

Department of Economics

Working Paper Series

## Bartlett-type Correction of Distance Metric Test

10-003 | Wanling Huang  
Concordia University  
Artem Prokhorov  
Concordia University and CIREQ



UNIVERSITÉ

**Concordia**

UNIVERSITY

# BARTLETT-TYPE CORRECTION OF DISTANCE METRIC TEST\*

Wanling Huang<sup>†</sup>     Artem Prokhorov<sup>‡§</sup>

March 2012

## Abstract

We derive a corrected distance metric (DM) test of general restrictions. The correction factor is a function of the uncorrected statistic, and the new statistic is Bartlett-type. In the setting of covariance structure models, we show using simulations that the quality of the new approximation is good and often remarkably good. Especially at around the 95th percentile, the distribution of the corrected test statistic is strikingly close to the relevant asymptotic distribution. This is true for various sample sizes, distributions, and degrees of freedom of the model. As a by-product we provide an intuition for the well-known observation in labor economic applications that using longer panels results in a reversal of the original inference.

*JEL Classification:* C12

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

---

\*Helpful comments from Gordon Fisher, Nikolay Gospodinov, Jorgen Hansen, Lynda Khalaf, Eric Renault, Paul Rilstone, seminar participants at University of Manitoba (Department of Economics) and Ryerson University (Department of Economics) and participants of the 26th annual meeting of Canadian Econometric Study Group (CESG), the 43rd Conference of the Canadian Economics Association (CEA), and 17th International Panel Data Conference are gratefully acknowledged. Research for this paper was supported by the FQRSC Doctoral Scholarship (Huang), the J. W. McConnell Memorial Graduate Fellowship (Huang) and an FQRSC Research Grant (Prokhorov).

<sup>†</sup>Department of Economics, McGill University and CIRANO, Montreal, PQ H3A2T7 Canada.

<sup>‡</sup>Department of Economics, Concordia University and CIREQ, Montreal, PQ H3G1M8 Canada.

<sup>§</sup>Corresponding author. E-mail: artem.prokhorov@concordia.ca (A. Prokhorov).

# 1 Introduction

The Distance Metric (DM) test of [Newey and West \(1987\)](#) is commonly used in econometrics to assess competing specifications. This is a simple test – the DM test statistic is usually calculated as the sample size times the difference in the criterion function evaluated at the restricted and the unrestricted estimate. At the same time, the test has several advantages over other classical tests. It is invariant to different but equivalent formulations of the restriction unlike, e.g, the Wald test (see, e.g., [Breusch and Schmidt, 1988](#)), and robust to autocorrelation and heteroskedasticity of unknown form provided that the criterion function uses a heteroskedasticity-consistent estimate of the covariance matrix (see, e.g., [Newey and McFadden, 1994](#)). This makes the test popular among applied researchers. For example, this test has been widely used in covariance structure analysis in the context of asymptotic distribution-free estimation (see, e.g., [Browne, 1984](#); [Satorra and Bentler, 2001](#), for the theory of ADF estimation).

It is well known that the DM test statistic asymptotically has the chi-square distribution with  $r$  degrees of freedom, where  $r$  is the number of restrictions (see, e.g., [Newey and McFadden, 1994](#)). However, the sampling distribution of the test statistic is poorly approximated by the asymptotic distribution if samples are small (see, e.g., [Clark, 1996](#)). Edgeworth expansions can deal with this problem by expanding the sampling density of test statistics around the asymptotic density in decreasing powers of  $N^{-\frac{1}{2}}$ , with  $N$  being the sample size. This may improve the accuracy of the asymptotic approximation. Surveys of Edgeworth expansion methods, including the theory of their validity, are provided by [Phillips \(1977, 1978\)](#); [Kallenberg \(1993\)](#); [Rothenberg \(1984\)](#); [Reid \(1991\)](#); [Sargan and Satchell \(1986\)](#), among others.

However, Edgeworth expansion methods have not yet been applied to the most general version of the DM test. Most of known results concern the LR, Wald and the score test (see, e.g., [Cribari-Neto and Cordeiro, 1996](#); [Phillips and Park, 1988](#); [Magee, 1989](#); [Linton, 2002](#); [Hausman and Kuersteiner, 2008](#)). [Hansen \(2006\)](#) is the only application (known to us) of

the Edgeworth correction to the DM test but it is restricted to the setting of a normal linear regression with a single constraint. Moreover, it is well known that Edgeworth expansions do not always improve the quality of first-order asymptotic approximations (see, e.g., [Phillips, 1983](#)). The main contribution of the paper is that we derive the Edgeworth correction, also known as the Bartlett-type correction, for the DM test in its general form and illustrate in simulations that this corrected approximation does work better, often surprisingly better, than the uncorrected test.

We do not consider alternative ways to remedy the inaccuracy of first-order asymptotic approximations. Such alternatives include resampling techniques and other types of asymptotic approximations, e.g., saddle-point (tilted Edgeworth) or Cornish-Fisher expansions. Validity of the former is usually based on existence of an asymptotic approximation in the first place (see, e.g., [Hall, 1992](#)) and the various forms of the latter are substantially more complicated than the classical Edgeworth expansion (see, e.g., [Barndorff-Nielsen and Cox, 1979](#)).

The paper can be viewed as a generalization of the results by [Hansen \(2006\)](#), who obtained the DM test correction in the setting of linear regressions with one restriction, to most of the extremum and minimum distance estimators and to multiple linear and nonlinear restrictions. We also draw on the results by [Phillips and Park \(1988\)](#) and [Kollo and Rosen \(2005\)](#). [Phillips and Park \(1988\)](#) investigate how higher-order terms in the asymptotic approximation of the Wald test are affected by various formulations of the null hypothesis. The DM test is invariant to such reformulations, however, their theorem on asymptotic expansion of the distribution provides a useful shortcut that substantially facilitates our proof. [Kollo and Rosen \(2005\)](#) provide general forms of Taylor series expansions for vector-valued functions, applicable in our setting.

In the application section, we consider a covariance structure model of [Abowd and Card \(1989\)](#). We address the question at what sample sizes would the proposed asymptotic correction make a difference for the empirical conclusions of that paper. It turns out that this

happens at sample sizes as large as 900-1,000 observations. An interesting by-product of the application is that it explains the old puzzle in labor economics that longer panels reverse the original inference.

The DM test statistic is defined in Section 2. In Section 3 we derive the asymptotic expansion to order  $O_p(N^{-1})$  of the DM test statistic, and in Section 4 we give the higher-order approximation of its distribution. Simple simulations are provided in Section 5, and an empirical illustration is presented in Section 6. Section 7 contains brief concluding remarks.

## 2 Distance Metric Test

For a family of distributions  $\{P_\theta, \theta \in \Theta \subset \mathbb{R}^p\}$ ,  $\Theta$  compact, consider the test

$$H_0 : g(\theta) = 0,$$

$$H_1 : g(\theta) \neq 0,$$

where  $g : \mathbb{R}^p \rightarrow \mathbb{R}^r$  is a continuously differentiable function with the first derivative defined by

$$A(\theta) \equiv \frac{dg(\theta)}{d\theta}.$$

$p \times r$

Let  $A(\theta_0)$  be denoted by  $A$ .

We assume that underlying the test is a parametric model that can be written in terms of the moment condition

$$\mathbb{E}m(Z_i, \theta) = 0 \quad \text{iff } \theta = \theta_0, \tag{1}$$

where  $m(\cdot, \cdot)$  is a continuous  $k$ -valued function,  $Z_i$  is a vector of data, independently distributed over  $i = 1, \dots, N$ , and  $\theta_0$  is the true value of the parameter vector. We assume

that the moments identify  $\theta_0$ . In covariance structure models, for example,  $m(Z_i, \theta) = \text{vech}Z_iZ_i' - \text{vech}\Sigma(\theta)$ , where *vech* denotes vertical vectorization of the lower triangle of a matrix and  $\Sigma(\theta)$  is a model for the covariance matrix, in which  $k \geq p$ .

For some positive definite weighting matrix  $W_N$ , define the criterion function

$$-Q_N(\theta) \equiv \frac{1}{2}m'_N(\theta)W_N m_N(\theta), \quad (2)$$

where  $m_N(\theta) \equiv \frac{1}{N} \sum_{i=1}^N m(Z_i, \theta)$ . In covariance structure literature, the estimator that minimizes this function is known as the asymptotically distribution free (ADF) or weighted least squared (WLS) estimator (see, e.g., [Browne, 1984](#)). It is well known that efficient weighting of  $m(\cdot, \cdot)$  requires that

$$W_N \xrightarrow{p} W \equiv \{\mathbb{E}[m(Z_i, \theta_0)m'(Z_i, \theta_0)]\}^{-1}.$$

We assume efficient weighting. What this means for our expansions will be clarified below.

The test statistic we consider is based on the value of  $Q_N(\theta)$  for two competing models, one that satisfies  $H_0$  and the other that is unrestricted. Let  $\bar{\theta}$  and  $\hat{\theta}$  denote the corresponding estimators:

$$\bar{\theta} = \arg \max_{\theta \in \Theta} Q_N(\theta), \text{ subject to } g(\theta) = 0;$$

$$\hat{\theta} = \arg \max_{\theta \in \Theta} Q_N(\theta).$$

Then, the DM test statistic is defined (see, e.g., [Newey and McFadden, 1994](#), p. 2222) as

$$DM \equiv -2N[Q_N(\bar{\theta}_N) - Q_N(\hat{\theta}_N)]. \quad (3)$$

Throughout, we assume that the standard regularity conditions are satisfied (see, e.g., [Newey](#)

and McFadden, 1994, conditions of Theorems 2.6, 3.4, 4.5, and 9.1).

### 3 Stochastic Expansion of DM Test Statistic

Let

$$\mathbb{M}_N(\theta) = W_N^{1/2} m_N(\theta).$$

Assume that  $\mathbb{M}_N(\theta)$  is three-times continuously differentiable. We follow Kollo and Rosen (2005, Definition 1.4.1) and define the derivative matrices recursively as follows

$$\begin{aligned} G_N(\theta) &\equiv \frac{\partial \mathbb{M}'_N(\theta)}{\partial \theta}, \\ D_N(\theta) &\equiv \frac{\partial \text{vec}' G_N(\theta)}{\partial \theta}, \\ C_N(\theta) &\equiv \frac{\partial \text{vec}' D_N(\theta)}{\partial \theta}. \end{aligned}$$

Let  $G = \mathbb{E}[G_N(\theta_0)]$ ,  $D = \mathbb{E}[D_N(\theta_0)]$ , and  $C = \mathbb{E}[C_N(\theta_0)]$ . In simulations, our focus is on covariance structure models for which the moment conditions have the form  $m(Z_i, \theta) = r(Z_i) + h(\theta)$ , for some functions  $r(\cdot)$  and  $h(\cdot)$ . In this case,  $G_N(\theta_0)$ ,  $D_N(\theta_0)$ , and  $C_N(\theta_0)$  are nonrandom matrices.

The quadratic form in (2) becomes

$$-Q_N(\theta) = \frac{1}{2} \mathbb{M}'_N(\theta) \mathbb{M}_N(\theta),$$

and the DM test statistic in (3) can be written as follows

$$DM = N[\mathbb{M}'_N(\bar{\theta}) \mathbb{M}_N(\bar{\theta}) - \mathbb{M}'_N(\hat{\theta}) \mathbb{M}_N(\hat{\theta})]. \quad (4)$$

Note that, due to the efficient weighting,

$$-\sqrt{N}\mathbb{M}_N(\theta_0) \equiv q_N \xrightarrow[k \times 1]{d} \bar{q} \sim N(0, \mathbb{I}). \quad (5)$$

Following Hansen (2006) and Phillips and Park (1988), we derive higher order expansions of the DM test under the stronger assumption that we have carried out the standardizing transformation and that

$$-\sqrt{N}\mathbb{M}_N(\theta_o) \equiv \bar{q} \sim N(0, \mathbb{I}). \quad (6)$$

We further assume that

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \equiv \tilde{q} \sim N(0, \Omega_1), \quad (7)$$

$$\sqrt{N}(\bar{\theta}_N - \hat{\theta}_N) \equiv \hat{q} \sim N(0, \Omega_2). \quad (8)$$

The usual first order asymptotic expansions of constrained and unconstrained GMM Newey and McFadden (1994, p. 2219) imply that

$$\tilde{q} = B^{-1}G\bar{q},$$

$$\hat{q} = -\mathbb{H}G\bar{q},$$

where  $\mathbb{H} \equiv B^{-1}A(A'B^{-1}A)^{-1}A'B^{-1}$  and  $B^{-1} = (GG')^{-1}$ .

Assumptions (6)-(8) substantially simplify derivations by disregarding possibly important higher order terms of  $\bar{q}$ ,  $\tilde{q}$  and  $\hat{q}$ . It is in principle possible to generalize our results as in Phillips and Park (1988, Appendix B) to the more general case of only (5), by carrying additional higher order terms involved in  $\bar{q}$  and in the transformations using  $W_N$ ,  $B$ ,  $G$  and  $\mathbb{H}$ . That is, in principle  $\bar{q}$ ,  $\tilde{q}$  and  $\hat{q}$  can come from any distribution that admits a valid Edgeworth expansion. This would account for the well known higher order biases of GMM (see, e.g., Newey and



Smith, 2004) and would allow  $W_N$  to depend on  $\theta$  as in the CU-GMM estimator of Hansen et al. (1996) or a two-step GMM procedure. However, the expansions for this general case quickly become hard to manage using matrix notation. Moreover, we focus on covariance structure models with relatively small deviations of the sampling distributions from the first-order asymptotics and it is unclear if the benefit of this generalization outweighs the costs in this setting. For example, in our simulations we consider other distributions of  $\bar{q}$  and find that our correction still works well. We leave such generalizations for future research.

Using the above notation and Theorem 3.1.1 of Kollo and Rosen (2005, p. 280), which we provide in Appendix A for reference, the Taylor expansion of  $\mathbb{M}_N(\bar{\theta}_N)$  about  $\hat{\theta}_N$  can be written as follows

$$\mathbb{M}_N(\bar{\theta}_N) = \mathbb{M}_N(\hat{\theta}_N) + G'_N(\hat{\theta}_N)(\bar{\theta}_N - \hat{\theta}_N) + \frac{1}{2}[I_k \otimes (\bar{\theta}_N - \hat{\theta}_N)'] D'_N(\hat{\theta}_N)(\bar{\theta}_N - \hat{\theta}_N) + o_p(N^{-1}). \quad (9)$$

Substituting (9) into (4), we obtain

$$\begin{aligned} DM &= \bar{q}' G' \mathbb{H} G_N(\hat{\theta}_N) G'_N(\hat{\theta}_N) \mathbb{H} G \bar{q} \\ &\quad + \mathbb{M}'_N(\hat{\theta}_N) (I_k \otimes \bar{q}' G' \mathbb{H}) D'_N(\hat{\theta}_N) \mathbb{H} G \bar{q} \\ &\quad - N^{-1/2} \bar{q}' G' \mathbb{H} G_N(\hat{\theta}_N) (I_k \otimes \bar{q}' G' \mathbb{H}) D'_N(\hat{\theta}_N) \mathbb{H} G \bar{q} \\ &\quad + \frac{1}{4} N^{-1} \bar{q}' G' \mathbb{H} D_N(\hat{\theta}_N) (I_k \otimes \mathbb{H} G \bar{q}) (I_k \otimes \bar{q}' G' \mathbb{H}) D'_N(\hat{\theta}_N) \mathbb{H} G \bar{q} + o_p(N^{-2}). \end{aligned} \quad (10)$$

We will now expand at  $\theta_0$  all functions of  $\hat{\theta}_N$  contained in (10). We wish to use Theorem 3.1.1 of Kollo and Rosen (2005) to do that. So we will transform the current representation into the one based on vector functions. Specifically, we need the vectorized versions of matrices  $G_N(\hat{\theta}_N)$  and  $D_N(\hat{\theta}_N)$ . Using the facts that

$$\text{vec}(ABC) = (C' \otimes A) \text{vec} B,$$

$$(A \otimes B)' = A' \otimes B',$$

we obtain the following equations

$$\begin{aligned}\bar{q}'G'\mathbb{H}G_N(\hat{\theta}_N) &= \text{vec}'G_N(\hat{\theta}_N)(I_k \otimes \mathbb{H}G\bar{q}), \\ D'_N(\hat{\theta}_N)\mathbb{H}G\bar{q} &= (I_{pk} \otimes \bar{q}'G'\mathbb{H})\text{vec}D_N(\hat{\theta}_N).\end{aligned}$$

Equation (10) can now be rewritten as

$$\begin{aligned}DM &= \text{vec}'G_N(\hat{\theta}_N)M_1\text{vec}G_N(\hat{\theta}_N) \\ &\quad + \mathbb{M}'_N(\hat{\theta}_N)M_2\text{vec}D_N(\hat{\theta}_N) \\ &\quad - N^{-1/2}\text{vec}'G_N(\hat{\theta}_N)M_3\text{vec}D_N(\hat{\theta}_N) \\ &\quad + N^{-1}\frac{1}{4}\text{vec}'D_N(\hat{\theta}_N)M_4\text{vec}D_N(\hat{\theta}_N) + o_p,\end{aligned}\tag{11}$$

where

$$\begin{aligned}M_1 &= (I_k \otimes \mathbb{H}G\bar{q})(I_k \otimes \bar{q}'G'\mathbb{H}), \\ M_2 &= I_k \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H}, \\ M_3 &= (I_k \otimes \mathbb{H}G\bar{q})(I_k \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H}), \\ M_4 &= I_k \otimes \mathbb{H}G\bar{q}\bar{q}'G'\mathbb{H} \otimes \mathbb{H}G\bar{q}\bar{q}'G'\mathbb{H}.\end{aligned}$$

Substituting the Taylor expansions at  $\theta_0$  of  $\mathbb{M}_N(\hat{\theta}_N)$ ,  $\text{vec}G_N(\hat{\theta}_N)$  and  $\text{vec}D_N(\hat{\theta}_N)$  into (11) gives the asymptotic expansion of the DM test statistic, which is summarized in the following theorem.

**Theorem 1.** *The asymptotic expansion of the DM test statistic is given by*

$$DM = \bar{q}'P\bar{q} + N^{-1/2}u(\bar{q}) + N^{-1}v(\bar{q}) + o_p,\tag{12}$$

where

$$P_{k \times k} \equiv G'HG, \quad (13)$$

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

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

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

$$u_1(\bar{q}) = 2\bar{q}'G'B^{-1}DM_1vecG, \quad (14)$$

$$u_2(\bar{q}) = \bar{q}'(G'B^{-1}G - I_k)M_2vecD, \quad (15)$$

$$u_3(\bar{q}) = -vec'GM_3vecD; \quad (16)$$

$$v_1(\bar{q}) = \bar{q}'G'B^{-1}DM_1D'B^{-1}G\bar{q} + \bar{q}'G'B^{-1}C(I_{pk} \otimes B^{-1}G\bar{q})M_1vecG, \quad (17)$$

$$v_2(\bar{q}) = \bar{q}'(G'B^{-1}G - I_k)M_2C'B^{-1}G\bar{q} + \frac{1}{2}\bar{q}'G'B^{-1}D(I_k \otimes B^{-1}G\bar{q})M_2vecD, \quad (18)$$

$$v_3(\bar{q}) = -\bar{q}'G'B^{-1}CM_3'vecG - \bar{q}'G'B^{-1}DM_3vecD, \quad (19)$$

$$v_4(\bar{q}) = \frac{1}{4}vec'DM_4vecD. \quad (20)$$

**Proof.** See Appendix B for all proofs.

## 4 Distribution of DM Test Statistic

In this section we follow [Phillips and Park \(1988\)](#) and use the Taylor expansion of  $DM$  to derive the Edgeworth expansion of its distribution to order  $O(N^{-1})$ . Theorem 2.4 of [Phillips and Park \(1988\)](#) allows us to skip intermediate steps in deriving the expansion for the distribution from the expansion of the test statistics. [Hansen \(2006\)](#) used this approach for a single restriction DM test in a normal linear regression with known error variance.

In order to use [Phillips and Park](#)'s results, we first show that  $u(\bar{q})$  and  $v(\bar{q})$  can be written in terms of Kronecker products of  $\bar{q}$  and  $\bar{q}\bar{q}'$ . This is done in the following lemma.

**Lemma 1.**  $u(\bar{q})$  and  $v(\bar{q})$  in [Theorem 1](#) can be written as

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

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

where  $\text{vec}J = \text{vec}J_1 + \text{vec}J_2 + \text{vec}J_3$  with

$$\text{vec}J_1 = 2(G'\mathbb{H}G \otimes G'\mathbb{H} \otimes G'B^{-1})\text{vec}D,$$

$$\text{vec}J_2 = [(G'B^{-1}G - I_k) \otimes G'\mathbb{H} \otimes G'\mathbb{H}]\text{vec}D,$$

$$\text{vec}J_3 = -(G'\mathbb{H}G \otimes G'\mathbb{H} \otimes G'\mathbb{H})\text{vec}D;$$

and

$$L = L_1 + L_2 + L_3 + L_4, \tag{21}$$

with

$$L_1 = (G'\mathbb{H} \otimes G'B^{-1})V_D(\mathbb{H}G \otimes B^{-1}G) + (G'\mathbb{H} \otimes G'B^{-1})M_V(I_k \otimes \mathbb{H}G), \tag{22}$$

$$L_2 = (G'\mathbb{H} \otimes G'\mathbb{H})M_{VI} + \frac{1}{2}(G'\mathbb{H} \otimes G'\mathbb{H})V_D(B^{-1}G \otimes B^{-1}G), \tag{23}$$

$$L_3 = -(G'\mathbb{H} \otimes G'\mathbb{H})M_V(I_k \otimes \mathbb{H}G) - (G'\mathbb{H} \otimes G'\mathbb{H})V_D(\mathbb{H}G \otimes B^{-1}G), \tag{24}$$

$$L_4 = \frac{1}{4}(G'\mathbb{H} \otimes G'\mathbb{H})V_D(\mathbb{H}G \otimes \mathbb{H}G), \tag{25}$$

where  $V_D$ ,  $M_V$  and  $M_{VI}$  are given in [Appendix B](#).

We can now follow [Hansen \(2006, Theorem 3\)](#) and apply the result of [Phillips and Park \(1988, p. 1069-1072\)](#). Specifically, we can obtain the characteristic function of the DM test

statistic:

$$\begin{aligned}
C_{DM}(t) &= (1 - 2it)^{-r/2} \left\{ 1 + \frac{1}{N} \left[ (a_0 - \frac{1}{4}b_1)it \right. \right. \\
&\quad + (a_1 + \frac{1}{4}b_1 - \frac{1}{4}b_2)it(1 - 2it)^{-1} \\
&\quad + (a_2 + \frac{1}{4}b_2 - \frac{1}{4}b_3)(1 - 2it)^{-2} \\
&\quad \left. \left. + \frac{1}{4}b_3it(1 - 2it)^{-3} \right] \right\} + o_p(N^{-1}),
\end{aligned}$$

where  $a_i$ ,  $i = 0, 1, 2$ , and  $b_j$ ,  $j = 1, 2, 3$ , are defined in Appendix B. Note that the first term  $(1 - 2it)^{-r/2}$  is the characteristic function for a  $\chi_r^2$  variate, reflecting the first order asymptotics. Then, using the Fourier transform, we can derive the distribution of the DM test statistic. This is done in Theorem 2.

**Theorem 2.** *The asymptotic expansion to  $O(N^{-1})$  of the distribution function of DM is given by*

$$F_{DM}(x) = F_r(x - N^{-1}(\alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3)) + o(N^{-1}) \quad (26)$$

where  $F_r$  denotes the distribution function of a  $\chi_r^2$  variate and

$$\begin{aligned}
\alpha_1 &= (4a_1 - b_2)/4r, \\
\alpha_2 &= (4a_2 + b_2 - b_3)/4r(r + 2), \\
\alpha_3 &= b_3/4r(r + 2)(r + 4),
\end{aligned}$$

with  $a_i$  ( $i = 1, 2$ ) and  $b_i$  ( $i = 1, 2, 3$ ) defined in Appendix B.

The Edgeworth correction factor that follows from (26) can be written as

$$1 - N^{-1}(\alpha_1 + \alpha_2 DM + \alpha_3 DM^2) \quad (27)$$

where  $DM$  is the original (uncorrected) DM test statistic. If multiplied by the correction factor, the DM test statistic should be better approximated by the  $\chi_r^2$  distribution than the uncorrected statistic. Strictly speaking, the correction cannot be called “Bartlett” because it depends on the uncorrected statistic  $DM$ . However, it is common to call such corrections Bartlett-type due to their similarity to the classical [Bartlett \(1937\)](#) correction (see, e.g., [Cribari-Neto and Cordeiro, 1996](#), for a review of Bartlett and Bartlett-type corrections of common tests).

Note that increasing the number of restrictions  $r$  does not necessarily result in a bigger correction factor because  $\alpha_i$  ( $i = 1, 2, 3$ ) may be negative. Moreover, it is important to note that, even if the restrictions are linear, the Bartlett-type correction factor in (27) will be different from one so long as  $M_N(\theta)$  is nonlinear in parameters. The theorem imposes no constraint on the number of restrictions tested or on the specific estimator represented by the moment condition (1).

Edgeworth expansions do not always improve the quality of asymptotic approximations. It has been documented that their performance is parameter dependent and that they fail when deviations of the sampling distribution from the first order asymptotic distribution is large (see, e.g., [Phillips, 1983](#)). We cannot expect the correction in (27) to work in all circumstances but when it does work, the quality of the correction can be expected to depend on nonlinearities (through matrices  $J$  and  $L$ ), the size of the model (through the number of restrictions  $r$ ), the sample size  $N$  and the true distribution (through  $\bar{q}$ ). We now demonstrate the behavior of the correction along some of these dimensions.

## 5 Illustrative Simulations

In this section, we use simulations to illustrate the theoretical results obtained in Section 4 in the settings of a simple covariance structure model. Consider a random vector  $Z \in \mathcal{Z} \subset \mathbb{R}^q$  from  $P_{\theta_0}$ ,  $\theta_0 \in \Theta \subset \mathbb{R}^p$ . Assume that  $\mathbb{E}[Z] = 0$ ,  $\mathbb{E}\{\|Z\|^4\} < \infty$  and  $\mathbb{E}[ZZ'] = \Sigma(\theta_0)$ . The

matrix function  $\Sigma(\theta)$  may come from a variety of models, e.g., LISREL, MIMIC, factor model, random effects or simultaneous equations model. For a random sample  $(Z_1, \dots, Z_N)$ , let

$$S_i \equiv Z_i Z_i'$$

and

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

Then,  $S$  satisfies a central limit theorem:

$$\sqrt{N}(\text{vech}S - \text{vech}\Sigma(\theta_0)) \rightarrow N(0, \Delta(\theta_0)),$$

where

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

Assume  $p \leq \frac{1}{2}q(q+1)$ . Then, in terminology of covariance structure literature, the degrees of freedom of the model are equal to  $\frac{q(q+1)}{2} - p$ , and they will be increased by one for each independent restriction imposed on  $\Sigma(\theta)$  by the model. We can write all distinct sample moment functions as follows

$$m_N(\theta) \equiv \frac{1}{N} \sum_{i=1}^N m(Z_i, \theta) = \text{vech}S - \text{vech}\Sigma(\theta)$$

$\frac{1}{2}q(q+1) \times 1$

where

$$m(Z_i, \theta) = \text{vech}S_i - \text{vech}\Sigma(\theta).$$

$\frac{1}{2}q(q+1) \times 1$

The sample covariance matrix of the moments is

$$\begin{aligned}
W_N^{-1}(\theta) &= \frac{1}{N} \sum_{i=1}^N [m(Z_i, \theta)m'(Z_i, \theta)] \\
&= \frac{1}{N} \sum_{i=1}^N [\text{vech}S_i \text{vech}'S_i - \text{vech}S_i \text{vech}'\Sigma(\theta) \\
&\quad - \text{vech}\Sigma(\theta) \text{vech}'S_i + \text{vech}\Sigma(\theta) \text{vech}'\Sigma(\theta)].
\end{aligned}$$

In practice, either the restricted or the unrestricted estimate of  $\theta$  will be used in these infeasible expressions.

We are interested in testing  $H_0 : \Sigma(\theta_o) = \Sigma(c)$  against  $H_1 : \Sigma(\theta_o) \neq \Sigma(c)$ , where  $c$  is a constant vector. This type of test is fundamental in covariance structure analysis. Known as the ADF test, it has been studied by [Korin \(1968\)](#); [Sugiura \(1969\)](#); [Nagarsenker and Pillai \(1973\)](#); [Browne \(1984\)](#); [Chou et al. \(1991\)](#); [Muthen and Kaplan \(1992\)](#); [Yuan and Bentler \(1997\)](#); [Satorra and Bentler \(2001\)](#); [Yanagihara et al. \(2004\)](#), among others. [Ogasawara \(2009\)](#) provides an asymptotic expansion similar to ours for the ADF test statistic in the setting of covariance structure models. The literature has focused on three dimensions of the test behavior: (1) what is the effect of the sample size; (2) how the sample size requirements change for different nonnormal distributions; (3) how the sample size requirements change for models of different size. We wish to apply our Bartlett-type correction to the DM test of this restriction and study its behavior along the same dimensions.

For simplicity, we consider a bivariate problem (i.e.  $q = 2$ ) in which

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

$\theta' = (\sigma_1, \sigma_{12}, \sigma_2)$ ,  $c' = (1, 0, 1)$  and  $p = k = r = 3$ . So the parameter vector is completely specified under the null and there are no parameters to estimate in the restricted model.



Write the null hypothesis as

$$H_0 : g(\theta) = 0, \quad \text{where} \quad g(\theta) = \text{vech}\Sigma(\theta) - \text{vech}\Sigma(c) = \begin{bmatrix} \sigma_1^2 - 1 \\ \sigma_{12} - 0 \\ \sigma_2^2 - 1 \end{bmatrix}.$$

In order to demonstrate the effect of the Bartlett-type correction, we generate a sample of varying size from normal, Student-t and uniform distributions and compute the uncorrected and corrected versions of the DM test statistics. This is done 1,000 times. Then we plot the quantiles of the resulting bootstrap distributions. These are displayed on Figures 1-3. The quantile curve of the chi-square distribution, marked “chi<sup>2</sup>”, is drawn as a reference. The uncorrected and corrected versions of the DM test statistic are marked “DM” and “DM\_star,” respectively.

All figures show severe over-rejection of the uncorrected DM test statistic. The fact that the size of the DM test is substantially greater in small samples than the asymptotic size is well documented (see, e.g., Clark, 1996), and our results agree with that. Our corrected statistic performs much better for all distributions and all sample sizes. Of course, the corrected distribution is not identical to the chi-square distribution and the corrected test exhibits over- and under-rejection at times, but the deviations are substantially smaller than for the uncorrected test. It is notable how much improvement one can obtain using the corrected statistic in the area close to the 95th percentile, which corresponds to the commonly used 5% significance level. At that level, the correction is almost perfect.

Figure 1 shows the quantiles for various sample sizes from  $\mathcal{N}(0, 1)$ . One can clearly see from the figure how the uncorrected curve deviates from the chi-square quantiles as the sample size decreases while the degree of model complexity does not change ( $q = 2$ ). At the same time, the corrected curve consistently provides a great deal of improvement.

In Figure 2 we show the behavior of the corrected and uncorrected test statistics for two

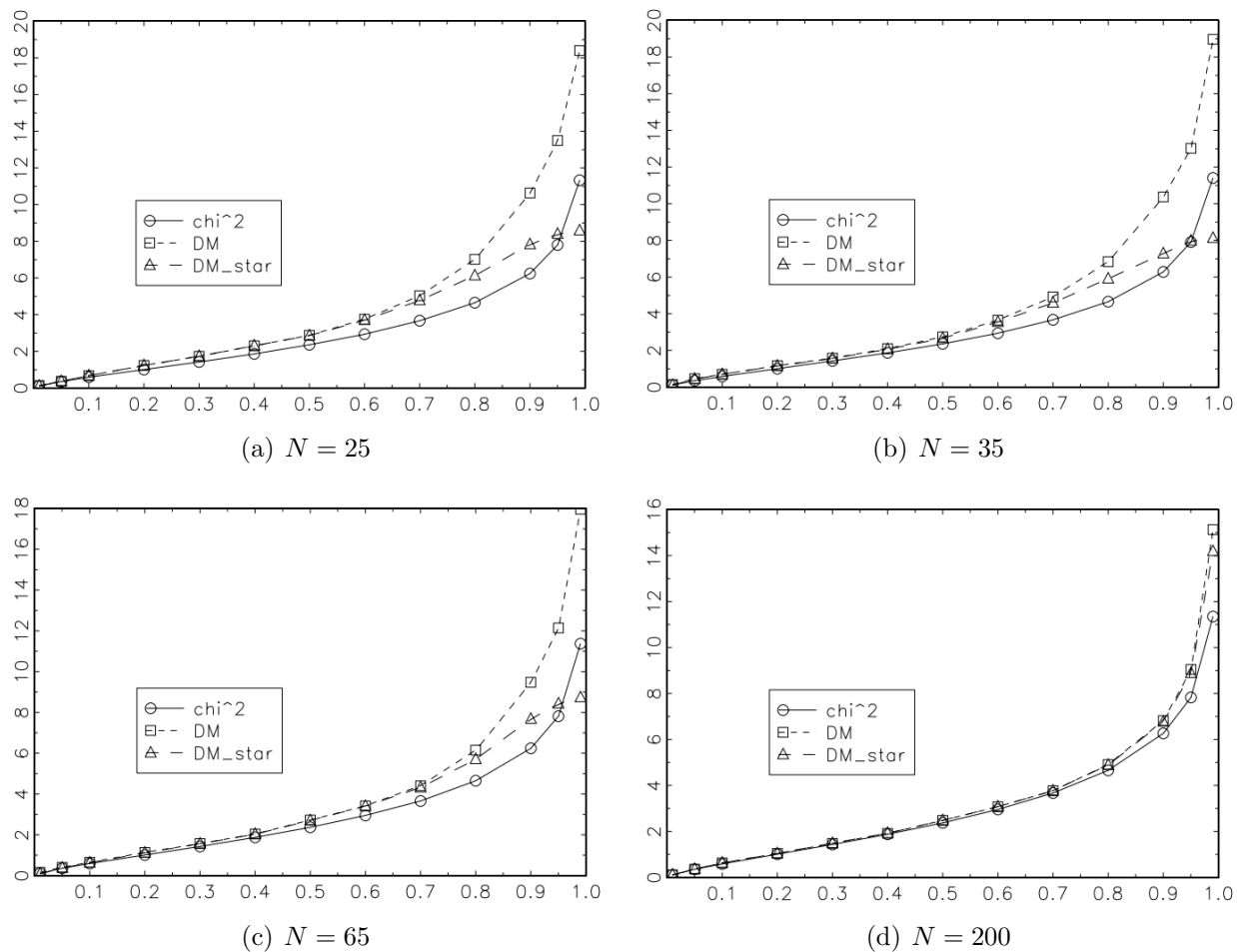


Figure 1: Quantiles of chi-square and bootstrap distribution of uncorrected and corrected DM test statistics for various sample sizes;  $q = 2$ .

distributions, Student-t and uniform, and two sample sizes,  $N = 25$  and  $N = 65$ . As expected, the test (and its correction), being distribution-free, exhibits similar behavior under the two distributions. The figures also show that the benefit of a larger sample size varies for the two distributions. For other distributions (not reported here), the sample size needed to obtain a similar level of approximation accuracy as in panel (d) was several hundred observations. For some distributions, the correction may be trivial even when samples are small while for others it may produce a large correction even when samples are large.

In Figure 3, in addition to the bivariate case, we consider a univariate ( $q = 1$ ) model in which  $\Sigma(\theta) = \sigma^2$ . The null is  $\sigma = c$ , and the restricted model has one degree of freedom.

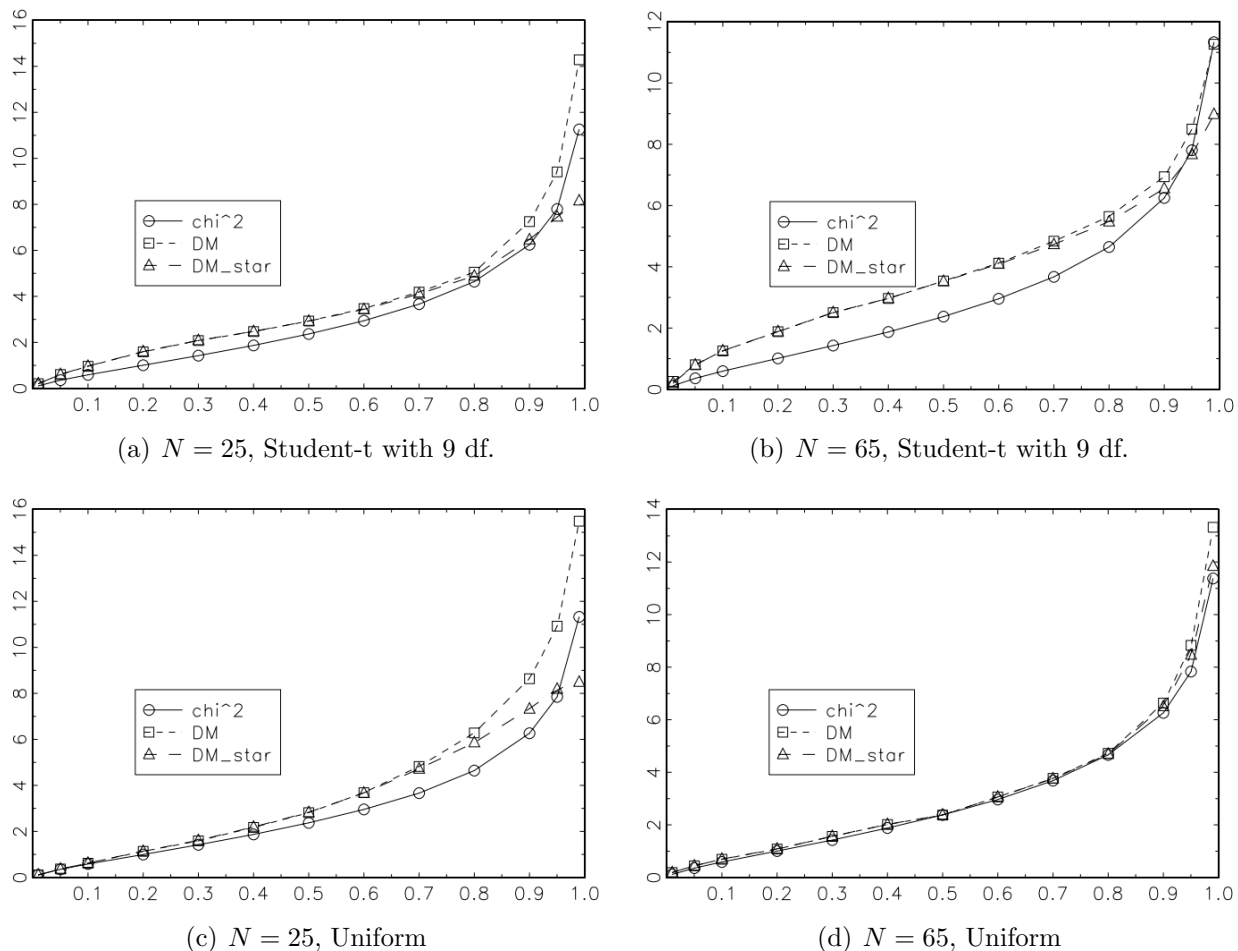


Figure 2: Quantiles of chi-square and bootstrap distribution of uncorrected and corrected DM test statistics for two data distributions and two sample sizes;  $q = 2$ .

This is done to show how model size (as measured by the degrees of freedom of the model) affects the performance of the test statistics. In the larger model ( $q = 2$ ), the gap between the sampling and the asymptotic  $\chi_3^2$  distribution is much larger than between the sampling and the asymptotic  $\chi_1^2$  distribution in the smaller model. It is interesting to note that the model size plays as important a role in accuracy of asymptotic approximations as the sample size: we more than double the sample size between panel (b) and panel (d), and this has a similar effect on the larger model accuracy as replacing it by a model with 2 fewer degrees of freedom. This is consistent with the findings of [Hoogland and Boomsma \(1998\)](#) that the chi-square statistics are sensitive to model size (as measured by the degrees of freedom of

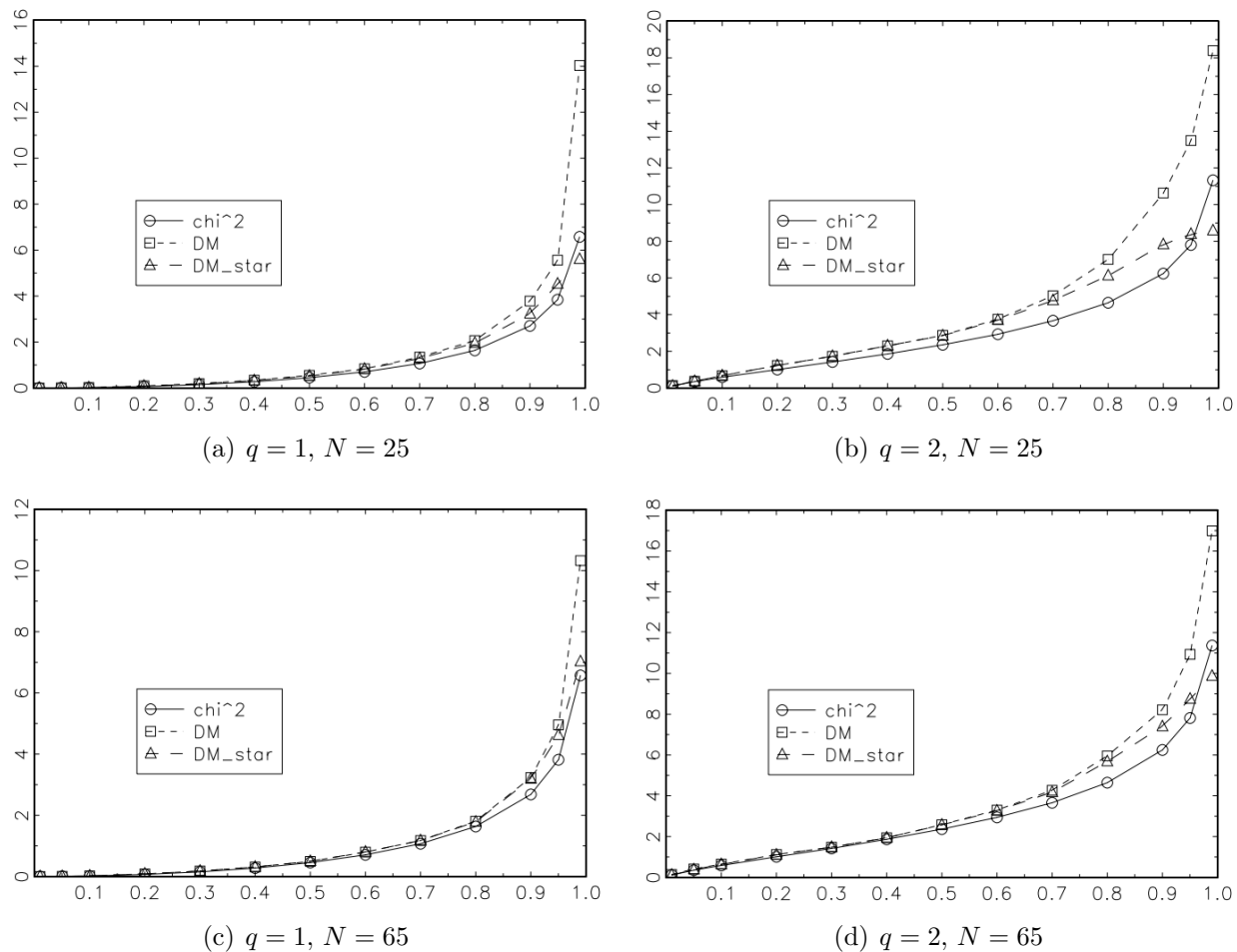


Figure 3: Quantiles of chi-square and bootstrap distribution of uncorrected and corrected DM test statistics for two values of  $q$  and two sample sizes.

the model). A bigger model requires a larger sample size to ensure good behavior of the statistics. At the same time, for the smaller models (panels (a) and (c)), larger sample sizes do not improve the asymptotic approximation by much – the approximation error is small to start with. The corrected statistic displays an improved behavior for both model sizes and both sample sizes.

## 6 Empirical Illustration

In this section, we study applicability of the Bartlett-type correction to a covariance structure model of earnings. This type of model has been a focus of many papers in labor economics (see, e.g., [MaCurdy, 1982](#); [Abowd and Card, 1987, 1989](#); [Topel and Ward, 1992](#); [Baker, 1997](#); [Baker and Solon, 2003](#)). Among other things, the literature has been concerned with the puzzling observation that the use of longer panels results in a reversal of the original inference (see, e.g., [Baker, 1997](#), p. 358). Longer panels are usually used to estimate higher-order autocovariances. However, the cost of longer balanced panels is a smaller number of individuals. For example, the sample sizes used by [Baker \(1997\)](#) in 10-year panels are 992 and 1,331 individuals for the periods 1967-76 and 1977-86, respectively; his 20-year panel contains only 534. On the other hand, as the panel gets longer ( $q$  increases), degrees of freedom grow. As mentioned earlier, this generally requires larger sample sizes for the DM statistic to remain close to the asymptotic approximation. In this section, we use parts of the sample of earnings used by [Abowd and Card \(1989\)](#) to demonstrate how the Bartlett-type correction affects the outcomes of a hypothesis test for various sample sizes.

The earnings data are from the Panel Study of Income Dynamics (PSID), conducted by Survey Research Center at University of Michigan. The sample consists of male heads of household, who were between the ages of 21 and 64 in the period 1969 to 1974 and who reported positive earnings in each year. The sample we use – a subsample of the data used by [Abowd and Card \(1989\)](#) – contains 1,578 individuals. Individuals with average hourly earnings greater than \$100 or those who reported annual hours worked greater than 4,680 were excluded. A detailed description of the PSID variables is given in [Appendix C](#). Covariances and correlations between demeaned changes in log of real annual earnings (in 1967 dollars) are displayed in [Table 1](#). Covariances are presented below the diagonal, while correlations and their two-tailed p-values are presented above the diagonal.

Table 1: Covariances (below diagonal) and correlations (above diagonal) between changes in log-earnings: PSID Males 1967-1974

with:	Covariance/Correlation(with two-tailed p-value) of:				
	$\Delta \ln e$ 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)	-0.006 (0.827)	0.018 (0.463)	-0.006 (0.823)
$\Delta \ln e$ 70-71	-0.04418	0.205	-0.415 (0)	-0.082 (0.001)	0 (0.994)
$\Delta \ln e$ 71-72	-0.00117	-0.08345	0.197	-0.347 (0)	-0.041 (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 generic population covariance matrix for Table 1 can be written as

$$\Sigma(\theta) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} & \sigma_{25} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} & \sigma_{35} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 & \sigma_{45} \\ \sigma_{51} & \sigma_{52} & \sigma_{53} & \sigma_{54} & \sigma_5^2 \end{bmatrix}, \quad (28)$$

where  $\theta = (\sigma_1, \sigma_{21}, \sigma_{31}, \sigma_{41}, \sigma_{51}, \sigma_2, \sigma_{32}, \sigma_{42}, \sigma_{52}, \sigma_3, \sigma_{43}, \sigma_{53}, \sigma_4, \sigma_{54}, \sigma_5)'$ .

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 an MA(2) process (possibly nonstationary) can serve as the model. Indeed, there are very few covariances (correlations) that are large or statistically significant at lags greater than two. 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 are zero. The null is  $H_0$  : changes in earnings are nonstationary MA(2), and

the alternative is  $H_1$  : changes in earnings are not nonstationary MA(2). Equivalently, the null can be rewritten as

$$H_0 : \begin{bmatrix} \sigma_{41} \\ \sigma_{51} \\ \sigma_{52} \end{bmatrix} = \underset{3 \times 1}{\mathbf{0}}. \quad (29)$$

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 (i)  $cov(\Delta \ln e_t, \Delta \ln e_{t-j})$  depends only on  $j$  and does not change over  $t$ , and (ii)  $cov(\Delta \ln e_t, \Delta \ln e_{t-j})$  is zero for  $|j| > 2$ . The null is  $H_0$  : changes in earnings are stationary MA(2), and the alternative is  $H_1$  : changes in earnings are not stationary MA(2). Equivalently, the null can be rewritten as

$$H_0 : \begin{bmatrix} \sigma_{41} \\ \sigma_{51} \\ \sigma_{52} \end{bmatrix} = \underset{3 \times 1}{\mathbf{0}}, \quad (30)$$

$$\sigma_1 = \sigma_2 = \sigma_3 = \sigma_4 = \sigma_5,$$

$$\sigma_{21} = \sigma_{32} = \sigma_{43} = \sigma_{54},$$

$$\sigma_{31} = \sigma_{42} = \sigma_{53}.$$

The test results are presented in Table 2. The values of the uncorrected and corrected DM test statistic (and the corresponding p-values) are very close for both tests. Not surprisingly, the corrections for this relatively large sample are minor to none. We now demonstrate the effect of the Bartlett-type correction as the sample size becomes smaller.

As expected, when the sample size becomes smaller the Bartlett-type correction becomes more important. Consider the second test as an example. The results for that test are presented in Table 3. We randomly select increasingly smaller subsamples of data. As the sample size decreases from  $N = 1,400$  to 900, the correction becomes larger to the point at

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

Goodness-of-Fit Test N=1,578	DM Test Statistic		Asy. P-Value	
	Uncorrected	Corrected	Uncorrected	Corrected
I. Nonstationary MA(2) ( $df = 3$ )	0.3325	0.3320	0.9538	0.9539
II. Stationary MA(2) ( $df = 12$ )	19.9889	19.6262	0.0673	0.0745

which the outcome of the test is reversed at conventional significance levels. For example, if  $N = 900$ , the corrected test does not reject at the 5% level while the uncorrected test does.

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=1,400	22.21	21.64	0.035	0.042
N=1,200	24.15	22.83	0.019	0.029
N=1,000	25.46	22.12	0.012	0.036
N=900	25.99	20.35	0.010	0.061

Assuming that the correction does bring the sampling distribution closer to its asymptotic approximation, we conclude from this table that, for the current number of degrees of freedom, cross sections as large as 900 are not large enough to justify application of the uncorrected first-order asymptotic approximation to this covariance structure model. If used against the asymptotic critical values, the uncorrected DM test severely over-rejects.

## 7 Concluding Remarks

This paper provides the Bartlett-type correction of the DM test statistic. Our setting covers linear and nonlinear restrictions and all extremum and minimum distance estimators that can be stated in terms of moment conditions. The expansions used to obtain the correction are based on several normality assumptions that can be relaxed using methods similar to [Phillips and Park \(1988, Appendix B\)](#). The correction may work better if we do so but we leave this general case for future work.



We also provide simple simulation evidence about the behavior of the corrected test statistic in a fairly general class of covariance structure models. Given the poor performance of Edgeworth approximations documented in settings when the error in the first order asymptotics is large, we use simulations where the errors are relatively small. We find that the correction works very well in such settings. In practice, it is often necessary to consider a very large (as measured by the degrees of freedom of the model) covariance structure model (see, e.g., [Herzog et al., 2007](#); [Kenny and McCoach, 2003](#)), which makes it difficult to maintain good properties of the DM test and of our correction even in large samples. Moreover, large samples are not always possible to obtain and the available data are often non-normal. We show that the correction still performs well for the sample sizes and non-normal distributions considered.

## A Theorem 3.1.1 of Kollo and Rosen (2005)

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  $\varepsilon_n \rightarrow 0$  as  $n \rightarrow \infty$ . If the function  $f(x)$  from  $\mathbb{R}^p$  to  $\mathbb{R}^s$  has continuous partial derivatives up to the order  $(\mathcal{M} + 1)$  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

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

where the Kroneckerian power  $A^{\otimes k}$  for any matrix  $A$  is given by  $A^{\otimes k} = \underbrace{A \otimes \dots \otimes A}_{k \text{ times}}$  with  $A^{\otimes 0} = 1$ ,  $\rho(\cdot, \cdot)$  is the

Euclidean distance in  $\mathbb{R}^p$ , and the matrix derivative for any matrices  $Y$  and  $X$  is given by  $\frac{d^k Y}{dX^k} = \frac{d}{dX} \left( \frac{d^{k-1} Y}{dX^{k-1}} \right)$

with  $\frac{dY}{dX} \equiv \frac{dvec'Y}{dvecX}$ ; and

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

## B Proofs

*Proof of Theorem 1:* Write (11) as

$$DM \cong 1_{DM} + 2_{DM} + 3_{DM} + 4_{DM}, \quad (31)$$

where,

$$\begin{aligned} 1_{DM} &= vec' G_N(\hat{\theta}_N) M_1 vec G_N(\hat{\theta}_N), \\ 2_{DM} &= \mathbb{M}'_N(\hat{\theta}_N) M_2 vec D_N(\hat{\theta}_N), \\ 3_{DM} &= -N^{-1/2} vec' G_N(\hat{\theta}_N) M_3 vec D_N(\hat{\theta}_N), \\ 4_{DM} &= N^{-1} \frac{1}{4} vec' D_N(\hat{\theta}_N) M_4 vec D_N(\hat{\theta}_N). \end{aligned}$$

Taking Taylor expansions of  $\mathbb{M}_N(\hat{\theta}_N)$ ,  $vec G_N(\hat{\theta}_N)$  and  $vec D_N(\hat{\theta}_N)$  about  $\theta_0$  and using (5) and (7), we have

$$\begin{aligned} \mathbb{M}_N(\hat{\theta}_N) &= \mathbb{M}_N(\theta_0) + G'(\hat{\theta}_N - \theta_0) + \frac{1}{2} [I_k \otimes (\hat{\theta}_N - \theta_0)'] D'(\hat{\theta}_N - \theta_0) + o_p(N^{-1}) \\ &= -N^{-1/2} \bar{q} + N^{-1/2} G' B^{-1} G \bar{q} + N^{-1} \frac{1}{2} (I_k \otimes \bar{q}' G' B^{-1}) D' B^{-1} G \bar{q} + o_p(N^{-1}), \end{aligned}$$

$$\begin{aligned}
\text{vec}G_N(\hat{\theta}_N) &= \text{vec}G + D'(\hat{\theta}_N - \theta_0) + \frac{1}{2}[I_{pk} \otimes (\hat{\theta}_N - \theta_0)']C'(\hat{\theta}_N - \theta_0) + o_p(N^{-1}) \\
&= \text{vec}G + N^{-1/2}D'B^{-1}G\bar{q} + N^{-1}\frac{1}{2}(I_{pk} \otimes \bar{q}'G'B^{-1})C'B^{-1}G\bar{q} + o_p(N^{-1}),
\end{aligned}$$

$$\begin{aligned}
\text{vec}D_N(\hat{\theta}_N) &= \text{vec}D + C'(\hat{\theta}_N - \theta_0) + o_p(N^{-1/2}) \\
&= \text{vec}D + N^{-1/2}C'B^{-1}G\bar{q} + o_p(N^{-1/2}).
\end{aligned}$$

Note that we do not need to expand  $\text{vec}D_N(\hat{\theta}_N)$  further for our purpose. Substituting these expressions into the terms of (31) gives:

$$\begin{aligned}
1_{DM} &= \text{vec}'G_N(\hat{\theta}_N)M_1\text{vec}G_N(\hat{\theta}_N) \\
&= \text{vec}'GM_1\text{vec}G + N^{-1/2}2\bar{q}'G'B^{-1}DM_1\text{vec}G \\
&\quad + N^{-1}[\bar{q}'G'B^{-1}DM_1D'B^{-1}G\bar{q} + \bar{q}'G'B^{-1}C(I_{pk} \otimes B^{-1}G\bar{q})M_1\text{vec}G] \\
&\quad + o_p(N^{-1}) \\
&= \bar{q}'P\bar{q} + N^{-1/2}u_1(\bar{q}) + N^{-1}v_1(\bar{q}) + o_p(N^{-1}),
\end{aligned} \tag{32}$$

where

$$P \equiv G'HG$$

is a projection matrix, and

$$\begin{aligned}
u_1(\bar{q}) &= 2\bar{q}'G'B^{-1}DM_1\text{vec}G, \\
v_1(\bar{q}) &= \bar{q}'G'B^{-1}DM_1D'B^{-1}G\bar{q} + \bar{q}'G'B^{-1}C(I_{pk} \otimes B^{-1}G\bar{q})M_1\text{vec}G;
\end{aligned}$$

$$\begin{aligned}
2_{DM} &= \mathbb{M}'_N(\hat{\theta}_N) M_2 \text{vec} D_N(\hat{\theta}_N) \\
&= -N^{-1/2} \bar{q}' M_2 \text{vec} D - N^{-1} \bar{q}' M_2 C' B^{-1} G \bar{q} \\
&\quad + N^{-1/2} \bar{q}' G' B^{-1} G M_2 \text{vec} D + N^{-1} \bar{q}' G' B^{-1} G M_2 C' B^{-1} G \bar{q} \\
&\quad + N^{-1} \frac{1}{2} \bar{q}' G' B^{-1} D(I_k \otimes B^{-1} G \bar{q}) M_2 \text{vec} D + o_p(N^{-1}) \\
&= N^{-1/2} (\bar{q}' G' B^{-1} M_2 \text{vec} D - \bar{q}' M_2 \text{vec} D) \\
&\quad + N^{-1} [\bar{q}' G' B^{-1} G M_2 C' B^{-1} G \bar{q} - \bar{q}' M_2 C' B^{-1} G \bar{q} \\
&\quad\quad + \frac{1}{2} \bar{q}' G' B^{-1} D(I_k \otimes B^{-1} G \bar{q}) M_2 \text{vec} D] + o_p(N^{-1}) \\
&= N^{-1/2} u_2(\bar{q}) + N^{-1} v_2(\bar{q}) + o_p(N^{-1}),
\end{aligned} \tag{33}$$

where

$$\begin{aligned}
u_2(\bar{q}) &= \bar{q}' G' B^{-1} G M_2 \text{vec} D - \bar{q}' M_2 \text{vec} D \\
&= \bar{q}' (G' B^{-1} G - I_k) M_2 \text{vec} D, \\
v_2(\bar{q}) &= \bar{q}' G' B^{-1} G M_2 C' B^{-1} G \bar{q} - \bar{q}' M_2 C' B^{-1} G \bar{q} + \frac{1}{2} \bar{q}' G' B^{-1} D(I_k \otimes B^{-1} G \bar{q}) M_2 \text{vec} D \\
&= \bar{q}' (G' B^{-1} G - I_k) M_2 C' B^{-1} G \bar{q} + \frac{1}{2} \bar{q}' G' B^{-1} D(I_k \otimes B^{-1} G \bar{q}) M_2 \text{vec} D;
\end{aligned}$$

$$\begin{aligned}
3_{DM} &= -N^{-1/2} \text{vec}' G_N(\hat{\theta}_N) M_3 \text{vec} D_N(\hat{\theta}_N) \\
&= -N^{-1/2} \text{vec}' G M_3 \text{vec} D - N^{-1} \text{vec}' G M_3 C' B^{-1} G \bar{q} - N^{-1} \bar{q}' G' B^{-1} D M_3 \text{vec} D + o_p(N^{-1}) \\
&= N^{-1/2} u_3(\bar{q}) + N^{-1} v_3(\bar{q}) + o_p(N^{-1}),
\end{aligned} \tag{34}$$

and

$$\begin{aligned}
u_3(\bar{q}) &= -\text{vec}' G M_3 \text{vec} D, \\
v_3(\bar{q}) &= -\text{vec}' G M_3 C' B^{-1} G \bar{q} - \bar{q}' G' B^{-1} D M_3 \text{vec} D \\
&= -\bar{q}' G' B^{-1} C M_3' \text{vec} G - \bar{q}' G' B^{-1} D M_3 \text{vec} D;
\end{aligned}$$

$$\begin{aligned}
4_{DM} &= N^{-1} \frac{1}{4} \text{vec}' D_N(\hat{\theta}_N) M_4 \text{vec} D_N(\hat{\theta}_N) \\
&= N^{-1} \frac{1}{4} \text{vec}' D M_4 \text{vec} D + o_p(N^{-1}) \\
&= N^{-1} v_4(\bar{q}) + o_p(N^{-1}),
\end{aligned} \tag{35}$$

where

$$v_4(\bar{q}) = \frac{1}{4} \text{vec}' D M_4 \text{vec} D.$$

Finally, collecting the terms (32)-(35) gives equation (12).  $\square$

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

$$u_i(\bar{q}) = \text{vec}' J_i(\bar{q} \otimes \bar{q} \otimes \bar{q}), \tag{36}$$

$$v_i(\bar{q}) = \text{tr}[L_i(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')], \tag{37}$$

then,

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

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

where

$$\text{vec} J = \text{vec} J_1 + \text{vec} J_2 + \text{vec} J_3,$$

and

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

Therefore, the proof is reduced to showing (36) and (37).

Using

$$\begin{aligned}(A \otimes C)(B \otimes D) &= (AB) \otimes (CD), \\ K_{p,q} \text{vec} A &= \text{vec}(A'), \\ A \otimes B &= K_{p,r}(B \otimes A)K_{s,q},\end{aligned}$$

for  $A : p \times q$  and  $B : r \times s$  where  $K$  is the commutation matrix, we can rewrite (14):

$$\begin{aligned}u_1(\bar{q}) &= 2\bar{q}'G'B^{-1}D(I_k \otimes \mathbb{H}G\bar{q})\text{vec}(\bar{q}'G'\mathbb{H}G) \\ &= 2\bar{q}'G'\mathbb{H}G(I_k \otimes \bar{q}'G'\mathbb{H})(\bar{q}'G'B^{-1} \otimes I_{pk})\text{vec}(D') \\ &= 2\bar{q}'G'\mathbb{H}G(I_k \otimes \bar{q}'G'\mathbb{H})(I_{pk} \otimes \bar{q}'G'B^{-1})\text{vec}D \\ &= 2(\bar{q}'G'\mathbb{H}G \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'B^{-1})\text{vec}D \\ &= 2(\bar{q}' \otimes \bar{q}' \otimes \bar{q}')(G'\mathbb{H}G \otimes G'\mathbb{H} \otimes G'B^{-1})\text{vec}D \\ &= \text{vec}'J_1(\bar{q} \otimes \bar{q} \otimes \bar{q}),\end{aligned}\tag{38}$$

where

$$\text{vec}J_1 = 2(G'\mathbb{H}G \otimes G'\mathbb{H} \otimes G'B^{-1})\text{vec}D.\tag{39}$$

Let

$$R_1 = (\mathbb{H}G \otimes B^{-1}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')(G'\mathbb{H} \otimes G'B^{-1}),\tag{40}$$

partition  $\text{vec}D$  as

$$\text{vec}D = \begin{bmatrix} V_{D1} \\ V_{D2} \\ \vdots \\ V_{Dk} \end{bmatrix}_{p^2k \times 1}\tag{41}$$

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

$$V_D = V_{D1}V'_{D1} + V_{D2}V'_{D2} + \cdots + V_{Dk}V'_{Dk}.\tag{42}$$

Then, since

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

the first term of  $v_1(\bar{q})$  in (17) becomes

$$\begin{aligned}
&\bar{q}'G'B^{-1}D(I_k \otimes \mathbb{H}G\bar{q})(I_k \otimes \bar{q}'G'\mathbb{H})D'B^{-1}G\bar{q} \\
&= \text{vec}'D(I_k \otimes \mathbb{H}G\bar{q} \otimes B^{-1}G\bar{q})(I_k \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'B^{-1})\text{vec}D \\
&= \text{vec}'D(I_k \otimes R_1)\text{vec}D \\
&= \begin{bmatrix} V'_{D1} & V'_{D2} & \cdots & V'_{Dk} \end{bmatrix} \begin{bmatrix} R_1 & & & 0 \\ & R_1 & & \\ & & \ddots & \\ 0 & & & R_1 \end{bmatrix} \begin{bmatrix} V_{D1} \\ V_{D2} \\ \vdots \\ V_{Dk} \end{bmatrix} \tag{43} \\
&= V'_{D1}R_1V_{D1} + V'_{D2}R_1V_{D2} + \cdots + V'_{Dk}R_1V_{Dk} \\
&= \text{tr}[(V_{D1}V'_{D1} + V_{D2}V'_{D2} + \cdots + V_{Dk}V'_{Dk})R_1] \\
&= \text{tr}[V_D(\mathbb{H}G \otimes B^{-1}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')(G'\mathbb{H} \otimes G'B^{-1})] \\
&= \text{tr}[(G'\mathbb{H} \otimes G'B^{-1})V_D(\mathbb{H}G \otimes B^{-1}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')].
\end{aligned}$$

Similarly, let

$$R_2 = (\mathbb{H}G \otimes B^{-1}G)(\bar{q} \otimes \bar{q}), \tag{44}$$

$$R_3 = \bar{q}'G'\mathbb{H}, \tag{45}$$

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

$$G'B^{-1}C = \begin{bmatrix} M_{GC1} & M_{GC2} & \cdots & M_{GCk} \end{bmatrix}, \quad (46)$$

$$vecG = \begin{bmatrix} V_{G1} \\ V_{G2} \\ \vdots \\ V_{Gk} \end{bmatrix}, \quad (47)$$

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

$$M_V = M'_{GC1} \otimes V'_{G1} + M'_{GC2} \otimes V'_{G2} + \cdots + M'_{GCk} \otimes V'_{Gk}. \quad (48)$$

Then, since

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



for some vector  $m$  and matrix  $M$  of appropriate sizes, the second term of  $v_1(\bar{q})$  in (17) becomes

$$\begin{aligned}
& \bar{q}' G' B^{-1} C (I_{pk} \otimes B^{-1} G \bar{q}) (I_k \otimes \mathbb{H} G \bar{q}) (I_k \otimes \bar{q}' G' \mathbb{H}) \text{vec} G \\
&= \bar{q}' G' B^{-1} C (I_k \otimes R_2) (I_k \otimes R_3) \text{vec} G \\
&= \bar{q}' \begin{bmatrix} M_{GC1} & M_{GC2} & \cdots & M_{GCk} \end{bmatrix} \begin{bmatrix} R_2 & & & 0 \\ & R_2 & & \\ & & \ddots & \\ 0 & & & R_2 \end{bmatrix} \begin{bmatrix} R_3 & & & 0 \\ & R_3 & & \\ & & \ddots & \\ 0 & & & R_3 \end{bmatrix} \begin{bmatrix} V_{G1} \\ V_{G2} \\ \vdots \\ V_{Gk} \end{bmatrix} \\
&= \sum_{i=1}^k (\bar{q}' M_{GCi} R_2 R_3 V_{Gi}) \\
&= \text{tr} \sum_{i=1}^k [\bar{q}' M_{GCi} (\mathbb{H} G \otimes B^{-1} G) (\bar{q} \otimes \bar{q}) \bar{q}' G' \mathbb{H} V_{Gi}] \\
&= \text{tr} \sum_{i=1}^k [\bar{q}' G' \mathbb{H} V_{Gi} \bar{q}' M_{GCi} (\mathbb{H} G \otimes B^{-1} G) (\bar{q} \otimes \bar{q})] \\
&= \text{tr} \sum_{i=1}^k \{ [(G' \mathbb{H} \otimes G' B^{-1}) M'_{GCi}] \otimes V'_{Gi} \mathbb{H} G \} (\bar{q} \bar{q}' \otimes \bar{q} \bar{q}') \} \\
&= \text{tr} \sum_{i=1}^k [(G' \mathbb{H} \otimes G' B^{-1}) (M'_{GCi} \otimes V'_{Gi}) (I_k \otimes \mathbb{H} G) (\bar{q} \bar{q}' \otimes \bar{q} \bar{q}')] \\
&= \text{tr} [(G' \mathbb{H} \otimes G' B^{-1}) M_V (I_k \otimes \mathbb{H} G) (\bar{q} \bar{q}' \otimes \bar{q} \bar{q}')].
\end{aligned} \tag{49}$$

From (43) and (49), (17) can be rewritten as

$$v_1(\bar{q}) = \text{tr} [L_1 (\bar{q} \bar{q}' \otimes \bar{q} \bar{q}')], \tag{50}$$

where

$$L_1 = (G' \mathbb{H} \otimes G' B^{-1}) V_D (\mathbb{H} G \otimes B^{-1} G) + (G' \mathbb{H} \otimes G' B^{-1}) M_V (I_k \otimes \mathbb{H} G). \tag{51}$$

Similar to  $u_1(\bar{q})$ ,  $u_2(\bar{q})$  in (15) can be rewritten as

$$\begin{aligned}
u_2(\bar{q}) &= \bar{q}' (G' B^{-1} G - I_k) (I_k \otimes \bar{q}' G' \mathbb{H} \otimes \bar{q}' G' \mathbb{H}) \text{vec} D \\
&= (\bar{q}' \otimes \bar{q}' \otimes \bar{q}') [(G' B^{-1} G - I_k) \otimes G' \mathbb{H} \otimes G' \mathbb{H}] \text{vec} D \\
&= \text{vec}' J_2 (\bar{q} \otimes \bar{q} \otimes \bar{q}),
\end{aligned} \tag{52}$$

where

$$vecJ_2 = [(G'B^{-1}G - I_k) \otimes G'\mathbb{H} \otimes G'\mathbb{H}]vecD. \quad (53)$$

The first term of  $v_2(\bar{q})$  in (18) can be written as

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

Since

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

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

$$vec(G'B^{-1}G - I_k) = \begin{bmatrix} V_{GI1} \\ V_{GI2} \\ \vdots \\ V_{GIk} \end{bmatrix} \quad (54)$$

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

$$\begin{aligned} tr \sum_{i=1}^k [\bar{q}' M_{GCi} (\mathbb{H}G \otimes \mathbb{H}G) (\bar{q} \otimes \bar{q}) \bar{q}' V_{GIi}] \\ = tr[(G'\mathbb{H} \otimes G'\mathbb{H}) M_{VI} (\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')], \end{aligned} \quad (55)$$

where

$$M_{VI} = M'_{GC1} \otimes V'_{GI1} + M'_{GC2} \otimes V'_{GI2} + \cdots + M'_{GCk} \otimes V'_{GIk}. \quad (56)$$

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

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

the second term of  $v_2(\bar{q})$  in (18) can be rewritten as

$$\begin{aligned}
& \frac{1}{2} \text{vec}' D(I_k \otimes B^{-1}G\bar{q} \otimes B^{-1}G\bar{q})(I_k \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H}) \text{vec} D \\
&= \frac{1}{2} \text{tr}[V_D(B^{-1}G \otimes B^{-1}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')(G'\mathbb{H} \otimes G'\mathbb{H})] \\
&= \text{tr}\left[\frac{1}{2}(G'\mathbb{H} \otimes G'\mathbb{H})V_D(B^{-1}G \otimes B^{-1}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')\right].
\end{aligned} \tag{57}$$

From (55) and (57), we have

$$v_2(\bar{q}) = \text{tr}[L_2(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')], \tag{58}$$

where

$$L_2 = (G'\mathbb{H} \otimes G'\mathbb{H})M_{VI} + \frac{1}{2}(G'\mathbb{H} \otimes G'\mathbb{H})V_D(B^{-1}G \otimes B^{-1}G). \tag{59}$$

Since

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

(16) becomes

$$\begin{aligned}
u_3(\bar{q}) &= -\bar{q}'G'\mathbb{H}G(I_k \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H}) \text{vec} D \\
&= -(\bar{q}' \otimes \bar{q}' \otimes \bar{q}')(G'\mathbb{H}G \otimes G'\mathbb{H} \otimes G'\mathbb{H}) \text{vec} D \\
&= \text{vec}' J_3(\bar{q} \otimes \bar{q} \otimes \bar{q}),
\end{aligned} \tag{60}$$

where

$$\text{vec} J_3 = -(G'\mathbb{H}G \otimes G'\mathbb{H} \otimes G'\mathbb{H}) \text{vec} D. \tag{61}$$

Similar to the second term of  $v_1(\bar{q})$ , the first term of  $v_3(\bar{q})$  in (19) can be rewritten as

$$\begin{aligned}
& -\bar{q}'G'B^{-1}C(I_k \otimes \mathbb{H}G\bar{q} \otimes \mathbb{H}G\bar{q})(I_k \otimes \bar{q}'G'\mathbb{H})\text{vec}G \\
& = \text{tr} \sum_{i=1}^k [-\bar{q}'M_{GC_i}(\mathbb{H}G \otimes \mathbb{H}G)(\bar{q} \otimes \bar{q})\bar{q}'G'\mathbb{H}V_{G_i}] \\
& = \text{tr}[-(G'\mathbb{H} \otimes G'\mathbb{H})M_V(I_k \otimes \mathbb{H}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')].
\end{aligned} \tag{62}$$

Similar to the second term of  $v_2(\bar{q})$ , the second term of  $v_3(\bar{q})$  in (19) can be rewritten as

$$\begin{aligned}
& -\bar{q}'G'B^{-1}D(I_k \otimes \mathbb{H}G\bar{q})(I_k \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H})\text{vec}D \\
& = -\text{vec}'D(I_{pk} \otimes B^{-1}G\bar{q})(I_k \otimes \mathbb{H}G\bar{q})(I_k \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H})\text{vec}D \\
& = -\text{vec}'D(I_k \otimes \mathbb{H}G\bar{q} \otimes B^{-1}G\bar{q})(I_k \otimes \bar{q}'G'\mathbb{H} \otimes \bar{q}'G'\mathbb{H})\text{vec}D \\
& = \text{tr}[-V_D(\mathbb{H}G \otimes B^{-1}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')(G'\mathbb{H} \otimes G'\mathbb{H})] \\
& = \text{tr}[-(G'\mathbb{H} \otimes G'\mathbb{H})V_D(\mathbb{H}G \otimes B^{-1}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')].
\end{aligned} \tag{63}$$

From (62) and (63), we have

$$v_3(\bar{q}) = \text{tr}[L_3(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')], \tag{64}$$

where

$$L_3 = -(G'\mathbb{H} \otimes G'\mathbb{H})M_V(I_k \otimes \mathbb{H}G) - (G'\mathbb{H} \otimes G'\mathbb{H})V_D(\mathbb{H}G \otimes B^{-1}G). \tag{65}$$

Similar to the first term of  $v_1(\bar{q})$ ,  $v_4(\bar{q})$  in (20) can be easily rewritten as

$$\begin{aligned}
v_4(\bar{q}) & = \frac{1}{4}\text{tr}[V_D(\mathbb{H}G \otimes \mathbb{H}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')(G'\mathbb{H} \otimes G'\mathbb{H})] \\
& = \text{tr}[\frac{1}{4}(G'\mathbb{H} \otimes G'\mathbb{H})V_D(\mathbb{H}G \otimes \mathbb{H}G)(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')] \\
& = \text{tr}[L_4(\bar{q}\bar{q}' \otimes \bar{q}\bar{q}')],
\end{aligned} \tag{66}$$

where

$$L_4 = \frac{1}{4}(G'\mathbb{H} \otimes G'\mathbb{H})V_D(\mathbb{H}G \otimes \mathbb{H}G). \tag{67}$$

By using (38), (50), (52), (58), (60), (64) and (66), we obtain (36) and (37), thus finishing the proof.  $\square$

*Proof of Theorem 2:* First,  $a_i$  and  $b_i$  are defined (Phillips and Park, 1988) as

$$a_i = \text{tr}(A_i) \quad (i = 0, 1, 2), \quad (68)$$

where

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

$$b_i = \text{vec}'JB_i\text{vec}J \quad (i = 1, 2, 3), \quad (69)$$

where

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

with

$$H = I + K_{k,k^2} + K_{k^2,k},$$

$$\bar{P} \equiv I - P.$$

Secondly, from (68),

$$\begin{aligned} a_0 = \text{tr}(A_0) &= \text{tr}\{L[(I + K_{k,k})(\bar{P} \otimes \bar{P}) + \text{vec}\bar{P}\text{vec}'\bar{P}]\} \\ &= \text{tr}[(\bar{P} \otimes \bar{P})L(I + K_{k,k}) + \text{vec}'\bar{P}L\text{vec}\bar{P}] \\ &= \text{tr}[(\bar{P} \otimes \bar{P})L(I + K_{k,k})] + \text{tr}(\text{vec}'\bar{P}L\text{vec}\bar{P}). \end{aligned} \quad (70)$$

Using (13) and  $\bar{P} \equiv I - P$ , we have

$$(A'B^{-1}G)\bar{P} = 0, \quad (71)$$

$$\bar{P}(G'B^{-1}A) = 0. \quad (72)$$

Therefore, by (21)-(25),

$$(\bar{P} \otimes \bar{P})L = 0, \quad (73)$$

and

$$(\mathbb{H}G \otimes B^{-1}G)\text{vec}\bar{P} = \text{vec}(B^{-1}G\bar{P}G\mathbb{H}) = 0, \quad (74)$$

$$(I_k \otimes \mathbb{H}G)\text{vec}\bar{P} = \text{vec}(\mathbb{H}G\bar{P}) = 0. \quad (75)$$

Combining (74) and (75) with (22) yields

$$L_1\text{vec}\bar{P} = 0. \quad (76)$$

Similarly,

$$L_3\text{vec}\bar{P} = 0, \quad (77)$$

$$L_4\text{vec}\bar{P} = 0, \quad (78)$$

and

$$vec' \bar{P} L_2 = (L_2' vec \bar{P})' = 0. \quad (79)$$

From (76)-(79),

$$tr(vec' \bar{P} L vec \bar{P}) = 0. \quad (80)$$

Substituting (73) and (80) into (70) gives

$$a_0 = 0. \quad (81)$$

Also, from (69),

$$\begin{aligned} b_1 &= vec' J B_1 vec J \\ &= vec' J H(P \otimes \bar{P} \otimes \bar{P}) H vec J \\ &\quad + vec' J H(P \otimes vec \bar{P} vec' \bar{P} + \bar{P} \otimes vec P vec' \bar{P} + \bar{P} \otimes vec \bar{P} vec' P) H vec J \\ &\quad + vec' J [P \otimes K_{k,k}(\bar{P} \otimes \bar{P}) + \bar{P} \otimes K_{k,k}(P \otimes \bar{P})] vec J \\ &\quad + vec' J [\bar{P} \otimes K_{k,k}(\bar{P} \otimes P) + K_{k,k}(P \otimes \bar{P}) \otimes \bar{P}] vec J \\ &\quad + vec' J [K_{k,k}(\bar{P} \otimes \bar{P}) \otimes \bar{P} + K_{k,k}(\bar{P} \otimes \bar{P}) \otimes P] vec J \\ &\quad + vec' J K_{k,k^2} \{ [P \otimes K_{k,k}(\bar{P} \otimes \bar{P})] + [\bar{P} \otimes K_{k,k}(P \otimes \bar{P})] \\ &\quad \quad + [\bar{P} \otimes K_{k,k}(\bar{P} \otimes P)] \} K_{k^2,k} vec J. \end{aligned} \quad (82)$$

Using

$$K_{p,q} vec A = vec(A'),$$

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

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

$$K_{k,k^2} \text{vec} J_1 = 2(G' B^{-1} \otimes G' \mathbb{H} G \otimes G' \mathbb{H}) \text{vec}(D'), \quad (83)$$

$$K_{k,k^2} \text{vec} J_2 = [G' \mathbb{H} \otimes (G' B^{-1} G - I_k) \otimes G' \mathbb{H}] \text{vec}(D'), \quad (84)$$

$$K_{k,k^2} \text{vec} J_3 = -(G' \mathbb{H} \otimes G' \mathbb{H} G \otimes G' \mathbb{H}) \text{vec}(D'); \quad (85)$$

$$K_{k^2,k} \text{vec} J_1 = 2(G' \mathbb{H} \otimes G' B^{-1} \otimes G' \mathbb{H} G) K_{p^2,k} \text{vec} D, \quad (86)$$

$$K_{k^2,k} \text{vec} J_2 = [G' \mathbb{H} \otimes G' \mathbb{H} \otimes (G' B^{-1} G - I_k)] K_{p^2,k} \text{vec} D, \quad (87)$$

$$K_{k^2,k} \text{vec} J_3 = -(G' \mathbb{H} \otimes G' \mathbb{H} \otimes G' \mathbb{H} G) K_{p^2,k} \text{vec} D. \quad (88)$$

Then, substituting (83)-(88) into (82), and using

$$\text{vec}(ABC) = (C' \otimes A) \text{vec} B,$$

$$(A \otimes B)' = A' \otimes B',$$

$$(A \otimes C)(B \otimes D) = (AB) \otimes (CD),$$

together with (71) and (72) yield

$$b_1 = 0. \quad (89)$$

Given (81) and (89), the proof of Theorem 2.4 in Phillips and Park (1988) establishes the conclusion of Theorem 2.  $\square$

## C 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 from 1969 to 1974, between the ages of 21 (not inclusive) and 64 (not inclusive), 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.

Variables V7492, V7490, V0313, V0794, V7460, V7476, V7491 listed on p.443 of Abowd and Card (1989)



are not available now on the PSID website. The variables for sex listed on that page are not consistent with those on the PSID website. The following are the PSID variables used here:

- ANNUAL EARNINGS: V1196, V1897, V2498, V3051, V3463, V3863;
- ANNUAL HOURS: V1138, V1839, V2439, V3027, V3423, V3823;
- SEX: ER32000;
- AGE: ER30046.

## 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.
- BARNDORFF-NIELSEN, O. AND D. R. COX (1979): “Edgeworth and Saddle-Point Approximations with Statistical Applications,” *Journal of the Royal Statistical Society. Series B (Methodological)*, 41, 279–312.
- BARTLETT, M. S. (1937): “Properties of Sufficiency and Statistical Tests,” *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 160, 268–282.
- BREUSCH, T. S. AND P. SCHMIDT (1988): “Alternative forms of the wald test: how long is a piece of string?” *Communications in Statistics - Theory and Methods*, 17, 2789 – 2795.
- BROWNE, M. W. (1984): “Asymptotically distribution-free methods for the analysis of covariance structures.” *British Journal of Mathematical and Statistical Psychology*, 37, 62–83.
- CHOU, C.-P., P. M. BENTLER, AND A. SATORRA (1991): “Scaled test statistics and robust standard errors for non-normal data in covariance structure analysis: A Monte Carlo study,” *British Journal of Mathematical and Statistical Psychology*, 44, 347–357.
- CLARK, T. E. (1996): “Small-Sample Properties of Estimators of Nonlinear Models of Covariance Structure,” *Journal of Business and Economic Statistics*, 14, 367–73.
- CRIBARI-NETO, F. AND G. M. CORDEIRO (1996): “On Bartlett and Bartlett-type corrections,” *Econometric Reviews*, 15, 339 – 367.
- HALL, P. (1992): *The Bootstrap and the Edgeworth Expansion*, New York: Springer-Verlag.

- 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., J. HEATON, AND A. YARON (1996): “Finite-sample properties of some alternative GMM estimators,” *Journal of Business and Economic Statistics*, 14, 262–280.
- HAUSMAN, J. AND G. KUERSTEINER (2008): “Difference in difference meets generalized least squares: Higher order properties of hypotheses tests,” *Journal of Econometrics*, 144, 371–391.
- HERZOG, W., A. BOOMSMA, AND S. REINECKE (2007): “The Model-Size Effect on Traditional and Modified Tests of Covariance Structures,” *Structural Equation Modeling: A Multidisciplinary Journal*, 14, 361 – 390.
- HOOGLAND, J. J. AND A. BOOMSMA (1998): “Robustness Studies in Covariance Structure Modeling: An Overview and a Meta-Analysis,” *Sociological Methods Research*, 26, 329–367.
- KALLENBERG, W. C. M. (1993): “Interpretation and manipulation of Edgeworth expansions,” *Annals of the Institute of Statistical Mathematics*, 45, 341–351.
- 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: A Multidisciplinary Journal*, 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.
- LINTON, O. (2002): “Edgeworth approximations for semiparametric instrumental variable estimators and test statistics,” *Journal of Econometrics*, 106, 325–368.
- 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.
- MAGEE, L. (1989): “An Edgeworth Test Size Correction for the Linear Model with AR(1) Errors,” *Econometrica*, 57, 661–674.
- MUTHEN, B. AND D. KAPLAN (1992): “A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model,” *British Journal of Mathematical and Statistical Psychology*, 45, 19–30.

- 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.
- NEWBY, W. AND D. MCFADDEN (1994): "Large sample estimation and hypothesis testing," in *Handbook of Econometrics*, vol. IV, 2113–2241.
- NEWBY, W. K. AND R. J. SMITH (2004): "Higher order properties of GMM and Generalized Empirical Likelihood estimators," *Econometrica*, 72, 219–255.
- NEWBY, W. K. AND K. D. WEST (1987): "Hypothesis Testing with Efficient Method of Moments Estimation," *International Economic Review*, 28, 777–787.
- OGASAWARA, H. (2009): "Asymptotic expansions of the distributions of the chi-square statistic based on the asymptotically distribution-free theory in covariance structures," *Journal of Statistical Planning and Inference*, 139, 3246–3261.
- PHILLIPS, P. C. B. (1977): "A General Theorem in the Theory of Asymptotic Expansions as Approximations to the Finite Sample Distributions of Econometric Estimators," *Econometrica*, 45, 1517–1534.
- (1978): "Edgeworth and Saddlepoint Approximations in the First-Order Noncircular Autoregression," *Biometrika*, 65, 91–98.
- (1983): "ERA's: A New Approach to Small Sample Theory," *Econometrica*, 51, 1505–1525.
- PHILLIPS, P. C. B. AND J. Y. PARK (1988): "On the Formulation of Wald Tests of Nonlinear Restrictions," *Econometrica*, 56, 1065–1083.
- REID, N. (1991): "Approximations and asymptotics," in *Statistical Theory and Modeling: In Honour of Sir David Cox, FRS*, ed. by D. Hinkley, N. Reid, and E. Snell, London: Chapman and Hall.
- ROTHENBERG, T. (1984): "Approximating the distributions of econometric estimators and test statistics," in *Handbook of Econometrics*, ed. by Z. Griliches and M. D. Intriligator, vol. II, 881–935.
- SARGAN, J. D. AND S. E. SATCHELL (1986): "A Theorem of Validity for Edgeworth Expansions," *Econometrica*, 54, 189–213.
- SATORRA, A. AND P. BENTLER (2001): "A scaled difference chi-square test statistic for moment structure analysis," *Psychometrika*, 66, 507–514.

- 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 (1997): "Mean and Covariance Structure Analysis: Theoretical and Practical Improvements," *Journal of the American Statistical Association*, 92, 767–774.
- (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.