Econometric Theory, **26**, 2010, 1838–1845.
doi:10.1017/S0266466609990818

GENERAL SPECIFICATION TESTING WITH LOCALLY MISSPECIFIED MODELS

ANIL K. BERA
University of Illinois

GABRIEL MONTES-ROJAS
City University London

WALTER SOSA-ESCUDERO
Universidad de San Andrés

A well known result is that many of the tests used in econometrics, such as the Rao score (RS) test, may not be robust to misspecified alternatives, that is, when the alternative model does not correspond to the underlying data generating process. Under this scenario, these tests spuriously reject the null hypothesis too often. We generalize this result to generalized method of moments–based (GMM-based) tests. We also extend the method proposed in Bera and Yoon (1993, *Econometric Theory* 9, 649–658) for constructing RS tests that are robust to local misspecification to GMM-based tests. Finally, a further generalization for general estimating and testing functions is developed. This framework encompasses both likelihood and GMM-based results.

1. INTRODUCTION

The standard Rao's score (RS) test based on the maximum-likelihood (ML) framework has been extensively used to derive tests for misspecification, especially when the estimation of a restricted model is computationally convenient. Nevertheless, Davidson and MacKinnon (1987) and Saikkonen (1989) showed that the RS test is adversely affected when the alternative hypothesis is incorrectly specified, that is, when the true model does not correspond to the alternative postulated by the researcher. Bera and Yoon (1993) proposed a modified RS test (BY test) that, albeit still based on a fully restricted ML estimator, is immune to local misspecification. This principle has been successfully implemented in many econometric "model search" problems; see, for example, Anselin, Bera, Florax,

We are grateful to the coeditor Jinyong Hahn and two anonymous referees for many pertinent comments and suggestions. However, we retain the responsibility for any remaining shortcomings. Address correspondence to Anil Bera, Department of Economics, University of Illinois, Urbana-Champaign, 1407 W. Gregory Drive, Urbana, IL 61801 USA, e-mail: abera@ad.uiuc.edu.

  0266-4666/10 $15.00

and Yoon (1996), Godfrey and Veall (2000), Bera, Sosa-Escudero, and Yoon (2001), and Baltagi and Li (2001).

An obviously restrictive feature of likelihood-based procedures is that they require complete specification of the underlying probabilistic structure of the model and that limits the scope of the BY procedure. This paper derives a BY adjusted type test in the generalized method of moments (GMM) framework that requires specification of some moment conditions only. The proposed test can be viewed as a BY type modification of the Newey and West (1987) formulation of the RS test under GMM setup. A further generalization provides robust BY tests in a general estimating and testing functions setup. Bera and Yoon (1993) showed that for local misspecification, their adjusted test is asymptotically equivalent to the Neyman (1959) $C(\alpha)$ test and therefore the BY procedure shares its optimality properties. We can expect the tests suggested in this paper to possess certain optimality properties.

2. RS TEST AND LOCAL MISSPECIFICATION

Let us denote the log-likelihood of n independent and identically distributed (i.i.d.) random variables $z_1, z_2, \ldots, z_n$ by $\ell_n(\theta)$ and consider the following partition of the parameter space $\theta = (\theta_1', \theta_2', \theta_3')'$. Here θ_1, θ_2, and θ_3 are, respectively, vectors in open subsets of $\Re^{p_1}$, $\Re^{p_2}$, and $\Re^{p_3}$, and thus the dimension of θ is $p_1 + p_2 + p_3 = p$. Let $d_{j,n}(\theta)$ denote the score vectors $n^{-1}\partial\ell_n(\theta)/\partial\theta_j$, $j = 1, 2, 3$. The information matrix $J(\theta)$ is given by

$$J(\theta) = -\mathrm{E}\left[\frac{1}{n}\frac{\partial^2 l_n(\theta)}{\partial\theta\partial\theta'}\right] = \begin{bmatrix} J_{11}(\theta) & J_{12}(\theta) & J_{13}(\theta) \\ J_{12}(\theta) & J_{22}(\theta) & J_{23}(\theta) \\ J_{13}(\theta) & J_{23}(\theta) & J_{33}(\theta) \end{bmatrix}.$$

Consider the problem of testing $H_0^2 : \theta_2 = \theta_{20}$ when $H_0^3 : \theta_3 = \theta_{30}$ holds. Let $\hat{\theta} = (\hat{\theta}_1', \theta_{20}', \theta_{30}')'$, where $\hat{\theta}_1$ is the ML estimator of θ_1 under the joint null H_0^{23} : $\theta_2 = \theta_{20}$, $\theta_3 = \theta_{30}$. A standard result is that under the local alternative $H_A^2 : \theta_2 = \theta_{20} + \delta_2/\sqrt{n}$, $0 < \delta_2 < \infty$, and H_0^3,

$$RS_{2\cdot 1}(\hat{\theta}) = n\, d_{2,n}(\hat{\theta})' J_{2\cdot 1}^{-1}(\hat{\theta}) d_{2,n}(\hat{\theta}) \ \xrightarrow{d}\ \chi^2_{p_2}(\lambda_{2\cdot 1}),$$

with $J_{2\cdot 1}(\theta) = J_{22}(\theta) - J_{21}(\theta)J_{11}^{-1}(\theta)J_{12}(\theta)$, and the noncentrality parameter $\lambda_{2\cdot 1} = \delta_2' J_{2\cdot 1}(\theta_0)\delta_2$, where $\theta_0 = (\theta_{10}', \theta_{20}', \theta_{30}')'$, θ_{10} being the true value of θ_1. Therefore, under H_0^2, $RS_{2\cdot 1}(\hat{\theta})$ has, asymptotically, a *central* chi-squared distribution and hence asymptotically correct size.

Davidson and MacKinnon (1987) and Saikkonen (1989) showed that when the alternative hypothesis is locally *misspecified*, that is, when $H_A^3 : \theta_3 = \theta_{30} + \delta_3/\sqrt{n}$ holds, $RS_{2\cdot 1}(\hat{\theta})$ no longer has a *central* asymptotic distribution. In fact, they showed that

$$RS_{2\cdot 1}(\hat{\theta}) \xrightarrow{d} \chi^2_{p_2}(\lambda_{2/3\cdot 1}),$$

with $\lambda_{2/3\cdot 1} = \delta_3' J_{32\cdot 1}(\theta_0) J_{2\cdot 1}^{-1}(\theta_0) J_{23\cdot 1}(\theta_0)\delta_3$ and $J_{23\cdot 1}(\theta) = J_{23}(\theta) - J_{21}(\theta) J_{11}^{-1}(\theta) J_{13}(\theta) = J_{32\cdot 1}'(\theta)$. In particular, when $J_{23\cdot 1}(\theta_0)$, which measures the partial covariance between $d_{2,n}(\hat{\theta})$ and $d_{3,n}(\hat{\theta})$ after controlling for the linear effect of $d_{1,n}(\hat{\theta})$, is a null matrix, then $\lambda_{2/3\cdot 1} = 0$; that is, the local misspecification of θ_3 has no asymptotic effect on the performance of $RS_{2\cdot 1}(\hat{\theta})$.

Bera and Yoon (1993) proposed a locally size-robust version of $RS_{2\cdot 1}(\hat{\theta})$, given by

$$RS_{2\cdot 1}^*(\hat{\theta}) = n\, d_{2\cdot 1,n}^*(\hat{\theta})' \left[J_{2\cdot 1}(\hat{\theta}) - J_{23\cdot 1}(\hat{\theta}) J_{2\cdot 1}^{-1}(\hat{\theta}) J_{32\cdot 1}(\hat{\theta}) \right]^{-1} d_{2\cdot 1,n}^*(\hat{\theta}),$$

where $d_{2\cdot 1,n}^*(\theta) \equiv d_{2,n}(\theta) - J_{23\cdot 1}(\theta) J_{3\cdot 1}^{-1}(\theta) d_{3,n}(\theta)$, and showed that when H_0^2 is true and H_0^3 or H_A^3 holds, $RS_{2\cdot 1}^*(\hat{\theta}) \xrightarrow{d} \chi_{p_2}^2(0)$, and so the test is robust to local misspecification because it preserves the *central* χ^2 asymptotic distribution even under local departures of θ_3 away from θ_{30}.

3. GMM-BASED ROBUST TESTS

An obvious restrictive feature of likelihood-based procedures is that they require full specification of the underlying probabilistic model, which limits the scope of the BY procedure. In this section we derive BY type adjustments to GMM-based RS tests, requiring moment conditions only.

We will assume that there is a vector of m functions $g(Z,\theta)$ satisfying the following moment conditions:

$$\mathrm{E}\left[g(Z,\theta)\right] = 0 \quad \text{if and only if } \theta = \theta_0,$$

where θ and θ_0 are vectors in open subsets of $\Re^p$ and for identification purposes we require $m \geq p$. The sample analog of the left-hand side of the preceding equation is

$$g_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} g(z_i,\theta).$$

Let $\Omega_n(\theta)$ be a $m \times m$ positive definite symmetric matrix. We consider the continuous updating estimator (CUE) version of GMM. See Hansen, Heaton, and Yaron (1996). Our (unrestricted) GMM estimator of θ_0 will be $\operatorname{argmax}_\theta Q_n(\theta)$, with $Q_n(\theta) = -\frac{1}{2} g_n(\theta)' \Omega_n^{-1}(\theta) g_n(\theta)$, which can be viewed as a counterpart of the log-likelihood function $\ell_n(\theta)$. Let $\Omega(\theta) = \mathrm{E}[g(Z,\theta)g(Z,\theta)']$. For asymptotic efficiency we will assume $\Omega_n(\theta) \xrightarrow{p} \Omega(\theta)$ (see Hansen, 1982; Newey and McFadden, 1994).

Let $\nabla_\theta g(z,\theta) = \partial g(z,\theta)/\partial\theta'$ be the $m \times p$ Jacobian matrix of $g(z,\theta)$, $G(\theta) = \mathrm{E}[\nabla_\theta g(Z,\theta)]$ and $G_n(\theta) = 1/n \sum_{i=1}^{n} \nabla_\theta g(z_i,\theta)$. Define the counterpart of the score (pseudo-score) as $q_n(\theta) = -G_n(\theta)\Omega_n^{-1}(\theta) g_n(\theta)$ and define $q_{j,n}(\theta)$ as the

$p_j \times 1$ subvector, $j = 1, 2, 3$. Also, let $B(\theta) = G(\theta)'\Omega^{-1}(\theta)G(\theta)$ and $B_n(\theta) = G_n(\theta)'\Omega_n^{-1}(\theta)G_n(\theta)$. We partition $B(\theta)$ and $B_n(\theta)$ the same way we partitioned the information matrix $J(\theta)$ for the three parameter vectors, θ_1, θ_2, and θ_3. The GMM estimator for θ under the joint null H_0^{23} is given by $\hat{\theta}^g = \text{argmax}_\theta \; Q_n(\theta)$ subject to $\theta_2 = \theta_{20}$, $\theta_3 = \theta_{30}$.

The equivalent of the score test for H_0^2 in the GMM framework is given by

$$LM_{2\cdot 1}(\hat{\theta}^g) = n\, q_{2,n}(\hat{\theta}^g)'\, B_{2\cdot 1,n}^{-1}(\hat{\theta}^g)\, q_{2,n}(\hat{\theta}^g),$$

where $B_{2\cdot 1}(\theta) = B_{22}(\theta) - B_{21}(\theta)B_{11}(\theta)^{-1}B_{12}(\theta)$. Under H_0^3 and H_A^2, $\sqrt{n}\, q_{2,n}(\hat{\theta}^g) \xrightarrow{d} N(B_{2\cdot 1}(\theta_0)\delta_2, B_{2\cdot 1}(\theta_0))$, and therefore,

$$LM_{2\cdot 1}(\hat{\theta}^g) \xrightarrow{d} \chi^2_{p_2}(\lambda^g_{2\cdot 1}),$$

with $\lambda^g_{2\cdot 1} = \delta_2' B_{2\cdot 1}(\theta_0)\delta_2$.

As expected, the presence of local misspecification in θ_3 adversely affects $LM_{2\cdot 1}(\hat{\theta}^g)$. The argument follows Saikkonen (1989) closely. Under the regularity conditions in Newey and MacFaden (1994), $G_n(\hat{\theta}^g) \xrightarrow{p} G(\theta_0)$ and $\Omega_n(\hat{\theta}^g) \xrightarrow{p} \Omega(\theta_0)$ (see Newey and MacFadden, 1994, Thm. 3.2), and then, by Slutsky's theorem, $G_n(\hat{\theta}^g)'\Omega_n^{-1}(\hat{\theta}^g)\sqrt{n}\; g_n(\hat{\theta}^g) \xrightarrow{d} G(\theta_0)'\Omega^{-1}(\theta_0)\sqrt{n}\; g_n(\hat{\theta}^g)$. Consider the Taylor expansions of $q_{1,n}(\hat{\theta}^g)$ and $q_{2,n}(\hat{\theta}^g)$ evaluated at $\theta^* = (\theta_{10}', \theta_{20}' + \delta_2'/\sqrt{n}, \theta_{30}' + \delta_3'/\sqrt{n})'$ and note that $G(\theta^*)\Omega^{-1}(\theta^*) = G(\theta_0)\Omega^{-1}(\theta_0) + o_p(1)$. Then,

$$\sqrt{n}\, q_{1,n}(\hat{\theta}^g) = \sqrt{n}\, q_{1,n}(\theta^*) - G_1(\theta_0)'\Omega^{-1}(\theta_0)G_1(\theta_0)\sqrt{n}\;\left(\hat{\theta}_1^g - \theta_{10}\right)$$
$$+ G_1(\theta_0)'\Omega^{-1}(\theta_0)G_2(\theta_0)\,\delta_2 + G_1(\theta_0)'\Omega^{-1}(\theta_0)G_3(\theta_0)\delta_3 + o_p(1),$$

and

$$\sqrt{n}\, q_{2,n}(\hat{\theta}^g) = \sqrt{n}\, q_{2,n}(\theta^*) - G_2(\theta_0)'\Omega^{-1}(\theta_0)G_1(\theta_0)\sqrt{n}\;\left(\hat{\theta}_1^g - \theta_{10}\right)$$
$$+ G_2(\theta_0)'\Omega^{-1}(\theta_0)G_2(\theta_0)\delta_2 + G_2(\theta_0)'\Omega^{-1}(\theta_0)G_3(\theta_0)\,\delta_3 + o_p(1).$$

By the first-order conditions of GMM, $\sqrt{n}\, q_{1,n}(\hat{\theta}^g) = 0$. Rearranging terms and using the definition of B, we have

$$\sqrt{n}\, q_{2,n}(\hat{\theta}^g) = \left(-G_2'\Omega^{-1} + B_{21}B_{11}^{-1}G_1'\Omega^{-1}\right)\sqrt{n}\, g_n(\theta^*)$$
$$+ \left(B_{22} - B_{21}B_{11}^{-1}B_{12}\right)\delta_2 + \left(B_{23} - B_{21}B_{11}^{-1}B_{13}\right)\delta_3 + o_p(1),$$

where the matrices G and Ω are evaluated at θ_0, but the latter is excluded for notational convenience. Finally, note that $G_2'\Omega^{-1}\sqrt{n}g_n(\theta^*) \xrightarrow{p} 0$ and $G_1'\Omega^{-1}\sqrt{n}g_n(\theta^*) \xrightarrow{p} 0$. Thus, $\sqrt{n}\;\, q_{2,n}(\hat{\theta}^g) \xrightarrow{d} N(B_{2\cdot 1}\delta_2 + B_{23\cdot 1}\delta_3, B_{2\cdot 1})$, where

$B_{23\cdot1} = B_{23} - B_{21}B_{11}^{-1}B_{13}$. Therefore, the asymptotic noncentral χ^2 distribution of $LM_{2\cdot1}(\hat{\theta}^g)$ under $H_A^2 : \theta_2 = \theta_{20} + \delta_2/\sqrt{n}$, and $H_A^3 : \theta_3 = \theta_{30} + \delta_3/\sqrt{n}$ is a direct consequence of the nonzero mean of the asymptotic normal distribution. We summarize this result as follows.

THEOREM 1. *Under H_0^2, but when H_A^3 holds, $\sqrt{n}q_{2,n}(\hat{\theta}^g) \xrightarrow{d} N(B_{23\cdot1}(\theta_0)\delta_3, B_{2\cdot1}(\theta_0))$ and $LM_{2\cdot1}(\hat{\theta}^g) \xrightarrow{d} \chi^2_{p_2}(\lambda^g_{2/3\cdot1})$, with $\lambda^g_{2/3\cdot1} = \delta_3' B_{32\cdot1}(\theta_0)B_{2\cdot1}^{-1}(\theta_0) B_{23\cdot1}(\theta_0)\delta_3$, where $B_{23\cdot1}(\theta_0) = B_{32\cdot1}(\theta_0)'$.*

This result can be seen as an extension of Davidson and MacKinnon (1987) and Saikkonen (1989) to the GMM framework, and it has the same implications, as in the ML estimation case, that the test $LM_{2\cdot1}(\hat{\theta}_g)$ will overreject H_0^2 and not provide any information regarding the source(s) of departure from the tested model.

The procedure for constructing a GMM-based locally size-robust test is as follows. Using Theorem 1 under H_0^2 and H_A^3,

$$\sqrt{n}\, q_{2,n}(\hat{\theta}^g) - B_{23\cdot1}(\theta_0)\delta_3 \xrightarrow{d} N(0, B_{2\cdot1}(\theta_0)).$$

The asymptotic distribution of the GMM score corresponding to θ_3 can be shown to be

$$\sqrt{n}\, q_{3,n}(\hat{\theta}^g) \xrightarrow{d} N(B_{3\cdot1}(\theta_0)\delta_3, B_{3\cdot1}(\theta_0)),$$

where $B_{3\cdot1}(\theta) = B_{33}(\theta) - B_{31}(\theta)B_{11}^{-1}(\theta)B_{13}(\theta)$. Therefore,

$$\sqrt{n}\, B_{3\cdot1}(\theta_0)^{-1}q_{3,n}(\hat{\theta}^g) \xrightarrow{d} N\left(\delta_3, B_{3\cdot1}^{-1}(\theta_0)\right).$$

Consequently, we have the asymptotic distribution of the effective GMM score, under H_0^2 and irrespective of H_0^3 or H_A^3, as

$$\sqrt{n}\left[q_{2,n}(\hat{\theta}^g) - B_{23\cdot1,n}(\hat{\theta}^g)B_{3\cdot1,n}^{-1}(\hat{\theta}^g)q_{3,n}(\hat{\theta}^g)\right] \xrightarrow{d} N\left[0, B_{2\cdot1}(\theta_0) - B_{23\cdot1}(\theta_0)B_{3\cdot1}^{-1}(\theta_0)B_{32\cdot1}(\theta_0)\right].$$

Because it has mean zero, an asymptotically robust BY type test $LM_{2\cdot1}^*(\hat{\theta}^g)$ for the GMM framework can be constructed as follows.

THEOREM 2. *Under H_0^2, irrespective of whether H_0^3 or H_A^3 holds, $LM_{2\cdot1}^*(\hat{\theta}_g) = n\, q_{2,n}^*(\hat{\theta}_g)' \left[B_{2\cdot1,n}(\hat{\theta}_g) - B_{23\cdot1,n}(\theta)B_{3\cdot1,n}^{-1}(\hat{\theta}_g)B_{32\cdot1,n}(\hat{\theta}_g)\right]^{-1} q_{2,n}^*(\hat{\theta}_g) \xrightarrow{d} \chi^2_{p_2}(0)$, where $q_{2,n}^*(\theta) \equiv q_{2,n}(\theta) - B_{23\cdot1}(\theta)B_{3\cdot1}^{-1}(\theta)\, q_{3,n}(\theta)$ is the adjusted pseudo-score for θ_2.*

Note that under H_A^2 and H_0^3, $LM_{2\cdot1}^*(\hat{\theta}_g) \xrightarrow{d} \chi^2_{p_2}(\lambda_{2\cdot1}^{g*})$, where $\lambda_{2\cdot1}^{g*} = \delta_2'\big(B_{2\cdot1}(\theta_0) - B_{23\cdot1}(\theta_0)B_{3\cdot1}^{-1}(\theta_0)B_{32\cdot1}(\theta_0)\big)\delta_2$. It follows that $\lambda_{2\cdot1}^{g} - \lambda_{2\cdot1}^{g*} \geq 0$. Hence when there is no local misspecification, the asymptotic power of $LM^*_{2\cdot1}(\hat{\theta}^g)$ is less than (or equal to) that of $LM_{2\cdot1}(\hat{\theta}^g)$. This magnitude can be seen as the cost of insuring against possible local misspecification, that is, the loss of power incurred by robustifying the test unnecessarily.

4. GENERALIZATION TO ESTIMATING AND TESTING FUNCTIONS

The test statistics presented previously can be extended to a more general framework. Let $w(Z,\theta)$ be an r-dimensional vector of functions and let $w_n = n^{-1} \sum_{i=1}^n w(z_i,\theta)$. The vector $w(Z,\theta)$ can be viewed as a general inference function, and it will be used both for estimation and testing under the framework of Newey (1985).

Let Γ_n be a $\gamma \times r$ matrix with $\gamma \geq p_1$ and $\Gamma_n = \Gamma + o_p(1)$. Assume that the following estimating equations for θ_1 hold:

$$\Gamma \mathrm{E}[w(Z,(\theta_1,\theta_{20},\theta_{30}))] = 0 \quad \text{only if } \theta_1 = \theta_{10}.$$

Let Π_n be a $\pi \times r$ matrix and $\Pi_n = \Pi + o_p(1)$. Assume that a specification test can be based on the testing equations

$$\Pi \mathrm{E}[w(Z,(\theta_1,\theta_{20},\theta_{30}))] = 0 \quad \text{only if } \theta_2 = \theta_{20} \quad \text{and} \quad \theta_3 = \theta_{30}.$$

Let $K = \mathrm{E}[\partial w(Z,\theta_0)/\partial\theta_1]$, $V = \mathrm{E}[w(Z,\theta_0)\ w(Z,\theta_0)']$, $D = \mathrm{E}[w(Z,\theta_0)d_{23}(Z,\theta_0)']$, and $P = I - K(\Gamma K)^{-1}\Gamma$. Assume that V and ΓK are nonsingular and that the regularity conditions in Newey (1985) hold. Then under $H_A^{23}: \theta_2 = \theta_{20} + \delta_2/\sqrt{n}, \theta_3 = \theta_{30} + \delta_3/\sqrt{n}$, and $\sqrt{n}\ \Pi_n w_n \xrightarrow{d} N\left(\Pi P D\delta_{23}, \Pi P V P'\Pi'\right)$. Hence

$$n\ w_n' \Pi_n' (\Pi P V P' \Pi')^{-1} \Pi_n w_n \xrightarrow{d} \chi^2_\pi(\lambda_\pi),$$

with $\delta_{23} = [\delta_2', \delta_3']'$ and $\lambda_\pi = (\Pi P D\delta_{23})'(\Pi P V P'\Pi')^{-1}(\Pi P D\delta_{23})$. In terms of estimation, the ML approach is a special case with scores as estimating functions, and the RS tests correspond to the case where the scores are used as test functions (see Bera and Bilias, 2001). GMM-based estimators and tests can also be constructed using the same setup with pseudo-scores in place of scores.

To derive locally size-robust tests for H_0^2 in the presence of local misspecification of θ_3, let $\Pi P D\delta_{23} \equiv \Delta\delta_{23} \equiv \Delta_2\delta_2 + \Delta_3\delta_3$. The BY approach can be restated as finding $\hat{\Delta}_3 = \Delta_3 + o_p(1)$ and $\hat{\delta}_3 = \delta_3 + O_p(1/\sqrt{n})$, such that $\sqrt{n}\ \Pi_n w_n - \hat{\Delta}_3\hat{\delta}_3 \xrightarrow{d} N(\beta(\delta_2), \Sigma)$, where $\beta(\cdot)$ depends on δ_2 but not on δ_3 and Σ denotes the asymptotic variance of $\sqrt{n}\ \Pi_n w_n - \hat{\Delta}_3\hat{\delta}_3$. The following result offers a general device to construct locally size-robust asymptotic tests.

THEOREM 3. *Assume that two different specification test statistics for H_0^{23} are available (say, test statistics A and B), satisfying the assumptions of Newey (1985). For each test, let Δ_2^A, Δ_3^A, Δ_2^B, Δ_3^B be defined as before and let $\hat{\Delta}$ denote* *their consistent estimators. Define* $m_{2(3)\cdot 1,n} = \Pi_{A,n} w_{A,n} - \hat{\Delta}_3^A (\hat{\Delta}_3^B)^M \Pi_{B,n} w_{B,n}$, *where the superscript M denotes the Moore–Penrose generalized inverse of a (not necessarily square) matrix. Then, under* H_A^2 *and when* H_0^3 *or* H_A^3 *holds,* $\sqrt{n}\, m_{2(3)\cdot 1,n} \xrightarrow{d} N\left[\left(\Delta_2^A - \Delta_3^A (\Delta_3^B)^M \Delta_2^B\right)\delta_2, \Sigma_{2(3)\cdot 1}\right]$, *where* $\Sigma_{2(3)\cdot 1}$ *denotes the asymptotic variance of* $\sqrt{n}\, m_{2(3)\cdot 1,n}$. *Moreover,*

$$n\, m'_{2(3)\cdot 1,n}\, \hat{\Sigma}_{2(3)\cdot 1}^{-1}\, m_{2(3)\cdot 1,n} \xrightarrow{d} \chi^2_{m_a}(\lambda_{2(3)\cdot 1}),$$

where $\lambda_{2(3)\cdot 1} = \delta_2' \left(\Delta_2^A - \Delta_3^A (\Delta_3^B)^M \Delta_2^B\right)' \Sigma_{2(3)\cdot 1}^{-1} \left(\Delta_2^A - \Delta_3^A (\Delta_3^B)^M \Delta_2^B\right)\delta_2$ *and* $\hat{\Sigma}_{2(3)\cdot 1}$ *is a consistent estimator of* $\Sigma_{2(3)\cdot 1}$.

Proof. The result follows from the Newey (1985) result and the fact that $m_{2(3)\cdot 1,n}$ is a linear combination of two asymptotically normal statistics. ■

This result provides a very general framework for testing under locally misspecified alternatives, which encompasses the RS and GMM-based test derived previously. For example the robust version of the RS test can be derived in this general framework by considering as "test *A*" the RS test for H_0^2, and $w_{A,n}(\theta) = [d_{1,n}(\theta), d_{2,n}(\theta)]$, $\Pi_A = [0, I]$, $\Gamma_A = [I, 0]$, $\Delta_2^A = J_{2\cdot 1}(\theta_0)$, $\Delta_3^A = J_{23\cdot 1}(\theta_0)$, and as "test B" the RS test for H_0^3, and $w_{B,n}(\theta) = [d_{1,n}(\theta), d_{3,n}(\theta)]$, $\Pi_B = [0, I]$, $\Gamma_B = [I, 0]$, $\Delta_3^B = J_{3\cdot 1}(\theta_0)$, $\Delta_2^B = J_{32\cdot 1}(\theta_0)$. Then, $n\, m'_{2(3)\cdot 1,n}\, \hat{\Sigma}_{2(3)\cdot 1}^{-1} m_{2(3)\cdot 1,n} = RS^*_{2\cdot 1}(\widehat{\theta})$, where $m_{2(3)\cdot 1,n}$ and $\hat{\Sigma}_{2(3)\cdot 1}$ are defined as in Theorem 3 and $RS^*_{2\cdot 1}(\widehat{\theta})$ is the adjusted RS statistic defined in Section 2. The modified version of the RS statistic in the GMM framework can be obtained similarly.

5. CONCLUSIONS

This paper provides a generalization of the Bera–Yoon principle to GMM-based tests and to general estimating and testing functions. The simplicity of this extension and the potentially vast usefulness of the aforementioned principle suggest that further developments would be desirable.

For instance, the idea can be extended to *nonparametric* scores as developed in González-Rivera and Ullah (2001). Another extension would be to consider White type (White, 1982) distributional misspecification into the likelihood-based BY adjusted RS tests, as in Bera, Bilias, and Yoon (2007). Under certain conditions choosing a convenient distributional form, although not necessarily the "true" one, is a valid alternative to the GMM-based testing framework developed here. Finally, it would be interesting to extend the principle to handle "moment function misspecification" as discussed in Hall (2005, Ch. 4).

Although we mentioned some applications of the BY principle, its generalization to the GMM framework opens up many potential applications. For instance, additional features of the econometric model can be incorporated into the Saavedra (2003) testing framework for spatial dependence based on the method of moments. Anselin et al. (1996) used the BY principle to identify the exact source of spatial dependence (through the error term of the lag of the dependent variable) in spatial regression models. Such spatial models are increasingly being estimated by method of moments. It would be interesting to explore the proposed GMM strategy of this paper in this context. Additionally, our approach can be used to develop specification tests in any setup where GMM estimators are preferred to ML estimation, for example, in the context of dynamic panel data models (Arellano and Bond, 1991) and selection models (Heckman, 1976).

REFERENCES

Anselin, L., A. Bera, R. Florax, & M. Yoon (1996) Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics* 26, 77–104.

Arellano, M. & S. Bond (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58, 277–297.

Baltagi, B. & D. Li (2001) LM tests for functional form and spatial error correlation. *International Regional Science Review* 24, 194–225.

Bera, A. & Y. Bilias (2001) On some optimality properties of Fisher-Rao score function in testing and estimation. *Communications in Statistics—Theory and Methods* 30, 1533–1559.

Bera, A., Y. Bilias, & M. Yoon (2007) Adjustments of Rao's Score Test for Distributional and Local Parametric Misspecifications. Working paper, University of Illinois at Urbana-Champaign.

Bera, A., W. Sosa-Escudero, & M. Yoon (2001) Tests for the error component model in the presence of local misspecification. *Journal of Econometrics* 101, 1–23.

Bera, A. & M. Yoon (1993) Specification testing with locally misspecified alternatives. *Econometric Theory* 9, 649–658.

Davidson, J. & J. MacKinnon (1987) Implicit alternatives and the local power of test statistics. *Econometrica* 55, 1305–1329.

Godfrey, L. & M.R. Veall (2000) Alternative approaches to testing by variable addition. *Econometric Reviews* 19, 241–261.

González-Rivera, G. & A. Ullah (2001) Rao's score test with nonparametric density estimators. *Journal of Statistical Planning and Inference* 97, 85–100.

Hall, A. (2005) *Generalized Method of Moments*. Oxford University Press.

Hansen, L.P. (1982) Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029–1054.

Hansen, L.P., J. Heaton, & A. Yaron (1996) Finite-sample properties of some alternative GMM estimators. *Journal of Business & Economic Statistics* 14, 262–280.

Heckman, J.J. (1976) The common structure of statistical models of truncation, sample selection and limited dependent variables and simple estimator for such models. *Annals of Economic and Social Measurement* 5, 475–492.

Newey, W. (1985) Maximum likelihood specification testing and conditional moment tests. *Econometrica* 53, 1047–1070.

Newey, W. & D. McFadden (1994) Large sample estimation and hypothesis testing. In R. Engle & D. McFadden (eds.), *Handbook of Econometrics*, vol. 4, pp. 2111–2245. North-Holland.

Newey, W. & K. West (1987) Hypothesis testing with efficient method of moments estimation. *International Economic Review* 28, 777–787.

Neyman, J. (1959) Optimal asymptotic tests of composite statistical hypothesis. In U. Grenander (ed.), *Probability and Statistics*, pp. 213–234. Wiley.

Saavedra, L. (2003) Tests for spatial lag dependence based on method of moments estimation. *Regional Science and Urban Economics* 33, 27–58.

Saikkonen, P. (1989) Asymptotic relative efficiency of the classical tests under misspecification. *Journal of Econometrics* 42, 351–369.

White, H. (1982) Maximum likelihood estimation of misspecified models. *Econometrica* 50, 1–25.