AN ASYMPTOTIC EXPANSION OF THE DISTRIBUTION OF THE DM TEST STATISTIC*

Wanling Huang† Artem Prokhorov‡

January 14, 2009

Abstract

This paper has three parts: theoretical results, simulations and an economic application. In the theoretical part, after the Distance Metric (DM) test statistic is expanded to the second order, the Edgeworth approximation of the distribution of the DM test statistic is derived and the Bartlett-type correction factor is obtained based on the results of Phillips and Park (1988) and Hansen (2006); in the simulation part, simple examples are given to illustrate the theoretical results; in the application part, the theoretical results are applied to study the covariance structures of earnings.

JEL Classification: C12

Keywords: Distance Metric, GMM, Asymptotic expansion, Bartlett-type correction

*Helpful comments from Nikolay Gospodinov, Gordon Fisher and Jorgen Hansen are gratefully acknowledged. Financial support from FQRSC and the J. W. McConnell Memorial Graduate Fellowship are highly appreciated.

†Department of Economics, Concordia University, 1455 de Maisonneuve Blvd W, Montreal QC H3G1M8 Canada; email: whuang@live.concordia.ca

‡Department of Economics, Concordia University, 1455 de Maisonneuve Blvd W, Montreal QC H3G1M8 Canada; email: artem.prokhorov@concordia.ca

1 Introduction

Traditionally in maximum likelihood models, one of the trio of Wald, likelihood ratio, and Lagrange multiplier test statistics is used to test a hypothesis. However, in many cases such as in rational expectations models, since the disturbances are serially correlated and heteroskedastic in complicated ways, the computational burden of maximum likelihood is enormous. This motivates the development of the Distance Metric (DM) test statistic (Newey and West, 1987; Newey and McFadden, 1994). Generally, the DM test statistic allows for autocorrelation and heteroskedasticity of possibly unknown forms, and it does not require any assumptions for the distribution of the data.

Exact finite-sample probability distributions of test statistics are available in convenient form only for simple functions of the data and when the likelihood function is completely specified, however, these conditions are often not satisfied in econometrics and inference is based on approximations (Rothenberg, 1984). Among various methods of approximation, asymptotic expansion is a popular one. There are two well-known methods for obtaining higher-order approximate distribution functions based on Fourier inversion of the approximate characteristic function: Edgeworth approximation and Saddlepoint approximation. Since the latter requires knowledge of the cumulant function which is rare in econometric applications, Edgeworth approximation becomes a common method.

Asymptotically, the DM test statistic has a chi-squared distribution (Newey and McFadden, 1994, page 2226). In practice, however, this is infeasible because the sample size is finite. It is expected that after Edgeworth expansion, the distribution of the corrected DM test statistic is closer to a chi-squared distribution than the uncorrected one.

Based on the results of Phillips and Park (1988) and Hansen (2006), the present paper derives the Edgeworth expansion of the DM test statistic and further derives the Edgeworth approximation of its distribution. Phillips and Park (1988) investigate the Wald test of nonlinear restrictions, and some of their formulae for the asymptotic expansion of the distribution of the Wald statistic can be used in our expansion. To our knowledge, Hansen (2006) is the only expansion of the DM test statistic in the existing literature. The results obtained by Hansen (2006) are for nonlinear restrictions, and the number of hypothesis restrictions is one. Besides, the expansion in Hansen (2006) is for only one specific model (linear regression), which strongly limits its application. Our results can be viewed as complementary to both Phillips and Park (1988) and Hansen (2006) in that they apply to both nonlinear and linear restrictions, furthermore, they cover broader test hypotheses and apply to various models and multiple restrictions. However, they are applicable only to moment conditions which are nonlinear in parameters.

The main theoretical results are stated in Section 2, illustrative simulations are performed in Section 3, and an empirical illustration is presented in Section 4. Some brief conclusions are given in Section 5.

2 Asymptotic Expansions for the Distribution of the DM Test Statistic

For a family of distributions $\{P_\theta, \theta \in \Theta \subset \mathbb{R}^p, \Theta\ \texttt{compact}\}$, suppose there is a test

$$\begin{aligned} &H_0 : \underset{r\times 1}{g(\theta)} = 0, \\ &H_1 : g(\theta) \neq 0, \end{aligned} \tag{2.1}$$

where $g : R^p \to R^r$. Assume $g(\theta)$ is continuously differentiable, then the first derivative of $g(\theta)$ can be defined as

$$\underset{p\times r}{A(\theta)} \equiv \frac{dg(\theta)}{d\theta} = \frac{dvec'g(\theta)}{d\theta}, \tag{2.2}$$

where the second equality follows definition 1.4.1 of Kollo and Rosen (2005) and *vec* denotes vertical vectorization of a matrix. Denote $A = A(\theta_0)$.

Let $\bar{\theta}_N$ and $\hat{\theta}_N$ be the constrained and unconstrained generalized method of moments (GMM) estimators (Hansen, 1982) of θ_0 respectively:

$$\begin{aligned} &\bar{\theta}_N = \arg\max_{\theta\in\Theta} Q_N(\theta),\ \texttt{subject to}\ g(\theta) = 0, \\ &\hat{\theta}_N = \arg\max_{\theta\in\Theta} Q_N(\theta), \end{aligned}$$

where

$$-Q_N(\theta) \equiv \frac{1}{2} m'_N(\theta) W_N^{-1} m_N(\theta) \tag{2.3}$$

with $\underset{m\times 1}{m_N(\theta)} \equiv \frac{1}{N}\sum_{i=1}^{N} m(Z_i, \theta)$ the sample moment function and W_N a positive definite symmetric matrix. Define the covariance matrix of the moments

$$\underset{m\times m}{W} \equiv \mathbb{E}[m(Z_i, \theta_0) m'(Z_i, \theta_0)].$$

Efficient weighting of a given set of m moments requires

$$W_N \xrightarrow{p} W.$$

Then, the DM test statistic is defined (Newey and McFadden, 1994, page 2222) as

$$DM_N \equiv -2N[Q_N(\bar{\theta}_N) - Q_N(\hat{\theta}_N)]. \tag{2.4}$$

We want to derive the Edgeworth approximation of the distribution of the DM test statistic. In order to do so, the DM test statistic needs to be expanded first. If the expanded DM test statistic is a function of some variates with known densities, then it is straightforward to obtain the characteristic function of the DM test statistic. The distribution of the DM test statistic will be derived using a Fourier transform.

2.1 Expansion for the DM Test Statistic

We expand the test statistic using the Taylor expansion of a vector function, which is stated in Lemma 2.1.

Lemma 2.1. *Let $\{x_n\}$ and $\{\varepsilon_n\}$ be sequences of random p-vectors and positive numbers, respectively, and let $x_n - x_0 = o_p(\varepsilon_n)$, where $\varepsilon_n \to 0$ as $n \to \infty$. If the function $f(x)$ from $\mathbb{R}^p$ to $\mathbb{R}^s$ has continuous partial derivatives up to the order $(n+1)$ in a neighborhood $\mathcal{D}$ of a point x_0, then the function $f(x)$ can be expanded at the point x_0 into the Taylor series*

$$f(x) = f(x_0) + \sum_{k=1}^{m} \frac{1}{k!} \left(I_s \otimes (x - x_0)^{\otimes(k-1)}\right)' \left(\frac{d^k f(x)}{dx^k}\right)'_{x=x_0} (x - x_0) + o(\rho^m(x, x_0)),$$

where the Kroneckerian power $A^{\otimes k}$ for any matrix A is given by $A^{\otimes k} = \underbrace{A \otimes \cdots \otimes A}_{k \text{ times}}$ with $A^{\otimes 0} = 1$, $\rho(.,.)$ is the Euclidean distance in $\mathbb{R}^p$, and the matrix derivative for any matrices Y and X is given by $\frac{d^kY}{dX^k} = \frac{d}{dX}\left(\frac{d^{k-1}Y}{dX^{k-1}}\right)$ with $\frac{dY}{dX} \equiv \frac{dvec'Y}{dvecX}$; and

$$f(x_n) = f(x_0) + \sum_{k=1}^{m} \frac{1}{k!} \left(I_s \otimes (x_n - x_0)^{\otimes(k-1)}\right)' \left(\frac{d^k f(x_n)}{dx_n^k}\right)'_{x_n=x_0} (x_n - x_0) + o_p(\varepsilon_n^m). \tag{2.5}$$

See Appendix A for all proofs that are not given in the main text.

Assume the moment function in (2.3) is three-times continuously differentiable, then its first, second and third derivatives can be defined as

$$\underset{p\times m}{G_N(\theta)} \equiv \frac{dm_N'(\theta)}{d\theta},$$
$$\underset{p\times pm}{D_N(\theta)} \equiv \frac{dvec'G_N(\theta)}{d\theta},$$
$$\underset{p\times p^2m}{C_N(\theta)} \equiv \frac{dvec'D_N(\theta)}{d\theta}.$$

Denote $G = \mathbb{E}[G(\theta_0)]$, $D = \mathbb{E}[D(\theta_0)]$, and $C = \mathbb{E}[C(\theta_0)]$. Note that G is the transpose of the Jacobian matrix defined on page 2216 of Newey and McFadden (1994). The process of taking derivatives here follows Definition 1.4.1 of Kollo and Rosen (2005), which facilitates Taylor expansions for vector functions.

For simplicity, set $W = I$, where I is an identity matrix. Note that this could always be done by standardizing transformations, i.e., assuming W is known and pre-multiplying the moments by $W^{-1/2}$ to standardize. After such standardization, $G_N(\theta)$, $D_N(\theta)$ and $C_N(\theta)$ are changed accordingly, the GMM quadratic form (2.3) becomes

$$-Q_N(\theta) \equiv \frac{1}{2}m_N'(\theta)m_N(\theta), \tag{2.6}$$

the DM test statistic in (2.4) becomes

$$DM_N = N[m_N'(\bar{\theta}_N)m_N(\bar{\theta}_N) - m_N'(\hat{\theta}_N)m_N(\hat{\theta}_N)], \tag{2.7}$$

and a central limit theorem implies

$$-\sqrt{N}m_N(\theta_0) \equiv \bar{q}_N \xrightarrow{d} \underset{m\times 1}{\bar{q}} \sim N(0, I). \tag{2.8}$$

Also, we have (Newey and McFadden, 1994, page 2219)

$$\sqrt{N}(\hat{\theta}_N - \theta_0) = B^{-1}G\bar{q}_N + o_p, \tag{2.9}$$
$$\sqrt{N}(\bar{\theta}_N - \hat{\theta}_N) = -\mathbb{H}G\bar{q}_N + o_p, \tag{2.10}$$

where

$$B^{-1} = (GG')^{-1}$$

is the asymptotic covariance of $\hat{\theta}_N$, and

$$\underset{p\times p}{\mathbb{H}} \equiv B^{-1}A(A'B^{-1}A)^{-1}A'B^{-1}.$$

By Lemma 2.1, Taylor expansion of $m_N(\bar{\theta}_N)$ about $\hat{\theta}_N$ to the second order yields

$$m_N(\bar{\theta}_N) = m_N(\hat{\theta}_N) + G'_N(\hat{\theta}_N)(\bar{\theta}_N - \hat{\theta}_N) + \frac{1}{2}[I_m \otimes (\bar{\theta}_N - \hat{\theta}_N)']D'_N(\hat{\theta}_N)(\bar{\theta}_N - \hat{\theta}_N) + o_p.$$

Substituting this into (2.7) and using (2.10) together with the unconstrained first-order condition

$$G(\hat{\theta}_N)m_N(\hat{\theta}_N) = 0$$

give

$$\begin{aligned} DM = {} & \bar{q}'G'\mathbb{H}G_N(\hat{\theta}_N)G'_N(\hat{\theta}_N)\mathbb{H}G\bar{q} \\ & + m'_N(\hat{\theta}_N)(I_m \otimes \bar{q}'G'\mathbb{H})D'_N(\hat{\theta}_N)\mathbb{H}G\bar{q} \\ & - N^{-1/2}\bar{q}'G'\mathbb{H}G_N(\hat{\theta}_N)(I_m \otimes \bar{q}'G'\mathbb{H})D'_N(\hat{\theta}_N)\mathbb{H}G\bar{q} \\ & + \frac{1}{4}N^{-1}\bar{q}'G'\mathbb{H}D_N(\hat{\theta}_N)(I_m \otimes \mathbb{H}G\bar{q})(I_m \otimes \bar{q}'G'\mathbb{H})D'_N(\hat{\theta}_N)\mathbb{H}G\bar{q} + o_p. \end{aligned} \tag{2.11}$$

The functions of $\hat{\theta}_N$ in (2.11) need to be expanded at θ_0. In order to use Lemma 2.1, it is necessary to do some transformations. The transformations are needed so that we can apply the Taylor representation of Lemma 2.1 to matrix functions $G_N(\hat{\theta}_N)$ and $D_N(\hat{\theta}_N)$. Note that $G'_N(\hat{\theta}_N)\mathbb{H}G\bar{q}$ and $D'_N(\hat{\theta}_N)\mathbb{H}G\bar{q}$ are vectors. Thus, using

$$\begin{aligned} & vec(ABC) = (C' \otimes A)vecB, \\ & (A \otimes B)' = A' \otimes B', \end{aligned}$$

we obtain

$$\begin{aligned} & \bar{q}'G'\mathbb{H}G_N(\hat{\theta}_N) = vec'G_N(\hat{\theta}_N)(I_m \otimes \mathbb{H}G\bar{q}), \\ & D'_N(\hat{\theta}_N)\mathbb{H}G\bar{q} = (I_{pm} \otimes \bar{q}'G'\mathbb{H})vecD_N(\hat{\theta}_N). \end{aligned}$$

Then, (2.11) becomes

$$\begin{aligned} DM = {} & vec'G_N(\hat{\theta}_N)M_1vecG_N(\hat{\theta}_N) \\ & + m'_N(\hat{\theta}_N)M_2vecD_N(\hat{\theta}_N) \\ & - N^{-1/2}vec'G_N(\hat{\theta}_N)M_3vecD_N(\hat{\theta}_N) \\ & + N^{-1}\frac{1}{4}vec'D_N(\hat{\theta}_N)M_4vecD_N(\hat{\theta}_N) + o_p, \end{aligned} \tag{2.12}$$

where

$$M_1 = (I_m \otimes \mathbb{H}G\bar{q})(I_m \otimes \bar{q}'G'\mathbb{H}), \tag{2.13}$$

$$M_2 = I_m \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H}, \tag{2.14}$$

$$M_3 = (I_m \otimes \mathbb{H}G\bar{q})(I_m \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H}), \tag{2.15}$$

$$M_4 = I_m \otimes \mathbb{H}G\bar{q}\bar{q}'G'\mathbb{H} \otimes \mathbb{H}G\bar{q}\bar{q}'G'\mathbb{H}. \tag{2.16}$$

Substituting Taylor expansions of $m_N(\hat{\theta}_N)$, $vecG_N(\hat{\theta}_N)$ and $vecD_N(\hat{\theta}_N)$ about θ_0 into (2.12) gives the asymptotic expansion of the DM test statistic, which is summarized in the following theorem.

Theorem 2.2. *The asymptotic expansion up to the second order of the DM test statistic is given by*

$$DM = \bar{q}'P\bar{q} + N^{-1/2}u(\bar{q}) + N^{-1}v(\bar{q}) + o_p, \tag{2.17}$$

where

$$\underset{m\times m}{P} \equiv G'\mathbb{H}G, \tag{2.18}$$

$$u(\bar{q}) = u_1(\bar{q}) + u_2(\bar{q}) + u_3(\bar{q}), \tag{2.19}$$

$$v(\bar{q}) = v_1(\bar{q}) + v_2(\bar{q}) + v_3(\bar{q}) + v_4(\bar{q}), \tag{2.20}$$

with $u_i(\bar{q})$ $(i = 1, 2, 3)$ *and* $v_i(\bar{q})$ $(i = 1, 2, 3, 4)$ *specified as*

$$u_1(\bar{q}) = 2\bar{q}'G'B^{-1}DM_1vecG, \tag{2.21}$$

$$u_2(\bar{q}) = \bar{q}'(G'B^{-1}G - I_m)M_2vecD, \tag{2.22}$$

$$u_3(\bar{q}) = -vec'GM_3vecD; \tag{2.23}$$

$$v_1(\bar{q}) = \bar{q}'G'B^{-1}DM_1D'B^{-1}G\bar{q} + \bar{q}'G'B^{-1}C(I_{pm} \otimes B^{-1}G\bar{q})M_1vecG, \tag{2.24}$$

$$v_2(\bar{q}) = \bar{q}'(G'B^{-1}G - I_m)M_2C'B^{-1}G\bar{q} + \frac{1}{2}\bar{q}'G'B^{-1}D(I_m \otimes B^{-1}G\bar{q})M_2vecD, \tag{2.25}$$

$$v_3(\bar{q}) = -\bar{q}'G'B^{-1}CM_3'vecG - \bar{q}'G'B^{-1}DM_3vecD, \tag{2.26}$$

$$v_4(\bar{q}) = \frac{1}{4}vec'DM_4vecD. \tag{2.27}$$

By the theorem for the distribution of an idempotent quadratic form in a standard normal vector (Greene, 2003), it can be seen from (2.17) that asymptotically the DM test statistic has a chi-squared distribution. This is consistent with Newey and McFadden (1994).

2.2 Expansion for the Distribution of the DM Test Statistic

We wish to use the results of Phillips and Park (1988) to derive the Edgeworth expansion of the distribution of the DM test statistic. In order to do so, it is necessary to rewrite $u(\bar{q})$ and $v(\bar{q})$ in Theorem 2.2 into appropriate forms, which is summarized in the following lemma.

Lemma 2.3. *$u(\bar{q})$ and $v(\bar{q})$ in Theorem 2.2 could be rewritten as*

$$u(\bar{q}) = vec'J(\bar{q} \otimes \bar{q} \otimes \bar{q}), \tag{2.28}$$

$$v(\bar{q}) = tr[L(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')], \tag{2.29}$$

where

$$vecJ = vecJ_1 + vecJ_2 + vecJ_3, \tag{2.30}$$

with

$$vecJ_1 = 2(G'\mathbb{H}G \otimes G'\mathbb{H} \otimes G'B^{-1})vecD, \tag{2.31}$$

$$vecJ_2 = [(G'B^{-1}G - I_m) \otimes G'\mathbb{H} \otimes G'\mathbb{H}]vecD, \tag{2.32}$$

$$vecJ_3 = -(G'\mathbb{H}G \otimes G'\mathbb{H} \otimes G'\mathbb{H})vecD; \tag{2.33}$$

and

$$L = L_1 + L_2 + L_3 + L_4, \tag{2.34}$$

with

$$L_1 = (G'\mathbb{H} \otimes G'B^{-1})V_D(\mathbb{H}G \otimes B^{-1}G) + (G'\mathbb{H} \otimes G'B^{-1})M_V(I_m \otimes \mathbb{H}G), \tag{2.35}$$

$$L_2 = (G'\mathbb{H} \otimes G'\mathbb{H})M_{VI} + \frac{1}{2}(G'\mathbb{H} \otimes G'\mathbb{H})V_D(B^{-1}G \otimes B^{-1}G), \tag{2.36}$$

$$L_3 = -(G'\mathbb{H} \otimes G'\mathbb{H})M_V(I_m \otimes \mathbb{H}G) - (G'\mathbb{H} \otimes G'\mathbb{H})V_D(\mathbb{H}G \otimes B^{-1}G), \tag{2.37}$$

$$L_4 = \frac{1}{4}(G'\mathbb{H} \otimes G'\mathbb{H})V_D(\mathbb{H}G \otimes \mathbb{H}G), \tag{2.38}$$

where V_D, M_V and M_{VI} are given in Appendix A.

We can now apply the results on page 1069-p1072 of Phillips and Park (1988) (see also Hansen, 2006, Theorem 3) to obtain the following theorem:

Theorem 2.4. *The asymptotic expansion up to $O(N^{-1})$ of the distribution of the DM test statistic as $N \to \infty$ is given by*

$$F_{DM}(x) = F_r\left(x - N^{-1}(\alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3)\right) + o(N^{-1}) \tag{2.39}$$

where r denotes the rank of the projection matrix P in (2.18), F_r denotes the distribution function of a χ_r^2 variate and where

$$\alpha_1 = (4a_1 - b_2)/4r, \tag{2.40}$$

$$\alpha_2 = (4a_2 + b_2 - b_3)/4r(r+2), \tag{2.41}$$

$$\alpha_3 = b_3/4r(r+2)(r+4), \tag{2.42}$$

with a_i $(i = 1, 2)$ and b_i $(i = 1, 2, 3)$ defined in Appendix A, in calculation of which L and $vecJ$ are from (2.34) and (2.30) respectively.

From (2.39), a correction factor to $O(N^{-1})$ is derived:

$$1 - N^{-1}(\alpha_1 + \alpha_2 DM + \alpha_3 DM^2) \tag{2.43}$$

where DM denotes the DM test statistic. This correction factor is a function of the uncorrected DM test statistic. Therefore, the correction is not a 'Bartlett correction' in the classical sense. It is called Bartlett-type correction (Cribari-Neto and Cordeiro, 1996). It is obvious that asymptotically the distribution of the uncorrected DM test statistic matches the distribution of the corrected one because the second term in (2.43) approaches zero as N goes to infinity.

Note that increasing the value of r does not necessarily result in larger correction because α_i $(i = 1, 2, 3)$ may be negative. Also note that throughout the analytical section, no specific distributions for the data are assumed. In other words, the theoretical results obtained in this section are distribution-free.

Moreover, it is important to note that, even if the hypothesis restrictions are linear, the Bartlett-type correction factor in (2.43) might still be working. However, our results are applicable only to moment conditions which are nonlinear in parameters. One condition for the theory in Phillips and Park (1988) and Hansen (2006) to apply is that restrictions of a hypothesis test should be nonlinear. In this sense, our results complement Phillips and Park (1988) and Hansen (2006). Since in many economic applications the moments are nonlinear, our results are not meaningless.

Another aspect that needs to be noted is that the number of restrictions is not restricted. This makes it possible to apply the current results to structural equation models where the number of restrictions is usually larger than one and the number of degrees of freedom is possibly large. To our knowledge, Hansen (2006) is the only expansion of the distribution of the DM test statistic, however, the results derived by Hansen (2006) are for only one restriction ($g(\theta)$ is 1×1) and one specific model (linear regression). Therefore, the present paper complements Hansen (2006) in several aspects.

3 Illustrative Simulations

In this section, simulations for covariance structure models (see, e.g., Prokhorov, 2008) are used to illustrate the theoretical results obtained in Section 2.

Consider a random vector $Z \in \mathcal{Z} \subset \mathbb{R}^q$ from P_{θ_0}, $\theta_0 \in \Theta$. It is assumed that $\mathbb{E}[Z] = 0$, $\mathbb{E}\{\|Z\|^4\} < \infty$ and $\mathbb{E}[ZZ'] = \Sigma(\theta_0)$. The matrix function $\Sigma(\theta)$ comes from a structural model, e.g., LISREL, MIMIC, factor analysis, random effects or simultaneous equations model. For a random sample $(Z_1, \cdots, Z_N)$, denote

$$S_i \equiv Z_i Z_i' \tag{3.1}$$

and

$$S \equiv \frac{1}{N}\sum_{i=1}^{N} S_i. \tag{3.2}$$

S satisfies a central limit theorem:

$$\sqrt{N}(vechS - vech\Sigma(\theta_0)) \to N(0, \Delta(\theta_0)), \tag{3.3}$$

where

$$\Delta(\theta_0) = \mathbb{V}(vechS_i) = \mathbb{E}[vechS_i vech'S_i] - vech\Sigma(\theta_0)vech'\Sigma(\theta_0)$$

and $vech$ denotes vertical vectorization of the lower triangle of a matrix. In general, $p \leq \frac{1}{2}q(q+1)$. Then, the moment function in (2.3) is

$$\underset{\frac{1}{2}q(q+1)\times 1}{m_N(\theta)} \equiv \frac{1}{N}\sum_{i=1}^{N} m(Z_i, \theta) = vechS - vech\Sigma(\theta), \tag{3.4}$$

where

$$\underset{\frac{1}{2}q(q+1)\times 1}{m(Z_i,\theta)} = vechS_i - vech\Sigma(\theta).$$

We are interested in the test of $H_0 : \Sigma(\theta) = \Sigma(c)$ against $H_1 : \Sigma(\theta) \neq \Sigma(c)$, where c is a constant vector. This type of test is fundamental in covariance literature and has been richly researched in the literature (Korin, 1968; Sugiura, 1969; Nagarsenker and Pillai, 1973; Yanagihara et al., 2004; etc.). We want to see how the Bartlett-type correction factor in (2.43) works.

To make the simulations simple, consider the problem of testing the covariance of two uncorrelated random variables (i.e. $q = 2$) from the standard normal population $\mathcal{N}(0, 1)$. The null and alternative hypotheses are, respectively:

$$\begin{aligned} H_0 &: \Sigma(\theta) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\ H_1 &: \Sigma(\theta) \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned} \tag{3.5}$$

Note that $\Sigma(\theta)$ is the generic population covariance matrix before imposing any restrictions such as 'uncorrelated' and 'standard normal' in this case. Denote

$$\Sigma(\theta) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

where

$$\theta = \begin{bmatrix} \sigma_1 \\ \sigma_{12} \\ \sigma_2 \end{bmatrix}.$$

Here $p = 3$ and $m = 3$. Therefore,

$$G(\theta) = -\frac{dvech'\Sigma(\theta)}{d\theta} = \begin{bmatrix} -2\sigma_1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2\sigma_2 \end{bmatrix},$$

$$D(\theta) = \frac{dvec'G(\theta)}{d\theta} = \begin{bmatrix} -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 \end{bmatrix},$$

$$C(\theta) = \frac{dvec'D(\theta)}{d\theta} = \underset{3\times 27}{0}.$$

Note that here the derivatives of the moment function do not depend on the data. Equivalently, the null and alterative could be rewritten as:

$$\begin{aligned} &H_0 : \underset{3\times 1}{g(\theta)} = 0, \\ &H_1 : g(\theta) \neq 0, \end{aligned} \tag{3.6}$$

where

$$g(\theta) = vech\Sigma(\theta) - vech \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 \\ \sigma_{12} \\ \sigma_2^2 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 - 1 \\ \sigma_{12} - 0 \\ \sigma_2^2 - 1 \end{bmatrix}.$$

Note that $r = 3$. Then,

$$A(\theta) = \frac{dg'(\theta)}{d\theta} = \begin{bmatrix} 2\sigma_1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2\sigma_2 \end{bmatrix}.$$

Under the null, $\sigma_1 = 1$, $\sigma_{12} = 0$, $\sigma_2 = 1$. Therefore,

$$G = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \tag{3.7}$$

$$D = \begin{bmatrix} -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 \end{bmatrix}, \tag{3.8}$$

$$C = \underset{3\times 27}{0}, \tag{3.9}$$

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}. \tag{3.10}$$

Using theorem 2.4 gives $\alpha_1 = -0.04167$, $\alpha_2 = -0.04167$, and $\alpha_3 = 0.01786$.

Generate a sample $(Z_1, \cdots, Z_N)$ from $\mathcal{N}(0, I_2)$. The forth moment of the data is

$$\Delta(\theta_0) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \tag{3.11}$$

which is used to standardize the GMM moments. Compute the DM test statistic before correction and further compute the Bartlett-type corrected DM test statistic using the correction factor (2.43).

The procedure described above is repeated 1000 times for each value of the sample size $N = 25, 50, 75, 100$. Then, plot the histograms of the corrected statistic (see Figure 1). It can be seen that the corrected DM test statistic is well approximated by the chi-squared distribution.

We also did simulations for other hypothesis tests such as $H_0 : \theta_i = \theta_j$ against $H_1 : \theta_i \neq \theta_j$ and $H_0 : \theta_i\theta_j = a$ against $H_1 : \theta_i\theta_j \neq a$ for some constant a. The former could be used to test stationary processes in economic application, see Section 4; the latter is similar to the one in Phillips and Park (1988) and Hansen (2006). The results for both tests are similar to the test of $H_0 : \Sigma(\theta) = \Sigma(c)$ against $H_1 : \Sigma(\theta) \neq \Sigma(c)$, so they are not presented here.

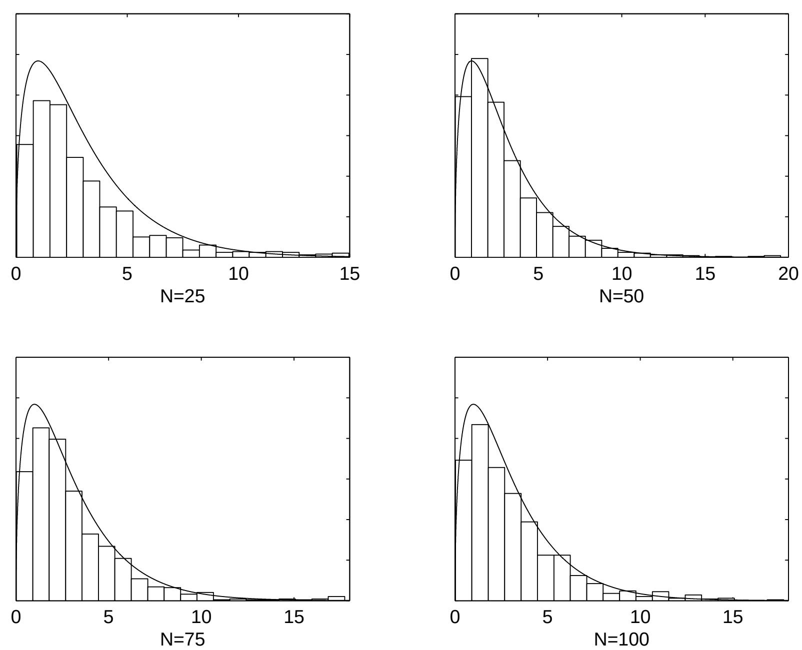


Figure 1: Histograms of the Bartlett-type corrected DM test statistic for the test of H_0 : $\Sigma(\theta) = \Sigma(c)$ against $H_1 : \Sigma(\theta) \neq \Sigma(c)$ with $q = 2$. The density of χ^2_3 is superimposed on the histograms.

An important question is: how does the Bartlett-type correction work? In order to demonstrate the effect of the Bartlett-type correction, we draw data from various distributions and various dimensions, and the quantiles as a function of probabilities are plotted. See Figure

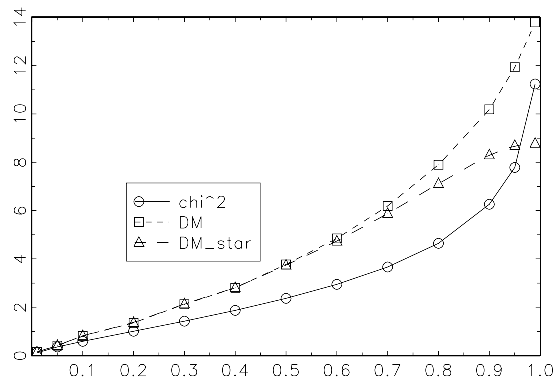
2, 3, 4 and 5. The quantile curve of chi-squared distribution, marked “chiˆ2”, is set as the benchmark in each panel. The uncorrected DM test statistic and corrected DM test statistic are marked “DM” and “DM_star” respectively. All Figures 2, 3, 4 and 5 show severe over-rejection of the uncorrected DM test statistic which is improved by the corrected DM test statistic, especially when sample sizes are small. It can be seen that in each panel, the corrected DM test statistic is closer to the chi-squared distribution than the uncorrected one. This confirms our expectation in the beginning of this paper.

Figure 2 shows the quantiles of the uncorrected and corrected DM test statistics as compared with chi-squared quantiles for the test of $H_0 : \Sigma(\theta) = \Sigma(c)$ against $H_1 : \Sigma(\theta) \neq \Sigma(c)$ with $q = 2$ for various sample sizes. The data are from $\mathcal{N}(0, 1)$, and the number of restrictions is $r = \frac{1}{2}q(q + 1)$. In Figure 2, as the value of sample size N increases, the DM test statistic before correction is closer to chi-square and the Bartlett-type correction is becoming less. This is theoretically reasonable.

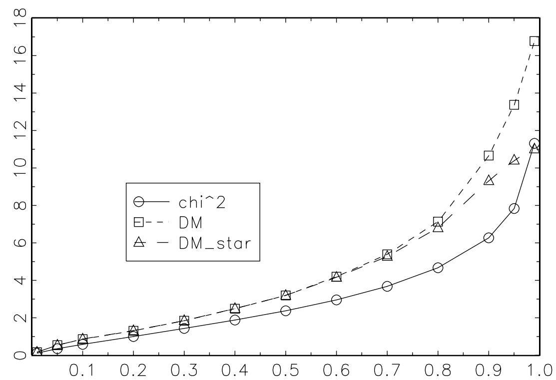
Figure 3 displays the quantiles of the uncorrected and corrected DM test statistics as compared with chi-squared quantiles for the test of $H_0 : \Sigma(\theta) = \Sigma(c)$ against $H_1 : \Sigma(\theta) \neq \Sigma(c)$ with $q = 1$ and $q = 2$. The data are from normal distribution, student's t distribution for 3 degrees of freedom and uniform distribution. The sample size is $N = 25$, and the number of restrictions is $r = \frac{1}{2}q(q+1)$. Figure 3 confirms that our theoretical results obtained in Section 2 are distribution-free.

Figure 4 presents the quantiles of the uncorrected and corrected DM test statistics as compared with chi-squared quantiles for the test of $H_0 : \Sigma(\theta) = \Sigma(c)$ against $H_1 : \Sigma(\theta) \neq \Sigma(c)$ with $q = 2$. The data are from normal distribution, student's t distribution for 3 degrees of freedom and uniform distribution. The sample size is $N = 25$, and the numbers of restrictions are $r = \frac{1}{2}q(q + 1) - 1 = 2$ and $r = \frac{1}{2}q(q + 1) - 2 = 1$. As analyzed in Section 2, increasing the number of restrictions does not necessarily make the correction larger. This is revealed in Figure 4.

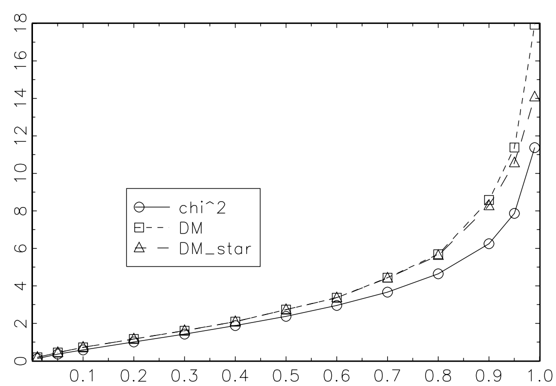
Figure 5 plots the quantiles of the uncorrected and corrected DM test statistics as compared with chi-squared quantiles for the test of $H_0 : \Sigma(\theta) = \Sigma(c)$ against $H_1 : \Sigma(\theta) \neq \Sigma(c)$ with $q = 1$ and $q = 2$. The data are from standard normal distribution. Figure 5 shows that DM test statistic before correction for $q = 1$ ($df = 1$) is much closer to chi-square distribution than $q = 2$ ($df = 3$), which holds for each of $N = 15$, $N = 25$ and $N = 50$. Note that the y-axis for $q = 2$ is compressed, otherwise, the deviation of the uncorrected DM test statistic for $q = 2$ from chi-square would be much more than that for $q = 1$. This is consistent with the findings of Hoogland and Boomsma (1998), who show that chi-square statistics are sensitive to the model size (i.e., the number of degrees of freedom)[1]: larger number of degrees of freedom requires larger sample sizes to obtain good behavior of the statistics. It is important to note that the Bartlett-type correction effect for $q = 2$ is almost the same as that for $q = 1$. This can be seen from the percentage correction. For example, for $N = 25$ and on a 5% significance level, the correction is about half for both $q = 1$ and $q = 2$.



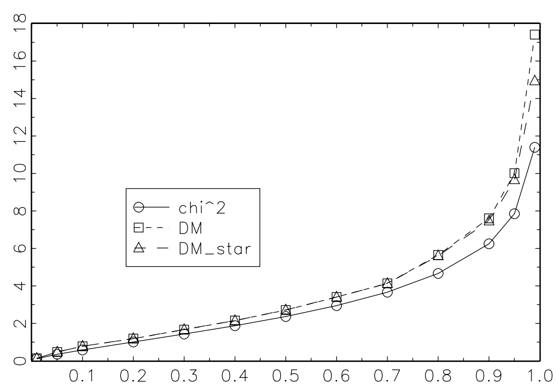
(a) $N = 15$



(b) $N = 25$



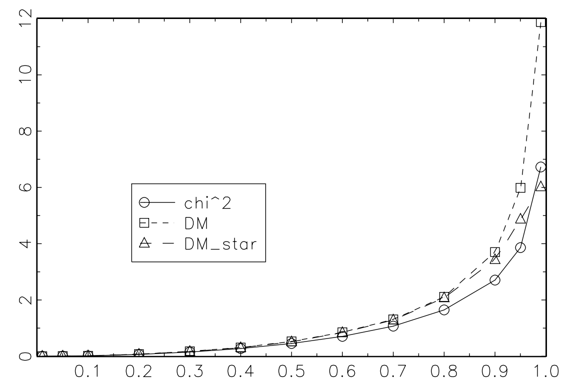
(c) $N = 50$



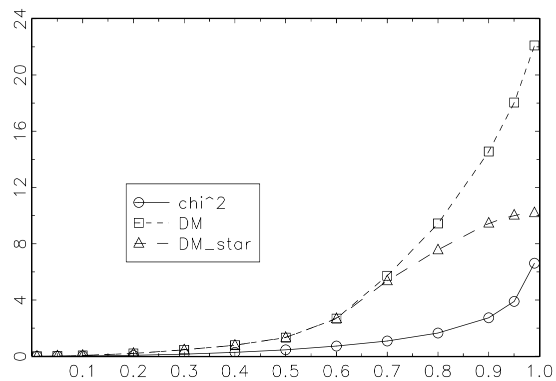
(d) $N = 75$

Figure 2: Plot of quantiles: chi-squared vs the uncorrected and corrected DM test statistics for the test of $H_0 : \Sigma(\theta) = \Sigma(c)$ against $H_1 : \Sigma(\theta) \neq \Sigma(c)$ with $q = 2$. The data are from $\mathcal{N}(0, 1)$, and the number of restrictions is $r = \frac{1}{2}q(q + 1)$.

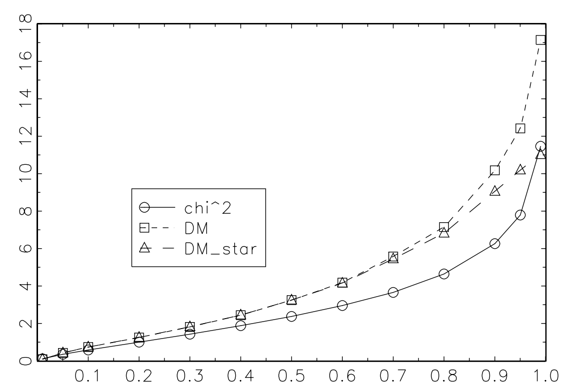
[1]For structural equation modelling, a model's degrees of freedom equals to the number of distinct elements in the covariance minus the number of parameters.



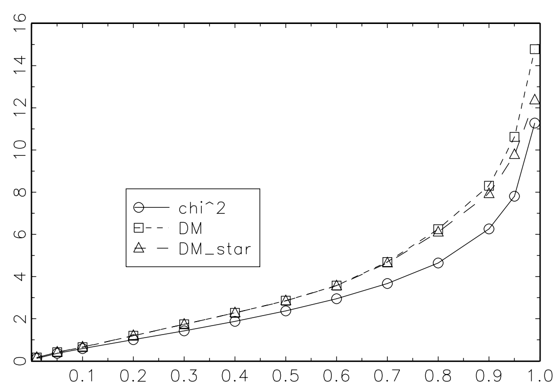
(a) $q = 1$, $\mathcal{N}(0, 1)$



(b) $q = 1$, Student's t Distribution for 3 df.

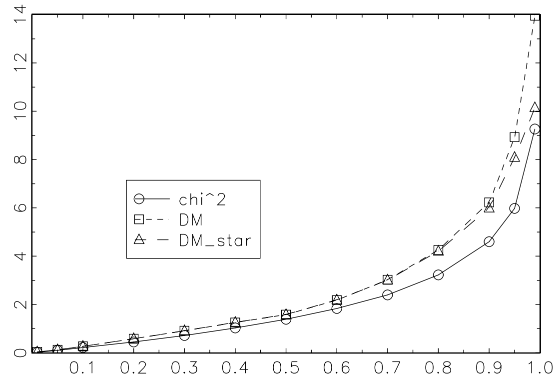


(c) $q = 2$, $\mathcal{N}(0, 0.25)$

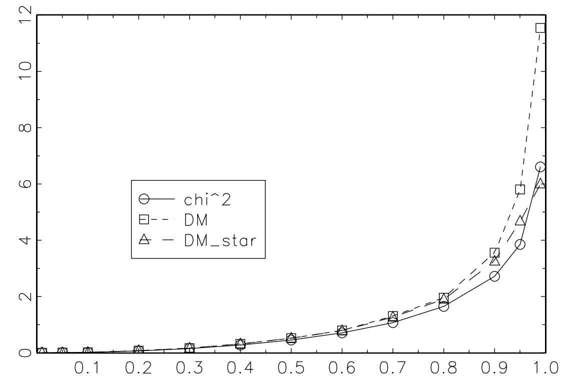


(d) $q = 2$, Uniform Distribution

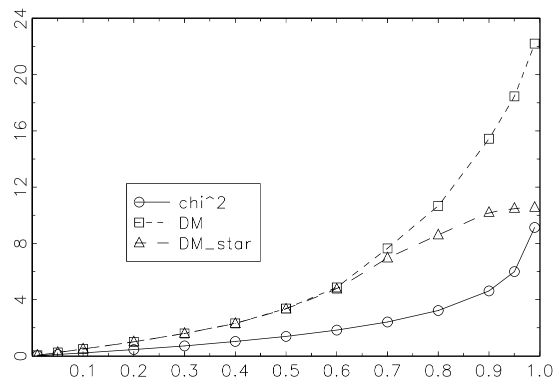
Figure 3: Plot of quantiles: chi-squared vs the uncorrected and corrected DM test statistics for the test of $H_0 : \Sigma(\theta) = \Sigma(c)$ against $H_1 : \Sigma(\theta) \neq \Sigma(c)$ with $q = 1$ and $q = 2$. The data are from normal distribution, student's t distribution for 3 degrees of freedom and uniform distribution. The sample size is $N = 25$, and the number of restrictions is $r = \frac{1}{2}q(q + 1)$.



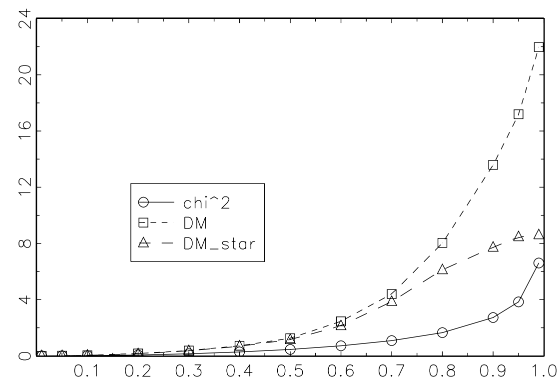
(a) $\mathcal{N}(0,1)$, $r=2$



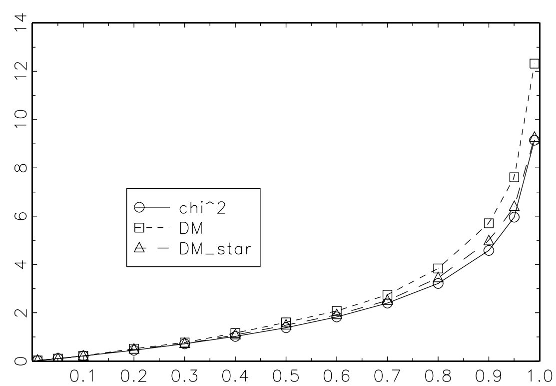
(b) $\mathcal{N}(0,1)$, $r=1$



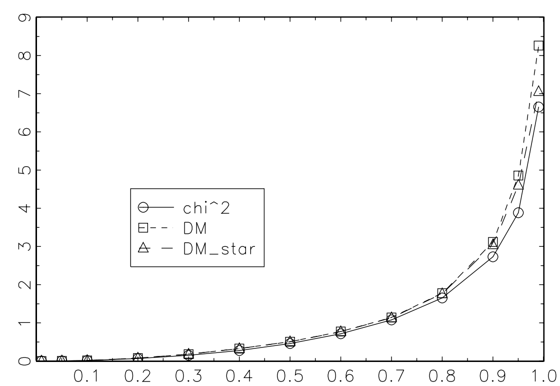
(c) Student's t Distribution for 3 df., $r=2$



(d) Student's t Distribution for 3 df., $r=1$

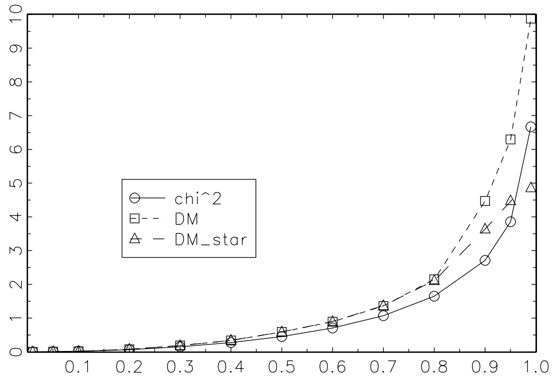


(e) Uniform Distribution, $r=2$

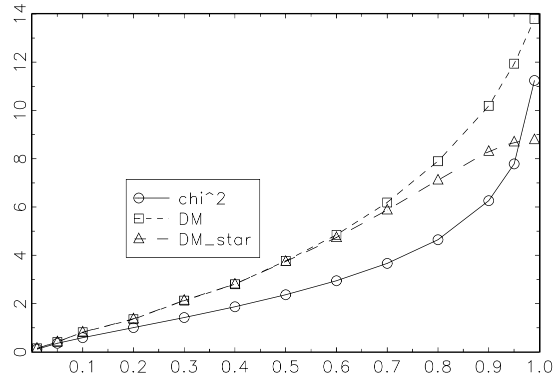


(f) Uniform Distribution, $r=1$

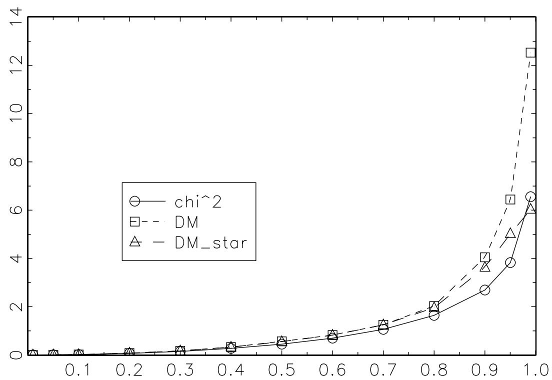
Figure 4: Plot of quantiles: chi-squared vs the uncorrected and corrected DM test statistics for the test of $H_0 : \Sigma(\theta) = \Sigma(c)$ against $H_1 : \Sigma(\theta) \neq \Sigma(c)$ with $q = 2$. The data are from normal distribution, student's t distribution for 3 degrees of freedom and uniform distribution. The sample size is $N = 25$, and the numbers of restrictions are $r = \frac{1}{2}q(q+1) - 1 = 2$ and $r = \frac{1}{2}q(q+1) - 2 = 1$.



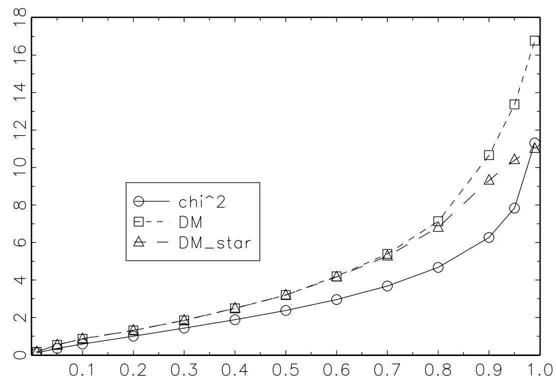
(a) $q = 1$, $N = 15$



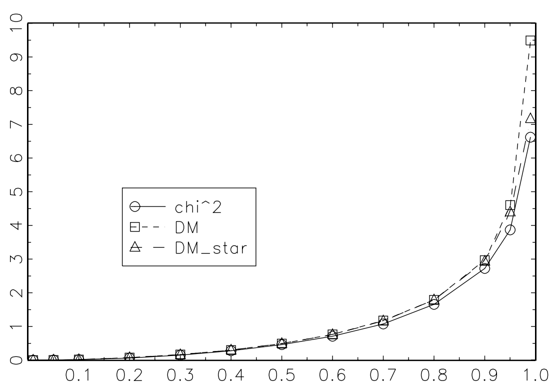
(b) $q = 2$, $N = 15$



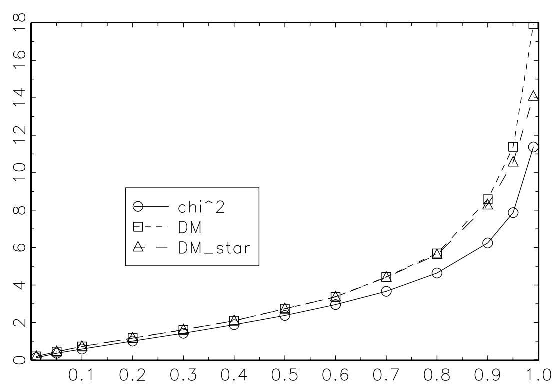
(c) $q = 1$, $N = 25$



(d) $q = 2$, $N = 25$



(e) $q = 1$, $N = 50$



(f) $q = 2$, $N = 50$

Figure 5: Plot of quantiles: chi-squared vs the uncorrected and corrected DM test statistics for the test of $H_0 : \Sigma(\theta) = \Sigma(c)$ against $H_1 : \Sigma(\theta) \neq \Sigma(c)$ with $q = 1$ and $q = 2$. The data are from $\mathcal{N}(0, 1)$.

4 Empirical Illustration

In labor economics literature, many articles study covariance structures of earnings (MaCurdy, 1982; Abowd and Card, 1987, 1989; Topel and Ward, 1992; Baker, 1997; Baker and Solon, 2003; etc.). A longer panel provides observations on higher-order autocovariances; however, in order to obtain a balanced panel, the cost for a longer panel is a smaller number of observations. For example, in Baker (1997), the sample sizes available for the 10-year panels of 1967-76 and 1977-86 are 992 and 1331 respectively, but the sample size available for the whole 20-year panel is only 534. On the other hand, as the panel gets longer, the number of degrees of freedom under usual null becomes larger, which, as mentioned earlier, generally requires larger sample sizes for chi-square statistics to maintain good behavior. This may explain the "troublesome" discrepancy puzzle on p358 of Baker (1997) that a longer panel reverses the original inference. In this section, we use part of the first sample in Abowd and Card (1989), earnings from 1969 to 1974, to demonstrate how the Bartlett-type correction has an effect on the outcomes of a hypothesis test.

The earnings data used in this section are drawn from the Panel Study of Income Dynamics (PSID), conducted at the Survey Research Center, Institute for Social Research, University of Michigan. The sample consists of men who were heads of household in every year from 1969 to 1974, who were between the ages of 21 and 64 in each year, and who reported positive earnings in each year. The sample used in the current paper contains a total of 1578 individuals. Individuals with average hourly earnings greater than \$100 or reported annual hours greater than 4680 were excluded. A detailed description of the PSID variables used is given in Appendix B. The covariances and correlations between the demeaned changes in log real annual earnings (in the 1967 dollars) are displayed in Table 1. Covariances are presented below diagonal, while correlations and their associated two-tailed prob-values are presented above diagonal. A correlation coefficient is statistically significant at 10% two-tailed significance level if its prob-value is less than 0.10.

The generic population covariance matrix for Table 1 can be denoted as

$$\Sigma(\theta) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} & \sigma_{25} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} & \sigma_{35} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 & \sigma_{45} \\ \sigma_{51} & \sigma_{52} & \sigma_{53} & \sigma_{54} & \sigma_5^2 \end{bmatrix}, \tag{4.1}$$

where

$$\theta = vec \begin{bmatrix} \sigma_1 & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} \\ \sigma_{21} & \sigma_2 & \sigma_{23} & \sigma_{24} & \sigma_{25} \\ \sigma_{31} & \sigma_{32} & \sigma_3 & \sigma_{34} & \sigma_{35} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4 & \sigma_{45} \\ \sigma_{51} & \sigma_{52} & \sigma_{53} & \sigma_{54} & \sigma_5 \end{bmatrix}. \tag{4.2}$$

Table 1: Covariances Between Changes in Earnings : PSID Males 1967-1974

	Covariance (correlation)[a] of :				
with :	$\Delta \ln e$ 69-70	$\Delta \ln e$ 70-71	$\Delta \ln e$ 71-72	$\Delta \ln e$ 72-73	$\Delta \ln e$ 73-74
$\Delta \ln e$ 69-70	0.228	-0.204	-0.006	0.018	-0.006
		(0)	(0.827)	(0.463)	(0.823)
$\Delta \ln e$ 70-71	-0.04418	0.205	-0.415	-0.082	0
			(0)	(0.001)	(0.994)
$\Delta \ln e$ 71-72	-0.00117	-0.08345	0.197	-0.347	-0.041
				(0)	(0.101)
$\Delta \ln e$ 72-73	0.003442	-0.01447	-0.06	0.152	-0.305
					(0)
$\Delta \ln e$ 73-74	-0.00102	-0.0000303	-0.00697	-0.04518	0.144

[a]Covariances are below diagonal, while correlations and their associated two-tailed prob-values are above diagonal.

The question Abowd and Card (1989) ask is whether the information in the covariance matrix in Table 1 could be adequately summarized by some relatively simple statistical model. Specifically, they ask whether a (possibly nonstationary) MA(2) process can serve as a model. Indeed, there are very few covariances (correlations) that are large or statistically significant at lags greater than two periods. In order to address this concern, two tests were performed using the DM test statistic.

The first one is to test for a nonstationary MA(2) representation of the changes in earnings. The changes in earnings have a nonstationary MA(2) representation if the covariances at lags greater than two periods are zero. The null is H_0 : changes in earnings is a nonstationary MA(2), and the alternative is H_1 : changes in earnings is not a nonstationary MA(2).

Equivalently, the null can be rewritten as

$$H_0 : \begin{bmatrix} \sigma_{41} \\ \sigma_{51} \\ \sigma_{52} \end{bmatrix} = \underset{3\times 1}{0}. \tag{4.3}$$

The second one is to test for a stationary MA(2) representation of the changes in earnings. By a stationary MA(2) representation, we mean (1) $cov(\Delta \ln e_t, \Delta \ln e_{t-j})$ depends only on j and does not change over t, and (2) $cov(\Delta \ln e_t, \Delta \ln e_{t-j})$ is zero for $|j| > 2$. The null is H_0 : changes in earnings is a stationary MA(2), and the alternative is H_1 : changes in earnings is not a stationary MA(2). Equivalently, the null can be rewritten as

$$\begin{aligned} H_0 : &\begin{bmatrix} \sigma_{41} \\ \sigma_{51} \\ \sigma_{52} \end{bmatrix} = \underset{3\times 1}{0}, \\ &\sigma_1 = \sigma_2 = \sigma_3 = \sigma_4 = \sigma_5, \\ &\sigma_{21} = \sigma_{32} = \sigma_{43} = \sigma_{54}, \\ &\sigma_{31} = \sigma_{42} = \sigma_{53}. \end{aligned} \tag{4.4}$$

The test results are presented in Table 2. Calculation shows that the values of the uncorrected DM test statistic for the two tests are 0.33247975 and 19.988882 respectively. Then, α_1, α_2, α_3 are computed and the Bartlett-type correction is conducted. For the first test, $\alpha_1 = 2.1219335$, $\alpha_2 = 0.052738023$, $\alpha_3 = 0$, and the value of the corrected DM test statistic is 0.33202897. For the second test, $\alpha_1 = -1.8877232$, $\alpha_2 = 0.96340811$, $\alpha_3 = 0.028180808$, and the value of DM test statistic after correction is 19.626226. The corrections are minor. The reason for the minor corrections is that the sample size $N = 1578$ is quite large for the current number of degrees of freedom in both tests. It is very important to note that we are using only part of the sample in Abowd and Card (1989). If the full sample ($df = 112$ for the first test and $df = 87$ for the second one) is used, the sample size $N = 1578$ might not be large enough and the DM test statistic might need to be corrected. Detailed analysis of this is beyond the scope of the current study, but is a useful direction for future research. Our goal rather is to demonstrate the effect of the Bartlett-type correction as the sample size becomes smaller.

What would happen if the sample size decreases? The answer is: if smaller sample sizes are used, Bartlett-type corrections may change the outcomes of the tests. Take the second test as an example. The results are presented in Table 3. As sample size decreases from N =1578 to 1400, 1200, 1000 and 900 (first N individuals), Bartlett-type corrections are

Table 2: Goodness-of-Fit Tests for Changes in Earnings: PSID Males 1967-1974

Goodness-of-Fit Test	DM Test Statistic		Asy. P-Value	
N=1578	Uncorrected	Corrected	Uncorrected	Corrected
I. Nonstationary MA(2) ($df = 3$)	0.33247975	0.33202897	0.95380852	0.95389632
II. Stationary MA(2) ($df = 12$)	19.988882	19.626226	0.067296564	0.074495171

becoming larger and larger. This is expected by the theoretical analysis in Section 2. Table 3 also reveals how the outcomes of the test change at various significance levels. For example, if $N = 900$, the value of the DM test statistic before correction is rejected at 5% significance level, however, we fail to reject the value after correction.

Table 3: Testing Stationary MA(2) for Changes in Earnings: PSID Males 1967-1974

Sample Size	DM Test Statistic		Asy. P-Value	
	Uncorrected	Corrected	Uncorrected	Corrected
N=1400	22.213672	21.638984	0.035193684	0.041770848
N=1200	24.153230	22.830842	0.019386430	0.029196854
N=1000	25.459123	22.119035	0.012790968	0.036207173
N=900	25.997606	20.348330	0.010742264	0.060777036

5 Concluding Remarks

This paper provides a study on the asymptotic expansions of the distribution of the DM test statistic. It covers all kinds of cases with nonlinear moments. If the moments are nonlinear, the nonlinearity requirement for the restrictions of a hypothesis test in Phillips and Park (1988) and Hansen (2006) can be relaxed. Besides, the theoretical results apply to possibly large number of degrees of freedom. In these aspects, the present theoretical results complement both Phillips and Park (1988) and Hansen (2006). After the theoretical results are obtained, simple simulations are performed to illustrate the seemingly complicated theory. Finally, the theoretical results are applied in the covariance structures of earnings to demonstrate how the Bartlett-type correction can play a role in the outcome of the hypothesis tests when the finite sample sizes are not large enough.

Edgeworth expansion for the DM test statistic is promising. In practice, an increasing number of large structural equation models (measured by large number of degrees of freedom) are estimated nowadays (Herzog et al., 2007; Kenny and McCoach, 2003). A large sample size is a remedy (Hoogland and Boomsma, 1998), but large sample sizes are not always possible to get. Another aspect in the practice is that data are often nonnormal. The present paper attacks these problems simultaneously. The advantage of our approach over other finite sample corrections used in structural equation modeling (e.g., Yuan and Bentler, 1998, 1999) is that our results are applicable to a much wider range of models.

Simulations and applications in the extensively existing nonlinear moment cases remain for future research.

References

ABOWD, J. M. AND D. CARD (1987): "Intertemporal Labor Supply and Long-Term Employment Contracts," *The American Economic Review*, 77, 50–68.

——— (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57, 411–445.

BAKER, M. (1997): "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," *Journal of Labor Economics*, 15, 338–375.

BAKER, M. AND G. SOLON (2003): "Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records," *Journal of Labor Economics*, 21, 289–321.

CRIBARI-NETO, F. AND G. M. CORDEIRO (1996): "On Bartlett and Bartlett-Type Corrections," *Econometric Reviews*, 15, 339–367.

GREENE, W. H. (2003): *Econometric Analysis*, Upper Saddle River, NJ : Prentice Hall, 5th ed.

HANSEN, B. E. (2006): "Edgeworth Expansions for the Wald and GMM Statistics for Nonlinear Restrictions," in *Econometric theory and practice: frontiers of analysis and applied research*, ed. by D. Corbae, S. N. Durlauf, and B. E. Hansen, Cambridge; New York: Cambridge University Press, 9–35.

HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

HERZOG, W., A. BOOMSMA, AND S. REINECKE (2007): "The Model-Size Effect on Traditional and Modified Tests of Covariance Structures," *Structural Equation Modeling*, 14, 361–390.

HOOGLAND, J. J. AND A. BOOMSMA (1998): "Robustness Studies in Covariance Structure Modeling," *Sociological Methods & Research*, 26, 329–367.

Kenny, D. A. and D. B. McCoach (2003): "Effect of the Number of Variables on Measures of Fit in Structural Equation Modeling," *Structural Equation Modeling*, 10, 333–351.

Kollo, T. and D. v. Rosen (2005): *Advanced Multivariate Statistics with Matrices*, Dordrecht : Springer.

Korin, B. P. (1968): "On the Distribution of a Statistic Used for Testing a Covariance Matrix," *Biometrika*, 55, 171–178.

MaCurdy, T. E. (1982): "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, 18, 83–114.

Nagarsenker, B. N. and K. C. S. Pillai (1973): "Distribution of the Likelihood Ratio Criterion for Testing a Hypothesis Specifying a Covariance Matrix," *Biometrika*, 60, 359–364.

Newey, W. K. and D. McFadden (1994): "Chapter 36 Large Sample Estimation and Hypothesis Testing," Elsevier, vol. 4 of *Handbook of Econometrics*, 2111–2245.

Newey, W. K. and K. D. West (1987): "Hypothesis Testing with Efficient Method of Moments Estimation," *International Economic Review*, 28, 777–787.

Phillips, P. and J. Y. Park (1988): "On The Formulation of Wald Tests of Nonlinear Restrictions," *Econometrica*, 56, 1065–1083.

Prokhorov, A. (2008): "On Relative Efficiency of Quasi-MLE and GMM Estimators of Covariance Structure Models," *Economics Letters*, In Press, Corrected Proof, –.

Rothenberg, T. J. (1984): "Chapter 15 Approximating the Distributions of Econometric Estimators and Test Statistics," Elsevier, vol. 2 of *Handbook of Econometrics*, 881–935.

Sugiura, N. (1969): "Asymptotic Expansions of the Distributions of the Likelihood Ratio Criteria for Covariance Matrix," *The Annals of Mathematical Statistics*, 40, 2051–2063.

Topel, R. H. and M. P. Ward (1992): "Job Mobility and the Careers of Young Men," *The Quarterly Journal of Economics*, 107, 439–479.

Yanagihara, H., C. Matsumoto, and T. Tonda (2004): "Asymptotic Expansion of the Null Distribution of the Modified normal likelihood ratio criterion for testing $\Sigma = \Sigma_0$ under Nonnormality," *Hiroshima Mathematical Journal*, 34, 81–100.

Yuan, K.-H. and P. M. Bentler (1998): "Normal Theory Based Test Statistics in Structural Equation Modeling," *British Journal of Mathematical and Statistical Psychology*, 51, 289–309.

——— (1999): "F Tests for Mean and Covariance Structure Analysis," *Journal of Educational and Behavioral Statistics*, 24, 225–243.

A Proofs

Proof of Lemma 2.1: See Kollo and Rosen (2005), p151 and p280. □

Proof of Theorem 2.2: From (2.12),

$$DM \cong 1_{DM} + 2_{DM} + 3_{DM} + 4_{DM}, \tag{A.1}$$

where,

$$1_{DM} = vec' G_N(\hat{\theta}_N) M_1 vec G_N(\hat{\theta}_N), \tag{A.2}$$

$$2_{DM} = m'_N(\hat{\theta}_N) M_2 vec D_N(\hat{\theta}_N), \tag{A.3}$$

$$3_{DM} = -N^{-1/2} vec' G_N(\hat{\theta}_N) M_3 vec D_N(\hat{\theta}_N), \tag{A.4}$$

$$4_{DM} = N^{-1}\frac{1}{4} vec' D_N(\hat{\theta}_N) M_4 vec D_N(\hat{\theta}_N). \tag{A.5}$$

Taking Taylor expansions of $m_N(\hat{\theta}_N)$, $vecG_N(\hat{\theta}_N)$ and $vecD_N(\hat{\theta}_N)$ about θ_0 and using (2.8) and (2.9) yield:

$$\begin{aligned}
\underset{m\times 1}{m_N(\hat{\theta}_N)} &= m_N(\theta_0) + G'(\hat{\theta}_N - \theta_0) + \frac{1}{2}[I_m \otimes (\hat{\theta}_N - \theta_0)']D'(\hat{\theta}_N - \theta_0) + o_p \\
&= -N^{-1/2}\bar{q} + N^{-1/2} G' B^{-1} G \bar{q} + N^{-1}\frac{1}{2}(I_m \otimes \bar{q}' G' B^{-1}) D' B^{-1} G \bar{q} + o_p, \\
\underset{pm\times 1}{vecG_N(\hat{\theta}_N)} &= vecG + D'(\hat{\theta}_N - \theta_0) + \frac{1}{2}[I_{pm} \otimes (\hat{\theta}_N - \theta_0)']C'(\hat{\theta}_N - \theta_0) + o_p \\
&= vecG + N^{-1/2} D' B^{-1} G \bar{q} + N^{-1}\frac{1}{2}(I_{pm} \otimes \bar{q}' G' B^{-1}) C' B^{-1} G \bar{q} + o_p, \\
\underset{p^2m\times 1}{vecD_N(\hat{\theta}_N)} &= vecD + C'(\hat{\theta}_N - \theta_0) + o_p \\
&= vecD + N^{-1/2} C' B^{-1} G \bar{q} + o_p.
\end{aligned}$$

Note that we do not need to expand $vecD_N(\hat{\theta}_N)$ further for our purpose. Substituting these expressions into (A.2)-(A.5) gives:

$$\begin{aligned}
1_{DM} &= vec' G_N(\hat{\theta}_N) M_1 vec G_N(\hat{\theta}_N) \\
&= vec' G M_1 vec G + N^{-1/2} 2\bar{q}' G' B^{-1} D M_1 vec G \\
&\quad + N^{-1}[\bar{q}' G' B^{-1} D M_1 D' B^{-1} G \bar{q} + \bar{q}' G' B^{-1} C (I_{pm} \otimes B^{-1} G \bar{q}) M_1 vec G] \\
&\quad + o_p \\
&= \bar{q}' P \bar{q} + N^{-1/2} u_1(\bar{q}) + N^{-1} v_1(\bar{q}) + o_p,
\end{aligned} \tag{A.6}$$

where

$$\underset{m\times m}{P} \equiv G' H G$$

is the projection matrix, and

$$\begin{aligned}
u_1(\bar{q}) &= 2\bar{q}' G' B^{-1} D M_1 vec G, \\
v_1(\bar{q}) &= \bar{q}' G' B^{-1} D M_1 D' B^{-1} G \bar{q} + \bar{q}' G' B^{-1} C (I_{pm} \otimes B^{-1} G \bar{q}) M_1 vec G;
\end{aligned}$$

$$
\begin{aligned}
2_{DM} &= m_N'(\hat{\theta}_N)M_2vecD_N(\hat{\theta}_N) \\
&= -N^{-1/2}\bar{q}'M_2vecD - N^{-1}\bar{q}'M_2C'B^{-1}G\bar{q} \\
&\quad + N^{-1/2}\bar{q}'G'B^{-1}GM_2vecD + N^{-1}\bar{q}'G'B^{-1}GM_2C'B^{-1}G\bar{q} \\
&\quad + N^{-1}\frac{1}{2}\bar{q}'G'B^{-1}D(I_m \otimes B^{-1}G\bar{q})M_2vecD + o_p \\
&= N^{-1/2}(\bar{q}'G'B^{-1}M_2vecD - \bar{q}'M_2vecD) \\
&\quad + N^{-1}[\bar{q}'G'B^{-1}GM_2C'B^{-1}G\bar{q} - \bar{q}'M_2C'B^{-1}G\bar{q} \\
&\qquad + \frac{1}{2}\bar{q}'G'B^{-1}D(I_m \otimes B^{-1}G\bar{q})M_2vecD] + o_p \\
&= N^{-1/2}u_2(\bar{q}) + N^{-1}v_2(\bar{q}) + o_p,
\end{aligned} \tag{A.7}
$$

where

$$
\begin{aligned}
u_2(\bar{q}) &= \bar{q}'G'B^{-1}GM_2vecD - \bar{q}'M_2vecD \\
&= \bar{q}'(G'B^{-1}G - I_m)M_2vecD, \\
v_2(\bar{q}) &= \bar{q}'G'B^{-1}GM_2C'B^{-1}G\bar{q} - \bar{q}'M_2C'B^{-1}G\bar{q} + \frac{1}{2}\bar{q}'G'B^{-1}D(I_m \otimes B^{-1}G\bar{q})M_2vecD \\
&= \bar{q}'(G'B^{-1}G - I_m)M_2C'B^{-1}G\bar{q} + \frac{1}{2}\bar{q}'G'B^{-1}D(I_m \otimes B^{-1}G\bar{q})M_2vecD;
\end{aligned}
$$

$$
\begin{aligned}
3_{DM} &= -N^{-1/2}vec'G_N(\hat{\theta}_N)M_3vecD_N(\hat{\theta}_N) \\
&= -N^{-1/2}vec'GM_3vecD - N^{-1}vec'GM_3C'B^{-1}G\bar{q} - N^{-1}\bar{q}'G'B^{-1}DM_3vecD + o_p \\
&= N^{-1/2}u_3(\bar{q}) + N^{-1}v_3(\bar{q}) + o_p,
\end{aligned} \tag{A.8}
$$

where

$$
\begin{aligned}
u_3(\bar{q}) &= -vec'GM_3vecD, \\
v_3(\bar{q}) &= -vec'GM_3C'B^{-1}G\bar{q} - \bar{q}'G'B^{-1}DM_3vecD \\
&= -\bar{q}'G'B^{-1}CM_3'vecG - \bar{q}'G'B^{-1}DM_3vecD;
\end{aligned}
$$

$$
\begin{aligned}
4_{DM} &= N^{-1}\frac{1}{4}vec'D_N(\hat{\theta}_N)M_4vecD_N(\hat{\theta}_N) \\
&= N^{-1}\frac{1}{4}vec'DM_4vecD + o_p \\
&= N^{-1}v_4(\bar{q}) + o_p,
\end{aligned} \tag{A.9}
$$

where

$$
v_4(\bar{q}) = \frac{1}{4}vec'DM_4vecD.
$$

Finally, substituting (A.6)-(A.9) back into (A.1) gives equation (2.17). □

Proof of Lemma 2.3: From Theorem 2.2, if $u_i(\bar{q})$ $(i = 1, 2, 3)$ and $v_i(\bar{q})$ $(i = 1, 2, 3, 4)$ could be rewritten as

$$
u_i(\bar{q}) = vec'J_i(\bar{q} \otimes \bar{q} \otimes \bar{q}), \tag{A.10}
$$

$$
v_i(\bar{q}) = tr[L_i(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')], \tag{A.11}
$$

then,

$$u(\bar{q}) = vec'J(\bar{q}\otimes\bar{q}\otimes\bar{q}),$$
$$v(\bar{q}) = tr[L(\bar{q}\bar{q}'\otimes\bar{q}\bar{q}')]$$

where

$$vecJ = vecJ_1 + vecJ_2 + vecJ_3,$$

and

$$L = L_1 + L_2 + L_3 + L_4.$$

Therefore, the proof is reduced to show (A.10) and (A.11).

Using

$$(A\otimes C)(B\otimes D) = (AB)\otimes(CD),$$
$$K_{p,q}vecA = vec(A'),$$
$$A\otimes B = K_{p,r}(B\otimes A)K_{s,q},$$

for $A: p\times q$ and $B: r\times s$ where K is the commutation matrix, we can rewrite (2.21):

$$\begin{aligned} u_1(\bar{q}) &= 2\bar{q}'G'B^{-1}D(I_m\otimes\mathbb{H}G\bar{q})vec(\bar{q}'G'\mathbb{H}G) \\ &= 2\bar{q}'G'\mathbb{H}G(I_m\otimes\bar{q}'G'\mathbb{H})(\bar{q}'G'B^{-1}\otimes I_{pm})vec(D') \\ &= 2\bar{q}'G'\mathbb{H}G(I_m\otimes\bar{q}'G'\mathbb{H})(I_{pm}\otimes\bar{q}'G'B^{-1})vecD \\ &= 2(\bar{q}'G'\mathbb{H}G\otimes\bar{q}'G'\mathbb{H}\otimes\bar{q}'G'B^{-1})vecD \\ &= 2(\bar{q}'\otimes\bar{q}'\otimes\bar{q}')(G'\mathbb{H}G\otimes G'\mathbb{H}\otimes G'B^{-1})vecD \\ &= vec'J_1(\bar{q}\otimes\bar{q}\otimes\bar{q}), \end{aligned} \tag{A.12}$$

where

$$vecJ_1 = 2(G'\mathbb{H}G\otimes G'\mathbb{H}\otimes G'B^{-1})vecD. \tag{A.13}$$

Let

$$R_1 = (\mathbb{H}G\otimes B^{-1}G)(\bar{q}\bar{q}'\otimes\bar{q}\bar{q}')(G'\mathbb{H}\otimes G'B^{-1}),$$

partition $vecD$ as

$$\underset{p^2m\times 1}{vecD} = \begin{bmatrix} V_{D1} \\ V_{D2} \\ \vdots \\ V_{Dm} \end{bmatrix} \tag{A.14}$$

where each subvector V_{Di} is $p^2\times 1$, and let

$$V_D = V_{D1}V'_{D1} + V_{D2}V'_{D2} + \cdots + V_{Dm}V'_{Dm}. \tag{A.15}$$

Then, since

$$\begin{aligned}(I_m \otimes \bar{q}'G'\mathbb{H})D'B^{-1}G\bar{q} &= (I_m \otimes \bar{q}'G'\mathbb{H})(\bar{q}'G'B^{-1} \otimes I_{pm})vec(D') \\ &= (I_m \otimes \bar{q}'G'\mathbb{H})(I_{pm} \otimes \bar{q}'G'B^{-1})vecD \\ &= (I_m \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'B^{-1})vecD,\end{aligned}$$

the first term of $v_1(\bar{q})$ in (2.24) becomes

$$\begin{aligned}&\bar{q}'G'B^{-1}D(I_m \otimes \mathbb{H}G\bar{q})(I_m \otimes \bar{q}'G'\mathbb{H})D'B^{-1}G\bar{q} \\ &\quad = vec'D(I_m \otimes \mathbb{H}G\bar{q} \otimes B^{-1}G\bar{q})(I_m \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'B^{-1})vecD \\ &\quad = vec'D(I_m \otimes R_1)vecD \\ &\quad = \begin{bmatrix} V'_{D1} & V'_{D2} & \cdots & V'_{Dm} \end{bmatrix} \begin{bmatrix} R_1 & & & 0 \\ & R_1 & & \\ & & \ddots & \\ 0 & & & R_1 \end{bmatrix} \begin{bmatrix} V_{D1} \\ V_{D2} \\ \vdots \\ V_{Dm} \end{bmatrix} \qquad \text{(A.16)} \\ &\quad = V'_{D1}R_1V_{D1} + V'_{D2}R_1V_{D2} + \cdots + V'_{Dm}R_1V_{Dm} \\ &\quad = tr[(V_{D1}V'_{D1} + V_{D2}V'_{D2} + \cdots + V_{Dm}V'_{Dm})R_1] \\ &\quad = tr[V_D(\mathbb{H}G \otimes B^{-1}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')(G'\mathbb{H} \otimes G'B^{-1})] \\ &\quad = tr[(G'\mathbb{H} \otimes G'B^{-1})V_D(\mathbb{H}G \otimes B^{-1}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')].\end{aligned}$$

Similarly, let

$$\begin{aligned}R_2 &= (\mathbb{H}G \otimes B^{-1}G)(\bar{q} \otimes \bar{q}), \\ R_3 &= \bar{q}'G'\mathbb{H},\end{aligned}$$

partition $G'B^{-1}C$ and $vecG$ as

$$\underset{m \times p^2 m}{G'B^{-1}C} = \begin{bmatrix} M_{GC1} & M_{GC2} & \cdots & M_{GCm} \end{bmatrix}, \qquad \text{(A.17)}$$

$$\underset{pm \times 1}{vecG} = \begin{bmatrix} V_{G1} \\ V_{G2} \\ \vdots \\ V_{Gm} \end{bmatrix} \qquad \text{(A.18)}$$

where M_{GCi} and V_{Gi} are $m \times p^2$ and $p \times 1$ respectively, and let

$$M_V = M'_{GC1} \otimes V'_{G1} + M'_{GC2} \otimes V'_{G2} + \cdots + M'_{GCm} \otimes V'_{Gm}. \qquad \text{(A.19)}$$

Then, since

$$\begin{aligned}\bar{q}'m\bar{q}'M(\bar{q} \otimes \bar{q}) &= m'\bar{q}\bar{q}'M(\bar{q} \otimes \bar{q}) \\ &= [(\bar{q} \otimes \bar{q})'M' \otimes m']vec(\bar{q}\bar{q}') \\ &= (\bar{q} \otimes \bar{q})'(M' \otimes m')(\bar{q} \otimes \bar{q}) \\ &= tr[(M' \otimes m')(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')]\end{aligned} \qquad \text{(A.20)}$$

for some vector m and matrix M of appropriate sizes, the second term of $v_1(\bar{q})$ in (2.24) becomes

$$\begin{aligned}
&\bar{q}'G'B^{-1}C(I_{pm}\otimes B^{-1}G\bar{q})(I_m\otimes \mathbb{H}G\bar{q})(I_m\otimes \bar{q}'G'\mathbb{H})vecG\\
&\quad= \bar{q}'G'B^{-1}C(I_m\otimes R_2)(I_m\otimes R_3)vecG\\
&\quad= \bar{q}'\begin{bmatrix} M_{GC1} & M_{GC2} & \cdots & M_{GCm}\end{bmatrix}\begin{bmatrix} R_2 & & & 0\\ & R_2 & & \\ & & \ddots & \\ 0 & & & R_2\end{bmatrix}\begin{bmatrix} R_3 & & & 0\\ & R_3 & & \\ & & \ddots & \\ 0 & & & R_3\end{bmatrix}\begin{bmatrix} V_{G1}\\ V_{G2}\\ \vdots\\ V_{Gm}\end{bmatrix}\\
&\quad= \sum_{i=1}^{m}(\bar{q}'M_{GCi}R_2R_3V_{Gi})\\
&\quad= tr\sum_{i=1}^{m}[\bar{q}'M_{GCi}(\mathbb{H}G\otimes B^{-1}G)(\bar{q}\otimes\bar{q})\bar{q}'G'\mathbb{H}V_{Gi}]\\
&\quad= tr\sum_{i=1}^{m}[\bar{q}'G'\mathbb{H}V_{Gi}\bar{q}'M_{GCi}(\mathbb{H}G\otimes B^{-1}G)(\bar{q}\otimes\bar{q})]\\
&\quad= tr\sum_{i=1}^{m}\left\{\{[(G'\mathbb{H}\otimes G'B^{-1})M'_{GCi}]\otimes V'_{Gi}\mathbb{H}G\}(\bar{q}\bar{q}'\otimes\bar{q}\bar{q}')\right\}\\
&\quad= tr\sum_{i=1}^{m}[(G'\mathbb{H}\otimes G'B^{-1})(M'_{GCi}\otimes V'_{Gi})(I_m\otimes\mathbb{H}G)(\bar{q}\bar{q}'\otimes\bar{q}\bar{q}')]\\
&\quad= tr[(G'\mathbb{H}\otimes G'B^{-1})M_V(I_m\otimes\mathbb{H}G)(\bar{q}\bar{q}'\otimes\bar{q}\bar{q}')].
\end{aligned} \tag{A.21}$$

From (A.16) and (A.21), (2.24) could be rewritten as

$$v_1(\bar{q}) = tr[L_1(\bar{q}\bar{q}'\otimes\bar{q}\bar{q}')], \tag{A.22}$$

where

$$L_1 = (G'\mathbb{H}\otimes G'B^{-1})V_D(\mathbb{H}G\otimes B^{-1}G) + (G'\mathbb{H}\otimes G'B^{-1})M_V(I_m\otimes\mathbb{H}G). \tag{A.23}$$

Similar to $u_1(\bar{q})$,

$$\begin{aligned}
u_2(\bar{q}) &= \bar{q}'(G'B^{-1}G - I_m)(I_m\otimes\bar{q}'G'\mathbb{H}\otimes\bar{q}'G'\mathbb{H})vecD\\
&= (\bar{q}'\otimes\bar{q}'\otimes\bar{q}')[(G'B^{-1}G - I_m)\otimes G'\mathbb{H}\otimes G'\mathbb{H}]vecD\\
&= vec'J_2(\bar{q}\otimes\bar{q}\otimes\bar{q}),
\end{aligned} \tag{A.24}$$

where

$$vecJ_2 = [(G'B^{-1}G - I_m)\otimes G'\mathbb{H}\otimes G'\mathbb{H}]vecD. \tag{A.25}$$

The first term of $v_2(\bar{q})$ in (2.25) can be written as

$$\bar{q}'G'B^{-1}C(I_m\otimes\mathbb{H}G\bar{q}\otimes\mathbb{H}G\bar{q})(G'B^{-1}G - I_m)\bar{q}.$$

Since

$$\begin{aligned}
(G'B^{-1}G - I_m)\bar{q} &= vec[\bar{q}'(G'B^{-1}G - I_m)]\\
&= (I_m\otimes\bar{q}')vec(G'B^{-1}G - I_m),
\end{aligned}$$

and $vec(G'B^{-1}G - I_m)$ could be partitioned as

$$vec(G'B^{-1}G - I_m) = \begin{bmatrix} V_{GI1} \\ V_{GI2} \\ \vdots \\ V_{GIm} \end{bmatrix} \tag{A.26}$$

where V_{GIi} is $m \times 1$, we may mimic the second term of $v_1(\bar{q})$ and rewrite the first term of $v_2(\bar{q})$ further as

$$\begin{aligned} & tr\sum_{i=1}^{m}[\bar{q}' M_{GCi}(\mathbb{H}G \otimes \mathbb{H}G)(\bar{q} \otimes \bar{q})\bar{q}' V_{GIi}] \\ & \quad = tr[(G'\mathbb{H} \otimes G'\mathbb{H})M_{VI}(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')], \end{aligned} \tag{A.27}$$

where

$$M_{VI} = M'_{GC1} \otimes V'_{GI1} + M'_{GC2} \otimes V'_{GI2} + \cdots + M'_{GCm} \otimes V'_{GIm}. \tag{A.28}$$

Similar to the first term of $v_1(\bar{q})$, since

$$\bar{q}'G'B^{-1}D = vec'(\bar{q}'G'B^{-1}D) = vec'D(I_{pm} \otimes B^{-1}G\bar{q}),$$

the second term of $v_2(\bar{q})$ in (2.25) could be rewritten as

$$\begin{aligned} & \frac{1}{2}vec'D(I_m \otimes B^{-1}G\bar{q} \otimes B^{-1}G\bar{q})(I_m \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H})vecD \\ & \quad = \frac{1}{2}tr[V_D(B^{-1}G \otimes B^{-1}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')(G'\mathbb{H} \otimes G'\mathbb{H})] \\ & \quad = tr[\frac{1}{2}(G'\mathbb{H} \otimes G'\mathbb{H})V_D(B^{-1}G \otimes B^{-1}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}'))]. \end{aligned} \tag{A.29}$$

From (A.27) and (A.29), we have

$$v_2(\bar{q}) = tr[L_2(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')], \tag{A.30}$$

where

$$L_2 = (G'\mathbb{H} \otimes G'\mathbb{H})M_{VI} + \frac{1}{2}(G'\mathbb{H} \otimes G'\mathbb{H})V_D(B^{-1}G \otimes B^{-1}G). \tag{A.31}$$

Since

$$\begin{aligned} & vec'G(I_m \otimes \mathbb{H}G\bar{q}) \\ & \quad = [(I_m \otimes \bar{q}'G'\mathbb{H})vecG]' \\ & \quad = \bar{q}'G'\mathbb{H}G, \end{aligned}$$

(2.23) becomes

$$\begin{aligned} u_3(\bar{q}) &= -\bar{q}'G'\mathbb{H}G(I_m \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H})vecD \\ &= -(\bar{q}' \otimes \bar{q}' \otimes \bar{q}')(G'\mathbb{H}G \otimes G'\mathbb{H} \otimes G'\mathbb{H})vecD \\ &= vec'J_3(\bar{q} \otimes \bar{q} \otimes \bar{q}), \end{aligned} \tag{A.32}$$

where

$$vecJ_3 = -(G'\mathbb{H}G \otimes G'\mathbb{H} \otimes G'\mathbb{H})vecD. \tag{A.33}$$

Similar to the second term of $v_1(\bar{q})$, the first term of $v_3(\bar{q})$ in (2.26) could be rewritten as

$$\begin{aligned}
&-\bar{q}'G'B^{-1}C(I_m \otimes \mathbb{H}G\bar{q} \otimes \mathbb{H}G\bar{q})(I_m \otimes \bar{q}'G'\mathbb{H})vecG \\
&\qquad = tr\sum_{i=1}^{m}[-\bar{q}'M_{GCi}(\mathbb{H}G \otimes \mathbb{H}G)(\bar{q} \otimes \bar{q})\bar{q}'G'\mathbb{H}V_{Gi}] \\
&\qquad = tr[-(G'\mathbb{H} \otimes G'\mathbb{H})M_V(I_m \otimes \mathbb{H}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')].
\end{aligned} \tag{A.34}$$

Similar to the second term of $v_2(\bar{q})$, the second term of $v_3(\bar{q})$ in (2.26) could be rewritten as

$$\begin{aligned}
&-\bar{q}'G'B^{-1}D(I_m \otimes \mathbb{H}G\bar{q})(I_m \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H})vecD \\
&\qquad = -vec'D(I_{pm} \otimes B^{-1}G\bar{q})(I_m \otimes \mathbb{H}G\bar{q})(I_m \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H})vecD \\
&\qquad = -vec'D(I_m \otimes \mathbb{H}G\bar{q} \otimes B^{-1}G\bar{q})(I_m \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H})vecD \\
&\qquad = tr[-V_D(\mathbb{H}G \otimes B^{-1}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')(G'\mathbb{H} \otimes G'\mathbb{H})] \\
&\qquad = tr[-(G'\mathbb{H} \otimes G'\mathbb{H})V_D(\mathbb{H}G \otimes B^{-1}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')].
\end{aligned} \tag{A.35}$$

From (A.34) and (A.35), we have

$$v_3(\bar{q}) = tr[L_3(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')], \tag{A.36}$$

where

$$L_3 = -(G'\mathbb{H} \otimes G'\mathbb{H})M_V(I_m \otimes \mathbb{H}G) - (G'\mathbb{H} \otimes G'\mathbb{H})V_D(\mathbb{H}G \otimes B^{-1}G). \tag{A.37}$$

Similar to the first term of $v_1(\bar{q})$, $v_4(\bar{q})$ in (2.27) could be easily rewritten as

$$\begin{aligned}
v_4(\bar{q}) &= \frac{1}{4}tr[V_D(\mathbb{H}G \otimes \mathbb{H}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')(G'\mathbb{H} \otimes G'\mathbb{H})] \\
&= tr[\frac{1}{4}(G'\mathbb{H} \otimes G'\mathbb{H})V_D(\mathbb{H}G \otimes \mathbb{H}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')] \\
&= tr[L_4(\bar{q}\bar{q}' \otimes \bar{q}\bar{q})],
\end{aligned} \tag{A.38}$$

where

$$L_4 = \frac{1}{4}(G'\mathbb{H} \otimes G'\mathbb{H})V_D(\mathbb{H}G \otimes \mathbb{H}G). \tag{A.39}$$

By (A.12), (A.22), (A.24), (A.30), (A.32), (A.36) and (A.38), (A.10) and (A.11) are obtained, thus finishing the proof. □

Proof of Theorem 2.4: First, a_i and b_i are defined (Phillips and Park, 1988) as

$$a_i = tr(A_i) \ \ (i = 0, 1, 2), \tag{A.40}$$

where

$$\begin{aligned}
A_0 &= L[(I + K_{m,m})(\bar{P} \otimes \bar{P}) + vec\bar{P}vec'\bar{P}], \\
A_1 &= L[(I + K_{m,m})(\bar{P} \otimes P + P \otimes \bar{P}) + vec\bar{P}vec'P + vecPvec'\bar{P}], \\
A_2 &= L[(I + K_{m,m})(P \otimes P) + vecPvec'P];
\end{aligned}$$

$$b_i = vec'JB_i vecJ \ \ (i = 1, 2, 3), \tag{A.41}$$

where

$$\begin{aligned}
B_0 &= H(\bar{P} \otimes \bar{P} \otimes \bar{P}) + H(\bar{P} \otimes vec\bar{P}vec'\bar{P})H \\
&\quad + \bar{P} \otimes K_{m,m}(\bar{P} \otimes \bar{P}) + K_{m,m}(\bar{P} \otimes \bar{P}) \otimes \bar{P} \\
&\quad + K_{m,m^2}[\bar{P} \otimes K_{m,m}(\bar{P} \otimes \bar{P})]K_{m^2,m} = C_0(\bar{P}), \ say, \\
B_1 &= H(P \otimes \bar{P} \otimes \bar{P})H \\
&\quad + H(P \otimes vec\bar{P}vec'\bar{P} + \bar{P} \otimes vecPvec'\bar{P} + \bar{P} \otimes vec\bar{P}vec'P)H \\
&\quad + P \otimes K_{m,m}(\bar{P} \otimes \bar{P}) + \bar{P} \otimes K_{m,m}(P \otimes \bar{P}) \\
&\quad + \bar{P} \otimes K_{m,m}(\bar{P} \otimes P) + K_{m,m}(P \otimes \bar{P}) \otimes \bar{P} \\
&\quad + K_{m,m}(\bar{P} \otimes \bar{P}) \otimes \bar{P} + K_{m,m}(\bar{P} \otimes \bar{P}) \otimes P \\
&\quad + K_{m,m^2}\{[P \otimes K_{m,m}(\bar{P} \otimes \bar{P})] + [\bar{P} \otimes K_{m,m}(P \otimes \bar{P})] \\
&\qquad\qquad + [\bar{P} \otimes K_{m,m}(\bar{P} \otimes P)]\}K_{m^2,m} = C_1(\bar{P}, P), \ say, \\
B_2 &= C_1(P, \bar{P}), \\
B_3 &= C_0(P),
\end{aligned}$$

with

$$\begin{aligned}
H &= I + K_{m,m^2} + K_{m^2,m}, \\
\bar{P} &\equiv I - P.
\end{aligned}$$

Secondly, from (A.40),

$$\begin{aligned}
a_0 = tr(A_0) &= tr\{L[(I + K_{m,m})(\bar{P} \otimes \bar{P}) + vec\bar{P}vec'\bar{P}]\} \\
&= tr[(\bar{P} \otimes \bar{P})L(I + K_{m,m}) + vec'\bar{P}Lvec\bar{P}] \\
&= tr[(\bar{P} \otimes \bar{P})L(I + K_{m,m})] + tr(vec'\bar{P}Lvec\bar{P}).
\end{aligned} \tag{A.42}$$

Using (2.18) and $\bar{P} \equiv I - P$ gives,

$$(A'B^{-1}G)\bar{P} = 0, \tag{A.43}$$

$$\bar{P}(G'B^{-1}A) = 0. \tag{A.44}$$

Therefore, by (2.34)-(2.38),

$$(\bar{P} \otimes \bar{P})L = 0, \tag{A.45}$$

and

$$(\mathbb{H}G \otimes B^{-1}G)vec\bar{P} = vec(B^{-1}G\bar{P}G\mathbb{H}) = 0, \tag{A.46}$$

$$(I_m \otimes \mathbb{H}G)vec\bar{P} = vec(\mathbb{H}G\bar{P}) = 0. \tag{A.47}$$

Combining (A.46) and (A.47) with (2.35) yields

$$L_1 vec\bar{P} = 0. \tag{A.48}$$

Similarly,

$$L_3 vec\bar{P} = 0, \tag{A.49}$$
$$L_4 vec\bar{P} = 0, \tag{A.50}$$

and

$$vec'\bar{P}L_2 = (L_2' vec\bar{P})' = 0. \tag{A.51}$$

From (A.48)-(A.51),

$$tr(vec'\bar{P}Lvec\bar{P}) = 0. \tag{A.52}$$

Substituting (A.45) and (A.52) into (A.42) gives

$$a_0 = 0. \tag{A.53}$$

Also, from (A.41),

$$\begin{aligned}
b_1 &= vec'JB_1vecJ \\
&= vec'JH(P \otimes \bar{P} \otimes \bar{P})HvecJ \\
&\quad + vec'JH(P \otimes vec\bar{P}vec'\bar{P} + \bar{P} \otimes vecPvec'\bar{P} + \bar{P} \otimes vec\bar{P}vec'P)HvecJ \\
&\quad + vec'J[P \otimes K_{m,m}(\bar{P} \otimes \bar{P}) + \bar{P} \otimes K_{m,m}(P \otimes \bar{P})]vecJ \\
&\quad + vec'J[\bar{P} \otimes K_{m,m}(\bar{P} \otimes P) + K_{m,m}(P \otimes \bar{P}) \otimes \bar{P}]vecJ \\
&\quad + vec'J[K_{m,m}(\bar{P} \otimes \bar{P}) \otimes \bar{P} + K_{m,m}(\bar{P} \otimes \bar{P}) \otimes P]vecJ \\
&\quad + vec'JK_{m,m^2}\{[P \otimes K_{m,m}(\bar{P} \otimes \bar{P})] + [\bar{P} \otimes K_{m,m}(P \otimes \bar{P})] \\
&\qquad\qquad + [\bar{P} \otimes K_{m,m}(\bar{P} \otimes P)]\}K_{m^2,m}vecJ.
\end{aligned} \tag{A.54}$$

Using

$$K_{p,q}vecA = vec(A'),$$
$$A \otimes B = K_{p,r}(B \otimes A)K_{s,q},$$

for $A : p \times q$ and $B : r \times s$ where K is the commutation matrix, the following equations are obtained:

$$K_{m,m^2}vecJ_1 = 2(G'B^{-1} \otimes G'\mathbb{H}G \otimes G'\mathbb{H})vec(D'), \tag{A.55}$$
$$K_{m,m^2}vecJ_2 = [G'\mathbb{H} \otimes (G'B^{-1}G - I_m) \otimes G'\mathbb{H}]vec(D'), \tag{A.56}$$
$$K_{m,m^2}vecJ_3 = -(G'\mathbb{H} \otimes G'\mathbb{H}G \otimes G'\mathbb{H})vec(D'); \tag{A.57}$$

$$K_{m^2,m}vecJ_1 = 2(G'\mathbb{H} \otimes G'B^{-1} \otimes G'\mathbb{H}G)K_{p^2,m}vecD, \tag{A.58}$$
$$K_{m^2,m}vecJ_2 = [G'\mathbb{H} \otimes G'\mathbb{H} \otimes (G'B^{-1}G - I_m)]K_{p^2,m}vecD, \tag{A.59}$$
$$K_{m^2,m}vecJ_3 = -(G'\mathbb{H} \otimes G'\mathbb{H} \otimes G'\mathbb{H}G)K_{p^2,m}vecD. \tag{A.60}$$

Then, substituting (A.55)-(A.60) into (A.54), and using

$$vec(ABC) = (C' \otimes A)vecB,$$
$$(A \otimes B)' = A' \otimes B',$$
$$(A \otimes C)(B \otimes D) = (AB) \otimes (CD),$$

together with (A.43) and (A.44) yield

$$b_1 = 0. \tag{A.61}$$

Combining (A.53), (A.61) and the proof of Theorem 2.4 in Phillips and Park (1988) gives Theorem 2.4 in the current paper. □

B Data Description

The earnings data used are drawn from the Panel Study of Income Dynamics (PSID), available at *http : //psidonline.isr.umich.edu/*. The sample consists of men who were heads of household in every year from 1969 to 1974, who were between the ages of 21 (not inclusive) and 64 (not inclusive) in each year, and who reported positive earnings in each year. Individuals with average hourly earnings greater than $100 or reported annual hours greater than 4680 were excluded.

Some variables such as V7492, V7490, V0313, V0794, V7460, V7476, V7491 listed on p443 of Abowd and Card (1989) are not available now on the PSID website. The variables for sex listed on that page are incorrect. The following is the PSID variables used in the present paper:

ANNUAL EARNINGS: V1196, V1897, V2498, V3051, V3463, V3863;

ANNUAL HOURS: V1138, V1839, V2439, V3027, V3423, V3823;

SEX: ER32000;

AGE: ER30046.