Hypothesis Testing via Multivariate Kernel Density with an Application to the Information Matrix Test

Muhammad Akram, **Xibin Zhang**, and **Maxwell L. King**[1]

Department of Econometrics and Business Statistics
Monash University, Australia

Abstract: The information matrix (IM) test, introduced by White (1982), is well known as a general test for misspecification of a parametric likelihood. This test is based on the fact that if the model is correct then the information matrix is equal to the covariance of the score vector. However, the use of information matrix test in applied econometrics is limited because the actual size of the test derived according to asymptotic critical values often differs considerably from its nominal size. This paper investigates the application of the multiple testing procedure, proposed by King et al. (2008), to the information matrix test, where the test is conducted through the bootstrapping procedure due to Horowitz (1994). This procedure is used to obtain the critical values, which are used for computing p-value of the test. To examine the size- and power-performance of the information matrix test based on this approach, we employed the Tobit model to simulate data. The proposed testing procedure is compared with the Lancaster (1984) version of information matrix test. We find that both tests have approximately correct sizes, while the power of our test is higher than that of Lancaster's version of information matrix test.

Key words: information matrix, bootstrapping, Hessian, multivariate kernel density, score vector, Tobit model.

1 Introduction

It is often important to test whether a model is correctly specified. When the model is correctly specified and estimated by maximizing the likelihood, the information matrix should be asymptotically equal to the negative Hessian matrix. The information matrix test, introduced by

[1]Corresponding Author. Email: max.king@adm.monash.edu.au; telephone: 61-3-99052449; fax: 61-3-99058039; address: Monash University, Victoria 3800, Australia.

White (1982), aims to test the significance of the discrepancy between the negative Hessian and the outer product of the score vector, where the lower triangular components of matrix of such differences are organized into a vector called the test vector. Chesher (1984) demonstrated that it can be viewed as a Lagrange multiplier (LM) test for specification error against the alternative of parameter heterogeneity. As a by product of this analysis, Chesher (1983) and Lancaster (1984) provided an nR^2 version of the IM test, where n is the sample size and R^2 is obtained through the ordinary least squares regression of a column of ones on a matrix whose elements are functions of 1st and 2nd derivatives of the log-density function. For the normal fixed regressor linear model, Hall (1987) showed that the test decomposed asymptotically into the sum of three components, while one is White's general test for heteroscedasticity and the other two test some forms of normality. An important finding of Hall (1987) is that the components of the IM test are insensitive to serial correlation.

However, the use of information matrix test in applied econometrics is limited because the actual size of the test obtained according to asymptotic critical values often differs greatly from its nominal size, as evidence by the Monte Carlo experiments reported in Taylor (1987), Orme (1990), Chesher and Spady (1991), Davidson and MacKinnon (1992). Davidson and MacKinnon (1992) have proposed dealing with this problem by using double-length artificial regressions to compute a variant of the IM test statistic, but the models for discrete, censored, or truncated data cannot be dealt with their method. Chesher and Spady (1991) have proposed obtaining the critical value for the IM test from the Edgeworth expansion through order $O(n^{-1})$ of the finite-sample distribution of the test statistic, where n is the sample size. They provided Monte Carlo evidence indicating that such critical values provide a good approximation to the exact finite-sample distribution of the IM test statistics and found to be superior to the usual χ^2 approximation in some special cases. In the examples considered by Chesher and Spady (1991), the Edgeworth expansions are independent of the parameters of the models being tested, but this is not the case in general. Horowitz (1994) proposed a bootstrapping procedure to obtain critical values for the information matrix test and illustrated the ability of bootstrapping to overcome the incorrect-size problem in finite samples. The main purpose of Horowitz paper is to point out that in many circumstances that are important in application, good finite-sample critical values for the IM test can be obtained easily through the use of Monte Carlo simulations or the bootstrap. It is not necessary to derive Edgeworth expansions or to carry out other algebraically complicated manipulations. The another purpose of Horowitz's paper is to provide some insights into the power of the IM test and to encourage further investigation of power.

All existing versions of the information matrix test rely on the estimate of the asymptotic covariance matrix of the test vector. The original formula for the asymptotic covariance of the test vector is analytically complicated and involves the third derivative of the log-likelihood function. Lancaster (1984) pointed out how the covariance matrix of the White's information matrix test can be estimated without computing the third derivative of the log-likelihood function. Dhaene and Hoorelbeke (2004) indicated that the incorrect-size problem stems from the inaccurate variance-covariance matrix of the test vector, and proposed the covariance matrix of the test vector be estimated using a parametric bootstrap. Nonetheless, with the multivariate kernel density estimation technique of Zhang et al. (2006), we are able to estimate the joint density of the test vector through a bootstrapping procedure other than to construct a scalar statistic in the way that the current information matrix test does. According to the estimated joint density of the test vector, we are able to calculate the p-value for the test.

In this paper, we propose an information matrix test using a multiple testing procedure introduced by King et al. (2008), which involves estimating the density of multiple statistics such as the test vector in the information matrix test under the null hypothesis, and the bootstrapping procedure of Horowitz (1994) is used in our test. The variance-covariance matrix of the test vector is computed according to Lancaster (1984). The proposed testing procedure is then compared with the Lancaster (1984) version of IM test. The reason of selecting the Lancaster's version for comparison is as follows. First, the estimation of variance-covariance matrix of the vector of indicators is analytically simple and does not require third derivatives of the log likelihood function. Secondly, it has similar or better size- and power-performances then the other IM tests considered by Horowitz (1994).

The rest of the paper is organized as follows. Section 2 briefly describes the information matrix test. In Section 3, we provide simulation methodology, where we examine the performance of the proposed testing procedure and the Lancaster's version of the information matrix test in terms of size and power. We conclude the results of the paper in Section 4.

2 Information Matrix Test

Let $f(y|\theta)$ denote the density for a postulated model and θ an $p \times 1$ vector of parameters. Let $\mathbf{y} = (y_1, y_2, \cdots, y_n)'$ be the vector of observations, and $\ell(y|\theta) = \log f(y|\theta)$ the logarithmic density. We need the following notations.

$$A(\theta) = E\left[\frac{\partial^2 \ell(y|\theta)}{\partial\theta\partial\theta'}\right], \qquad A_n(\mathbf{y}, \theta) = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2 \ell(y_i|\theta)}{\partial\theta\partial\theta'},$$

$$B(\theta) = E\left[\frac{\partial \ell(y|\theta)}{\partial\theta}\frac{\partial \ell(y|\theta)}{\partial\theta'}\right], \qquad B_n(\mathbf{y}, \theta) = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial \ell(y_i|\theta)}{\partial\theta}\frac{\partial \ell(y_i|\theta)}{\partial\theta'},$$

where expectations are taken with respect to the true density. When the model is correctly specified, the true density is $f(y|\theta_0)$, where θ_0 is the true value of θ.

The information matrix procedure is based on the information-matrix equality, which states that $A(\theta_0) + B(\theta_0) = 0$ when the model is correctly specified. Given the vector of n independent observations, $\mathbf{y}$, the information-matrix test investigates the statistical significance of $A_n(\mathbf{y}, \hat{\theta}) + B_n(\mathbf{y}, \hat{\theta})$, where $\hat{\theta}$ is the maximum likelihood estimator of θ.

Let

$$d_{ij} = \frac{1}{n}\sum_{t=1}^{n}\left[\frac{\partial \ell(y_t|\theta)}{\partial\theta_i}\frac{\partial \ell(y_t|\theta)}{\partial\theta_j} + \frac{\partial^2 \ell(y_t|\theta)}{\partial\theta_i\partial\theta_j}\right], \tag{1}$$

which is evaluated at $\theta = \hat{\theta}$. Let D denote the test vector whose elements are d_{ij}, for $i = 1, \cdots, p$, $j = 1, \cdots, i$, and $\hat{V}$ the consistent estimator of the covariance matrix of D. White (1982) shows that under regularity conditions, the information-matrix test is

$$\xi_n = nD'\hat{V}^{-1}D. \tag{2}$$

Under the null hypothesis, ξ_n is distributed asymptotically χ^2_q with $q = p(p+1)/2$. Since $A_n(\mathbf{y}, \hat{\theta}) + B_n(\mathbf{y}, \hat{\theta})$ is a symmetric matrix, a test of the complete IM identity can be based on the lower triangular elements of $A_n(\mathbf{y}, \hat{\theta}) + B_n(\mathbf{y}, \hat{\theta})$. However, according to White (1982), in many situation it is inappropriate to base the test on all q indicators because some indicators may be identically zero, furthermore, some indicators may be linear combinations of other. In either case, it is appropriate to ignore such indicators.

The available Monte Carlo evidence shows that the finite-sample distribution of the IM statistic is poorly approximated by the χ^2_q distribution (see, for example, Dhaene and Hoorelbeke, 2004). Under certain regularity conditions, a central limit theorem can be applied to show that as n tends to infinity, $\sqrt{n}D$ converges to a multivariate normal distribution with mean zero if the model is correctly specified (see White, 1984, p113).

3 Simulation method

This section reports the results of a Monte Carlo investigation of the finite-sample size and power of the Information matrix test. The experiments use two forms of the IM test statistic: the proposed form of the information matrix test (introduced by King et al., 2008) denoted by IM_P, and the Lancaster form of IM test denoted by IM_L.

Bootstrapping procedure of Horowitz (1994) has been employed to compare the proposed form of the information matrix test with the Lancaster's form of information matrix test. This procedure is used to obtain the critical values of the test, where the test statistic is not pivotal. In all the experiments considered by Chesher and Spady (1991), the IM statistic is pivotal. That is, under the null hypothesis of correct specification, the finite-sample distribution of the test statistic is independent of the parameters of the model being tested. In the cases investigated by Horowitz (1994), the use of bootstrap-based critical values makes the empirical sizes of the IM test very close to its nominal sizes, whereas the empirical and nominal sizes can differ enormously when asymptotic critical values are used.

3.1 Tobit model

To evaluate the finite-sample size and power performance of the information matrix test based on these approaches, we employed the Tobit model to simulate data. The Tobit model is

$$y_i = \begin{cases} x_i'\beta + u_i & \text{if RHS } > 0 \\ 0 & \text{if RHS } \leq 0 \end{cases}, \tag{3}$$

where $u_i \sim N(0, \sigma^2)$. It will be convenient to re-parameterize the model as

$$hy_i = \begin{cases} x_i'b + v_i & \text{if RHS } > 0 \\ 0 & \text{if RHS } \leq 0 \end{cases}, \tag{4}$$

where $h = 1/\sigma$, $b = \beta/\sigma$, and $v_i \sim N(0, 1)$.

Given $y_i > 0$ and x_i, the conditional cumulative density of y_i is

$$\begin{aligned} F(y|y_i > 0, x_i, b, h) &= P\{y_i \leq y | y_i > 0, x_i\} \\ &= \frac{\Phi(hy_i - x_i'b) - \Phi(-x_i'b)}{\Phi(x_i'b)}, \end{aligned}$$

where $\Phi(\cdot)$ is the cumulative density function of the standard normal distribution. Given $y_i > 0$ and x_i, the conditional density of y_i is

$$f(y|y_i > 0) = \frac{\frac{h}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(hy_i - x_i'b)^2\right\}}{\Phi(x_i'b)}.$$

Let δ_i be an indicator variable, such that,

$$\delta_i = \begin{cases} 1 & \text{if } y_i > 0 \\ 0 & \text{if } y_i = 0 \end{cases}$$

then we have

$$P\{\delta_i = 1\} = P\{y_i > 0\} = \Phi(x_i'b),$$

$$P\{\delta_i = 0\} = 1 - \Phi(x_i'b).$$

Let $\theta = (b', h)'$ denote the parameter vector. The likelihood function for model (4) is

$$L(y_1, y_2, \cdots, y_n|\theta) = \prod_{i=1}^{n} [1 - \Phi(x_i'b)]^{1-\delta_i} \times [\Phi(x_i'b)]^{\delta_i} \times [f(y_i|y_i > 0)]^{\delta_i}.$$

When $y_i = 0$, only the first term, which is $P\{y_i = 0\}$, is kept, and when $y_i > 0$ the product of the second and third terms, which are $P\{y_i > 0\}f(y_i|y_i > 0)$, is kept. The logarithmic likelihood function is

$$\ell(y_1, y_2, \cdots, y_n|\theta) = \sum_{i=1}^{n}(1-\delta_i)\log\{1-\Phi(x_i'b)\} + \sum_{i=1}^{n}\delta_i \log h - \frac{1}{2}\sum_{i=1}^{n}\delta_i(hy_i - x_i'b)^2.$$

The first-order derivatives of $\ell(.)$ with respect to θ are, respectively,

$$\frac{\partial \ell(.)}{\partial b} = -\sum_{i=1}^{n}(1-\delta_i)\frac{\phi(x_i'b)}{1-\Phi(x_i'b)}x_i + \sum_{i=1}^{n}\delta_i(hy_i - x_i'b)x_i,$$

$$\frac{\partial \ell(.)}{\partial h} = \sum_{i=1}^{n}\delta_i/h - \sum_{i=1}^{n}\delta_i(hy_i - x_i'b)y_i,$$

where $\phi(.)$ is the density function of standard normal distribution. The second-order derivatives of $\ell(.)$ with respect to θ are, respectively,

$$\frac{\partial^2 \ell(.)}{\partial b \partial b'} = \sum_{i=1}^{n}(1-\delta_i)\frac{\phi(x_i'b)}{1-\Phi(x_i'b)}\left[x_i'b - \frac{\phi(x_i'b)}{1-\Phi(x_i'b)}\right]x_i x_i' - \sum_{i=1}^{n}\delta_i x_i x_i',$$

$$\frac{\partial^2 \ell(.)}{\partial b \partial h} = \sum_{i=1}^{n}\delta_i y_i x_i,$$

$$\frac{\partial^2 \ell(.)}{\partial h \partial h} = -\sum_{i=1}^{n}\delta_i/h^2 - \sum_{i=1}^{n}\delta_i y_i^2.$$

3.2 Design of experiment

The experiments consist of applying both forms of IM tests to tobit model, given in (4). In model (4), x_i is a vector of explanatory variables, b is a vector of parameters and $v_i \sim N(0,1)$. x_i consists of an intercept component and either one or two additional variables. The values of x_i are fixed in repeated samples. The values of b are (0.75,1) when x_i consists of an intercept and one regressor, and (0.75,1,1) when x_i consists of an intercept and two regressors. The non-intercept components of x_i are sampled independently either from the standard normal distribution or from the uniform distribution on (-1,1). The values of σ^2 is 1 in all of the experiments. The sample sizes are 50, 100, 200 and 300.

The null hypothesis underlying the class of IM tests is $H_0 : A(\theta_0) + B(\theta_0) = 0$, i.e., the model is correctly specified or, in other words, the IM test is a score test of model specification against the alternative of local random parameter heterogeneity. In order to test this hypothesis the proposed form of IM test, IM_P, aims to estimate the joint density of the vector of indicators (D) based on which we can derive the critical values, and therefore the null hypothesis can be tested. We use Monte Carlo simulation to investigate the size and power performance of both forms of the IM tests, where the size-corrected critical values were used for computing the sizes and powers of the tests. When examining the power of the two tests, the data were generated by the following two models, following Horowitz (1994).

$$y_i = max(0, x_i'\beta + u_i), \qquad u_i \sim N(0, exp(0.5\ x_i'\beta)) \tag{5}$$

and

$$y_i = max(0, x_i'\beta + 0.75x_2x_3 + u_i), \qquad u_i \sim N(0, 1) \tag{6}$$

where x_2 and x_3 are the two non-intercept components of x_i. In the experiment based on model (5) the null hypothesis is false because it ignores heteroscedasticity, and the experiment based on model (6) the null hypothesis uses an incorrect mean function.

3.3 Bootstrapping procedure

In this section, finite-sample sizes and powers of the IM_L and the IM_P tests are illustrated with bootstrap-based critical values. The bootstrap-based method, due to Horowitz (1994), can be used to obtain critical values that are more accurate than asymptotic ones. The experiments consist of applying these versions of IM tests to tobit model. In all experiments, the results based on 2000 Monte Carlo replications with 200 bootstrap samples in each replication. According to Horowitz (1994), the bootstrap samples beyond 100 had little effect on the results of the experiment. The bootstrap is very accurate with sample sizes as small as 50 to 100 in the cases investigated by Horowitz (1994). It follows from Hall (1986) that the error in the size of a test using bootstrap based critical values is $O[n^{-(j+1)/2}]$ regardless of the number of bootstrap samples used to estimate α level critical value.

3.3.1 Finite-sample sizes and powers

For the finite-sample sizes, the experiment consisted of following steps.

1) Generate an estimation data set of size n by random sampling from model (4). x_i is fixed under repeated samples. Estimate parameters (θ) by ML method and compute the IM_L test statistic using the vector of non-zero indicators, D. Denote this value by $\hat{f}_1(IM_L)$.

2) Generate a bootstrap sample of size n by random sampling from model (4) but using the parameter values estimated in step 1 instead of the true values. Using this sample, re-estimate parameters (θ) by ML method and compute the IM_L test statistic using the vector of non-zero indicators. Repeat this step for B=200 times, to estimate the empirical distribution function of the IM_L test statistic. Estimate the α-level critical value of IM_L test statistic from its empirical distribution. Let $f_\alpha(IM_L)$ denote the estimated α-level critical value.

3) The proposed testing procedure aims to estimate the joint density of the vector of indicators, D, through the multivariate kernel density (MKD) method. Using the "B" bootstrap vectors of indicators obtained in step 2, compute the proposed form of IM test statistic (IM_P) for each vector of non-zero indicators, through the MKD method. Consequently, we get the empirical distribution function of the IM_P test statistic. Estimate the α-level critical values from the empirical distribution of IM_P. Let $f_\alpha(IM_P)$ denote the estimated α-level bootstrap-based critical value of the IM_P test.

4) Using the density function of indicator vectors, obtained in step 3, compute the IM_P test statistic of the vector of non-zero indicators computed in step 1. Denote this value by $\hat{f}_1(IM_P)$.

5) Reject the model being tested at the nominal α level based on the bootstrap critical value if $\hat{f}_1(IM_L) > f_\alpha(IM_L)$ for the Lancaster form of IM test, and if $\hat{f}_1(IM_P) < f_\alpha(IM_P)$ for the proposed form of IM test.

6) Repeat steps 1-5 for m=2000 times, and compute the relative frequencies that $\hat{f}_j(IM_L) >$

$f_\alpha(IM_L)$ holds for the IM_L test, and $\hat{f}_j(IM_P) < f_\alpha(IM_P)$ holds for the IM_P test, for $j = 1, 2, \cdots, m$, and $\alpha = 0.01,\ 0.05,\ 0.10$. The p-value is approximated by these relative frequencies.

For the finite-sample powers of the IM_L and the IM_P tests, the values of x_i and parameters are the same as in the case of finite-sample sizes. In each experiment, the power of the IM tests were computed using bootstrap-based critical values. The experiment were carried out using the above 6-steps procedure with the modification that in step 1 the data were sampled from the true data generating process (one of models 5 or 6) not the model being tested.

The results, based on bootstrap procedure, are presented in Tables 1 and 2 for sizes and powers, respectively. From Table 1, we found that the sizes derived through the IM_P test are very close to the corresponding nominal sizes for both one-regressor and two-regressors models, and whether the design matrix x_i is generated through standard normal distribution or uniform distribution. Whereas, the sizes obtained through IM_L test have mixed behavior. At 1% level, the sizes seems to be over rejecting the null hypothesis, while at 5% and 10%, levels, the sizes are close to their nominal sizes. This behavior is consistent for both design matrices as well as both one-regressor and two-regressor models. In term of accuracy of the estimated sizes, the IM_P test perform better than the IM_L test, specially at 1% level.

Table 2 presents the probability values of rejecting a false null hypothesis, where model (5) is used as a true alternative hypothesis. We found that the IM_P test has larger probability vales then the IM_L test in rejecting the false null hypothesis for both one-regressor and two-regressor models. This result is consistent whether the design matrix x_i is generated through standard normal distribution or uniform distribution. Moreover, the difference between the estimated powers of IM_P and IM_L tests decrease by the increase of sample size, specially, in the case of two-regressor model. Thus, the simulation study shows that the proposed form of IM test, IM_P, produce correct sizes and has higher power probabilities than the Lancaster form of IM test, IM_L.

If we use model (6) as a true alternative, the behavior of probability values of rejecting a false null hypothesis is similar to those where model (5) is used under the true alternative hypothesis.

We haven't provided these results in this paper.

4 Conclusion

This paper presents an information matrix test using a new multiple testing procedure proposed by King et al. (2008), which involves estimating the density of the multiple test statistics, such as the score vector in the information matrix test, through bootstrapping rather than constructing a critical region through a scalar test statistic. According to the estimated density, the p-value has been derived using Monte Carlo procedure. This testing procedure is conducted through the bootstrapping procedure of Horowitz (1994).

The proposed form of information matrix test (IM_P) is then compared with the well-known Lancaster's version of information matrix test (IM_L).

The simulation studies have shown that the sizes derived through IM_P test are very close to the corresponding nominal sizes, whereas, the sizes obtained through IM_L test have mixed behavior. At 1% level, the sizes seems to be over rejecting the null hypothesis, while at 5% and 10%, levels, the sizes are close to their nominal sizes. The power of IM_P test is better than the IM_L test in almost all the cases. Moreover, the probability values obtained through one-regressor model are smaller than those derived through the two-regressors model for almost all sample sizes and all nominal sizes, except for sample size 50.

Acknowledgements

We acknowledge computational support from the Victorian Partnership for Advanced Computing (VPAC) and financial support from the Australian Research Council under the discovery project DP0664926.

Table 1: *Estimated sizes of the IM test through bootstrapping.*

IM test	Sample size	One regressor			Two regressor		
		0.01	0.05	0.10	0.01	0.05	0.10
$x_i \sim N(0,1)$							
	50	0.011	0.055	0.109	0.009	0.049	0.104
IM_P	100	0.009	0.048	0.097	0.009	0.055	0.112
	200	0.011	0.057	0.106	0.012	0.046	0.098
	300	0.013	0.047	0.087	0.011	0.054	0.107
	50	0.018	0.058	0.105	0.012	0.047	0.089
IM_L	100	0.016	0.056	0.108	0.015	0.051	0.092
	200	0.016	0.052	0.100	0.013	0.049	0.098
	300	0.014	0.049	0.100	0.017	0.065	0.124
$x_i \sim U(-1,1)$							
	50	0.011	0.048	0.092	0.011	0.048	0.100
IM_P	100	0.007	0.049	0.097	0.011	0.054	0.103
	200	0.007	0.050	0.097	0.014	0.051	0.106
	300	0.011	0.054	0.098	0.011	0.046	0.096
	50	0.017	0.058	0.109	0.015	0.051	0.099
IM_L	100	0.018	0.051	0.097	0.018	0.059	0.109
	200	0.013	0.051	0.098	0.020	0.064	0.114
	300	0.015	0.054	0.100	0.020	0.059	0.100

Table 2: *Estimated powers of the IM test through bootstrapping with H_1 specified by model 5.*

IM test	Sample size	One regressor 0.01	0.05	0.10	Two regressor 0.01	0.05	0.10
$x_i \sim N(0,1)$							
	50	0.171	0.400	0.532	0.135	0.320	0.455
IM_P	100	0.268	0.539	0.682	0.356	0.610	0.736
	200	0.531	0.796	0.886	0.743	0.917	0.953
	300	0.744	0.930	0.965	0.924	0.985	0.992
	50	0.035	0.112	0.183	0.039	0.093	0.173
IM_L	100	0.056	0.183	0.285	0.067	0.186	0.314
	200	0.152	0.389	0.557	0.318	0.589	0.720
	300	0.329	0.631	0.780	0.661	0.872	0.932
$x_i \sim U(-1,1)$							
	50	0.047	0.168	0.281	0.077	0.210	0.315
IM_P	100	0.091	0.247	0.363	0.128	0.311	0.438
	200	0.187	0.440	0.581	0.268	0.522	0.648
	300	0.274	0.562	0.692	0.424	0.697	0.790
	50	0.024	0.074	0.132	0.024	0.060	0.118
IM_L	100	0.030	0.088	0.158	0.036	0.114	0.195
	200	0.068	0.175	0.293	0.084	0.214	0.330
	300	0.100	0.251	0.403	0.172	0.371	0.536

References

Chesher, A. (1983) The information matrix test: simplified calculation via a score test interpretation, *Economic letters*, **13**, 45–48.

Chesher, A. (1984) Testing for neglected heterogeneity, *Econometrica*, **52**, 865–872.

Chesher, A. and R. Spady (1991) Asymptotic expansions of the information matrix test statistic, *Econometrica*, **59**, 787–815.

Davidson, R. and J. G. MacKinnon (1992) A new form of the information matrix test, *Econometrica*, **60**, 1455–157.

Dhaene, G. and D. Hoorelbeke (2004) The information matrix test with bootstrap-based covariance matrix estimation, *Economics Letters*, **82**, 341–347.

Hall, A. (1986) On the number of bootstrap simulations required to construct a confidence interval, *Annals of Statistics*, **14**, 1453–1462.

Hall, A. (1987) The information matrix test for the linear model, *Review of Economic Studies*, **54**, 257–263.

Horowitz, J. L. (1994) Bootstrap-based critical values for the information matrix test, *Journal of Econometrics*, **61**, 395–411.

King, M. L., X. Zhang and M. Akram (2008) A new procedure for multiple testing of econometric models, working paper.

Lancaster, T. (1984) The covariance matrix of the information matrix test, *Econometrica*, **52**, 1051–1053.

Orme, C. (1990) The small-sample performance of the information matrix test, *Journal of Econometrics*, **46**, 309–331.

Taylor, L. W. (1987) The size bias of white's information matrix test, *Economics Letters*, **24**, 63–68.

White, H. (1982) Maximum lilkelihood estimation of misspecified models, *Econometrica*, **50**, 1–26.

White, H. (1984) *Asymptotic theory for econometricians*, Academic press: NY.

Zhang, X., M. L. King and R. J. Hyndman (2006) Abayesian approach to bandwidth selection for multivariate kernel density estimation, *Computational Statistics and Data Analysis*, **50**, 3009–3031.