, Two-Stage Least Squares, Linear Simultaneous Equations, Reduced Rank Regression.

*KM08-12-26-4. This paper is a revision of Discussion Paper CIRJE-F-237 (Graduate School of Economics, University of Tokyo) under the title "Asymptotic Expansions of Distributions of Semi-Parametric Estimators in a Linear Simultaneous Equation System." We thank T.W.Anderson, Masafumi Akahira, Yasunori Fujikoshi and the referees of this journal for giving constructive comments to the previous versions of this paper.

†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 econometrics has led to develop several estimation methods as alternatives to the least squares estimation method. The classical examples are the limited information maximum likelihood (LIML) method and the instrumental variables (IV) method including the two-stage least squares (TSLS) method. See Anderson and Rubin (1949), Anderson, Kunitomo and Sawa (1982), Phillips (1983), and Anderson, Kunitomo and Morimune (1986) for their finite sample properties, for instance. The estimation problem of a single structural equation is the same as the reduced rank regression model originally developed by Anderson (1951). In addition to these methods the generalized method of moments (GMM) estimation method, which was originally proposed by Hansen (1982) in econometrics and is essentially the same as the estimating equation method (EEM) by Godambe (1960), has been often used in the past two decades. (We use the term GMM for convenience hereafter.) Also the maximum empirical likelihood (MEL) method has gotten attention recently because it gives an asymptotically efficient estimator in the semi-parametric sense and improves the serious bias problem known in the GMM method when the number of instruments is large. See Owen (2001), Qin and Lawless (1994), Kitamura and Stutzer (1997), and Kitamura, Tripathi and Ahn (2004) on the MEL method, for instance. Since we have two semi-parametric estimation methods and they are asymptotically equivalent, it is important to compare the finite sample properties of these estimation methods. There has been a growing interest on the related topics in econometrics and some relevant literatures in recent years are Newey and Smith (2004), Mittelhammer, Judge and Schoenberg (2005), Anderson, Kunitomo and Matsushita (2005, 2007, 2008) and their references, for instance.

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

In order to compare 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 distribution and density functions directly. Also it is important to note that 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 that the LIML estimator, for instance, does not possess any moments of positive integer order under a set of reasonable assumptions while some of recent literatures in econometrics seem to ignore this problem. 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 or the simultaneous equation models. Because of the semi-parametric features of our analysis, we develop the conditional expansion approach which has new technical problems.

Our formulation and method are intentionally similar to the earlier studies on the single equation estimation methods 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 estimators in the classical parametric framework as well as in the semi-parametric framework. The main results of our paper are related to the studies of higher order asymptotic efficiency estimation by Pfanzagl and Wefelmeyer (1978), Akahira and Takeuchi (1981, 1990), Pfanzagl (1990), Bickel et al. (1993) in the statistical literature, and Takeuchi and Morimune (1985) and Newey and Smith (2004) in the econometric literature.

In Section 2 we define the structural equation model and its estimation methods. Then in Section 3 we give the asymptotic expansions of the distribution functions of estimators in a simple case which illustrate the merit of our approach. In Section 4, we give the results on the asymptotic expansions of the density functions of estimators under a set of assumptions on the disturbances and compare the higher order properties of alternative estimators in a more general case. Some discussion on the higher order properties of estimators and concluding remarks are given in Section 5. The derivations of the asymptotic expansions, the proofs of Lemmas and Theorems and useful formulas will be given in Appendices.

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

Let a linear structural equation be given by

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

where y_{1i} and $\mathbf{y}_{2i}$ are a scalar and a vector of G_1 endogenous variables, $\mathbf{z}_{1i}$ is a vector of K_1 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 an equation in a system of simultaneous 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 $\{\mathbf{z}_i\} = (\mathbf{z}_{1i}^{'}, \mathbf{z}_{2i}^{'})^{'}$ including $\{\mathbf{z}_{1i}\}$. The set of exogenous variables $\{\mathbf{z}_i\}$ are often called the instrumental variables and we have the orthogonal condition $\mathbf{E}(u_i \mathbf{z}_i) = \mathbf{0}$ ($i = 1, \cdots, n; n > K, n > 3$). Because we do not specify the equations except (2.1), we consider the limited information estimation methods based on the set of instrumental variables (or instruments).

The reduced form equations for $\mathbf{y}_i^{'} = (y_{1i}, \mathbf{y}_{2i}^{'})$ are

$$\mathbf{y}_i = \boldsymbol{\Pi} \mathbf{z}_i + \mathbf{v}_i \quad (i = 1, \cdots, n) \;, \tag{2.2}$$

where $\mathbf{v}_i^{'} = (v_{1i}, \mathbf{v}_{2i}^{'})$ is a $1 \times (1 + G_1)$ disturbance terms with $\mathbf{E}[\mathbf{v}_i] = \mathbf{0}$, $\boldsymbol{\Pi}^{'} = (\boldsymbol{\pi}_1, \boldsymbol{\Pi}_2)$ is a $K \times (1 + G_1)$ partitioned matrix of the reduced form coefficients and $\boldsymbol{\Pi}_2$ is a $K \times G_1$

matrix. By multiplying $(1, -\boldsymbol{\beta}^{'})$ to (2.2) from the left-hand side, $(1, -\boldsymbol{\beta}^{'})\mathbf{\Pi} = (\boldsymbol{\gamma}^{'}, \mathbf{0}^{'})$ and $u_i = v_{1i} - \boldsymbol{\beta}^{'}\mathbf{v}_{2i}$ $(i = 1, \cdots, n)$, that is, the rank of $\mathbf{\Pi}$ is reduced.

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

$$L_n^*(\boldsymbol{\lambda}, \boldsymbol{\theta}) = \sum_{i=1}^{n} \log p_i - \nu \left[\sum_{i=1}^{n} p_i - 1\right] - n\boldsymbol{\lambda}^{'} \sum_{i=1}^{n} p_i \mathbf{z}_i \left[y_{1i} - (\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\boldsymbol{\theta})\right] , \tag{2.3}$$

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

$$L_n(\boldsymbol{\lambda}, \boldsymbol{\theta}) = -\sum_{i=1}^{n} \log \left\{1 + \boldsymbol{\lambda}^{'}\mathbf{z}_i \left[y_{1i} - (\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\boldsymbol{\theta})\right]\right\} , \tag{2.4}$$

where $\hat{\nu} = n$ and $[n\hat{p}_i]^{-1} = 1 + \lambda^{'}\mathbf{z}_i[y_{1i} - (\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\boldsymbol{\theta})]$. By differentiating (2.4) with respect to λ and combining the resulting equation with the restriction $\sum_{i=1}^{n} p_i = 1$, we have $\sum_{i=1}^{n} \hat{p}_i \mathbf{z}_i \left[y_{1i} - (\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\hat{\boldsymbol{\theta}})\right] = \mathbf{0}$ and

$$\hat{\boldsymbol{\lambda}} = \left[\sum_{i=1}^{n} \hat{p}_i u_i^2(\hat{\theta})\mathbf{z}_i\mathbf{z}_i^{'}\right]^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} u_i(\hat{\theta})\mathbf{z}_i\right] , \tag{2.5}$$

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 $\boldsymbol{\theta}$. Then the MEL estimator of $\boldsymbol{\theta}$ is the solution of

$$\begin{aligned} & \left[\sum_{i=1}^{n} \hat{p}_i \begin{pmatrix} \mathbf{y}_{2i} \\ \mathbf{z}_{1i} \end{pmatrix} \mathbf{z}_i^{'}\right] \left[\sum_{i=1}^{n} \hat{p}_i u_i(\hat{\theta})^2 \mathbf{z}_i\mathbf{z}_i^{'}\right]^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i \; y_{1i}\right] \qquad (2.6) \\ = \; & \left[\sum_{i=1}^{n} \hat{p}_i \begin{pmatrix} \mathbf{y}_{2i} \\ \mathbf{z}_{1i} \end{pmatrix} \mathbf{z}_i^{'}\right] \left[\sum_{i=1}^{n} \hat{p}_i u_i(\hat{\theta})^2 \mathbf{z}_i\mathbf{z}_i^{'}\right]^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i \left(\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'}\right)\right] \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\gamma} \end{pmatrix} . \end{aligned}$$

If we substitute $1/n$ for $\hat{p}_i$ $(i = 1, \cdots, n)$ in (2.6) and use an (efficient) initial estimator $\tilde{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ satisfying $\tilde{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}} = o_p(1/\sqrt{n})$ to replace $u_i(\hat{\boldsymbol{\theta}})$ in (2.6), we have a representation of the (optimal) generalized method of moments (GMM) estimator for $\boldsymbol{\theta}^{'} = (\boldsymbol{\beta}^{'}, \boldsymbol{\gamma}^{'})$. In this paper we forcus on the convergent (many-step) GMM estimator, which is a limit of iteration of $\boldsymbol{\theta}$ and $u_i(\hat{\boldsymbol{\theta}})$ because it agrees with the original idea of the GMM estimation. Although the GMM estimator here could be different from some of two-step GMM estimators, it is certainly possible, with some complications, to extend our analysis to the GMM with any consistent initial estimator. (See Hayashi (2000) on the standard GMM approach in econometrics for instance.) By generalizing the weights p_i $(i = 1, \cdots, n)$ in (2.6), we introduce a class of estimators. Let

$$n\hat{p}_i^* = \left[1 + a\; \boldsymbol{\lambda}^{'}\mathbf{z}_i u_i(\hat{\theta})\right]^{-1} , \tag{2.7}$$

where a is a non-negative constant $(0 \leq a \leq 1)$ and $\hat{\theta}$ is the MEL estimator of $\boldsymbol{\theta}$. Then we define the modification of the MEL estimator (MMEL) by substituting $\hat{p}_i^*$ $(i = 1, \cdots, n)$ into (2.5)-(2.6).

If we assume the homoscedasticity of disturbances and replace $u_i^2(\hat{\boldsymbol{\theta}})$ by $\hat{\sigma}^2$ in (2.6), we can regard that the MEL estimator and the GMM estimator correspond to the

LIML estimator and the TSLS estimator, respectively. (See Section 2 of Anderson et al. (2008).) The latter methods were originally developed as the parametric estimation methods by Anderson and Rubin (1949).

In the rest of this paper, we shall consider the standardized estimator as

$$\hat{\mathbf{e}} = \sqrt{n} \begin{bmatrix} \hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \\ \hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma} \end{bmatrix} , \tag{2.8}$$

where $\hat{\boldsymbol{\theta}}^{'} = (\hat{\boldsymbol{\beta}}^{'}, \hat{\boldsymbol{\gamma}}^{'})$. We sometimes denote $\hat{\mathbf{e}}$ for the MEL estimator and its modification when it causes no confusion. Under a set of regularity conditions, the asymptotic covariance matrix of any asymptotically (semi-parametric) efficient estimator is

$$\mathbf{Q} = \left[\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\right]^{-1} , \tag{2.9}$$

where $\mathbf{M}_n$ and $\mathbf{C}_n$ $(n > K, n > 3)$ and their (constant) probability limits are defined by

$$\mathbf{M}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i\mathbf{z}_i^{'} \xrightarrow{p} \mathbf{M} , \quad \mathbf{C}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i\mathbf{z}_i^{'}u_i^2 \xrightarrow{p} \mathbf{C} , \tag{2.10}$$

and

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

We assume that $\mathbf{M}$ and $\mathbf{C}$ are positive definite and $\mathbf{rank}\,[\mathbf{D}] = p\ (= G_1 + K_1)$. These conditions assure that the limiting covariance matrix $\mathbf{Q}$ is non-degenerate. The rank condition implies that the order condition $L = K - p \geq 0$ holds, which is the degree of over-identification. When the disturbance terms are (conditionally or unconditionally) homoscedastic random variables, then $\mathbf{C} = \sigma^2\mathbf{M}$, $\mathbf{E}(u_i^2) = \sigma^2$ and $\mathbf{Q} = \sigma^2[\mathbf{D}^{'}\mathbf{M}\mathbf{D}]^{-1}$.

In order to compare alternative efficient estimation methods in the finite sample sense, we shall derive the asymptotic expansions of the density functions of the standardized estimators (2.8) in the form of

$$f(\boldsymbol{\xi}) = \phi_{\mathbf{Q}}(\boldsymbol{\xi}) \left[1 + \frac{1}{\sqrt{n}}H_1(\boldsymbol{\xi}) + \frac{1}{n}H_2(\boldsymbol{\xi}) \right] + o(n^{-1}) , \tag{2.11}$$

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 covariance matrix $\mathbf{Q}$, and $H_i(\boldsymbol{\xi})$ $(i = 1, 2)$ are some polynomial functions of elements of $\boldsymbol{\xi}$. Then we shall use the mean operator $\mathbf{AM}_n(\hat{\mathbf{e}})$, which is defined by 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 (2.11). We write the asymptotic bias and the asymptotic MSE by $\mathbf{ABIAS}_n(\hat{\mathbf{e}}) = \mathbf{AM}_n(\hat{\mathbf{e}})$ and $\mathbf{AMSE}_n(\hat{\mathbf{e}}) = \mathbf{AM}_n(\hat{\mathbf{e}}\hat{\mathbf{e}}^{'})$. These quantities are useful because the asymptotic expansion of the distribution of estimators are quite complicated in the general case.

It should be noted, however, that they are not necessarily the same as the asymptotic expansions of the exact moments and some care should be taken. One important case is that the LIML estimator and its related statistics do not have any positive integer moments in our setting. This does not mean that the LIML estimator should be ruled out, but that we should use other criteria different from the exact bias, the exact MSE, and their analogues in Monte Carlo experiments. An illustrative example is the estimation problem of reciprocal of (non-zero) normal mean. Hence the results of previous Monte Carlo experiments without this consideration may have drawbacks and careful interpretation should be needed.

3 Asymptotic Expansions of Distributions of Estimators and Their Approximations in a Simple Case

The exact density functions of alternative estimators and their asymptotic expansions are quite complicated in the general case. For an illustration we present the asymptotic expansions of the distribution functions of estimators in the simple case when $G_1 = 1$ and the homoscedastic disturbances $\{u_i\}$ are normally distributed. In this case we partition a $[1 + K_1] \times [1 + K_1]$ matrix as

$$\mathbf{Q} = \begin{pmatrix} Q_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix} = \sigma^2 \left[\begin{pmatrix} \mathbf{\Pi}_2^{'} \\ (\mathbf{I}_{K_1}, \mathbf{O}) \end{pmatrix} \mathbf{M} \left(\mathbf{\Pi}_2, \begin{pmatrix} \mathbf{I}_{K_1} \\ \mathbf{O} \end{pmatrix} \right) \right]^{-1} .$$

Then the upper-left corner of $\mathbf{Q}$ is given by $Q_{11} = \sigma^2 (\mathbf{\Pi}_{22}^{'} \mathbf{M}_{22.1} \mathbf{\Pi}_{22})^{-1}$, where $\mathbf{\Pi}_{22}$ is a $K_2 \times 1$ vector of the lower corner of $\mathbf{\Pi}_2$ (a $K \times 1$ vector). We take the coefficient of an endogenous variable β in the right-hand side of (2.1) and consider

$$\mathbf{P} \left(\frac{\sqrt{n \mathbf{\Pi}_{22}^{'} \mathbf{M}_{22.1} \mathbf{\Pi}_{22}}}{\sigma} (\hat{\beta} - \beta) \leq x \right) , \tag{3.1}$$

provided that $\mathbf{M}_{22.1} = \operatorname{plim} n^{-1} [\sum_{i=1}^n \mathbf{z}_{2i} \mathbf{z}_{2i}^{'} - \sum_{i=1}^n \mathbf{z}_{2i} \mathbf{z}_{1i}^{'} (\sum_{i=1}^n \mathbf{z}_{1i} \mathbf{z}_{1i}^{'})^{-1} \sum_{i=1}^n \mathbf{z}_{1i} \mathbf{z}_{2i}^{'}]$ is a positive definite matrix and $Q_{11} > 0$.
From (2.8) and (2.9) in the standard large sample theory, the limiting distribution of (3.1) is the standard normal. In this form it is relatively easy to make comparison of alternative estimators and some useful information can be drawn.

When $G_1 = 1$, we can obtain simple formulas of the asymptotic expansion of the distribution function of estimators if we use *the key parameters* and the notations of Anderson, Kunitomo and Sawa (1982). From this reason, we define the 2×2 covariance matrix $\mathbf{\Omega} (= (\omega_{ij})) = \mathbf{E}[\mathbf{v}_i \mathbf{v}_i^{'}]$, the standardized coefficient (the degree of endogeneity) $\alpha = [\omega_{22} / |\mathbf{\Omega}|^{1/2}] [\beta - \omega_{12}/\omega_{22}]$ and the noncentrality (or concentration) parameter $\mu^2 = [(1 + \alpha^2)/\omega_{22}] \mathbf{\Pi}_{22}^{'} \mathbf{A}_{22.1} \mathbf{\Pi}_{22}$, where $\mathbf{A}_{22.1} = \sum_{i=1}^n \mathbf{z}_{2i} \mathbf{z}_{2i}^{'} - \sum_{i=1}^n \mathbf{z}_{2i} \mathbf{z}_{1i}^{'} (\sum_{i=1}^n \mathbf{z}_{1i} \mathbf{z}_{1i}^{'})^{-1} \sum_{i=1}^n \mathbf{z}_{1i} \mathbf{z}_{2i}^{'}]$ corresponds to $n \mathbf{M}_{22.1}$. Define an additional (semi-parametric) factor by

$$\tau = 2\sigma^2 \frac{(1 + \alpha^2)}{\omega_{22}} Q_{11}^{-1} \left[\mathbf{Q} \mathbf{D}^{'} \mathbf{F} \mathbf{D} \mathbf{Q} \right]_{11} Q_{11}^{-1} , \tag{3.2}$$

where $[\,\cdot\,]_{11}$ is the (1,1) element of matrix,

$$\mathbf{F} = \operatorname*{plim}_{n \to \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{z}_i (\mathbf{z}_i^{'} \mathbf{A} \mathbf{z}_i) \mathbf{z}_i^{'} , \tag{3.3}$$

and $\mathbf{A} = \mathbf{C}^{-1} - \mathbf{C}^{-1} \mathbf{M} \mathbf{D} \mathbf{Q} \mathbf{D}^{'} \mathbf{M} \mathbf{C}^{-1}$. In the large sample theory we assume that the noncentrality parameter μ^2 is proportional to the sample size n (see the conditions of (2.9), (2.10) and (3.3). However, alternative asymptotic theories can be developed. (See Anderson et al. (2005, 2007, 2008), for instance.) To be precise we first make a set of simple conditions.

Assumption I : (i) Suppose that $G_1 = 1$ and the sequences $\{\mathbf{v}_i\}$ $(i = 1, \cdots, n)$ (hence $\{u_i\}$) are independently and normally distributed with $\mathbf{E}[\mathbf{v}_i] = \mathbf{0}$, $\mathbf{E}[\mathbf{v}_i \mathbf{v}_i^{'}] = \mathbf{\Omega}$ (> 0)

and $\mathbf{E}[u_i^2] = \sigma^2$. (ii) The instrumental variables $\mathbf{z}_i$ are non-stochastic, the limits of (2.10) and (3.3) exist and there exists a (positive) constant c such that $\frac{\mu^2}{n} = c + o(n^{-1/2})$.

By using the asymptotic expansion of the density function of the MEL estimator in *Theorem 4.2* in Section 4 and setting $a = 1$, we obtain the result for the normalized form of distribution function when $G_1 = 1$ and the disturbances are homoscedastic and normally distributed. The derivation will be given in *Appendix B*.

Theorem 3.1 : Under *Assumption I*, an asymptotic expansion of the distribution function of the normalized MEL estimator as $\mu^2 \to \infty$ (and $n \to \infty$) is

$$\mathbf{P}\left(\frac{\sqrt{n\mathbf{\Pi}_{22}'\mathbf{M}_{22.1}\mathbf{\Pi}_{22}}}{\sigma}(\hat{\beta}_{MEL} - \beta) \leq x\right) \tag{3.4}$$

$$= \Phi(x) + \left\{-\frac{\alpha}{\mu}x^2 - \frac{1}{2\mu^2}\left[(\tau + L)x + (1 - 2\alpha^2)x^3 + \alpha^2 x^5\right]\right\}\phi(x) + o(\mu^{-2}),$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the cdf and the density function of the standard normal distribution, respectively.

Also by setting $a = 0$ for the GMM estimator, we have an asymptotic expansion of its distribution function.

Theorem 3.2 : Under *Assumption I*, an asymptotic expansion of the distribution function of the normalized GMM estimator as $\mu^2 \to \infty$ (and $n \to \infty$) is

$$\mathbf{P}\left(\frac{\sqrt{n\mathbf{\Pi}_{22}'\mathbf{M}_{22.1}\mathbf{\Pi}_{22}}}{\sigma}(\hat{\beta}_{GMM} - \beta) \leq x\right) \tag{3.5}$$

$$= \Phi(x) + \left\{-\frac{\alpha}{\mu}\left[x^2 - L\right] - \frac{1}{2\mu^2}\left[(\tau + L^2\alpha^2 - L)x + (1 - 2(L+1)\alpha^2)x^3 + \alpha^2 x^5\right]\right\}\phi(x) + o(\mu^{-2}),$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are defined as *Theorem 3.1*.

There is an interesting observation that if we set $\tau = 0$ in the above expressions, the resulting formulas in (3.4) and (3.5) 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. Hence τ could be interpreted as the semi-parametric (3rd order) *inefficiency* factor under the homoscedasticity assumption of disturbances. (See Section 4 and *Appendix A* for the detail.)

A Numerical Illustration

For an illustration on the use of the asymptotic expansion formulas, we give some figures and tables as Figures 1-3 and Tables 1-2 in *Appendix E* as typical cases. We computed the distribution functions of the MEL estimator and the GMM estimator of the coefficient β in the normalized terms (3.1) based on large number of simulations. When

$G_1 = 1$, we can easily generate the normalized probability (3.1), which depends on the key parameters and other factors as discussed by Anderson et al. (1982, 2005). We first generate the vectors of the normal disturbance terms and the exogenous variables $(\mathbf{v}_i, \mathbf{z}_i)$ $(i = 1, \cdots, n)$ and then generate the endogenous variables by utilizing (2.1) and (2.2). Then we can simulate the probability of (3.1) by iterating the calculations of (2.5) and (2.6) until we have stable convergences numerically. We denote the resulting values as *Exact* in Tables since they are very accurate in two decimal digits at least. (The number of replications in all simulations are basically 5,000 and we have confirmed their accuracy by comparing the exact distributions of the TSLS and LIML estimators. Our method of evaluating the distribution functions of estimators in numerical analysis is essentially the same as Anderson, Kunitomo and Matsushita (2005, 2008) which explain the details of our evaluation procedure and the accuracy of our computations.) In tables we have given the 5% and 95% percentiles, Lower (L.QT), Median (MEDN) and Upper (U.QT) quantiles, and the interquantile range (IQR). Also we have given the approximations based on the asymptotic expansions of the distribution functions of estimators in the forms of (3.4) and (3.5), which are denoted as *Approx* in tables and figures. *Difference* is defined by *Approx* minus *Exact* except the rounding errors. We did a large number of numerical calculations, but we have chosen only a small number of results.

First, we find that in most cases the approximations based on the asymptotic expansions of the distribution functions given by (3.4) and (3.5) are quite accurate in its middle range areas. There can be some discrepancy in the tail quantiles when K_2 is relatively large in particular. As we have expected from our discussions on the exact moments of estimators, we have confirmed that the exact bias and the exact MSE of the LIML estimator calculated from the simulations are sometimes not stable. Second, the distribution functions of the MEL and the LIML estimators are very similar while the distribution functions of the GMM and the TSLS estimators are also very similar. This finding is quite consistent with the asymptotic expansions of the distribution functions in (3.4) and (3.5) under the homoscedasticity and normality of disturbances. Thus we could interpret that the MEL estimator is a semi-parametric extension of the LIML estimator while the GMM estimator is a semi-parametric extension of the TSLS estimator.

However, we find that the distribution functions of the MEL and GMM estimators have some differences. As an illustration on this issue we show one typical case with $K_2 = 10$ (Figures 2 and 3 in *Appendix E*) which have been taken from Anderson et al. (2005, 2008). The most important finding is that the distribution function of the MEL estimator is almost median unbiased while the distribution function of the GMM estimator is biased significantly. It makes some doubts on the standard use of the GMM estimation when K_2 is nor very small. This issue has been investigated by Anderson et al. (2007) in more details.

A Heteroscedastic Case

When the disturbances are not conditionally homoscedastic, the above results still hold essentially. For instance, *Theorem 4.1* of Section 4 implies that for the MMEL estimator with arbitrary a $(0 \leq a \leq 1)$,

$$\mathbf{P}\left(\sqrt{n}(\hat{\beta} - \beta) \leq y \right) \tag{3.6}$$

$$
\begin{aligned}
= \quad & \Phi_{Q_{11}}(y) + \Bigg\{ (1-a) \left[\operatorname*{plim}_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathrm{tr}(\mathbf{z}_i \mathbf{z}_i^{'} \sigma_i^2 \mathbf{A}) [\frac{|\mathbf{\Omega}|^{1/2}}{\sigma^2}]_i \alpha_i \right] \\
& - \left[\operatorname*{plim}_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} [\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{z}_i\mathbf{z}_i^{'}\sigma_i^2\mathbf{C}^{-1}\mathbf{M}\mathbf{D}]_{11} [\frac{|\mathbf{\Omega}|^{1/2}}{\sigma^2}]_i \alpha_i Q_{11}^{-1} y^2 \right] \Bigg\} \frac{Q_{11}}{\sqrt{n}} \phi_{Q_{11}}(y) \\
& + o(n^{-1/2})
\end{aligned}
$$

provided that there exist limits in the right-hand side of (3.6), where $\mathrm{tr}(\cdot)$ is the trace of a matrix, $[\cdot]_{11}$ is the (1,1)-element of a matrix, $\sigma_i^2 = \mathbf{E}(u_i^2|\mathbf{z}_i)$, $[|\mathbf{\Omega}|^{1/2}/\sigma^2]_i$ and α_i are defined (as $[|\mathbf{\Omega}|^{1/2}/\sigma^2]$ and α) for $\mathbf{E}(\mathbf{v}_i\mathbf{v}_i^{'}|\mathbf{z}_i) = \mathbf{\Omega}_i = (\omega_{jk}^{(i)})$ $(i = 1, \cdots, n)$, and $\Phi_{Q_{11}}(\cdot)$ and $\phi_{Q_{11}}(\cdot)$ are the cdf and the density function of $N(0, Q_{11})$, respectively.

When σ_i^2, $[|\mathbf{\Omega}|^{1/2}/\sigma^2]_i$ and α_i are independent of i, (3.6) with $a = 1$ and $a = 0$ are the same as (3.4) and (3.5) up to $O(n^{-1/2})$, respectively, and $\mathrm{tr}(\mathbf{CA}) = L(= K - p)$. As we shall see further terms of $O(n^{-1})$ become substantially complicated.

To summarize our findings in this section, the results of asymptotic expansions of distributions give useful information on the finite sample properties of alternative estimators beyond their biases and MSEs when $G_1 = 1$ and the disturbances are normally distributed. In Section 4 we shall show that these observations on the finite sample properties are generally true even when $G_1 \geq 1$ and the distribution of disturbances are not necessarily normal in a more general setting.

4. Asymptotic Expansions of Densities and Higher Order Properties of Alternative Estimators

4.1 The method of Asymptotic Expansions and Assumptions

In order to derive the asymptotic expansions of the densities of estimators when the disturbances are not necessarily normally distributed, we need regularity conditions.

Assumption II : (i) The sequence $(\mathbf{z}_i^{'}, \mathbf{v}_i^{'}), i = 1, \cdots, n$, are mutually independent random vectors and $\mathbf{v}_i$ have the strictly positive density with respect to Lebesgue measure; $\mathbf{E}(\mathbf{v}_i|\mathbf{z}_i) = \mathbf{0}$, $\mathbf{E}(\mathbf{v}_i\mathbf{v}_i^{'}|\mathbf{z}_i) = \mathbf{\Omega}_i$ $(a.s.)$, $\mathbf{E}(u_i^2|\mathbf{z}_i) = \sigma_i^2$, $\mathbf{E}(u_i^3|\mathbf{z}_i) = \kappa_{3i}$, $\mathbf{E}(u_i^4|\mathbf{z}_i) = \kappa_{4i}$ and $\mathbf{E}[\|\mathbf{v}_i\|^6] < \infty$. (ii) The (constant) matrices $\mathbf{M}$ and $\mathbf{C}$ are positive definite, $\mathrm{rank}(\mathbf{D}) = p$, $n^{-1}\sum_{i=1}^{n} \mathbf{z}_i\mathbf{z}_i^{'}\sigma_i^2 = \mathbf{C} + o_p(n^{-1/2})$ and $n^{-1}\sum_{i=1}^{n} \mathbf{z}_i\mathbf{z}_i^{'} = \mathbf{M} + o_p(n^{-1/2})$. (iii) The sequence of vectors $\mathbf{z}_i = (z_{ij})$ $(i = 1, \cdots, n; j = 1, \cdots, K)$ are bounded or $n^{-1}\max_{1\leq i\leq n} \|\mathbf{z}_i\|^2 \xrightarrow{p} 0$ and $\mathbf{E}[\|\mathbf{z}_i\|^6] < \infty$ when they are stochastic. There exist finite $M_3(j_1, j_2, j_3)$ such that $n^{-1}\sum_{i=1}^{n} \kappa_{3i} z_{ij_1} z_{ij_2} z_{ij_3} = M_3(j_1, j_2, j_3) + o_p(n^{-1/2})$.

We need some moment conditions on disturbance terms to derive higher order stochastic expansions of the associated random variables up to $O(n^{-1})$. Conditions (ii) and (iii) of *Assumption II* could be weakened, but then the resulting formulas and their derivations become more complicated than those reported 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 $\sum_{i=1}^{n} \mathbf{z}_i u_i$. In order to avoid cumbersome arguments, however, we mostly treat $\{\mathbf{z}_i\}$ as if they were deterministic.

In our analysis we first use the consistency of the MEL estimator (Owen (1990) and Qin and Lawless (1994)). Since $n\hat{p}_i \xrightarrow{p} 1$, $\hat{\boldsymbol{\theta}}_{EL} \xrightarrow{p} \boldsymbol{\theta}_0$, ($\boldsymbol{\theta}_0$ is the true value of $\boldsymbol{\theta}$) and

$\sqrt{n}\hat{\lambda}$ converges to a random vector as $n \to \infty$, we represent $\hat{\mathbf{e}}$ as

$$
\begin{aligned}
&\left[\sum_{i=1}^{n} \hat{p}_i \left(\begin{array}{c} \mathbf{y}_{2i} \\ \mathbf{z}_{1i} \end{array} \right) \mathbf{z}_i^{'} \right] \left[\sum_{i=1}^{n} \hat{p}_i u_i(\hat{\theta})^2 \mathbf{z}_i \mathbf{z}_i^{'} \right]^{-1} \left[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{z}_i u_i \right] \qquad (4.1)\\
= &\left[\sum_{i=1}^{n} \hat{p}_i \left(\begin{array}{c} \mathbf{y}_{2i} \\ \mathbf{z}_{1i} \end{array} \right) \mathbf{z}_i^{'} \right] \left[\sum_{i=1}^{n} \hat{p}_i u_i(\hat{\theta})^2 \mathbf{z}_i \mathbf{z}_i^{'} \right]^{-1} \left[\frac{1}{n} \sum_{i=1}^{n} \mathbf{z}_i \left(\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'} \right) \right] \hat{\mathbf{e}} \ ,
\end{aligned}
$$

where $\hat{\boldsymbol{\theta}}$ for $\hat{\boldsymbol{\theta}}_{MMEL}$. As $n \to \infty$, we write the first order term of $\hat{\mathbf{e}}$ as $\tilde{\mathbf{e}}_0$, which is

$$
\tilde{\mathbf{e}}_0 = \left[\mathbf{D}^{'} \mathbf{M} \mathbf{C}^{-1} \mathbf{M} \mathbf{D} \right]^{-1} \mathbf{D}^{'} \mathbf{M} \mathbf{C}^{-1} \left[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{z}_i u_i \right] \ . \qquad (4.2)
$$

In the following derivation it is convenient to use the fact that $\mathbf{Q}_n = \left[\mathbf{D}^{'} \mathbf{M}_n \mathbf{C}_n^{-1} \mathbf{M}_n \mathbf{D} \right]^{-1}$, $\tilde{\mathbf{e}}_0 - \mathbf{e}_0 = o_p(1)$ and

$$
\mathbf{e}_0 = \left[\mathbf{D}^{'} \mathbf{M}_n \mathbf{C}_n^{-1} \mathbf{M}_n \mathbf{D} \right]^{-1} \mathbf{D}^{'} \mathbf{M}_n \mathbf{C}_n^{-1} \left[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{z}_i u_i \right] \ . \qquad (4.3)
$$

By applying a central limit theorem (CLT) to the last term of (4.2), we have a weak convergence $\mathbf{X}_n = n^{-1/2} \sum_{i=1}^{n} \mathbf{z}_i u_i \xrightarrow{w} N_p(\mathbf{0}, \mathbf{C})$. Then $\tilde{\mathbf{e}}_0 \xrightarrow{w} N_p(\mathbf{0}, \mathbf{Q})$, where $\mathbf{Q}$ is given by (2.9) and $\xrightarrow{w}$ means the weak convergence as $n \to \infty$. By using

$$
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{z}_i u_i(\hat{\theta}) = \mathbf{X}_n + \frac{1}{n} \left[-\sum_{i=1}^{n} \mathbf{z}_i (\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'}) \hat{\mathbf{e}} \right] = \mathbf{X}_n - \mathbf{M}_n \mathbf{D} \hat{\mathbf{e}} + O_p(n^{-1/2}) \ , \qquad (4.4)
$$

we find that $\sqrt{n}\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0 \xrightarrow{p} 0$ ($\hat{\boldsymbol{\lambda}}$ is $\boldsymbol{\lambda}$ with $\hat{\boldsymbol{\theta}}$) and

$$
\boldsymbol{\lambda}_0 = \mathbf{C}_n^{-1/2} \left[\mathbf{I}_K - \mathbf{C}_n^{-1/2} \mathbf{M}_n \mathbf{D} \mathbf{Q}_n \mathbf{D}^{'} \mathbf{M}_n \mathbf{C}_n^{-1/2} \right] \left[\mathbf{C}_n^{-1/2} \mathbf{X}_n \right] \ . \qquad (4.5)
$$

Because the limiting distribution of $\mathbf{B}_n = \mathbf{C}^{-1/2} \mathbf{X}_n$ is $N_K(\mathbf{0}, \mathbf{I}_K)$, $\mathbf{C}_n^{1/2} \sqrt{n} \hat{\boldsymbol{\lambda}} \xrightarrow{w} N_K(\mathbf{0}, \bar{\mathbf{P}}_{D^*})$ and $\bar{\mathbf{P}}_{D^*} = \mathbf{I}_K - \mathbf{D}^* (\mathbf{D}^{*'} \mathbf{D}^*)^{-1} \mathbf{D}^{*'}$ is constructed by a $K \times p$ matrix $\mathbf{D}^* = \mathbf{C}^{-1/2} \mathbf{M} \mathbf{D}$, where $\mathbf{D}_n^* = \mathbf{C}_n^{-1/2} \mathbf{M}_n \mathbf{D} \xrightarrow{p} \mathbf{D}^*$ as $n \longrightarrow +\infty$. Then the covariance matrix of the limiting distribution $\boldsymbol{\lambda}_0$ is given by $\mathbf{A} = \mathbf{C}^{-1} - \mathbf{C}^{-1} \mathbf{M} \mathbf{D} \mathbf{Q} \mathbf{D}^{'} \mathbf{M} \mathbf{C}^{-1}$, which plays important roles in our analysis.

We shall derive the asymptotic expansions of the density functions of estimators. Our method is the conditional expansion approach which is similar to the one in Fujikoshi et al. (1982) and Anderson et al. (1986). Because the early works could utilize aspects of the multivariate normal distributions directly which we cannot use, the derivations of asymptotic expansions become more complicated as explained in *Appendix A*. In our conditional expansion approach, first we expand $\hat{\mathbf{e}}$ by the perturbation method in each components of $\mathbf{X}_n = (X_j^{(n)})$, $\mathbf{Y}_n = (y_{jk}^{(n)})$, $\mathbf{Z}_n = (z_{jk}^{(n)})$ and $\mathbf{U}_n = (U_{jk}^{(n)})$, which are defined by

$$
\mathbf{Y}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\mathbf{z}_i \mathbf{z}_i^{'} u_i^2 - \mathbf{E}(\mathbf{z}_i \mathbf{z}_i^{'} u_i^2) \right] \ , \mathbf{Z}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{z}_i (\mathbf{v}_{2i}^{'}, \mathbf{0}^{'}), \mathbf{U}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{w}_i \mathbf{z}_i^{'} \ , \qquad (4.6)
$$

where $\mathbf{w}_i = (\mathbf{v}_{2i}^{'}, \mathbf{0}^{'})^{'} - \mathbf{q}_i u_i$ and $\mathbf{q}_i^{'} = (1/\sigma_i^2)\mathbf{E}[(\mathbf{v}_{2i}^{'}, \mathbf{0}^{'})u_i|\mathbf{z}_i]$ for $\mathbf{v}_i^{'} = (v_{1i}, \mathbf{v}_{2i}^{'})$ $(i = 1, \cdots, n)$. Then if $\mathbf{E}[\|\mathbf{v}_i\|^s] < \infty$ for $s \geq 3$, we can take a positive (bounded) constant $c_n(1, s)$ depending on n which satisfies

$$\mathbf{P}(\|\mathbf{X}_n\| > [\Lambda_n \log n]^{1/2}) \leq c_n(1, s)\frac{(1/\sqrt{n})^{s-2}}{(\log n)^{s/2}} \ , \tag{4.7}$$

where Λ_n as the maximum of the characteristic roots of $\mathbf{E}(\mathbf{C}_n)$. Also for $\mathbf{Y}_n, \mathbf{Z}_n$ and $\mathbf{U}_n$ we can also take positive (bounded) constants $c_n(i, s)$ $(i = 2, 3, 4)$ and similar inequalities for $s \geq 3$ under *Assumption II.* The basic arguments on the validity have been given by Bhattacharya and Ghosh (1978) (see Bhattacharya and Rao (1976) also) for the i.i.d. random vector sequences. They can be extended to our case while the derivations and resulting explanations become quite lengthy.

We shall derive the stochastic expansions of the estimators up to $O_p(n^{-1/2})$ under *Assumption II* and write $\mathbf{e} = \mathbf{e}_0 + n^{-1/2}\mathbf{e}_1 + o_p(n^{-1/2})$ (see *Theorem 4.1* in the next subsection). The resulting expressions of $O_p(n^{-1})$, however, become complicated in the expression as $\mathbf{e} = \mathbf{e}_0 + n^{-1/2}\mathbf{e}_1 + n^{-1}\mathbf{e}_2 + o_p(n^{-1})$. It is partly because the conditional expectations of some random variables of $O_p(1)$ with $\mathbf{e}_0$ and $O_p(n^{-1/2})$ with $\mathbf{e}_1$ lead to the terms of $O_p(n^{-1/2})$ as well as some further terms of $O_p(n^{-1})$. (See [A5] of *Appendix A.*) When we ignore the effects of the third order moments of the disturbances and they are homoscedastic, the asymptotic expansions of estimators with an arbitrary a $(0 \leq a \leq 1)$ can be simplified greatly. For *Theorem 4.2* in the next subsection we impose further conditions.

Assumption III : (i) The sequence of $(\mathbf{v}_i^{'}, \mathbf{z}_i^{'}), i = 1, \cdots, n,$ satisfy Condition (i) of *Assumption II*; $\mathbf{E}[\|\mathbf{v}_i\|^8] < \infty$, $\mathcal{E}(\mathbf{v}_i) = \mathbf{0}$, $\mathbf{E}(\mathbf{v}_i\mathbf{v}_i^{'}) = \mathbf{\Omega}$, $\mathbf{E}(u_i^2) = \sigma^2$, $\mathbf{E}(u_i^4) = \kappa_4$, $\mathbf{C}_2^* = \mathbf{E}(\mathbf{w}_i\mathbf{w}_i^{'})$, $\mathbf{q}_i = \mathbf{q}$ and $\kappa = \mathbf{E}(u_i^4)/\sigma^4 - 3$. (ii) Conditions (ii) and (iii) of Assumption II with $n^{-1}\sum_{i=1}^n \mathbf{z}_i\mathbf{z}_i^{'} = \mathbf{M} + o_p(n^{-1})$ and $\mathbf{E}[\|\mathbf{z}_i\|^8] < \infty$ when $\mathbf{z}_i$ are stochastic. (iii) $\mathbf{E}[u_i^3] = \kappa_3 = 0$ and $\mathbf{E}[u_i^2\mathbf{w}_i] = \mathbf{0}$ $(i = 1, \cdots, n)$.

It is immediate that Condition (ii) can be relaxed as $(1/n)\sum_{i=1}^n z_i^{(j)} z_i^{(k)} z_i^{(l)} u_i^3 = o_p(n^{-1/2})$ and the similar conditions on the third order moments on $\{u_i^2\mathbf{w}_i\}$ in *Assumption III.*

4.2 Asymptotic Expansions of Density Functions

Although there are many terms appeared in the stochastic expansion of $\hat{\mathbf{e}}$ in *Appendix A*, it is possible to obtain the explicit forms of the asymptotic expansions of the density functions of semi-parametric estimators. In order to derive the asymptotic expansions of their density functions, we consider a stochastic expansion $\hat{\mathbf{e}} = \mathbf{e}_0 + n^{-1/2}\mathbf{e}_1 + n^{-1}\mathbf{e}_2 + o_p(n^{-1})$ with $\mathbf{e}_0$ as the leading term. Because we use $\mathbf{e}_0^* = \tilde{\mathbf{e}}_0$ as the leading term, we rewrite $\mathbf{e}_0 = \tilde{\mathbf{e}}_0 + n^{-1/2}\mathbf{e}_0^{(1)} + n^{-1}\mathbf{e}_0^{(2)} + o_p(n^{-1})$. We apply the same arguments to $\mathbf{e}_1$ and $\mathbf{e}_2$ recursively. From the terms of the order $O_p(n^{-1/2})$, we define $\mathbf{e}_1^*(\mathbf{x})$ as the sum of constant order terms of the conditional expectation $\mathbf{E}[\mathbf{e}_1^{(0)} + \mathbf{e}_0^{(1)}|\tilde{\mathbf{e}}_0 = \mathbf{x}]$, where the explicit forms of $\mathbf{e}_1^{(0)}$ and $\mathbf{e}_0^{(1)}$ are given in *Appendix A*. From the terms of the order $O_p(n^{-1})$, we define $\mathbf{e}_2^*(\mathbf{x})$ as the sum of $O_p(n^{-1/2})$ terms of the conditional expectation $\mathbf{E}[\mathbf{e}_1^{(0)} + \mathbf{e}_0^{(1)}|\tilde{\mathbf{e}}_0 = \mathbf{x}]$ plus the conditional expectation $\mathbf{E}[\mathbf{e}_0^{(2)} + \mathbf{e}_1^{(1)} + \mathbf{e}_2|\tilde{\mathbf{e}}_0 = \mathbf{x}]$. As the cross-product terms, we define $\mathbf{e}_{11}^*(\mathbf{x})$ as the sum of the conditional expectation $\mathbf{E}[(\mathbf{e}_1^{(0)} + \mathbf{e}_0^{(1)})(\mathbf{e}_1^{(0)} + \mathbf{e}_0^{(1)})^{'}|\tilde{\mathbf{e}}_0 = \mathbf{x}]$, where the explicit expressions of $\mathbf{e}_0^{(2)}, \mathbf{e}_1^{(1)}$ and $\mathbf{e}_2$ are also given in *Appendix A*.

Then we consider the characteristic function of the standardized estimator $\hat{\mathbf{e}}$ in order to derive the asymptotic expansion of its distribution function and we calculate

$$
\begin{aligned}
C(t) &= \mathbf{E}[\exp(i\mathbf{t}^{'}\mathbf{x})] + \frac{1}{\sqrt{n}}\mathbf{E}[i\mathbf{t}^{'}\mathbf{e}_1^*(\mathbf{x})\exp(i\mathbf{t}^{'}\mathbf{x})] \qquad (4.8)\\
&\quad + \frac{1}{2n}\mathbf{E}\{2i\mathbf{t}^{'}\mathbf{e}_2^*(\mathbf{x})\exp(i\mathbf{t}^{'}\mathbf{x}) + i^2\mathbf{t}^{'}\mathbf{e}_{11}^*(\mathbf{x})\mathbf{t}\exp(i\mathbf{t}^{'}\mathbf{x})\} + o(n^{-1}) \ ,
\end{aligned}
$$

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 using the Fourier inversion formulas in *Appendix D*, we invert the characteristic function (4.8). Although the intermediate computations are quite tedious but they are straightforward. First we consider the asymptotic expansion of the density function of $\tilde{\mathbf{e}}_0$ and its limiting distribution is normal as $n \to +\infty$. By expanding its characteristic function $\mathbf{E}[exp(i\mathbf{t}^{'}\tilde{\mathbf{e}}_0)]$ and inverting it under *Assumption II*, we have

$$
\begin{aligned}
\phi_{\mathbf{Q}}^*(\boldsymbol{\xi}) &= \phi_{\mathbf{Q}}(\xi)\Big\{1 + \frac{1}{6\sqrt{n}}\sum_{l_1,l_2,l_3=1}^{p}\beta_{l_1l_2l_3}h_3(\xi_{l_1},\xi_{l_2},\xi_{l_3}) \qquad (4.9)\\
&+ \frac{1}{24n}\left[\sum_{l_1,l_2,l_3,l_4=1}^{p}\beta_{l_1l_2l_3l_4}h_4(\xi_{l_1},\xi_{l_2},\xi_{l_3},\xi_{l_4}) - 3\sum_{l_1,l_2,m_1,m_2=1}^{p}\beta_{l_1l_2}\beta_{m_1m_2}h_2(\xi_{l_1},\xi_{l_2})h_2(\xi_{m_1},\xi_{m_2})\right]\\
&+ \frac{1}{72n}\sum_{l_1,l_2,l_3,m_1,m_2,m_3=1}^{p}\beta_{l_1l_2l_3}\beta_{m_1m_2m_3}h_6(\xi_{l_1},\xi_{l_2},\xi_{l_3},\xi_{m_1},\xi_{m_2},\xi_{m_3})\Big\} + o(n^{-1}),
\end{aligned}
$$

where $\phi_{\mathbf{Q}}(\boldsymbol{\xi})$ is the p-dimensional normal density function with means $\mathbf{0}$ and the covariance matrix $\mathbf{Q}$. The coefficients in (4.9) are given by
$\beta_{l_1l_2} = \mathrm{plim}_{n\to\infty}(1/n)\sum_{i=1}^{n}\sigma_i^2 z_{il_1}^* z_{il_2}^*$, $\beta_{l_1l_2l_3} = \mathrm{plim}_{n\to\infty}(1/n)\sum_{i=1}^{n}\kappa_{3i}z_{il_1}^* z_{il_2}^* z_{il_3}^*$,
and $\beta_{l_1l_2l_3l_4} = \mathrm{plim}_{n\to\infty}(1/n)\sum_{i=1}^{n}\kappa_{4i}z_{il_1}^* z_{il_2}^* z_{il_3}^* z_{il_4}^*$,
where $\mathbf{z}_i^* = (z_{il}^*) = \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{z}_i$ $(i = 1, \cdots, n)$, and $\sum_{l_1,l_2,l_3,l_4}$ means the combinations of two pairs such as (l_1, l_2) and (l_3, l_4) (i.e., it is 3 when $l_1 = l_2 = l_3 = l_4$, for instance). We define $h_k(x_{l_1,\cdots,l_k})$ $(k = 2, \cdots, 6)$ by $h_k(x_{l_1,\cdots,l_k})\phi_{\mathbf{Q}}(\mathbf{x}) = (-1)^k\frac{\partial^k\phi_{\mathbf{Q}}(x)}{\partial x_{l_1}\cdots\partial x_{l_k}}$.
It is important to find that (4.9) is common for all asymptotically efficient estimators and then it does not make any effects on the comparisons of (asymptotically) efficient estimators.

Next by using the results of *Appendix A*, the conditional expectations of the second order terms ((A.10) and (A.26)) are summarized as

$$
\begin{aligned}
\mathbf{e}_1^*(\mathbf{x}) &= (1-a)\mathbf{Q}\left[\operatorname*{plim}_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathbf{q}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\sigma_i^2 - \mathbf{m}_3\right] \qquad (4.10)\\
&\quad -\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\left[\operatorname*{plim}_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\sigma_i^2\mathbf{q}_i^{'}\mathbf{x}\right]\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{x} \ ,
\end{aligned}
$$

where

$$
\mathbf{m}_3 = \mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\left[\operatorname*{plim}_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\kappa_{3i}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\right] \ . \qquad (4.11)
$$

It is important to note that the semi-parametric estimation has the effects through the terms associated with $\mathbf{Q}$ and $\mathbf{m}_3$, which disappear only when $a = 1$ (i.e. the MEL estimator). By using the inversion formulas (i) and (ii) given in Appendix D, we have the next result.

Theorem 4.1 : Suppose the limit of (4.11) exists. Under *Assumption II*, an 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

$$f(\boldsymbol{\xi}) = \phi^*_{\mathbf{Q}}(\boldsymbol{\xi}) + \frac{1}{\sqrt{n}}\phi_{\mathbf{Q}}(\boldsymbol{\xi})\left\{(1-a)\left[\operatorname*{plim}_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathrm{tr}(\mathbf{z}_i\mathbf{z}_i^{'}\sigma_i^2\mathbf{A})\mathbf{q}_i^{'}\boldsymbol{\xi} - \mathbf{m}_3^{'}\boldsymbol{\xi}\right]\right. \tag{4.12}$$
$$+\left[\operatorname*{plim}_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}[\mathrm{tr}(\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{z}_i\mathbf{z}_i^{'}\sigma_i^2\mathbf{q}_i^{'}\boldsymbol{\xi}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}) + \mathbf{q}_i^{'}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{z}_i\mathbf{z}_i^{'}\sigma_i^2\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\boldsymbol{\xi}]\right]$$
$$\left.-\left[\operatorname*{plim}_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\xi}^{'}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{z}_i\mathbf{z}_i^{'}\sigma_i^2\mathbf{q}_i^{'}\boldsymbol{\xi}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\boldsymbol{\xi}\right]\right\} + o(n^{-1/2})$$

provided that the limits in the right-hand side of (4.12) exist, where $\mathbf{q}_i^{'} = (1/\sigma_i^2)\mathbf{E}[(\mathbf{v}_{2i}^{'},\mathbf{0}^{'})u_i|\mathbf{z}_i]$ $(i=1,\cdots,n)$, $\boldsymbol{\xi}$ is a $p\times 1$ $(p = G_1 + K_1)$ vector, $\phi^*_{\mathbf{Q}}(\xi)$ is given by (4.9) and $\phi_{\mathbf{Q}}(\xi)$ is the density function of $N_p(\mathbf{0},\mathbf{Q})$.

It is possible to extend *Theorem 4.1* to the terms of $O_p(n^{-1})$ in principle, but the resulting expressions become quite complicated. When the third order moments of disturbances are zeros, however, it is manageable to evaluate many terms of $O_p(n^{-1})$ and then we have useful representations. Also in this situation some terms of (4.9) vanish (i.e. $\beta_{l_1l_2l_3} = 0$) and we only have some extra terms of n^{-1}. When $\mathbf{q}_i = \mathbf{q}, i = 1,\cdots,n$, (4.10) becomes

$$\mathbf{e}_1^*(\mathbf{x}) = (1-a)\mathbf{Q}\left[L\mathbf{q} - \mathbf{m}_3\right] - \mathbf{x}\mathbf{q}^{'}\mathbf{x}\,. \tag{4.13}$$

By collecting the conditional expectation formulas in *Appendix A* ((A.30), (A.31); (A.14), (A.15), (A.16); (A.18), (A.19), (A.20) and (A.21)) under *Assumption III*,

$$\begin{aligned}
\mathbf{e}_2^*(\mathbf{x}) &= -(2+\kappa)\mathbf{Q}\mathbf{D}^{'}\mathbf{F}\mathbf{D}\mathbf{x} + [2 + a(2+\kappa)]\mathbf{Q}\mathbf{D}^{'}\mathbf{F}\mathbf{D}\mathbf{x} + \mathbf{x}\mathbf{x}^{'}\mathbf{C}_1^*\mathbf{x} \qquad (4.14)\\
&\quad +\mathbf{Q}\mathbf{Q}^*\mathbf{Q}\mathbf{C}_2^*\mathbf{x} - (1-a)L[\mathbf{x}\,\mathbf{tr}(\mathbf{C}_1^*\mathbf{Q}) + 2\mathbf{Q}\mathbf{C}_1^*\mathbf{x}] - (1-a)\mathbf{Q}\mathbf{C}_2^*\mathbf{x}\,\mathbf{tr}(\mathbf{M}\mathbf{A})\\
&\quad +[-3a + a]\mathbf{Q}\mathbf{D}^{\prime}\mathbf{F}\mathbf{D}\mathbf{x}\\
&= (a-1)\kappa\mathbf{Q}\mathbf{D}^{'}\mathbf{F}\mathbf{D}\mathbf{x} + \mathbf{x}\mathbf{x}^{'}\mathbf{C}_1^*\mathbf{x} + \mathbf{Q}\mathbf{Q}^*\mathbf{Q}\mathbf{C}_2^*\mathbf{x}\\
&\quad -(1-a)L[\mathbf{x}\,\mathbf{tr}(\mathbf{C}_1^*\mathbf{Q}) + 2\mathbf{Q}\mathbf{C}_1^*\mathbf{x}] - (1-a)\mathbf{Q}\mathbf{C}_2^*\mathbf{x}\,\mathbf{tr}(\mathbf{M}\mathbf{A}),
\end{aligned}$$

where $\mathbf{C}_1^* = \mathbf{q}\mathbf{q}^{'}$, $\mathbf{C}_2^* = \mathbf{E}(\mathbf{w}_i\mathbf{w}_i^{'})$ and $\mathbf{Q}^* = \mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}$. Also the second order conditional moments of $\mathbf{e}_{11}^*(x)$ under *Assumption III* can be summarized ((A.33), (A.34) and (A.13)) as

$$\begin{aligned}
\mathbf{e}_{11}^*(\mathbf{x}) &= (2+\kappa)\mathbf{Q}\mathbf{D}^{'}\mathbf{F}\mathbf{D}\mathbf{Q} + \mathbf{x}^{'}\mathbf{C}_1^*\mathbf{x}\mathbf{x}\mathbf{x}^{'} + \mathbf{Q}\mathbf{Q}^*\mathbf{Q}\mathbf{x}^{'}\mathbf{C}_2^*\mathbf{x} + \mathbf{Q}\mathbf{C}_2^*\mathbf{Q}\mathbf{tr}(\mathbf{M}\mathbf{A}) \qquad (4.15)\\
&\quad +(1-a)^2L(L+2)\mathbf{Q}\mathbf{C}_1^*\mathbf{Q} - (1-a)L[\mathbf{Q}\mathbf{C}_1^*\mathbf{x}\mathbf{x}^{'} + \mathbf{x}\mathbf{x}^{'}\mathbf{C}_1^*\mathbf{Q})].
\end{aligned}$$

Although there are many terms it is important to note that the semi-parametric estimation has the effects only through the additional terms associated with $\mathbf{Q}\mathbf{D}^{'}\mathbf{F}\mathbf{D}$ as explained in *Appendix A*. When the disturbance terms satisfy *Assumption III*, $\mathbf{C} = \sigma^2\mathbf{M}$, $\mathbf{Q} = \sigma^2(\mathbf{D}^{'}\mathbf{M}\mathbf{D})^{-1}$, $\mathbf{Q}^* = \sigma^{-2}\mathbf{Q}^{-1}$ and $\mathbf{tr}(\mathbf{M}\mathbf{A}) = \sigma^{-2}L$. Also the characteristic function of $\tilde{\mathbf{e}}_0 = \mathbf{x}$ is asymptotically equivalent to $\mathbf{E}[\exp(i\mathbf{t}^{'}\mathbf{x}^*)](1 + o(n^{-1/2}))$, where $\mathbf{x}^*$ is the limiting vector of $\mathbf{x}$. By using the inversion formulas in *Appendix D* we obtain the main result after lengthy but straightforward computations.

Theorem 4.2 : Suppose that the limits of (3.3) and (4.11) exist. Then under *Assumption III*, an asymptotic expansion of the joint density function of $\hat{\mathbf{e}}$ for a class of the MMEL estimator as $n \to \infty$ is given by

$$
\begin{aligned}
& f(\boldsymbol{\xi}) = \phi^*_{\mathbf{Q}}(\boldsymbol{\xi}) \qquad (4.16)\\
& +\frac{1}{\sqrt{n}}\phi_{\mathbf{Q}}(\boldsymbol{\xi})(\mathbf{q}'\boldsymbol{\xi})\left[p+1+(1-a)L-\boldsymbol{\xi}'\mathbf{Q}^{-1}\boldsymbol{\xi}\right]\\
& +\frac{1}{2n}\phi_{\mathbf{Q}}(\boldsymbol{\xi})\Big(\boldsymbol{\xi}'\mathbf{C}_1\boldsymbol{\xi}\left\{[p+1+(1-a)L-\boldsymbol{\xi}'\mathbf{Q}^{-1}\boldsymbol{\xi}]^2+p+1-3\boldsymbol{\xi}'\mathbf{Q}^{-1}\boldsymbol{\xi}+2(1-a)^2L\right\}\\
& \quad +\mathbf{tr}(\mathbf{C}_1\mathbf{Q})[(1-a)L][2-(1-a)(L+2)]\\
& \quad +\boldsymbol{\xi}'\mathbf{C}_2\boldsymbol{\xi}\{L[1-2(1-a)]-p-2+\boldsymbol{\xi}'\mathbf{Q}^{-1}\boldsymbol{\xi}\}+\mathbf{tr}(\mathbf{C}_2\mathbf{Q})\{L[2(1-a)-1]\}\\
& \quad +[2+(2a-1)\kappa][\boldsymbol{\xi}'\mathbf{D}'\mathbf{F}\mathbf{D}\boldsymbol{\xi}-\mathbf{tr}(\mathbf{D}'\mathbf{F}\mathbf{D}\mathbf{Q})]\Big)+o(n^{-1}) \ ,
\end{aligned}
$$

where $\boldsymbol{\xi}$ is a $p\times 1$ ($p = G_1 + K_1$) vector, $\phi^*_{\mathbf{Q}}(\xi)$ and $\mathbf{F}$ are given by (4.8) and (3.3), respectively, $\phi_{\mathbf{Q}}(\xi)$ is the density function of $N_p(\mathbf{0},\mathbf{Q})$, $\mathbf{C}_1 = \mathbf{C}_1^*$ $(= \mathbf{q}\mathbf{q}')$, $\mathbf{C}_2 = \sigma^{-2}\mathbf{C}_2^*$ $(= \sigma^{-2}\mathbf{E}(\mathbf{w}_i\mathbf{w}_i'))$, $\sigma^2 = \mathbf{E}(u_i^2)$ and $\kappa = [\mathbf{E}(u_i^4) - 3\sigma^4]/\sigma^4$.

The leading 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 (4.9) and $\phi^*_{\mathbf{Q}}(\mathbf{x}) = \phi_{\mathbf{Q}}(\mathbf{x})$. There is an interesting observation in Theorem 4.2 that if we further drop the last term

$$
[2+(2a-1)\kappa]\left[\boldsymbol{\xi}'\mathbf{D}'\mathbf{F}\mathbf{D}\boldsymbol{\xi}-\mathbf{tr}(\mathbf{D}'\mathbf{F}\mathbf{D}\mathbf{Q})\right] \qquad (4.17)
$$

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). Hence this term could be interpreted as *the effect of semi-parametric factor* in the linear simultaneous equations as we have observed in *Theorem 3.1* and *Theorem 3.2*. This term comes from many terms associated with the semi-parametric covariance estimation, (See the detail in *Appendix A*), which gives the MEL estimation a more variability in the order $O(n^{-1})$ depending on the kurtosis of the underlying distribution. In the first and second orders there is no distinctive different features between the density functions of the standardized MEL estimator and LIML estimator as in *Theorem 4.1*, which implies the same asymptotic bias up to $O_p(n^{-1/2})$. In that sense we may call the term (4.17) as the semi-parametric (3rd order) *inefficiency* factor under the homoscedasticity assumption for disturbances.

By using the asymptotic expansion of the density function, we can evaluate the asymptotic mean and the asymptotic mean squared errors of the MMEL estimator.

Corollary 4.3 : Under the assumptions of *Theorem 4.2*, the asymptotic bias and the asymptotic mean squared errors of $\hat{\mathbf{e}}$ with the MMEL estimator (based on the asymptotic expansion) as $n \to \infty$ are $\mathbf{ABIAS}_n(\hat{\mathbf{e}}) = n^{-1/2}\left[(1-a)L-1\right]\mathbf{Q}\mathbf{q} + o(n^{-1/2})$ and

$$
\begin{aligned}
\mathbf{AMSE}_n(\hat{\mathbf{e}}) \ = \ & \mathbf{Q}+\frac{1}{n}\Big\{\mathbf{Q}\mathbf{C}_1\mathbf{Q}[6-6(1-a)L+(1-a)^2L(L+2)] \qquad (4.18)\\
& +\mathbf{Q}\mathbf{tr}(\mathbf{C}_1\mathbf{Q})[3-2(1-\delta)L]+\mathbf{Q}\mathbf{tr}(\mathbf{C}_2\mathbf{Q})+[L+2-2L(1-a)]\mathbf{Q}\mathbf{C}_2\mathbf{Q}\\
& +[2+(2a-1)\kappa]\mathbf{Q}\mathbf{D}'\mathbf{F}\mathbf{D}\mathbf{Q}\Big\}+o(n^{-1}) \ ,
\end{aligned}
$$

respectively.

4.3 Discussions on Higher Order Properties of Estimators

Under *Assumption II* it is straightforward to obtain the asymptotic expansion of the density function of the MEL and GMM estimators up to $O(n^{-1/2})$. In *Theorem 4.1* when $\mathbf{q}_i = \mathbf{q}$ $(i = 1, \cdots, n)$, for instance, 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 when $a = 1$. Let $\hat{e}_{i.MEL}$ $(i = 1, \cdots, p)$ be the i-th component of $\hat{e}$ for the MEL estimator. Then

$$\mathbf{P}(\hat{e}_i \geq 0) = \frac{1}{2} + o(n^{-1/2}) \tag{4.19}$$

when $\kappa_{3i} = 0$ (i.e. $\beta_{l_1 l_2 l_3} = 0$ $(l_1, l_2, l_3 = 1, \cdots, p)$ in (4.9)). Hence it is still near to 1/2 (almost median-unbiased) for the MEL estimator when κ_3 is small in many applications.

On the other hand, the asymptotic expansion of the density function of the GMM estimator has an additional term and the term of $O(n^{-1/2})$ is proportional to $L(1/\sqrt{n})$, where $L = K_2 - G_1$. Hence when K_2 (the number of excluded instruments) is large, the probability bias of the GMM (or the TSLS) estimator becomes substantial while the MEL (or the LIML) estimator concentrates its probability around the true parameter values. (See Tables 2 and 3 in *Appendix E*. By taking the expectation of (4.13) when $\mathbf{q}_i = \mathbf{q}$ $(i = 1, \cdots, n)$, the asymptotic (unconditional) bias of the MMEL estimator with respect to the approximate distribution based on the asymptotic expansions is given by

$$\mathbf{ABIAS}_n(\hat{\mathbf{e}}) = \frac{1}{\sqrt{n}}\Big\{[(1-a)L - 1]\mathbf{Qq} - (1-a)\mathbf{Qm}_3\Big\} + o(n^{-1/2}). \tag{4.20}$$

The result on the asymptotic bias may agree with the observation by Newey and Smith (2004), which have derived the asymptotic bias of the MEL and GMM estimators in the more general nonlinear setting for the estimating equation models.

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 $\mathbf{AM}(\hat{\mathbf{e}}\hat{\mathbf{e}}^{'})$ for the asymptotic MSE of the MMEL estimator in the general linear case. There are many terms for an arbitrary a $(0 \leq a \leq 1)$ 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.

The issue of comparing the finite sample distributions of alternative estimators based on their asymptotic expansions in the order $O(n^{-1})$ for the normalized estimators are closely related to the problem of higher order asymptotic efficiency and deficiency of in the statistical asymptotic theory. On the one hand, 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 after bias adjustments when the disturbances are normally distributed. Recently, Newey and Smith (2004) utilized the multinomial distribution case and concluded (in their *Theorem 6.1*) that the MEL estimator is third order asymptotically efficient after bias adjustments by using the arguments by Pfanzagl and Wefelmeyer (1978) in the more general nonlinear estimating equation framework. It could be interpreted as an application of the higher order efficiency of estimation developed by Pfanzagl and Wefelmeyer

(1978) and Akahira and Takeuchi (1981) for the statistical framework of *parametric models*. On the other hand, Akahira and Takeuchi (1990) have given several examples and suggested that the asymptotic (higher order) deficiency in *semi-parametric models* often become infinite, which is quite different from the estimation problem of standard parametric models. There is a subtle statistical problem remained on the meaning of the asymptotic bound, the (higher order) asymptotic efficiency and deficiency of estimation in semi-parametric models (see Pfanzagl (1990) and Bickel et al. (1993)). The related analysis should be important, but it is beyond of the scope of this paper.

5. Concluding Remarks

In this paper we have developed the asymptotic expansions of the density functions for a class of semi-parametric estimators including the MEL and the GMM estimators. Although the general forms of the asymptotic expansions look quite complicated, it is possible to obtain some explicit formulas which make possible to compare alternative estimation methods.

On the other hand, Anderson et al. (2005, 2008), for instance, have investigated the finite sample properties of the distribution functions of the MEL and GMM estimators and have given extensive tables when $G_1 = 1, 2$ in a systematic way. In the more general case, however, it would not be possible to investigate the finite sample properties directly and hence the asymptotic expansion method should be useful for comparing different estimators. The explicit formulas in Section 4 give some useful information on the exact distributions of alternative estimators in more general cases. They should be the basis of comparing higher order terms of the distribution functions of alternative estimators beyond their asymptotic biases and MSEs.

It is important to note that the finite sample differences between the distributions of the LIML and MEL estimators (and also those between the GMM and TSLS estimators) are often very small as we have discussed in Sections 3 and 4 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 practical purposes.

Finally, it is obvious that the results reported in this paper have implications on the general reduced rank regression models. This problem is currently under investigation.

Appendices

In Appendix A and Appendix B, we give the derivations of stochastic expansions of alternative estimators. In Appendix C we give the proofs of two lemmas and in Appendix D we gather some useful inversion formulas. We give tables and figures in Appendix E.

Appendix A : Derivations of asymptotic expansions

[A1] Conditional Stochastic Expansions

We derive the asymptotic expansions of estimators under *Assumption II* and then we shall show how *Assumption III* simplifies the resulting expressions. By expanding (4.1) with respect to $\mathbf{e}_0$, formally we write

$$\hat{\mathbf{e}} = \tilde{\mathbf{e}}_0 + [\mathbf{e}_0 - \tilde{\mathbf{e}}_0] + \frac{1}{\sqrt{n}}\mathbf{e}_1 + \frac{1}{n}\mathbf{e}_2 + o_p(n^{-1}) \tag{A.1}$$

and

$$\sqrt{n}\hat{\boldsymbol{\lambda}} = \boldsymbol{\lambda}_0 + \frac{1}{\sqrt{n}}\boldsymbol{\lambda}_1 + \frac{1}{n}\boldsymbol{\lambda}_2 + o_p(n^{-1}) \ . \tag{A.2}$$

By substituting these expansions and $u_i(\hat{\boldsymbol{\theta}}) = u_i - (1/\sqrt{n})(\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\hat{\mathbf{e}}$ into p_i $(i = 1, \cdots, n)$, we also write

$$n\,\hat{p}_i = 1 + \frac{1}{\sqrt{n}}p_i^{(1)} + \frac{1}{n}p_i^{(2)} + o_p(n^{-1}) \ , \tag{A.3}$$

where $p_i^{(1)} = -\boldsymbol{\lambda}_0^{'}\mathbf{z}_i u_i$, $p_i^{(2)} = -\boldsymbol{\lambda}_1^{'}\mathbf{z}_i u_i + \boldsymbol{\lambda}_0^{'}\mathbf{z}_i(\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})\mathbf{e}_0 + (\boldsymbol{\lambda}_0^{'}\mathbf{z}_i)^2 u_i^2$ and $(\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'}) = \mathbf{z}_i^{'}\mathbf{D} + \mathbf{w}_i^{'} + \mathbf{q}_i^{'}u_i$.
Then it is possible to show that $\max_{1\leq i\leq n}|\hat{p}_i - 1/n| = o_p(1/n)$ since $(n\hat{p}_i)^{-1} = 1 + \boldsymbol{\lambda}^{'}\mathbf{z}_i u_i$ (see Owen (1990)), and $\max_{1\leq i\leq n}|\hat{p}_i - 1/n - p_i^{(1)}/(n\sqrt{n}) - p_i^{(2)}/n^2| = o_p(1/n^2)$. By using the recursive substitution, we expand

$$\hat{\mathbf{C}}_n = \sum_{i=1}^{n}\hat{p}_i u_i^2(\hat{\theta})\mathbf{z}_i\mathbf{z}_i^{'} = \mathbf{C}_n + \frac{1}{\sqrt{n}}\mathbf{C}_n^{(1)} + \frac{1}{n}\mathbf{C}_n^{(2)} + o_p(n^{-1}) \ , \tag{A.4}$$

$$\hat{\mathbf{E}}_n = \sum_{i=1}^{n}\hat{p}_i\begin{pmatrix}\mathbf{y}_{2i}\\ \mathbf{z}_{1i}\end{pmatrix}\mathbf{z}_i^{'} = \mathbf{D}^{'}\mathbf{M}_n + \frac{1}{\sqrt{n}}\mathbf{E}_n^{(1)} + \frac{1}{n}\mathbf{E}_n^{(2)} + o_p(n^{-1}) \ , \tag{A.5}$$

where we define

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

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

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

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

By using (2.6) we write $\hat{\mathbf{E}}_n\hat{\mathbf{C}}_n^{-1}\mathbf{X}_n = \hat{\mathbf{E}}_n\hat{\mathbf{C}}_n^{-1}[n^{-1}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{y}_{2i}^{'}, \mathbf{z}_{1i}^{'})]\hat{\mathbf{e}}$. Then by substituting $\hat{\mathbf{e}}, \hat{\boldsymbol{\lambda}}, \hat{p}_i$ $(i = 1, \cdots, n)$ and $\mathbf{Z}_n$, we determine each terms of the stochastic expansions of $\hat{\mathbf{e}}$ in the recursive way. By using the relation $\hat{\mathbf{C}}_n^{-1} = \mathbf{C}_n^{-1} + n^{-1/2}[-\mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1}] + n^{-1}[-\mathbf{C}_n^{-1}\mathbf{C}_n^{(2)}\mathbf{C}_n^{-1} + \mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1}] + o_p(n^{-1})$, the leading two terms of $\hat{\mathbf{e}}$ are

$$\mathbf{e}_1 = -\mathbf{Q}_n\mathbf{D}^{'}\mathbf{M}_n\mathbf{C}_n^{-1}\mathbf{Z}_n\mathbf{e}_0 + \mathbf{Q}_n[\mathbf{A}_{1n}][\mathbf{X}_n - \mathbf{M}_n\mathbf{D}\mathbf{e}_0], \tag{A.6}$$

$$\begin{aligned}\mathbf{e}_2 = {} & \mathbf{Q}_n[\mathbf{A}_{2n}][\mathbf{X}_n - \mathbf{M}_n\mathbf{D}\mathbf{e}_0] - \mathbf{Q}_n[\mathbf{A}_{1n}][\mathbf{M}_n\mathbf{D}\mathbf{e}_1 + \mathbf{Z}_n\mathbf{e}_0] \\ & -\mathbf{Q}_n\mathbf{D}^{'}\mathbf{M}_n\mathbf{C}_n^{-1}\mathbf{Z}_n\mathbf{e}_1,\end{aligned} \tag{A.7}$$

where $\mathbf{Q}_n^{-1} = \mathbf{D}'\mathbf{M}_n\mathbf{C}_n^{-1}\mathbf{M}_n\mathbf{D}$, $\hat{\mathbf{E}}_n\hat{\mathbf{C}}_n^{-1} = \mathbf{D}'\mathbf{M}_n\mathbf{C}_n^{-1} + n^{-1/2}\mathbf{A}_{1n} + n^{-1}\mathbf{A}_{2n} + o_p(n^{-1})$,

$$\begin{aligned}\mathbf{A}_{1n} &= -\mathbf{D}'\mathbf{M}_n\mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1} + \mathbf{E}_n^{(1)}\mathbf{C}_n^{-1} ,\\ \mathbf{A}_{2n} &= \mathbf{D}'\mathbf{M}_n[-\mathbf{C}_n^{-1}\mathbf{C}_n^{(2)}\mathbf{C}_n^{-1} + \mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1}] - \mathbf{E}_n^{(1)}\mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1} + \mathbf{E}_n^{(2)}\mathbf{C}_n^{-1}.\end{aligned}$$

[A2] Effects of $\mathbf{C}_n$ (Covariance Estimation)

We need to investigate the effects of estimating $\mathbf{C}$ by $\hat{\mathbf{C}}_n$ in the semi-parametric estimation methods. 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

$$Cov(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i^{(j)}z_i^{(k)}u_i^2, \frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i^{(l)}u_i|\mathbf{z}_i) = \frac{1}{n}\sum_{i=1}^{n}\kappa_{3i}z_i^{(j)}z_i^{(k)}z_i^{(l)} .$$

Thus $\mathbf{X}_n$ and $\mathbf{Y}_n$ are asymptotically independent when $\kappa_{3i} = \mathbf{E}(u_i^3) = 0$ and in that case our analyses can be simplified considerably as we shall see in [**A6**] in particular. Because $\mathbf{C}_n^{-1} = \mathbf{C}^{-1} + [-\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}] + [\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}] + o_p(n^{-1})$ and $\mathbf{Q}_n^{-1} = \mathbf{D}'\mathbf{M}_n[\mathbf{C}^{-1} + \mathbf{C}_n^{-1}(\mathbf{C} - \mathbf{C}_n)\mathbf{C}^{-1}]\mathbf{M}_n\mathbf{D}$, which is $\mathbf{Q}_n^{-1} = \mathbf{Q}^{-1} - n^{1/2}[\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}] + n^{-1}[\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}] + O_p(n^{-3/2})$ and $\mathbf{D}'\mathbf{M}_n\mathbf{C}_n^{-1} = \mathbf{D}'\mathbf{M}\mathbf{C}^{-1} - n^{-1/2}[\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}] + n^{-1}[\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}] + O_p(n^{-3/2})$.

Then $\mathbf{Q}_n = \mathbf{Q} + \mathbf{Q}_n(\mathbf{Q}^{-1} - \mathbf{Q}_n^{-1})\mathbf{Q}$ is expanded as $\mathbf{Q} + n^{-1/2}[\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}] + n^{-1}[-\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{M}\mathbf{D}\mathbf{Q}] + O_p(n^{-3/2})$ and $\mathbf{Q}_n\mathbf{D}'\mathbf{M}_n\mathbf{C}_n^{-1}$ is expanded as $\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1} + n^{-1/2}[-\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}] + n^{-1}[\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{Y}_n\mathbf{A}] + O_p(n^{-3/2})$, where $\mathbf{A}\mathbf{C}\mathbf{A} = \mathbf{A}$. Then we can express $\mathbf{e}_0 = \tilde{\mathbf{e}}_0 + n^{-1/2}\mathbf{e}_0^{(1)} + n^{-1}\mathbf{e}_0^{(2)} + O_p(n^{-3/2})$, 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 $\mathbf{e}_0^{(2)} = \mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n$. By using the expansions of $\mathbf{C}_n$ and $\mathbf{Q}_n$, we find a representation for (4.5) as

$$\boldsymbol{\lambda}_0 = \mathbf{A}\mathbf{X}_n + \frac{1}{\sqrt{n}}[-\mathbf{A}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n] + O_p(n^{-1}) . \tag{A.8}$$

[A3] Conditional Expectations involving $\mathbf{e}_1$

We investigate the effects of $\mathbf{e}_1$ and decompose $\mathbf{e}_1$ as $\mathbf{e}_1 = \mathbf{e}_{1.1} + \mathbf{e}_{1.2} + \mathbf{e}_{1.3}$, where $\mathbf{e}_{1.1} = \mathbf{Q}_n[\mathbf{A}_{1n}][\mathbf{X}_n - \mathbf{M}\mathbf{D}\mathbf{e}_0]$, $\mathbf{e}_{1.2} = -\mathbf{Q}_n\mathbf{D}'\mathbf{M}\mathbf{C}_n^{-1}n^{-1/2}\sum_{i=1}^{n}\mathbf{z}_iu_i\mathbf{q}_i'\mathbf{e}_0$ 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 are evaluated easily and we treat them first. Rewrite $\mathbf{e}_{1.2} = \mathbf{e}_{1.2}^{(0)} + n^{-1/2}\mathbf{e}_{1.2}^{(1)} + O_p(n^{-1})$, $\mathbf{e}_{1.2}^{(0)} = -\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}n^{-1/2}\sum_{i=1}^{n}\mathbf{z}_iu_i\mathbf{q}_i'\tilde{\mathbf{e}}_0$ and $\mathbf{e}_{1.2}^{(1)} = -\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n(n^{-1/2}\sum_{i=1}^{n}\mathbf{z}_iu_i\mathbf{q}_i')\tilde{\mathbf{e}}_0 + \mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n(n^{-1/2}\sum_{i=1}^{n}\mathbf{z}_iu_i\mathbf{q}_i')\tilde{\mathbf{e}}_0\mathbf{A}\mathbf{X}_n$. Also we have $\mathbf{e}_{1.3} = \mathbf{e}_{1.3}^{(0)} + n^{-1/2}\mathbf{e}_{1.3}^{(1)} + O_p(n^{-1})$, where $\mathbf{e}_{1.3}^{(0)} = -\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n'\tilde{\mathbf{e}}_0$ and $\mathbf{e}_{1.3}^{(1)} = \mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n'\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n + \mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{U}_n'\tilde{\mathbf{e}}_0$.

The analysis of $\mathbf{e}_{1.1}$ becomes more complicated because there are some terms with $\mathbf{C}_n^{(1)}$ and $\mathbf{E}_n^{(1)}$. We rewrite $\mathbf{C}_n^{(1)} = \mathbf{C}_n^{(1.0)} + n^{-1/2}\mathbf{C}_n^{(1.1)}$ and $\mathbf{C}_n^{(1,0)} = -2(n^{-1}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i'u_i^2\mathbf{q}_i')\tilde{\mathbf{e}}_0 - \boldsymbol{\Theta}_{3n}^{\lambda_0}$ by defining $\boldsymbol{\Theta}_{3n}^{\lambda_0} = n^{-1}\sum_{i=1}^{n}\kappa_{3i}\mathbf{z}_i\mathbf{z}_i'(\mathbf{z}_i'\boldsymbol{\lambda}_0)$. Also we have $\mathbf{E}_n^{(1)} = \mathbf{E}_n^{(1.0)} + n^{-1/2}\mathbf{E}_n^{(1.1)}$, where $\mathbf{E}_n^{(1.0)} = \mathbf{U}_n + n^{-1/2}\sum_{i=1}^{n}\mathbf{q}_i\mathbf{z}_i'u_i + n^{-1/2}\sum_{i=1}^{n}\mathbf{q}_i(-\boldsymbol{\lambda}_0'\mathbf{z}_i\mathbf{z}_i'u_i^2)$.

Then $\mathbf{e}_{1.1} = [-\mathbf{Q}_n\mathbf{D}'\mathbf{M}_n\mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1} + \mathbf{Q}_n\mathbf{E}_n^{(1)}\mathbf{C}_n^{-1}]\mathbf{C}_n\mathbf{A}\mathbf{X}_n$ becomes

$$\begin{aligned}\mathbf{e}_{1.1} &= 2\mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\mathbf{C}\mathbf{A}\mathbf{X}_n + \mathbf{Q}\mathbf{D}'\mathbf{M}\mathbf{C}^{-1}\boldsymbol{\Theta}_{3n}^{\lambda_0}\mathbf{C}^{-1}\mathbf{A}\mathbf{X}_n \qquad (A.9)\\ &\quad + \mathbf{Q}\mathbf{U}_n\mathbf{A}\mathbf{X}_n + \mathbf{Q}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{q}_iu_i\mathbf{z}_i' - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{q}_i(\boldsymbol{\lambda}_0'\mathbf{z}_i)\mathbf{z}_i'u_i^2\right]\mathbf{A}\mathbf{X}_n + O_p(n^{-1/2}) .\end{aligned}$$

By collecting each terms of $\mathbf{e}_1$, we summarize $\mathbf{e}_1 = \mathbf{e}_1^{(0)} + n^{-1/2}\mathbf{e}_1^{(1)} + o_p(n^{-1/2})$, $\mathbf{e}_1^{(0)} = \mathbf{e}_{1.1}^{(0)} + \mathbf{e}_{1.2}^{(0)} + \mathbf{e}_{1.3}^{(0)}$ and $\mathbf{e}_1^{(1)} = \mathbf{e}_{1.1}^{(1)} + \mathbf{e}_{1.2}^{(1)} + \mathbf{e}_{1.3}^{(1)}$.

In order to derive the asymptotic expansions of the distributions of the MMEL estimator, we use $\hat{p}_i^*$ instead of $\hat{p}_i$ $(i = 1, 2)$. For an arbitrary (fixed) a $(0 \le a \le 1)$, we substitute $a\boldsymbol{\lambda}_0$ (and $a\boldsymbol{\lambda}_1$) into $\boldsymbol{\lambda}_0$ (and $\boldsymbol{\lambda}_1$). Since $\tilde{\mathbf{e}}_0$ is asymptotically uncorrelated with $\mathbf{AX}_n$, $\mathbf{E}[\mathbf{e}_{1.3}^{(0)}|\mathbf{x}] = o_p(1)$ and *Lemma A.3* in **[A5]**, the conditional expectation of $\mathbf{e}_1^{(0)}$, given $\tilde{\mathbf{e}}_0 = \mathbf{x}$, is

$$
\begin{aligned}
& \mathbf{E}[\mathbf{e}_1^{(0)}|\mathbf{x}] \qquad\qquad (A.10) \\
= \; & a\mathbf{Qm}_3 + (1-a)\mathbf{Q}\left[\frac{1}{n}\sum_{i=1}^{n}\mathbf{q}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\sigma_i^2\right] - \mathbf{QD}^{'}\mathbf{MC}^{-1}\left[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\sigma_i^2\mathbf{q}_i^{'}\mathbf{x}\right]\mathbf{C}^{-1}\mathbf{MDx},
\end{aligned}
$$

with the remainder terms of $o_p(1)$, where $n^{-1/2}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{E}[u_i|\mathbf{x}]\mathbf{q}_i^{'} = n^{-1}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\sigma_i^2\mathbf{q}_i^{'}\mathbf{x} + o_p(1)$ and $\mathbf{m}_3$ is given by (4.11).

Now we explicitly use the assumption $\mathbf{q}_i = \mathbf{q}$ $(i = 1, \cdots, n)$ and *Assumption III* in order to evaluate many terms in the order of $O_p(n^{-1})$. We write $\mathbf{C}_n^{(1)} = \mathbf{C}_n^{(1.0*)} + n^{-1/2}\mathbf{C}_n^{(1.1*)}$, $\mathbf{C}_n^{(1,0*)} = 2\mathbf{q}^{'}\tilde{\mathbf{e}}_0\mathbf{C}^{-1} - \boldsymbol{\Theta}_{3n}^{\lambda_0}$, $\boldsymbol{\Xi}_{3n}^{\lambda_0} = n^{-1/2}\sum_{i=1}^{n}(u_i^3 - \kappa_{3i})\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)$ and $\mathbf{C}_n^{(1.1*)} = -2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{Y}_n - 2n^{-1/2}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{\prime}u_i(\mathbf{z}_i^{'}\mathbf{D} + \mathbf{w}_i^{'})\tilde{\mathbf{e}}_0 + 2\mathbf{C}_n\mathbf{q}^{'}\mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{Y}_n\mathbf{AX}_n - a\boldsymbol{\Xi}_{3n}^{\lambda_0}$. Then we write $-\mathbf{Q}_n\mathbf{D}^{'}\mathbf{M}_n\mathbf{C}_n^{-1}\mathbf{C}_n^{(1)}\mathbf{C}_n^{-1} = B_1^{(1)} + n^{-1/2}B_1^{(2)} + o_p(n^{-1/2})$, $B_1^{(1)} = 2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{QD}^{'}\mathbf{MC}^{-1} + \mathbf{QD}^{'}\mathbf{MC}^{-1}a\boldsymbol{\Theta}_{3n}^{\lambda_0}\mathbf{C}^{-1}$ and

$$
\begin{aligned}
B_1^{(2)} &= -2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{Y}_n\mathbf{A} - \mathbf{QD}^{'}\mathbf{MC}^{-1}[a\boldsymbol{\Theta}_{3n}^{\lambda_0}\mathbf{C}^{-1}\mathbf{Y}_n + \mathbf{Y}_n\mathbf{A}a\boldsymbol{\Xi}_{3n}^{\lambda_0}]\mathbf{C}^{-1} \\
&- \mathbf{QD}^{'}\mathbf{MC}^{-1}[-a\boldsymbol{\Xi}_{3n}^{\lambda_0} - 2\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{\prime}u_i(\mathbf{z}_i^{'}\mathbf{D} + \mathbf{w}_i^{'})\tilde{\mathbf{e}}_0 + 2\mathbf{Cq}^{'}\mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{Y}_n\mathbf{AX}_n]\mathbf{C}^{-1}.
\end{aligned}
$$

We also write $\mathbf{E}_n^{(1)} = \mathbf{E}_n^{(1.0*)} + n^{-1/2}\mathbf{E}_n^{(1.1*)}$, $\mathbf{E}_n^{(1.0*)} = \mathbf{U}_n + \mathbf{q}(\mathbf{X}_n^{'} - a\boldsymbol{\lambda}_0^{'}\mathbf{C}_n)$ and $\mathbf{E}_n^{(1.1*)} = -\mathbf{D}^{\prime}n^{-1/2}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(a\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)u_i - n^{-1/2}\sum_{i=1}^{n}\mathbf{w}_i\mathbf{z}_i^{'}(a\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)u_i$. Then $\mathbf{Q}_n\mathbf{E}_n^{(1)}\mathbf{C}_n^{-1} = B_2^{(1)} + n^{-1/2}B_2^{(2)} + o_p(n^{-1/2})$, $B_2^{(1)} = \mathbf{QU}_n\mathbf{C}^{-1} + \mathbf{Qq}(\mathbf{X}_n^{'} - a\boldsymbol{\lambda}_0^{'}\mathbf{C})\mathbf{C}^{-1}$ and

$$
\begin{aligned}
B_2^{(2)} = \; & \mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{MDQ}[\mathbf{U}_n + \mathbf{q}(\mathbf{X}_n^{'} - a\boldsymbol{\lambda}_0^{'}\mathbf{C}_n)]\mathbf{C}^{-1} \\
& -\mathbf{Q}[\mathbf{U}_n + \mathbf{q}(\mathbf{X}_n^{'} - a\boldsymbol{\lambda}_0^{'}\mathbf{C}_n)]\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1} \\
& +\mathbf{Q}[-\mathbf{D}^{\prime}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(a\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)u_i - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{w}_i\mathbf{z}_i^{'}(a\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)u_i]\mathbf{C}^{-1} - \mathbf{Qq}a\boldsymbol{\lambda}_0^{'}\mathbf{Y}_n\mathbf{C}^{-1}\ .
\end{aligned}
$$

By using $\mathbf{X}_n - \mathbf{M}_n\mathbf{De}_0 = \mathbf{CAX}_n + n^{-1/2}[\mathbf{MDQD}^{'}\mathbf{MC}^{-1}\mathbf{Y}_n\mathbf{AX}_n] + O_p(n^{-1})$ and $\mathbf{C}^{-1}\mathbf{MDQD}^{'}\mathbf{MC}^{-1} = \mathbf{C}^{-1} - \mathbf{A}$, for an arbitrary a, $\mathbf{e}_{1.1} = \mathbf{e}_{1.1}^{(0)} + n^{-1/2}\mathbf{e}_{1.1}^{(1)} + o_p(n^{-1/2})$, $\mathbf{e}_{1.1}^{(0)} = \left[\mathbf{QD}^{'}\mathbf{MC}^{-1} + \mathbf{QD}^{'}\mathbf{MC}^{-1}a\boldsymbol{\Theta}_{3n}^{\lambda_0}\mathbf{C}^{-1} + \mathbf{QU}_n\mathbf{C}^{-1} + \mathbf{Qq}(\mathbf{X}_n^{'} - a\boldsymbol{\lambda}_0^{'}\mathbf{C})\mathbf{C}^{-1}\right]\mathbf{CAX}_n$ and $\mathbf{e}_{1.1}^{(1)} = (B_1^{(2)} + B_2^{(2)})\mathbf{AX}_n + (B_1^{(1)} + B_2^{(1)})\mathbf{MDQD}^{'}\mathbf{MC}^{-1}\mathbf{Y}_n\mathbf{AX}_n$, which is

$$
\begin{aligned}
\mathbf{e}_{1.1}^{(1)} = \; & \Big(a\mathbf{QD}^{'}\mathbf{MC}^{-1}\boldsymbol{\Theta}_{3n}^{\lambda_0}(\mathbf{C}^{-1} - \mathbf{A})\mathbf{Y}_n\mathbf{AX}_n \qquad (A.11) \\
& +\mathbf{QU}_n(\mathbf{C}^{-1} - \mathbf{A})\mathbf{Y}_n\mathbf{AX}_n + \mathbf{Qq}\tilde{\mathbf{e}}_0^{'}\mathbf{D}^{'}\mathbf{MC}^{-1}\mathbf{Y}_n\mathbf{AX}_n\Big) \\
& -a\mathbf{QD}^{'}\mathbf{MC}^{-1}[\boldsymbol{\Theta}_{3n}^{\lambda_0}\mathbf{C}_n\mathbf{Y}_n\mathbf{AX}_n + \mathbf{Y}_n\mathbf{A}\boldsymbol{\Xi}_{3n}^{\lambda_0}\mathbf{AX}_n]
\end{aligned}
$$

$$-\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[-2\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n u_i(\mathbf{z}_i^{'}\mathbf{D}+\mathbf{w}_i^{'})\tilde{\mathbf{e}}_0 - a\mathbf{\Xi}_{3n}^{\lambda_0}]$$

$$+\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}[\mathbf{U}_n+\mathbf{q}(\mathbf{X}_n^{'}-a\boldsymbol{\lambda}_0^{'}\mathbf{C}_n)]\mathbf{A}\mathbf{X}_n$$

$$-\mathbf{Q}[\mathbf{U}_n+\mathbf{q}(\mathbf{X}_n^{'}-a\boldsymbol{\lambda}_0^{'}\mathbf{C}_n)]\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n$$

$$+\mathbf{Q}[-a\mathbf{D}'\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2u_i - a\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{w}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2u_i]$$

$$+\left(-2\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{q}^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n\mathbf{A}\mathbf{X}_n - a\mathbf{Q}\mathbf{q}\mathbf{X}_n^{'}\mathbf{A}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n\right)\ .$$

We note that some terms are cancelled out and (A.11) will be needed in **[A6]** (two terms of $\mathbf{e}_{1.1}^{(1)}$ have important roles). Since the first term of $\mathbf{e}_{1.1}$ (i.e. $2\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{A}\mathbf{X}_n]$) is $o_p(1)$ when $\mathbf{q}_i=\mathbf{q}$ and $\mathbf{e}_1^{(0)}=\mathbf{e}_{1.1}^{(0)}+\mathbf{e}_{1.2}^{(0)}+\mathbf{e}_{1.3}^{(0)}$, then

$$\mathbf{e}_1^{(0)}=\left[a\mathbf{Q}\mathbf{D}^{'}\mathbf{C}^{-1}\mathbf{\Theta}_{3n}^{\lambda_0}+\mathbf{Q}\mathbf{U}_n+(1-a)\mathbf{Q}\mathbf{q}\mathbf{X}_n^{'}\right]\mathbf{A}\mathbf{X}_n-(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\tilde{\mathbf{e}}_0-\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n^{'}\tilde{\mathbf{e}}_0. \tag{A.12}$$

Then the conditional second moments of $\mathbf{e}_1^{(0)}$, given $\tilde{\mathbf{e}}_0=\mathbf{x}$, are calculated as

$$\mathbf{E}[\mathbf{e}_1^{(0)}\mathbf{e}_1^{(0)'}|\mathbf{x}] \tag{A.13}$$

$$= a^2\left\{\mathbf{Q}\mathbf{m}_3\cdot\mathbf{m}_3^{'}\mathbf{Q}+2\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}(\frac{1}{n})^2\sum_{i,j}\kappa_{3i}\kappa_{3j}\mathbf{z}_i\mathbf{z}_j^{'}(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_j)^2\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\right\}$$

$$+a\left\{\mathbf{Q}\mathbf{m}_3[(1-a)(L+2)\mathbf{Q}\mathbf{q}-\mathbf{x}\mathbf{x}^{'}\mathbf{q}]^{'}+[(1-a)(L+2)\mathbf{Q}\mathbf{q}-\mathbf{x}\mathbf{x}^{'}\mathbf{q}]\mathbf{m}_3^{'}\mathbf{Q}\right\}$$

$$+\Big\{(\mathbf{x}^{'}\mathbf{C}_1^{*}\mathbf{x}^{'}\mathbf{x}\mathbf{x}^{'}+\mathbf{Q}\mathbf{Q}^{*}\mathbf{Q}\mathbf{x}^{'}(\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(\mathbf{w}_i\mathbf{w}_i^{'}))\mathbf{x}+\mathbf{Q}^{'}(\frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(\mathbf{w}_i\mathbf{w}_i^{'}))\mathbf{Q}\mathbf{tr}(\mathbf{A}\mathbf{M})$$

$$+(1-a)^2L(L+2)\mathbf{Q}\mathbf{C}_1^{*}\mathbf{Q}-(1-a)L[\mathbf{Q}\mathbf{C}_1^{*}\mathbf{x}\mathbf{x}^{'}+\mathbf{x}\mathbf{x}^{'}\mathbf{C}_1^{*}\mathbf{Q}]\Big\}+o_p(1),$$

where $\mathbf{C}_1^{*}=\mathbf{q}\mathbf{q}^{'}$ and 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 the relations (by applying *Lemma A.2* in **[A5]**) 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 a consequence of the fact that $\tilde{\mathbf{e}}_0$ and $\mathbf{A}\mathbf{X}_n$ are asymptotically uncorrelated, $\mathbf{A}\mathbf{X}_n\mathbf{X}_n^{'}\mathbf{A}=\mathbf{A}\mathbf{C}\mathbf{A}+o_p(1)$, $\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n$ is approximately $\chi^2(\mathrm{tr}(\mathbf{C}\mathbf{A}))$ and $\mathrm{tr}(\mathbf{C}\mathbf{A})=L$.

[A4] Conditional Expectations of $\mathbf{e}_2$

We shall evaluate the terms of $\mathbf{e}_2$ and decompose $\mathbf{e}_2=\mathbf{e}_{2.1}+\mathbf{e}_{2.2}+\mathbf{e}_{2.3}$, where $\mathbf{e}_{2.i}$ $(i=1,2,3)$ correspond to each terms of (A.7). Because we can estimate $\mathbf{Q}$ and $\mathbf{C}$ consistently by using $\mathbf{Q}_n$ and $\mathbf{C}_n$, their estimations do not affect many terms involving $\mathbf{e}_2$ asymptotically. We consider $\mathbf{e}_{2.3}=-\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\mathbf{U}_n^{'}+\mathbf{X}_n\mathbf{q}^{'}]\mathbf{e}_1^{(0)}+o_p(1)$. Because $\tilde{\mathbf{e}}_0$ and $\mathbf{A}\mathbf{X}_n$ are asymptotically orthogonal, the conditional expectation, $\mathbf{e}_{2.3.1}=-\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n^{'}\mathbf{e}_1^{(0)}$, given $\tilde{\mathbf{e}}_0=\mathbf{x}$, is $\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{E}(\mathbf{w}_i\mathbf{w}_i^{'})\mathbf{x}+o_p(1)$. Because $\tilde{\mathbf{e}}_0=\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{X}_n$, the conditional expectation of the second term of $\mathbf{e}_{2.3}$ is

$$\mathbf{E}[\mathbf{e}_{2.3.2}|\mathbf{x}] = \mathbf{E}\Big[-(\tilde{\mathbf{e}}_0\mathbf{q}^{'})[\mathbf{Q}\mathbf{U}_n\mathbf{A}\mathbf{X}_n+(1-a)\mathbf{Q}\mathbf{q}\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n$$

$$+a\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\kappa_{3i}\mathbf{z}_i(\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n)^2-\tilde{\mathbf{e}}_0(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)-\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n\tilde{\mathbf{e}}_0]|\mathbf{x}\Bigg]$$

$$= -(1-a)L\mathbf{q}^{'}\mathbf{Q}\mathbf{q}\mathbf{x}+\mathbf{x}(\mathbf{q}^{'}\mathbf{x})^2-a\mathbf{x}\mathbf{q}^{'}\mathbf{Q}\mathbf{m}_3+o_p(1)\ ,$$

and then

$$\mathbf{E}[\mathbf{e}_{2.3}|\mathbf{x}]=\mathbf{Q}\mathbf{Q}^{*}\mathbf{Q}\mathbf{C}_2^{*}\mathbf{x}-(1-a)L\mathbf{tr}(\mathbf{C}_1^{*}\mathbf{Q})\mathbf{x}+\mathbf{x}\mathbf{x}^{'}\mathbf{C}_1^{*}\mathbf{x}-a\mathbf{x}\mathbf{q}^{'}\mathbf{Q}\mathbf{m}_3+o_p(1)\ . \quad (A.14)$$

Secondly, we evaluate $\mathbf{e}_{2.2}$, where $\mathbf{e}_{2.2.1}=-\mathbf{Q}[\mathbf{A}_{1n}]\mathbf{M}\mathbf{D}\mathbf{e}_1^{(0)}$, and $\mathbf{e}_{2.2.2}=-\mathbf{Q}[\mathbf{A}_{1n}][\mathbf{U}_n^{'}+\mathbf{X}_n\mathbf{q}^{'}]\tilde{\mathbf{e}}_0$ and $\mathbf{e}_{2.2}-\mathbf{e}_{2.2.1}-\mathbf{e}_{2.2.2}=o_p(1)$. The second term is rewritten as

$$\begin{aligned}\mathbf{e}_{2.2.2} &= -\Big\{2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}+\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}a\mathbf{\Theta}_{3n}^{\lambda_0}\mathbf{C}^{-1}\\ &\quad +\mathbf{Q}_n\mathbf{U}_n\mathbf{C}^{-1}+\mathbf{Q}\mathbf{q}(\mathbf{X}_n^{'}-a\boldsymbol{\lambda}_0^{'}\mathbf{C}_n)\mathbf{C}^{-1}\Big\}[\mathbf{U}_n^{'}+\mathbf{X}_n\mathbf{q}^{'}]\tilde{\mathbf{e}}_0+o_p(1)\end{aligned}$$

and its conditional expectation is

$$\begin{aligned}\mathbf{E}[\mathbf{e}_{2.2.2}|\mathbf{x}] &= -2(\mathbf{q}^{'}\mathbf{x})\mathbf{x}(\mathbf{q}^{'}\mathbf{x})-\mathbf{Q}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{C}_2^{*}\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{z}_i]\mathbf{x}-\mathbf{Q}\mathbf{q}\mathbf{E}[\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{X}_n|\mathbf{x}]\mathbf{q}^{'}\tilde{\mathbf{e}}_0\\ &\quad -a\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{n}\sum_{i=1}^{n}\kappa_{3i}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)](\mathbf{q}^{'}\mathbf{x})+a\mathbf{Q}\mathbf{q}\mathbf{E}[\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n|\mathbf{x}]\mathbf{q}^{'}\tilde{\mathbf{e}}_0+o_p(1)\ .\end{aligned}$$

By decomposing $\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{X}_n = \mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n+\tilde{\mathbf{e}}_0^{'}\mathbf{Q}^{-1}\tilde{\mathbf{e}}_0$ and using $\mathbf{E}[\mathbf{\Theta}_{3n}^{\lambda_0}\mathbf{C}^{-1}\mathbf{X}_n|\mathbf{x}] = an^{-1}\sum_{i=1}^{n}\kappa_{3i}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)+o_p(1)$, it is rewritten as

$$\begin{aligned}\mathbf{E}[\mathbf{e}_{2.2.2}|\mathbf{x}] &= -2\mathbf{x}\mathbf{x}^{'}\mathbf{C}_1^{*}\mathbf{x}-\mathbf{Q}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{C}_2^{*}\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{z}_i]\mathbf{x} \qquad (A.15)\\ &\quad -\mathbf{Q}\mathbf{C}_1^{*}\mathbf{x}[L\mathbf{I}_K+\mathbf{x}^{'}\mathbf{Q}^{-1}\mathbf{x}]+aL\mathbf{Q}\mathbf{C}_1^{*}\mathbf{x}-a\mathbf{Q}\mathbf{m}_3(\mathbf{q}^{'}\mathbf{x})+o_p(1)\ .\end{aligned}$$

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

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

We use the relations that $\mathbf{A}\mathbf{C}\mathbf{M}\mathbf{D}=\mathbf{O}$ and $\mathbf{E}[\mathbf{X}_n^{'}\mathbf{A}\tilde{\mathbf{e}}_0|\mathbf{x}]=\mathbf{O}_p(n^{-1/2})$,

$$\mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(a\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}|\mathbf{x}]=a\mathbf{E}\left[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{X}_n^{'}\mathbf{A}\mathbf{z}_i|\mathbf{x}\right]=o_p(1),$$

$$\mathbf{E}[\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\tilde{\mathbf{e}}_0|\mathbf{x}]=\mathbf{E}\left[\mathbf{X}_n^{'}\mathbf{A}\mathbf{M}\mathbf{D}\tilde{\mathbf{e}}_0+\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\tilde{\mathbf{e}}_0|\mathbf{x}\right]=\tilde{\mathbf{e}}_0^{'}\mathbf{Q}^{-1}\tilde{\mathbf{e}}_0.$$

Then, given $\tilde{\mathbf{e}}_0 = \mathbf{x}$, the conditional expectation $\mathbf{E}[\mathbf{e}_{2.2.1}|\mathbf{x}]$ is evaluated as

$$
\begin{aligned}
& \Big[-2(1-a)L(\mathbf{q}^{'}\mathbf{x})\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{q} - 2a(\mathbf{q}^{'}\mathbf{x})\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{m}_3 \\
& \qquad\qquad +2(\mathbf{q}^{'}\mathbf{x})^2\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{x}\Big] \\
& -a\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{E}[\mathbf{\Theta}_{3n}^{\lambda_0}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}|\mathbf{x}][(1-a)L\mathbf{Q}\mathbf{q} + a\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\kappa_{3i}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)] \\
& +\mathbf{Q}\frac{1}{n}\sum_{i=1}^{n}\mathbf{C}_2^{*}\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{z}_i\mathbf{x} \\
& -\mathbf{Q}\mathbf{q}\ \mathbf{E}\Big\{\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}[(1-a)\mathbf{Q}\mathbf{q}L + a\mathbf{Q}\mathbf{D}^{'}\mathbf{m}_3] - (\mathbf{q}^{'}\mathbf{e}_0)\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\tilde{\mathbf{e}}_0|\mathbf{x}\Big\} \\
= \ & 2(\mathbf{q}^{'}\mathbf{x})^2\mathbf{x} + \mathbf{Q}\frac{1}{n}\sum_{i=1}^{n}\mathbf{C}_2^{*}\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{z}_i\mathbf{x} \qquad \text{(A.16)} \\
& +\mathbf{Q}\mathbf{q}\mathbf{q}^{'}\mathbf{x}\mathbf{x}^{'}\mathbf{Q}^{-1}\mathbf{x} - 3(1-a)L\mathbf{Q}\mathbf{C}_1^{*}\mathbf{x} - 2a(\mathbf{q}^{'}\mathbf{x})\mathbf{Q}\mathbf{m}_3 - a\mathbf{Q}\mathbf{q}\mathbf{x}^{'}\mathbf{Q}\mathbf{m}_3 + o_p(1).
\end{aligned}
$$

Hence we have obtained the explicit form 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 $o_p(1)$. Next, we evaluate the terms involving $\mathbf{e}_{2.1}$, which is the first term of (A.7), and we need more complicated computations. We write $\mathbf{e}_{2.1}^{*} = \mathbf{e}_{2.1}(A) + \mathbf{e}_{2.1}(B) + \mathbf{e}_{2.1}(C) + \mathbf{e}_{2.1}(D)$, where $\mathbf{e}_{2.1}(A) = -\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{C}_n^{(2)}\mathbf{A}\mathbf{X}_n$, $\mathbf{e}_{2.1}(B) = \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{C}_n^{(1)}\mathbf{C}^{-1}\mathbf{C}_n^{(1)}\mathbf{A}\mathbf{X}_n$, $\mathbf{e}_{2.1}(C) = -\mathbf{Q}\mathbf{E}_n^{(1)}\mathbf{C}^{-1}\mathbf{C}_n^{(1)}\mathbf{A}\mathbf{X}_n$ and $\mathbf{e}_{2.1}(D) = \mathbf{Q}\mathbf{E}_n^{(2)}\mathbf{A}\mathbf{X}_n$. Because these terms depend on $p_i^{(2)}$ $(i = 1, \cdots, n)$ and $u_i(\hat{\theta}) = u_i - (1/\sqrt{n})[\mathbf{z}_i^{'}\mathbf{D} + (\mathbf{v}_{2i}^{'}, \mathbf{0}^{'})]\hat{\mathbf{e}}$, we need to use $\boldsymbol{\lambda}_1$ given by

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

Then by using $\mathbf{C}^{-1}\mathbf{X}_n = \mathbf{A}\mathbf{X}_n + \mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{X}_n$ and $2\mathbf{A}\mathbf{X}_n - \mathbf{C}^{-1}\mathbf{X}_n = \mathbf{A}\mathbf{X}_n - \mathbf{C}^{-1}\mathbf{M}\mathbf{D}\tilde{\mathbf{e}}_0$, we find

$$
\begin{aligned}
\boldsymbol{\lambda}_1 \ = \ & -\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{e}_1^{(0)} - \mathbf{C}^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{v}_{2i}^{'}, \mathbf{0}^{'})\mathbf{e}_0 - \mathbf{C}^{-1}\mathbf{C}_n^{(1)}\mathbf{A}\mathbf{X}_n + o_p(1) \\
= \ & -\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\Big[a\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{\Theta}_{3n}^{\lambda_0}\mathbf{A}\mathbf{X}_n + \mathbf{Q}\mathbf{U}_n\mathbf{A}\mathbf{X}_n + (1-a)\mathbf{Q}\mathbf{q}\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n - (\mathbf{q}^{'}\mathbf{x})\mathbf{x} \\
& \qquad -\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{U}_n^{'}\mathbf{x}\Big] - \mathbf{C}^{-1}\Big[\mathbf{U}_n^{'}\mathbf{x} + \mathbf{X}_n\mathbf{q}^{'}\Big] \\
& -\mathbf{C}^{-1}\Big[-2\mathbf{C}_n\mathbf{q}^{'}\mathbf{x}\mathbf{A}\mathbf{X}_n + \mathbf{\Theta}_{3n}^{\lambda_0}\mathbf{A}\mathbf{X}_n\Big] + o_p(1) \\
= \ & -\mathbf{A}\mathbf{U}_n^{'}\tilde{\mathbf{e}}_0 - \mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{U}_n\mathbf{A}\mathbf{X}_n + (\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{A}\mathbf{X}_n \\
& -(1-a)\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}\mathbf{q}\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n + a\mathbf{A}\mathbf{\Theta}_{3n}^{\lambda_0}\mathbf{A}\mathbf{X}_n + o_p(1)\ . \qquad \text{(A.17)}
\end{aligned}
$$

Although we could have used $\boldsymbol{\lambda}_1$ with $a = 1$, we used (A.17) in order to make no confusion. For the GMM estimator we could have set $\boldsymbol{\lambda}_1 = \mathbf{0}$ and $p_i^{(j)} = 0$ $(j = 1, 2)$, but then we need different notations. Then we can evaluate each terms by using $\mathbf{e}_1$ and $\boldsymbol{\lambda}_1$. By using the stochastic expansion of $p_i^{(1)}$ $(i=1,\cdots,n)$,

$$\mathbf{e}_{2.1}(A) = -\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\Big\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\Big[(\mathbf{z}_i^{'}\mathbf{D}\tilde{\mathbf{e}}_0+\mathbf{w}_i^{'}\tilde{\mathbf{e}}_0+u_i\mathbf{q}^{'}\tilde{\mathbf{e}}_0)^2$$

$$-2(u_i\mathbf{z}_i^{'}\mathbf{D}\mathbf{e}_1+u_i\mathbf{w}_i^{'}\mathbf{e}_1+u_i^2\mathbf{q}^{'}\mathbf{e}_1)+2u_i^2(a\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)(\mathbf{z}_i^{'}\mathbf{D}\tilde{\mathbf{e}}_0+\mathbf{w}_i^{'}\tilde{\mathbf{e}}_0+u_i\mathbf{q}^{'}\tilde{\mathbf{e}}_0)+u_i^2p_i^{(2)}\Big]\Big\}\mathbf{A}\mathbf{X}_n.$$

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

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

Here we illustrate our arguments and for the last term, we write

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

whose each terms are of $O_p(n^{-1/2})$. By taking the conditional expectations applying *Lemma A.3* in **[A5]**, the first term in the above equation is of $O_p(n^{-1/2})$. Also

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

Then by ignoring the terms of $o_p(1)$,

$$\mathbf{E}[\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}n^{-1}\sum_{i=1}^{n}\mathbf{z}_iu_i^2\mathbf{w}_i^{'}\tilde{\mathbf{e}}_0(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2|\mathbf{x}]=n^{-1}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{E}[u_i^2\mathbf{w}_i^{'}]\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i\mathbf{x}+o_p(1)\;,$$

$$\mathbf{E}[n^{-1}\sum_{i=1}^{n}\mathbf{z}_iu_i^3\mathbf{q}^{'}\tilde{\mathbf{e}}_0(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2|\mathbf{x}]=n^{-1}\sum_{i=1}^{n}\kappa_{3i}\mathbf{z}_i\mathbf{q}^{'}\mathbf{x}\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i+o_p(1)\;.$$

Then for the conditional expectations with an arbitrary a we find

$$\mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i u_i^2 p_i^{(2)}|\mathbf{x}] = a\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i[\mathbf{E}(u_i^2\mathbf{w}_i^{'})\mathbf{x}+\sigma_i^2\mathbf{z}_i^{'}\mathbf{D}\mathbf{x}]+o_p(1)\ .$$

When $\kappa_{3i}=0$ and $\mathbf{E}(u_i^2\mathbf{w}_i)=\mathbf{0}$, by gathering the conditional expectations of other terms, we have

$$\mathbf{E}[\mathbf{e}_{2.1}(A)|\mathbf{x}] = -3a\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\sigma_i^2(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}]\mathbf{D}\mathbf{x}+o_p(1)\ . \tag{A.18}$$

Similarly, the second term $\mathbf{e}_{2.1}(B)$ is $\mathbf{e}_{2.1}(B)=\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{C}_n^{(1,0)}\mathbf{C}^{-1}\mathbf{C}_n^{(1,0)}\mathbf{A}\mathbf{X}_n+o_p(1)$. Since $\mathbf{A}\mathbf{X}_n$ and $\tilde{\mathbf{e}}_0$ are asymptotically uncorrelated, the conditional expectation is reduced to

$$\mathbf{E}[\mathbf{e}_{2.1}(B)|\mathbf{x}] = 4a(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{Q}\mathbf{m}_3+\ o_p(1)\ , \tag{A.19}$$

which is $o_p(1)$. For the third term, we write $\mathbf{e}_{2.1}(C)=-\mathbf{Q}\mathbf{E}_n^{(1)}\mathbf{C}^{-1}\mathbf{C}_n^{(1)}\mathbf{A}\mathbf{X}_n$.
By using $\mathbf{Q}\mathbf{E}_n^{(1)}\mathbf{C}^{-1}=[\mathbf{Q}\mathbf{U}_n+\mathbf{Q}\mathbf{q}(\mathbf{X}_n^{'}-a\boldsymbol{\lambda}_0^{'}\mathbf{C}_n)]\mathbf{C}^{-1}+o_p(1)$, $\mathbf{C}_n^{(1)}\mathbf{A}\mathbf{X}_n=[-2(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)\mathbf{C}_n-an^{-1}\sum_{i=1}^{n}\kappa_{3i}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)]\mathbf{A}\mathbf{X}_n+o_p(1)$ and $\mathbf{X}_n^{'}-a\boldsymbol{\lambda}^{'}\mathbf{C}=\tilde{\mathbf{e}}_0^{'}\mathbf{D}^{'}\mathbf{M}+(1-a)\mathbf{X}_n^{'}\mathbf{A}\mathbf{C}$, the conditional expectation of $\mathbf{e}_{2.1}(C)$ is

$$\begin{aligned}\mathbf{E}[\mathbf{e}_{2.1}(C)|\mathbf{x}] &= 2\mathbf{Q}\mathbf{q}(\mathbf{q}^{'}\mathbf{x})(1-a)\mathbf{E}[\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n|\mathbf{x}]+\mathbf{Q}\mathbf{q}(\mathbf{X}_n^{'}-a\boldsymbol{\lambda}_0^{'}\mathbf{C})a\boldsymbol{\Theta}_{3n}^{\lambda_0}\mathbf{A}\mathbf{X}_n+o_p(1)\\ &= 2(1-a)L\mathbf{Q}\mathbf{q}\mathbf{q}^{'}\mathbf{x}+\mathbf{Q}\mathbf{q}\mathbf{e}_0^{'}\mathbf{m}_3+o_p(1)\ .\end{aligned} \tag{A.20}$$

The fourth term $\mathbf{e}_{2.1}(D)$ with an arbitrary a is
$\mathbf{e}_{2.1}(D)\ =\ a\mathbf{Q}\left\{\mathbf{D}^{'}n^{-1}\sum_{i=1}^{n}p_i^{(2)}\mathbf{z}_i\mathbf{z}_i^{'}+n^{-1}\sum_{i=1}^{n}p_i^{(2)}(\mathbf{v}_{2i}^{'},\mathbf{0}^{'})^{'}\mathbf{z}_i^{'}\right\}\mathbf{A}\mathbf{X}_n$. Since the first term of $\mathbf{e}_{2.1}(D)$ is similar to the last term of $\mathbf{e}_{2.1}(A)$, its conditional expectation with an arbitrary a is

$$\begin{aligned}&\mathbf{E}\left[\mathbf{Q}\mathbf{D}^{'}\frac{1}{n}\sum_{i=1}^{n}p_i^{(2)}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n|\mathbf{x}\right]\\ =\ &\mathbf{Q}\mathbf{D}^{'}\mathbf{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)(a\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)(\mathbf{z}_i^{'}\mathbf{D}\mathbf{e}_0+\mathbf{w}_i^{'}\mathbf{e}_0+u_i\mathbf{q}^{'}\mathbf{e}_0)\right.\\ &\quad\left.-a\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)\boldsymbol{\lambda}_1^{'}\mathbf{z}_iu_i+\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)(a\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)^2u_i^2|\mathbf{x}\right\}\\ =\ &\mathbf{E}\Big\{a\mathbf{Q}\mathbf{D}^{'}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}]\mathbf{D}\mathbf{x}+a\mathbf{Q}\mathbf{D}^{'}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2\mathbf{A}\mathbf{X}_n|\mathbf{x}\Big\}+o_p(1)\\ =\ &a\mathbf{Q}\mathbf{D}^{'}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}]\mathbf{D}\mathbf{x}\ +\ o_p(1)\ .\end{aligned}$$

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

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

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

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

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

$$\begin{aligned}\mathbf{E}[\mathbf{e}_{2.1.1}(D)|\mathbf{x}] &= a\mathbf{E}\left[\mathbf{Q}\frac{1}{n}\sum_{i=1}^{n}(\mathbf{w}_i\mathbf{w}_i^{'}+\mathbf{q}\mathbf{q}^{'}u_i^2)\tilde{\mathbf{e}}_0(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2|\mathbf{x}\right]\\ &= a\mathbf{Q}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{C}_2^*\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i]\mathbf{x}+aL\mathbf{Q}\mathbf{C}_1^*\mathbf{x}+o_p(1)\end{aligned}$$

by using that $a\mathbf{E}[\boldsymbol{\lambda}_0^{'}n^{-1}\sum_{i=1}^{n}u_i^2\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n|\mathbf{x}]\cong a\mathbf{E}[\mathbf{X}_n^{'}\mathbf{A}\mathbf{C}\mathbf{A}\mathbf{X}_n]=aL+o_p(1)$. Also

$$\begin{aligned}\mathbf{E}[\mathbf{e}_{2.1.3}(D)|\mathbf{x}] &= a^2\mathbf{E}\left[\mathbf{Q}\frac{1}{n}\sum_{i=1}^{n}\mathbf{q}u_i^3(\mathbf{z}_i^{'}\boldsymbol{\lambda}_0)^2\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n|\mathbf{x}\right]\\ &= a^2\mathbf{Q}\mathbf{q}\mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\kappa_{3i}(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^3|\mathbf{x}]+o_p(1)\ .\end{aligned}$$

But since $\mathbf{A}\mathbf{X}_n$ is asymptotically normal and uncorrelated with $\tilde{\mathbf{e}}_0$, $\mathbf{E}[\mathbf{e}_{2.1.3}(D)|\mathbf{x}]=o_p(1)$. For the conditional expectation of $\mathbf{e}_{2.1.2}(D)$, we use that the pairs of vectors $(\mathbf{w}_i^{'},u_i)$ are uncorrelated and $n^{-1}\sum_{i=1}^{n}\mathbf{w}_iu_iz_i^{(j)}z_i^{(k)}\xrightarrow{p}\mathbf{0}$. As for the remaining conditional expectation terms, by using $\boldsymbol{\lambda}_1$ we find

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

Hence we summarize

$$\begin{aligned}\mathbf{E}[\mathbf{e}_{2.1}(D)|\mathbf{x}] &= a\mathbf{Q}\mathbf{D}^{'}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}]\mathbf{D}\mathbf{x} \\ &\quad +a\mathbf{Q}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{C}_2^*\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i]\mathbf{x}+aL\mathbf{Q}\mathbf{C}_1^*\mathbf{x}-aL\mathbf{Q}\mathbf{q}\mathbf{q}^{'}\mathbf{x}+o_p(1)\ .\end{aligned}\tag{A.21}$$

Finally, we obtain $\mathbf{E}[\mathbf{e}_{2.1}|\tilde{\mathbf{e}}_0=\mathbf{x}]$ by collecting $\mathbf{E}[\mathbf{e}_{2.1}(A)|\mathbf{x}]$, $\mathbf{E}[\mathbf{e}_{2.1}(B)|\mathbf{x}]$, $\mathbf{E}[\mathbf{e}_{2.1}(C)|\mathbf{x}]$ and $\mathbf{E}[\mathbf{e}_{2.1}(D)|\mathbf{x}]$. The resulting formulas become relatively simple since we can ignore the third order moments and then many terms disappear in the formulas eventually.

[A5] Conditional Expectation Formulas

We prepare useful formulas on the conditional expectations and the proofs will be given in Appendix C. They are used repeatedly in our evaluations by setting $\mathbf{Z}=\tilde{e}_0$.

Lemma A.1 : Let the vectors $\tilde{\mathbf{e}}_0$,$\mathbf{X}_n=(x_l^{(n)})$, and $\mathbf{Y}_n=(y_{kl}^{(n)})$ be defined as in Section 4. Then $\mathbf{E}[y_{jk}^{(n)}|x_l^{(n)},\mathbf{z}_i]=n^{-1}\sum_{i=1}^{n}z_{ik}z_{ik}\kappa_{3i}z_{il}^{(n)}x_l^{(n)}/var(x_l^{(n)})+o_p(n^{-1/2})$ and

$$\mathbf{E}[\mathbf{Y}_n\mathbf{A}\mathbf{X}_n|\tilde{e}_0=\mathbf{x}]=\operatorname*{plim}_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\kappa_{3i}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)+O_p(n^{-1/2})\ .\tag{A.22}$$

Lemma A.2 : Let a set of vectors $\mathbf{X} = (X_i)$ and $\mathbf{T} = (t_i)$ be normally distributed. Then

$$\begin{aligned} & \mathbf{E}[X_iX_jX_k|\mathbf{T}] - \mathbf{E}(X_i|\mathbf{T})\mathbf{E}(X_j|\mathbf{T})\mathbf{E}(X_k|\mathbf{T}) \qquad (A.23) \\ = \; & Cov(X_i, X_j|\mathbf{T})\mathbf{E}(X_k|\mathbf{T}) + Cov(X_j, X_k|\mathbf{T})\mathbf{E}(X_i|\mathbf{T}) \\ & +Cov(X_k, X_i|\mathbf{T})\mathbf{E}(X_j|\mathbf{T}) \end{aligned}$$

and

$$\begin{aligned} & \mathbf{E}[X_iX_jX_kX_l|\mathbf{T}] - \mathbf{E}(X_i|\mathbf{T})\mathbf{E}(X_j|\mathbf{T})\mathbf{E}(X_k|\mathbf{T})\mathbf{E}(X_l|\mathbf{T}) \qquad (A.24) \\ = \; & Cov(X_i, X_j|\mathbf{T})Cov(X_k, X_l|\mathbf{T}) + Cov(X_i, X_k|\mathbf{T})Cov(X_j, X_l|\mathbf{T}) \\ & +Cov(X_i, X_l|\mathbf{T})Cov(X_j, X_k|\mathbf{T}) \\ + \; & Cov(X_i, X_j|\mathbf{T})\mathbf{E}(X_k|\mathbf{T})\mathbf{E}(X_l|\mathbf{T}) + Cov(X_i, X_k|\mathbf{T})\mathbf{E}(X_j|\mathbf{T})\mathbf{E}(X_l|\mathbf{T}) \\ & +Cov(X_i, X_l|\mathbf{T})\mathbf{E}(X_j|\mathbf{T})\mathbf{E}(X_k|\mathbf{T}) + Cov(X_j, X_k|\mathbf{T})\mathbf{E}(X_i|\mathbf{T})\mathbf{E}(X_l|\mathbf{T}) \\ & +Cov(X_j, X_l|\mathbf{T})\mathbf{E}(X_i|\mathbf{T})\mathbf{E}(X_k|\mathbf{T}) + Cov(X_k, X_l|\mathbf{T})\mathbf{E}(X_i|\mathbf{T})\mathbf{E}(X_j|\mathbf{T}) \, . \end{aligned}$$

Lemma A.3 : Let $\mathbf{u}_n = (u_i)$ and v_n be $p \times 1$ vector and a scalor with $\mathbf{E}(u_i) = 0, \mathbf{E}(v_n) = 0, \mathbf{E}(u_iu_j) = \delta(i,j), \mathbf{E}(v_n^2) = 1$ and they have finite fourth order moments. Assume that they are sums of i.i.d. (non-lattice) vectors and asymptotically normally distributed and admit the asymptotic expansion of their distribution function up to $O_p(n^{-1})$. Then

$$\begin{aligned} & \mathbf{E}[v_n|\mathbf{u}_n] = \boldsymbol{\rho}^{'}\mathbf{u}_n \qquad (A.25) \\ & +\frac{1}{6\sqrt{n}}\left\{3\sum_{l_1,l_2=1}^{p}\beta_{l_1l_2v}h_2(u_{l_1},u_{l_2}) - 3\sum_{l_2,l_3=1}^{p}[\sum_{l_1=1}^{p}\beta_{l_1l_2l_3}\rho_{l_1}]h_2(u_{l_2},u_{l_3})\right\} + O_p(n^{-1}) \, , \end{aligned}$$

where $\beta_{l_1l_2v} = \mathbf{E}(u_{l_1}u_{l_2}v_n)$, $\beta_{l_1l_2l_3} = \mathbf{E}(u_{l_1}u_{l_2}u_{l_3})$, $h_2(u_{l_1},u_{l_2}) = u_{l_1}u_{l_2} - \delta(l_1,l_2)$ ($\delta(l_1,l_2) = 1$ if $l_1 = l_2$ and $\delta(l_1,l_2) = 0$ if $l_1 \neq l_2$), and $\boldsymbol{\rho} = Cov(v, \mathbf{u}_n)$.
In particular, if $\mathbf{E}(u_iu_ju_k) = 0$ ($i \neq j \neq k$), then $\beta_{l_1l_2l_3} = 0$.

[A6] Higher Order Effects of $\mathbf{e}_0$ and $\mathbf{e}_1$

We need to evaluate the higher order effects of additional terms from $\mathbf{e}_0^{(1)}, \mathbf{e}_0^{(2)}$ and $\mathbf{e}_1^{(1)}$ up to $O_p(n^{-1})$. By applying a version of *Lemma A.3* to $\mathbf{e}_0^{(1)}$ and use $\mathbf{E}[y_{jk}^{(n)}\mathbf{X}_n\mathbf{C}^{-1}\mathbf{X}_n|\mathbf{X}_n] = n^{-1}\sum_{i=1}^{n}\kappa_{3i}z_i^{(j)}z_i^{(k)}\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{X}_n$ with $\mathbf{Y}_n = (y_{jk}^{(n)})$. By conditioning with respect to $\mathbf{X}_n$ and using $\mathbf{C}^{-1} = \mathbf{A} + \mathbf{C}^{-1}\mathbf{MDQD}^{'}\mathbf{MC}^{-1}$, the conditional expectation of $\mathbf{e}_0^{(1)}$ is

$$\begin{aligned} \mathbf{E}[\mathbf{e}_0^{(1)}|\mathbf{X}_n] \;=\; & -\mathbf{QD}^{'}\mathbf{MC}^{-1}[\frac{1}{n}\sum_{i=1}^{n}\kappa_{3i}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{AX}_n\mathbf{X}_n^{'}\mathbf{C}^{-1}\mathbf{z}_i] \\ & + \frac{1}{6\sqrt{n}}\{-3\mathbf{QD}^{'}\mathbf{MC}^{-1}[\frac{1}{n}\sum_{i=1}^{n}(\mathbf{E}(u_i^4-(\sigma_i^2)^2)\sum_{l_1,l_2}z_{il_1}z_{il_2}h_2(x_{l_1},x_{l_2})]\mathbf{AX}_n \end{aligned}$$

$$+3\mathbf{QD}^{'}\mathbf{MC}^{-1}[\frac{1}{n}\sum_{i=1}^{n}(\kappa_{3i})^2\mathbf{z}_i\mathbf{z}_i^{'}\sum_{l_1,l_2}z_{il_1}z_{il_2}h_2(x_{l_1},x_{l_2})]\mathbf{AX}_n\}+o_p(n^{-1/2}).$$

Hence under *Assumption III*

$$\sum_{l_1,l_2}\mathbf{E}[y_{jk}^{(n)}t_{l_1}t_{l_2}h_2(l_1,l_2)|\mathbf{X}_n]=(\frac{1}{\sqrt{n}})^3\sum_{i=1}^{n}\mathbf{E}(u_i^4-\sigma_i^4)z_i^{(j)}z_i^{(k)}\mathbf{z}_i^{'}(\mathbf{C}^{-1}\mathbf{X}_n\mathbf{X}_n^{'}-\mathbf{C}^{-1}]\mathbf{z}_i,$$

and

$$\begin{aligned}
&\mathbf{E}\left[\mathbf{z}_i^{'}(\mathbf{C}^{-1}\mathbf{X}_n\mathbf{X}_n^{'}-\mathbf{C}^{-1})\mathbf{z}_i\mathbf{AX}_n|\tilde{\mathbf{e}}_0\right]\\
=\;&\mathbf{E}\left[(\mathbf{z}_i^{'}\mathbf{AX}_n+\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{MD}\tilde{\mathbf{e}}_0)(\mathbf{X}_n^{'}\mathbf{Az}_i+\tilde{\mathbf{e}}_0^{'}\mathbf{D}^{'}\mathbf{MC}^{-1}\mathbf{z}_i)\mathbf{AX}_n-\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{z}_i\mathbf{AX}_n|\tilde{\mathbf{e}}_0\right]\\
=\;&2\tilde{\mathbf{e}}_0^{'}\mathbf{D}^{'}\mathbf{MC}^{-1}\mathbf{z}_i\mathbf{E}[\mathbf{AX}_n\mathbf{X}_n^{'}\mathbf{Az}_i]\;.
\end{aligned}$$

Then given $\tilde{\mathbf{e}}_0=\mathbf{x}$

$$\begin{aligned}
\mathbf{E}[\mathbf{e}_0^{(1)}|\mathbf{x}]=-\mathbf{Qm}_3+\frac{1}{6\sqrt{n}}\Big\{&-3\mathbf{QD}^{'}\mathbf{MC}^{-1}[\frac{1}{n}\sum_{i=1}^{n}(\kappa_{4i}-\sigma_i^4)2\mathbf{z}_i\mathbf{z}_i{'}\mathbf{C}^{-1}\mathbf{MD}\tilde{\mathbf{e}}_0\mathbf{z}_i^{'}\mathbf{Az}_i]\\
&+3\mathbf{QD}^{'}\mathbf{MC}^{-1}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\kappa_{3i}^2(2\mathbf{z}_i^{'}\mathbf{C}^{-1}\mathbf{MD}\tilde{\mathbf{e}}_0\mathbf{z}_i^{'}\mathbf{Az}_i)]\Big\}+o_p(n^{-1/2})\;.
\end{aligned}$$

Hence we summarize

$$\mathbf{E}[\mathbf{e}_1^{(0)}+\mathbf{e}_0^{(1)}|\mathbf{x}]=(1-a)\mathbf{Q}\left[\frac{1}{n}\sum_{i=1}^{n}\mathbf{q}_i(\mathbf{z}_i^{'}\mathbf{Az}_i)\sigma_i^2-\mathbf{m}_3\right]-\mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{C}_n\mathbf{q}^{'}\mathbf{x}\mathbf{C}^{-1}\mathbf{MDx}\;, \tag{A.26}$$

Next we evaluate the conditional expectation of $\mathbf{e}_1^{(1)}=\mathbf{e}_{1.1}^{(1)}+\mathbf{e}_{1.2}^{(1)}+\mathbf{e}_{1.3}^{(1)}$. This term plays an important role in $O_p(n^{-1})$. 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 by using *Lemma A.1* and *Lemma A.3*,

$$\begin{aligned}
\mathbf{E}[\mathbf{e}_{1.2}^{(1)}|\mathbf{x}]&=\mathbf{E}[\mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{Y}_n^{'}\mathbf{AX}_n(\mathbf{q}^{'}\tilde{\mathbf{e}}_0)+\tilde{\mathbf{e}}_0\mathbf{q}^{'}\mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{Y}_n^{'}\mathbf{AX}_n|\mathbf{x}]\\
&=(\mathbf{q}^{'}\mathbf{x})\mathbf{Qm}_3+\mathbf{xq}^{'}\mathbf{Qm}_3+o_p(1)
\end{aligned}$$

and

$$\mathbf{E}[\mathbf{e}_{1.3}^{(1)}|\mathbf{x}]=\mathbf{QD}^{'}\mathbf{MC}^{-1}\mathbf{E}[\mathbf{U}_n^{'}|\mathbf{x}]\mathbf{Qm}_3+\mathbf{QD}^{'}\mathbf{MC}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{Az}_i\mathbf{E}(u_i^2\mathbf{w}_i^{'})\mathbf{x}\;, \tag{A.27}$$

which are both of $o_p(1)$. Then we evaluate the conditional expectation of $\mathbf{e}_{1.1}^{(1)}$ associated with $\mathbf{C}_n^{(1)}$ and $\mathbf{E}_n^{(1)}$ have been cancelled out. We also evaluate remaining terms of $O_p(n^{-1/2})$ and the conditional expectation of the first two lines of (A.11) are

$$\begin{aligned}
&\mathbf{E}\left\{a\mathbf{QD}^{'}\mathbf{MC}^{-1}\frac{1}{n}\sum_{i=1}^{n}\kappa_{3i}\mathbf{z}_i\mathbf{z}_i^{'}\mathbf{AX}_n\mathbf{z}_i^{'}(\mathbf{C}^{-1}-\mathbf{A})\mathbf{E}(\mathbf{Y}_n|\mathbf{X}_n)\mathbf{AX}_n|\mathbf{x}\right\}\\
+\;&\mathbf{E}\left\{\mathbf{QU}_n(\mathbf{C}^{-1}-\mathbf{A})\mathbf{E}(\mathbf{Y}_n|\mathbf{X}_n)\mathbf{AX}_n+\mathbf{Qq}\tilde{\mathbf{e}}_0^{'}\mathbf{D}^{'}\mathbf{MC}^{-1}\mathbf{E}(\mathbf{Y}_n|\mathbf{X}_n)\mathbf{AX}_n|\mathbf{x}\right\}\;,
\end{aligned}$$

which are of $o_p(1)$ under *Assumption III*. Similarly, the terms in the third line of (A.11) leads to

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

which are of $o_p(1)$. The important terms of $O_p(n^{-1/2})$ are two terms appeared in the 4th and 7th lines of (A.11), which are dependent on the fourth order moments of $\{u_i\}$, which are

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

up to $O_p(n^{-1/2})$. It is straightforward to obtain the conditional expectation of the first term as

$$
2\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_iu_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\mathbf{D}\mathbf{x})\mathbf{A}\mathbf{z}_i\}+O_p(n^{-1/2})\ , \tag{A.28}
$$

but some careful evaluation is needed for the second term. (Under *Assumption III* the fourth order cumulant is $\kappa=[E(u_i^4)-3\sigma^4)]/\sigma^4$.) We use two steps and as the first step we take the conditional expectation, given $\mathbf{X}_n=\mathbf{r}$, $\mathbf{E}[n^{-1/2}\sum_{i=1}^{n}r_i(u_i^3-\kappa_{3i}-u_i\sigma_i^2)|\mathbf{r}]=n^{-1}\sum_{i=1}^{n}r_i\mathbf{E}(u_i^4-u_i^2\sigma_i^2)\mathbf{z}_i(\mathbf{E}(\mathbf{X}_n\mathbf{X}_n))^{-1}\mathbf{r}+o_p(1)$ (r_i are functions of $\mathbf{z}_i$). Then as the second step we take the conditional expectation given $\tilde{\mathbf{e}}_0=\mathbf{x}$ by using the decomposition $\mathbf{C}^{-1}=\mathbf{A}+\mathbf{C}^{-1}M\mathbf{D}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}$ and the asymptotic normality of the corresponding random variables. Then the conditional expectation of the second term can be evaluated as

$$
\begin{aligned}
& a\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{E}[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{X}_n)^2(u_i^3-\kappa_{3i}-\sigma_i^2u_i)|\mathbf{x}] \\
= \quad & a\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{E}[u_i^4-\sigma_i^2u_i^2]\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{x}+O_p(n^{-1/2}) \\
= \quad & a(2+\kappa)\mathbf{Q}\mathbf{D}^{'}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^{'}(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)]\mathbf{D}\mathbf{x}+O_p(n^{-1/2})\ .
\end{aligned} \tag{A.29}
$$

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

$$
\begin{aligned}
& \mathbf{E}\{-\mathbf{Q}[\mathbf{U}_n\mathbf{C}^{-1}+\mathbf{q}(\mathbf{X}_n^{'}\mathbf{C}^{-1}-a\mathbf{X}_n^{'}\mathbf{A})]\mathbf{E}(\mathbf{Y}_n|\mathbf{X}_n)\mathbf{A}\mathbf{X}_n|\mathbf{x}\} \\
+ \quad & \mathbf{E}\{\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q}[\mathbf{U}_n\mathbf{A}\mathbf{X}_n+(1-a)\mathbf{q}\mathbf{X}_n^{'}\mathbf{A}\mathbf{X}_n|\mathbf{x}\} \\
+ \quad & \mathbf{E}\{-2\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{q}^{'}\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n\mathbf{A}\mathbf{X}_n-a\mathbf{Q}\mathbf{q}\mathbf{X}_n^{'}\mathbf{A}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n|\mathbf{x}\},
\end{aligned}
$$

which are of $o_p(1)$.

Since we can ignore the effects of the third order moments of disturbances under *Assumption III*, many terms with third order moments disappear and we only have the above two terms involving $\mathbf{e}_1^{(1)}$. Thus the conditional expectation of $\mathbf{e}_1^{(1)}$ is rewritten as

$$\mathbf{E}[\mathbf{e}_1^{(1)}|\mathbf{x}] = [2 + a(2+\kappa)]\mathbf{Q}\mathbf{D}^{'}\mathbf{F}\mathbf{D}\mathbf{x} + o_p(1), \tag{A.30}$$

where $\mathbf{F} = \text{plim}_{n\to\infty} n^{-1}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}$. Similarly, under *Assumption III*, the conditional expectation of $\mathbf{e}_0^{(1)}$ is in $O_p(n^{-1/2})$

$$\mathbf{E}[\mathbf{e}_0^{(1)}|\mathbf{x}] = \frac{1}{\sqrt{n}}\{-(2+\kappa)\mathbf{Q}\mathbf{D}^{'}[\frac{1}{n}\sum_{i=1}^{n}\mathbf{z}_i(\mathbf{z}_i^{'}\mathbf{A}\mathbf{z}_i)\mathbf{z}_i^{'}]\mathbf{D}\mathbf{x}\} + o_p(n^{-1/2})\ . \tag{A.31}$$

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

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

which has been simplified under *Assumption III.* Then

$$\mathbf{E}[\mathbf{e}_0^{(2)}|\mathbf{x}] = \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{E}[\mathbf{Y}_n\mathbf{A}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n|\mathbf{x}] = O_p(n^{-1/2}) \tag{A.32}$$

because each components of $\mathbf{Y}_n$ and $\mathbf{X}_n$ are asymptotically normally distributed, the vector $\mathbf{A}\mathbf{X}_n$ is asymptotically uncorrelated with $\tilde{\mathbf{e}}_0$. We also

$$\begin{aligned}
\mathbf{E}[\mathbf{e}_0^{(1)}\mathbf{e}_0^{(1)'}|\mathbf{x}] &= \mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{E}\left\{\mathbf{E}[\mathbf{Y}_n\mathbf{A}\mathbf{X}_n\mathbf{X}_n^{'}\mathbf{A}\mathbf{Y}_n|\mathbf{X}_n]|\mathbf{x}\right\}\mathbf{C}^{-1}\mathbf{M}\mathbf{D}\mathbf{Q} \\
&= (2+\kappa)\mathbf{Q}\mathbf{D}^{'}\mathbf{F}\mathbf{D}\mathbf{Q} + O_p(n^{-1/2})
\end{aligned} \tag{A.33}$$

because $\mathbf{A}\mathbf{X}_n\mathbf{X}_n^{'}\mathbf{A} = \mathbf{A}\mathbf{C}_n\mathbf{A} + o_p(1) = \mathbf{A} + o_p(1)$. For $\mathbf{U}_n = (u_{jk}) = n^{-1/2}\sum_{i=1}^{n}\mathbf{w}_i\mathbf{z}_i^{'}$, we apply *Lemma A.1* and *Lemma A.3* and use the fact that $Cov(u_{jk}, \tilde{\mathbf{e}}_0) = \mathbf{0}$,

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

Because $\mathbf{e}_1^{(0)} = -\mathbf{Q}\mathbf{D}^{'}\mathbf{M}\mathbf{C}^{-1}\mathbf{Y}_n\mathbf{A}\mathbf{X}_n$, we find

$$\mathbf{E}[\mathbf{e}_0^{(1)}\mathbf{e}_1^{(0)'}|\mathbf{x}] = o_p(1) \tag{A.34}$$

after lengthy, but straightforward calculations of each terms in the left hand side under *Assumption III.*

Appendix B : Derivations of Theorem 3.1 and Theorem 3.2

In the univariate and homoscedastic case ($G_1 = 1, p = 1 + K_1$) we use the notation $\mathbf{Q}^{-1} = \sigma^{-2}\mathbf{D}'\mathbf{M}\mathbf{D}$ and $Q_{11} = \sigma^2(\mathbf{\Pi}_{22}'\mathbf{M}_{22.1}\mathbf{\Pi}_{22})^{-1}$ as the $(1,1)-$element of $\mathbf{Q}$. The right-hand side of $\phi^*(x)$ for the standardized estimator in (4.9) can be simplified and it is given by

$$\phi(x)\{1+\frac{1}{\sqrt{n}}[\beta_3(x^3-3x)]+\frac{1}{n}[\frac{\beta_4}{24}(x^4-6x^2+3)+\frac{\beta_3^2}{72}(x^6-15x^4+45x^2-15)]\}, \quad \text{(A.35)}$$

where $\beta_3 = \beta_{111}$ and $\beta_4 = \beta_{1111} - 3\beta_{11}^2$ are the third and fourth order cumulants in (4.9) 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. Under the normal disturbances $\beta_3 = \beta_4 = 0$.

We partition the $p-$dimensional ($p = 1 + K_1$) normal vector $\mathbf{x} = (x_1, \mathbf{x}_2')' \sim N_p(\mathbf{0}, \mathbf{Q})$ and

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \mathbf{x}_2 \end{pmatrix} = \mathbf{Q}\begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} Q_{11}^{-1}x_1 + \begin{pmatrix} 0 \\ \mathbf{x}_2 - \mathbf{Q}_{21}Q_{11}^{-1}x_1 \end{pmatrix}, \quad \text{(A.36)}$$

where two vectors on the right-hand side are independent under the normality. By using the notation of Section 3, we find the relations $1 + \alpha^2 = \sigma^2\omega_{22}/|\mathbf{\Omega}|$, $(1, \mathbf{0}')\mathbf{q} = [\omega_{21} - \omega_{22}\beta]/\sigma^2 = [-|\mathbf{\Omega}|^{1/2}/\sigma^2]\alpha$, $Q_{11}^{-1/2}(1/\sqrt{n})(-1)\left[-\alpha|\mathbf{\Omega}|^{1/2}/\sigma^2\right] = (1/\mu)[\alpha Q_{11}]$ and $\mu^2/n = [(1+\alpha^2)/\omega_{22}]\mathbf{\Pi}_{22}'\mathbf{A}_{22.1}\mathbf{\Pi}_{22}/n = [\sigma^2/|\mathbf{\Omega}|][\mathbf{\Pi}_{22}'\mathbf{A}_{22.1}\mathbf{\Pi}_{22}]/n$.
Now we set $e_1(z) = \mathbf{E}\left[[\mathbf{e}_1^*(x)]_1|z\right]$, $e_2(z) = \mathbf{E}\left[[\mathbf{e}_2^*(x)]_1|z\right]$, $e_{11}(z) = \mathbf{E}\left[[\mathbf{e}_{11}^*(x)]_{11}|z\right]$ and $z = Q_{11}^{-1/2}x_1$. Then since $m_3 = 0$ under the normality,

$$Q_{11}^{-1/2}\frac{1}{\sqrt{n}}[e_1(z)] = \frac{1}{\mu}\left\{-(1-a)L\alpha + \alpha(Q_{11}^{-1/2}x_1)^2\right\}$$

by ignoring the terms $o_p(\mu^{-2})$. Similarly, since $\kappa = 0$ under the normality,

$$\begin{aligned} Q_{11}^{-1/2}\frac{1}{n}[e_2(z)] = \frac{1}{\mu^2}\Big\{&\alpha^2[Q_{11}^{-1/2}x_1]^3 + [Q_{11}^{-1/2}x_1] - (1-a)L[3\alpha^2](Q_{11}^{-1/2}x_1) \\ &-(1-a)L(Q_{11}^{-1/2}x_1)\Big\}, \end{aligned}$$

$$\begin{aligned} Q_{11}^{-1}\frac{1}{2n}[e_{11}(z)] = \frac{1}{2\mu^2}\Bigg\{&2[\mathbf{Q}\mathbf{D}'\mathbf{F}\mathbf{D}\mathbf{Q}]_{11}Q_{11}^{-2}\frac{\sigma^4}{|\mathbf{\Omega}|} \\ &+\alpha^2[Q_{11}^{-1/2}x_1]^4 + [Q_{11}^{-1/2}x_1]^2 + L + (1-a)^2L(L+2)\alpha^2 - 2(1-a)L\alpha^2(Q_{11}^{-1/2}x_1)^2\Bigg\} \end{aligned}$$

by ignoring the terms of $o_p(\mu^{-2})$. We notice that under the normal disturbances we have $z = Q_{11}^{-1/2}x_1 \sim N(0,1)$, and then by using the inversion formula (for the distribution function) we only need to evaluate

$$\Phi(z)+\frac{1}{\sqrt{n}}\left\{-Q_{11}^{-1/2}e_1(z)\right\}\phi(z)+\frac{1}{2n}\left\{-2Q_{11}^{-1/2}e_2(z)+Q_{11}^{-1}\left[\frac{d}{dz}[e_{11}(z)]-ze_{11}(z)\right]\right\}\phi(z) \quad \text{(A.37)}$$

up to the orders of $O(n^{-1})$ or $O(\mu^{-2})$. Then by setting $a = 1$ for the MEL estimator and $a = 0$ for the GMM estimator, we have the results in *Theorem 3.1* and *Theorem 3.2*.

Appendix C : Proof of Lemmas

[C1] Proof of Lemma A.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, we have the first part. Also the conditional distribution of $(X_1, X_2)^{'}$ given X_3 is also asymptotically normal. Then

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

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

$$Cov(X_1, X_2) \cong \frac{1}{n} \sum_{\alpha=1}^{n} z_{\alpha i} z_{\alpha j} (\mathbf{A}\mathbf{z}_{\alpha k}) \mathbf{E}[u_{\alpha}^3], \tag{A.38}$$

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

[C2] Proof of Lemma A.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 $\mathbf{z}_j^{(n)}$ $(j = 1, \cdots, n)$: $\mathbf{z}_n = n^{-1/2} \sum_{j=1}^{n} \mathbf{z}_j^{(n)}$ and $\mathbf{E}[\mathbf{z}_j^{(n)}] = \mathbf{0}$, $\mathbf{E}[\mathbf{z}_j^{(n)} \mathbf{z}_j^{(n)'}] = \mathbf{\Sigma}$ (> 0). Then under a set of regularity conditions (see Bhattacharya and Rao (1976), for instance) the characteristic function of $\mathbf{z}_n$ can be expressed as

$$\varphi(\mathbf{t}) = \prod_{j=1}^{n} \mathbf{E}[e^{i \sum_{k=1}^{p+1} t_j z_{jk}^{(n)}}] = e^{-\frac{1}{2}\mathbf{t}^{'}\mathbf{\Sigma}\mathbf{t}} \left\{ 1 + \frac{1}{6\sqrt{n}} \sum_{l_1,l_2,l_3=1}^{p+1} \beta_{l_1 l_2 l_3} (it_{l_1})(it_{l_2})(it_{l_3}) \right\} + O(n^{-1}) \, ,$$

where β_{l_1,l_2,l_3} are the third order moments of $\mathbf{z}_j^{(n)}$. Then the density function of $\mathbf{z}_n$ has a representation

$$f_n(\mathbf{z}) = \phi_{\Sigma}(\mathbf{z}) \Big\{ 1 + \frac{1}{6\sqrt{n}} \sum_{l_1,l_2,l_3=1}^{p+1} \beta_{l_1 l_2 l_3} h_3(z_{l_1}, z_{l_2}, z_{l_3}) \Big\} + O(n^{-1}) \, , \tag{A.39}$$

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 is represented as

$$\begin{aligned} f_n(v_n|\mathbf{u}_n) &= \phi(v|\boldsymbol{\rho}^{'}\mathbf{u}_n, 1 - \boldsymbol{\rho}^{'}\boldsymbol{\rho}) \\ &\times \Big\{ 1 + \frac{1}{6\sqrt{n}} [\sum_{l_1,l_2,l_3=1}^{p} \beta_{l_1,l_2,l_3} h_{3,\cdot}(u_{l_1}, u_{l_2}, u_{l_3}) + 3 \sum_{l_1,l_2}^{p} \beta_{l_1,l_2,p+1} h_{3,\cdot}(u_{l_1}, u_{l_2}, v_n) \\ &\quad + 3 \sum_{l=1}^{p} \beta_{l,p+1,p+1} h_{3,\cdot}(u_l, v_n, v_n) + \beta_{p+1,p+1,p+1} h_{3,\cdot}(v_n, v_n, v_n) \\ &\quad - \sum_{l_1,l_2,l_3=1}^{p} \beta_{l_1,l_2,l_3} h_3(u_{l_1}, u_{l_2}, u_{l_3})] \Big\} + O_p(n^{-1}) \, , \end{aligned}$$

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 $(\mathbf{u}_n, v)$ and $h_3(\cdot)$ are the third order Hermitian polynomials for the $p-$dimensional random vector $\mathbf{u}_n$. Then the conditional expectation is

$$
\begin{aligned}
& \mathbf{E}[v_n|\mathbf{u}_n] \\
= \;& \boldsymbol{\rho}^{'}\mathbf{u}_n + \frac{1}{6\sqrt{n}}\Big\{ \sum_{l_1,l_2,l_3=1}^{p} \beta_{l_1,l_2,l_3} \int v(-1)^3 \frac{\partial^3 f_n(\mathbf{u}_n, v)}{\partial u_{l_1}\partial u_{l_2}\partial u_{l_3}} \frac{1}{f_n(\mathbf{u}_n)} dv \\
& +3 \sum_{l_1,l_2=1}^{p} \beta_{l_1,l_2,p+1}(-1)^3 \frac{\partial^2}{\partial u_{l_1}\partial u_{l_2}} \int v \frac{\partial f_n(\mathbf{u}_n, v)}{\partial v} \frac{1}{f_n(\mathbf{u}_n)} dv \\
& +3 \sum_{l_1=1}^{p} \beta_{l_1,p+1,p+1}(-1)^3 \frac{\partial}{\partial u_l} \int v \frac{\partial^2 f_n(\mathbf{u}_n, v)}{\partial v^2} \frac{1}{f_n(\mathbf{u}_n)} dv \\
& +\beta_{p+1,p+1,p+1}(-1)^3 \int v \frac{\partial^3 f_n(\mathbf{u}_n, v)}{\partial v^3} \frac{1}{f_n(\mathbf{u}_n)} dv \\
& -(\boldsymbol{\rho}^{'}\mathbf{u}_n) \sum_{l_1,l_2,l_3=1}^{p} \beta_{l_1,l_2,l_3} h_3(u_{l_1}, u_{l_2}, u_{l_3}) \Big\} + O_p(n^{-1})\ .
\end{aligned}
$$

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

$$
\begin{aligned}
& \mathbf{E}[v_n|\mathbf{u}_n] \\
& = \rho^{'}\mathbf{u}_n + \frac{1}{6\sqrt{n}}\Big\{(-1) \sum_{l_1,l_2,l_3=1}^{p} \beta_{l_1,l_2,l_3} [\frac{\partial^3}{\partial u_{l_1}\partial u_{l_2}\partial u_{l_3}} (\boldsymbol{\rho}^{'}\mathbf{u}_n f_n(\mathbf{u}_n))]/f_n(\mathbf{u}_n) \\
& \quad +3 \sum_{l_1,l_2=1}^{p} \beta_{l_1,l_2,p+1} \left[\frac{\partial^2}{\partial u_{l_1}\partial u_{l_2}} f_n(\mathbf{u}_n)\right] /f_n(\mathbf{u}_n) - \boldsymbol{\rho}^{'}\mathbf{u}_n \sum_{l_1,l_2,l_3=1}^{p} \beta_{l_1,l_2,l_3} h_3(u_{l_1}, u_{l_2}, u_{l_3})\Big\} \\
& \quad +O_p(n^{-1})\ . \\
& = \rho^{'}\mathbf{u}_n + \frac{1}{6\sqrt{n}}\Big\{3 \sum_{l_1,l_2=1}^{p} \beta_{l_1,l_2,p} h_2(u_{l_1}, u_{l_2}) - \sum_{l_1,l_2,l_3=1}^{p} \beta_{l_1,l_2,l_3} \left[\boldsymbol{\rho}^{'}\mathbf{u}_n h_3(u_{l_1}, u_{l_2}, u_{l_3})\right] \\
& \quad + \sum_{l_1,l_2,l_3=1}^{p} \beta_{l_1,l_2,l_3} \left[\rho^{'}\mathbf{u}_n h_3(u_{l_1}, u_{l_2}, u_{l_3}) - \rho_{l_1} h_2(u_{l_2}, u_{l_3}) - \rho_{l_2} h_2(u_{l_1}, u_{l_3}) - \rho_{l_3} h_2(u_{l_1}, u_{l_2})\right] \Big\} \\
& +O_p(n^{-1})\ ,
\end{aligned}
$$

where $h_2(u_{l_1}, u_{l_2})$ 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 D : Useful Inversion Formulas

This appendix gives the useful formulas, which correspond to the inversion of the characteristic function from the conditional expectations given $\mathbf{x}^*$ and $\mathbf{x}^*$ follows the p-dimensional normal distribution $N_p(\mathbf{0},\mathbf{Q})$. Let $\psi(\mathbf{t}) = \mathbf{E}[e^{i\mathbf{t}'\mathbf{x}^*}]$ be the characteristic function of $\mathbf{x}^*$. Then by using the integration-in-parts formula for $\mathbf{t} = (t_j)$ and $\boldsymbol{\xi} = (\xi_k)$,

$$(i\xi_j)\phi_{\mathbf{Q}}(\boldsymbol{\xi}) = (\frac{1}{2\pi})^p \int_{\mathbf{R}^p} e^{-i\mathbf{t}'\boldsymbol{\xi}} \left[\frac{\partial\psi(\mathbf{t})}{\partial t_j}\right] d\mathbf{t} \ , \tag{A.40}$$

for instance. By using integration-in-parts repeatedly with respect to $\mathbf{t} = (t_j)$ and differentiating with respect to $\boldsymbol{\xi} = (\xi_k)$, we have the Fourier inversion formulas

$$\mathcal{F}^{-1}\{h(-i\mathbf{t})\mathcal{E}[g(\mathbf{x})\exp(i\mathbf{t}'\mathbf{x}^*)]\} = h(\frac{\partial}{\partial\boldsymbol{\xi}})g(\boldsymbol{\xi})\phi_{\mathbf{Q}}(\boldsymbol{\xi}) \tag{A.41}$$

for any polynomials $h(\ \cdot\ )$ and $g(\ \cdot\ )$, where $i^2 = -1$ 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 originally developed by Fujikoshi et al. (1982) and Anderson et al. (1986). We present useful results including new formulas.

Lemma A.4 : Let $\boldsymbol{\eta}' = (\eta_1,\cdots,\eta_p)$ be a $1\times p$ constant vector, $\mathbf{B}$ 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) $\frac{\partial}{\partial\boldsymbol{\xi}'}[\boldsymbol{\eta}\phi_{\mathbf{Q}}(\boldsymbol{\xi})] = [-\boldsymbol{\eta}'\mathbf{Q}^{-1}\boldsymbol{\xi}]\phi_{\mathbf{Q}}(\boldsymbol{\xi})$,

(ii) $\frac{\partial}{\partial\boldsymbol{\xi}'}[\mathbf{B}\boldsymbol{\xi}(\boldsymbol{\eta}'\boldsymbol{\xi})\phi_{\mathbf{Q}}(\boldsymbol{\xi})] = \left[(\boldsymbol{\eta}'\boldsymbol{\xi})(\mathrm{tr}(\mathbf{B}) - \boldsymbol{\xi}'\mathbf{B}'\mathbf{Q}^{-1}\boldsymbol{\xi}) + \boldsymbol{\xi}'\mathbf{B}\boldsymbol{\eta}\right]\phi_{\mathbf{Q}}(\boldsymbol{\xi})$,

(iii) $\frac{\partial}{\partial\boldsymbol{\xi}'}[\mathbf{Q}\mathbf{B}\boldsymbol{\xi}\phi_{\mathbf{Q}}(\boldsymbol{\xi})] = \left[\mathbf{tr}(\mathbf{B}\mathbf{Q}) - \boldsymbol{\xi}'\mathbf{B}\boldsymbol{\xi}\right]\phi_{\mathbf{Q}}(\boldsymbol{\xi})$,

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

(v) $\mathbf{tr}\frac{\partial^2}{\partial\boldsymbol{\xi}\partial\boldsymbol{\xi}'}[\mathbf{Q}\mathbf{B}\mathbf{Q}\phi_{\mathbf{Q}}(\boldsymbol{\xi})] = \left[\boldsymbol{\xi}'\mathbf{B}\boldsymbol{\xi} - \mathbf{tr}(\mathbf{B}\mathbf{Q})\right]\phi_{\mathbf{Q}}(\boldsymbol{\xi})$,

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

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

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

Appendix E: Tables and Figures

In Tables 1-3 and Figures 1-2 the exact and approximate distributions based on the asymptotic expansions are presented in the standardized terms, that is, of (3.1). The basic procedure of simulations is to generate the vectors of the normal disturbance terms and the exogenous variables $\mathbf{v}_i, \mathbf{z}_i$ $(i = 1,\cdots,n)$ and generate the endogenous variables. Then we simulate the probability of (3.1) by utilizing (2.5) and (2.6) and do iterations until we have numerical convergence stably. We denote the resulting values as *Exact* in Tables 1 and 2 because they are very accurate in two decimal digits at least. Our method of evaluating the distribution functions of estimators in numerical analysis is essentially the same as Anderson et al. (2005, 2008) which explain its details and the accuracy of our computations.

The tables include three quartiles, the 5 and 95 percentiles and the interquartile range of the distribution for each case. Since the limiting distributions of (3.1) for the MEL and GMM estimators in the standard large sample theory are $N(0,1)$ as $n \to \infty$, we add the standard normal case as the bench mark. Figures 2 and 3 are taken from a case study of Anderson et al. (2005, 2008).

References

[1] Akahira, M. and Takeuchi, K. (1981), *Asymptotic Efficiency of Statistical Estimators :Concepts and Higher Order Asymptotic Efficiency*, Lecture Notes No.7, Springer-Verlag.

[2] Akahira, M. and Takeuchi, K. (1990), "First Order Asymptotic Efficiency in Semiparametric models implies Infinite Deficiency," *Publications de l'Institut de Statistique de l'Universite de Paris*, XXXV, facs. 1, 3-9.

[3] Anderson, T.W. (1951), "Estimating linear restrictions on regression coefficients for multivariate normal distributions," *Annals of Mathematical Statistics*, Vol. 22, 327-351.

[4] 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.

[5] 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.

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

[7] Anderson, T.W., N. Kunitomo, and Y. Matsushita (2005), "A New Light from Old Wisdoms : Alternative Estimation Methods 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).

[8] Anderson, T.W., N. Kunitomo, and Y. Matsushita (2007) "On the Asymptotic Optimality of the LIML Estimator with Possibly Many Instruments," Discussion Paper CIRJE-F-542, Graduate School of Economics, University of Tokyo.

[9] Anderson, T.W., N. Kunitomo, and Y. Matsushita (2008), "On Finite Sample Properties of Alternative Estimators of Coefficients in a Structural Equation with Many Instruments," Discussion Paper CIRJE-F-576, Graduate School of Economics, University of Tokyo, forthcoming in *Journal of Econometrics*.

[10] Anderson, T.W. and H. Rubin (1949), "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, Vol. 20, 46-63.

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

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

[13] Bickel, P. C. Klaassen, Y. Ritov and J. Wellner (1993), *Efficient and Adaptive Estimation for Semiparametric Models*, Johns-Hopkins University Press.

[14] 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.

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

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

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

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

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

[20] Mittelhammer, R., G. Judge, and R. Schoenberg (2004), "Empirical Evidence Concerning the Finite Sample Performance of EL-Type Structural Equation Estimation and Inference Methods," In *Identification and Inference for Econometric Models*, Cambridge University Press (edited by D.Andrews and J.Stock).

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

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

[23] Pfanzagl, J. (1990) *Estimation in Semiparametric Models*, Lecture Notes in Statistics 63, Springer, New York.

[24] Pfanzagl, J. and W. Wefelmeyer (1978) " A Third-Order Optimum Property of the Maximum Likelihood estimator," *Journal of Multivariate Analysis*, 8, 1-29.

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

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

[27] 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.

	$n-K=30, K_2=3, \alpha=1, \delta^2=30$					
	MEL			GMM		
x	Exact	Approx	Difference	Exact	Approx	Difference
-3	0.000	0.005	0.005	0.000	0.002	0.002
-2.5	0.002	0.007	0.005	0.002	0.003	0.001
-2	0.012	0.014	0.002	0.016	0.013	-0.002
-1.4	0.066	0.062	-0.004	0.092	0.087	-0.005
-1	0.146	0.146	0.000	0.194	0.200	0.006
-0.8	0.204	0.205	0.001	0.270	0.275	0.005
-0.6	0.273	0.274	0.000	0.356	0.357	0.001
-0.4	0.353	0.348	-0.005	0.444	0.442	-0.002
-0.2	0.429	0.425	-0.005	0.520	0.526	0.005
0	0.503	0.500	-0.003	0.597	0.603	0.006
0.2	0.573	0.571	-0.002	0.669	0.672	0.003
0.4	0.640	0.637	-0.003	0.732	0.733	0.001
0.6	0.698	0.696	-0.003	0.785	0.784	-0.001
0.8	0.752	0.747	-0.005	0.829	0.827	-0.002
1	0.793	0.792	-0.001	0.869	0.862	-0.006
1.4	0.866	0.862	-0.004	0.927	0.915	-0.013
2	0.931	0.930	0.000	0.968	0.959	-0.009
2.5	0.956	0.964	0.009	0.984	0.978	-0.006
3	0.974	0.985	0.010	0.994	0.990	-0.004
X05	-1.52	-1.49	0.03	-1.65	-1.58	0.07
L.QT	-0.67	-0.67	0.00	-0.85	-0.85	0.00
MEDN	-0.01	0.00	0.01	-0.25	-0.26	0.00
U.QT	0.79	0.81	0.02	0.47	0.45	-0.02
X95	2.36	2.26	-0.11	1.66	1.79	0.13
IQR	1.46	1.48	0.02	1.31	1.30	-0.01

Table 1: CDF of Standardized MEL and GMM estimators: $n-K=30, K_2=3, \alpha=1, \delta^2=50$

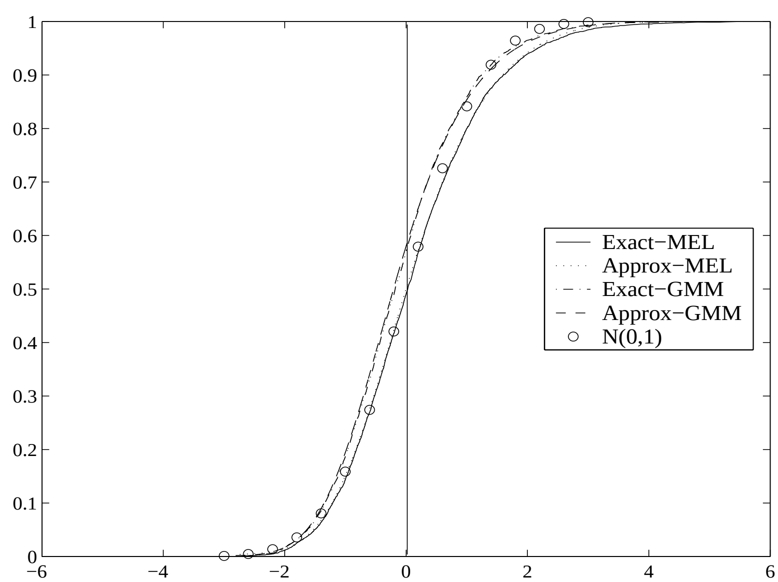

Figure 1: $n-K=30, K_2=3, \alpha=1, \delta^2=50$

	$n-K=100, K_2=10, \alpha=1, \delta^2=100$					
	MEL			GMM		
x	Exact	Approx	Difference	Exact	Approx	Difference
-3	0.000	0.002	0.002	0.002	0.002	0.000
-2.5	0.004	0.006	0.002	0.014	0.012	-0.002
-2	0.020	0.021	0.001	0.055	0.053	-0.002
-1.4	0.079	0.081	0.002	0.188	0.191	0.003
-1	0.165	0.165	-0.001	0.338	0.346	0.008
-0.8	0.224	0.221	-0.003	0.418	0.435	0.017
-0.6	0.289	0.285	-0.004	0.506	0.525	0.020
-0.4	0.360	0.355	-0.005	0.586	0.611	0.026
-0.2	0.428	0.427	-0.001	0.660	0.688	0.028
0	0.505	0.500	-0.005	0.726	0.754	0.028
0.2	0.577	0.571	-0.006	0.785	0.807	0.022
0.4	0.642	0.637	-0.005	0.832	0.849	0.018
0.6	0.697	0.698	0.001	0.869	0.882	0.013
0.8	0.748	0.753	0.005	0.901	0.907	0.006
1	0.793	0.801	0.008	0.926	0.928	0.001
1.4	0.869	0.877	0.008	0.969	0.958	-0.011
2	0.942	0.948	0.007	0.991	0.985	-0.006
2.5	0.970	0.978	0.008	0.998	0.995	-0.003
3	0.985	0.992	0.007	1.000	0.998	-0.001
X05	-1.61	-1.63	-0.03	-2.04	-2.02	0.02
L.QT	-0.72	-0.71	0.02	-1.22	-1.23	-0.01
MEDN	-0.01	0.00	0.01	-0.61	-0.66	-0.04
U.QT	0.81	0.79	-0.02	0.07	-0.01	-0.09
X95	2.12	2.02	-0.10	1.18	1.28	0.10
IQR	1.53	1.50	-0.04	1.30	1.22	-0.08

Table 2: CDF of Standardized MEL and GMM estimators: $n-K=100, K_2=10, \alpha=1, \delta^2=100$

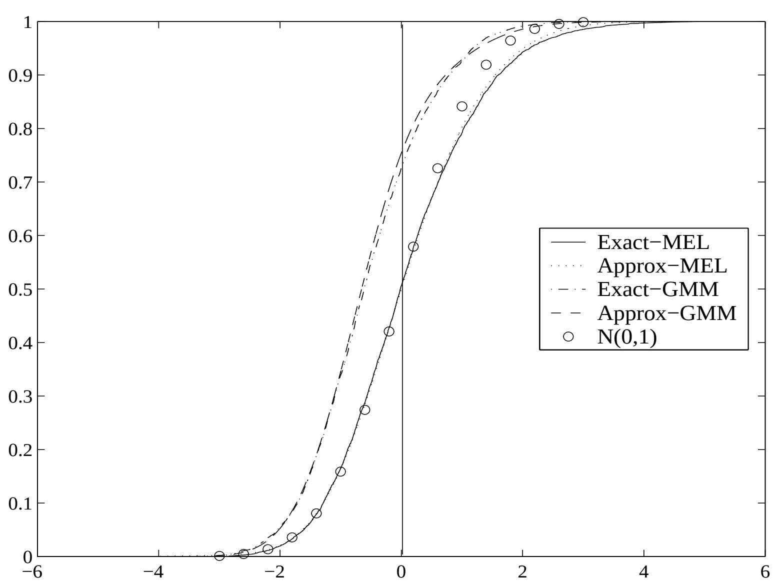

Figure 2: $n-K=100, K_2=10, \alpha=1, \delta^2=100$

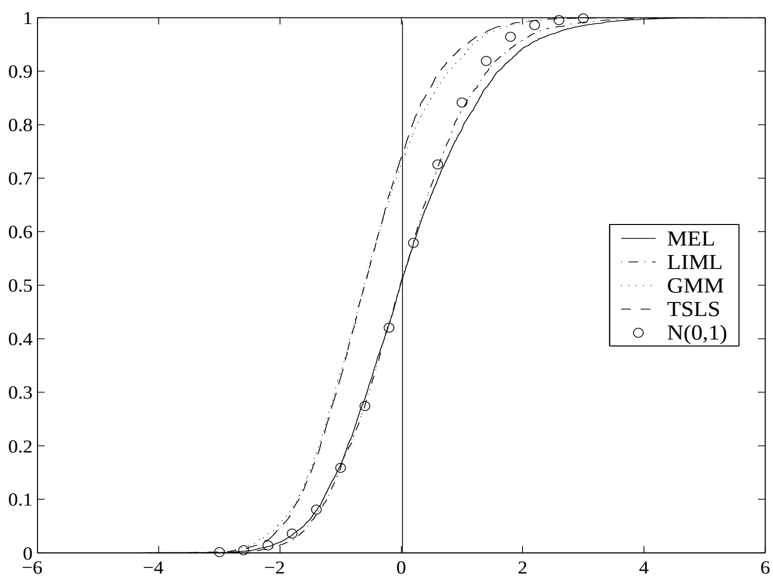

Figure 3: CDF of Standardized estimators: $n - K = 100, K_2 = 10, \alpha = 1, \delta^2 = 100$