Vidyodaya J. Soc., Sci Vol. 1 No. 2—January 1987 pp. 1—17

THE SMALL SAMPLE BEHAVIOUR OF SOME NON-NESTED TESTS

By

W. A. JAYATISSA

Department of Economics, University of Sri Jayawardenapura, Sri Lanka

1. Introduction

Empirical researchers are often confronted with the problem of making choise among alternative econometric models, of the process they are investigating. What variables should be included in a regression model? What functional form should the regression equation take? Is a log-linear functional form more appropriate than a linear regression model? Are the errors normally distributed or best characterised by some other distribution? Some of the issues, the econometrician has to answer, would normally be taken for granted, not because they are optimal from the point of view of economic theory but because they are extremely convenient for estimation and hypotheses testing purposes; however, problems of comparing alternative or rival theories also arise.

There are situations when comparisons among regression models with different variables and functional form are of interest. These comparisons arise in economics when competing economic theories give rise to different sets of relevant variables and the probability distributions (possibly in the form of conditional linear regression models), chosen to characterise these economic relatonships are likewise distinct. In these situations the data must be used to discriminate among competing hypotheses. Examples are the comparison of simple Quantity Theory models with simple Keynesian Theory models by Friedman and Meiselman (1963) and the comparison of alternative investment theories by Jorgenson and Siebert (1968).

In many economic applications the models that we eventually encounter are often non-nested in the sense that they have separate parametric families and one model cannot be obtained from the others as a limiting process. Most techniques for hypotheses testing in econometrics however, simply allow one to test restrictions on a model more general than one being valid; for such cases the usual likelihood ratio criterion or other appropriate test procedures can be utilised. Unfortunately, in the situation of non-nested hypotheses the application of such procedures cannot be directly employed and other suitable methods of testing have to be sought.

This paper was written while the author was visiting the University of Manchester under the Commonwealth Fellowship Programme. The support of the ACU is gratefully acknowledged.

In this respect the procedure suggested by Cox (1961) and Atkinson (1970) using the modified likelihood ratio is illuminating. In subsequent work Pesaran (1974) and Pesaran and Deaton (1978) developed tests using the results established by Cox to descriminate between linear as well as non-linear non-nested regression models.

Davidson and MackKinnon (1981) suggested a simple testing procedure to non-nested hypotheses by employing a different approach called the principle of artificial nesting. It is also worthy to mention the work of Fisher and McAleer (1981) and Godfrey (1983) on extnsions of the procedure suggested by Davidson and MackKinnon.

The aim of this paper is (a) to outline some testing procedures upon which tests of non-nested hypotheses may be based as well as their associated tests of significance, (*b*) to introduce some extensions to those tests and to explain the relationships among them and finally (*c*) to examine the small sample behaviour of these tests using Monte Carlo experiments.

2. Non-nested regression models

Consider the following two linear regression models:

$$H_0: \quad Y = Xb_0 + u_0, \quad u_0 \sim N(0, \sigma_0^2 I) \qquad (1)$$

$$H_1: \quad Y = Zb_1 + u_1, \quad u_1 \sim N(0, \sigma_1^2 I) \qquad (2)$$

where Y is the nxl vector of observations on the dependent variable, X and Z are nxk_0 and nxk_1 observation matrices for the regressors of H_0 and H_1, b_0 and b_1 are the k_0xl and k_1xl parmeter vectors, and u_o and u_1 are nxl disturbance vectors.

The ordinary least squares estimators of b_0 and b_1 will be deonoted by $\hat{b}_0$ and $\hat{b}_1$, and it will also be useful to introduce the OLS projection matrices.

$$M_0 = I - P_0 \quad \text{and} \quad M_1 = I - P_1$$

where

$$P_0 = X(X'X)^{-1}X' \text{ and } P_1 = Z(Z'Z)^{-1}Z'.$$

For purposes of discussing the tests, it is customary to assume that the elements of X and Z are uniformly bounded with

$$\lim \left[\frac{X'X}{n}\right] = \Sigma_{00} \quad \text{and} \quad \lim \left[\frac{Z'Z}{n}\right] = \Sigma_{11}$$

being finite non-singular matrices. These assumptions together with normality of the disturbances form the basis of the classical theory of a single equation, but need to be supplemented for the purposes of developing tests of (1) and (2) regarded as non-nested models. The conventional approach is to specify that the matrices.

$$\Sigma_{01} = \lim \left[\frac{X'Z}{n}\right] , \quad \Sigma_{10} = \Sigma'_{01},$$

$$\Sigma_{0} = \lim \left[\frac{X'M_1X}{n}\right] = \Sigma_{00} - \Sigma_{01}\ \Sigma_{11}^{-1}\ \Sigma_{10},$$

$$\Sigma_{1} = \lim \left[\frac{Z'M_0Z}{n}\right] = \Sigma_{11} - \Sigma_{10}\ \Sigma_{00}^{-1}\ \Sigma_{01},$$

exist as finite and non-null matrices.

3. Tests based on the centered likelihood ratio

Denoting $e_0' = (b_0', \sigma_0^2)$ and $e_1' = (b_1', \sigma_1^2)$, the maximised log–likelihood function under H_i is given by

$$\hat{l}_i = -\tfrac{n}{2}\log 2\pi - \tfrac{n}{2}\log \sigma_i^2 - \tfrac{n}{2} \quad (i = 0, 1)$$

where $\sigma_i^2 = n^{-1}Y'(I\text{-}P_i)\,Y$ is the maximum likelihood estimate of $\sigma_i^2 (i=0,1)$. Under H_o, $\hat{b}_0 = (X'X)^{-1}X^1Y$ is the maximum likelihood estimate of b_0 while $\hat{b}_1 = (Z^1Z)^{-1}Z^1Y$ is the maximum likelihood estimate of b_1 under H_1.

If H_0 were to be nested within H_1, the asymptotic expectation of $\hat{l}_0 - \hat{l}_1$, evaluated at the (restricted) maximum likelihood estimate $\theta_0 = \hat{\theta}_0$ would be zero. Since the two models are non-nested Cox suggested that, to test H_0, the mean of $\hat{l}_0 - \hat{l}_1$ evaluated under H_0 (i.e. $\theta_0 = \hat{\theta}_0$) should be subtracted from $\hat{l}_0 - \hat{l}_1$. Thus, the Cox test of H_0 against the alternative H_1 is based upon the statistic.

$$T_0 = (\hat{l}_0 - \hat{l}_1) - n[\text{plim}_0\, n^{-1}(\hat{l}_0 - \hat{l}_1)] \quad \theta_0 = \hat{\theta}_0 \qquad (3)$$

where plim_0 denotes the probability limit under H_0. Cox shows that under general conditions, given that H_0 is the true model, T_0 will be asymptotically normally distributed with mean zero and variance V_0. If a consistent estimate of V_0 is given by $\hat{V}$, the statistic $N_0 = T_0/(\hat{V}_0)^{\frac{1}{2}}$ will then be approximately distributed as a standard normal variate under H_0.

The numerator and the denominator of the Cox test for non-nested linear regression models are explicitly given by Pesaran (1974) as

$$T_0 = \tfrac{n}{2} \log (\hat{\sigma}_1^2/\hat{\sigma}_{10}^2) \tag{4}$$

and

$$\hat{V}_0 = (\hat{\sigma}_0^2/\hat{\sigma}_{10}^4)\ \hat{b}_0{}^1 X^1\ M_1 M_0 M_1 X\hat{b}_0 = (\hat{\sigma}_0^2/\hat{\sigma}_{10}^4)\ [Y' P_0 P_1\ (I-P_0)\ P_1 P_0 Y] \tag{5}$$

in which

$$\hat{\sigma}_{10}^2 = \hat{\sigma}_0^2 + n^{-1}\ \hat{b}_0^1\ X^1 M_1\ X\hat{b}_0 = \hat{\sigma}_0^2 + n^{-1}\ Y^1 P_0\ (I-P_1)\ P_0 Y \tag{6}$$

Unfortunately, the logarithmic form of expression (4) rstricts further simplication of this statistic. However, by examining the upper-bound linearisation of T_0 similar to those discussed by Cox (1961) and Fisher and McAleer (1981) it would be possible to simplify (4) as

$$TL_0 = \tfrac{n}{2}\ (\hat{\sigma}_1^2 - \hat{\sigma}_{10}^2)/\hat{\sigma}_{10}^2 = \tfrac{n}{2}[n^{-1}\ Y^1(I-P_1)\ Y - \hat{\sigma}_{10}^2]/\hat{\sigma}_{10}^2 \tag{7}$$

Since T_0 and TL_0 are asymptotically equivalent under H_0, they have the same asymptotic variance.

An unbiased estimate of $P_1Y=Zb_1$ under H_0 is $P_1P_0Y = Z\hat{b}_{10}$, where b_{10} is a consistent estimate under H_0 of $\text{plim}_0\ \hat{b}_1$. The Atkinson version of the Cox test is then based upon the statistic

$$TA_0 = (\hat{l}_0 - \hat{l}_{10}) - n[\text{plim}_0\ n^{-1}\ (\hat{l}_0 - \hat{l}_{10})] \qquad \theta_0 = \hat{\theta}_0 \tag{8}$$

in which

$$\hat{l}_{10} = -\tfrac{n}{2} \log 2\pi - \tfrac{n}{2} \log \hat{\sigma}_{10}^2 - \tfrac{n}{2}\ [n^{-1} Y'(I-P_1P_0)'(I-P_1P_0)Y]/\hat{\sigma}_{10}^2$$

Hence

$$TA_0 = \tfrac{n}{2}\ [n^{-1} Y^1\ (I-P_1P_0)^1\ (I-P_1P_0)Y - \hat{\sigma}_{10}^2]/\hat{\sigma}_{10}^2 \tag{9}$$

In (8) the entire statistic is evaluated under the null hypothesis. Since the second terms on the right-hand sides of (3) and (8) are asymptotically equivalent under H_0 and because $\text{plim}_0\ \hat{\sigma}_1^2 = \text{plim}\ \hat{\sigma}_{10}^2 = \text{plim}_0\ n^{-1} Y^1(I-P_1P_0)^1\ (I-P_1P_0)Y = \hat{\sigma}_{10}^2$ it follows that T_0 and TA_0 are asymptotically equivalent under H_0.

Since TA_0 and T_0 are asymptotically equivalent under H_0, a consistent estimate of the asymptotic variance of TA_0 is given in (5). Thus the statistic $NA_0 = TA_0/(\hat{V}_0)^{\frac{1}{2}}$ is also approximately distributed as a standard normal variate in large samples.

4. Tests based on artificial regressions

Davidson and Mackinnon (1981) proposed a procedure for testing H_0 by constructing an artificial regression from the models (1) and (2)

$$Y = Xb_0 + \alpha Zb_1 + U \qquad (10)$$

and testing whether $\alpha=0$ by using a conventional asymptotic t test or, equivalently, the likelihood ratio test.

Since α in (10) is not identifiable they suggested first replacing b_1 in (10) by $\hat{b}_1$, its least square estimate under H_1 leading to

$$Y = Xb_0 + \alpha Z\hat{b}_1 + U$$

or

$$Y = Xb_0 + \alpha P_1 Y + U \qquad (11)$$

where P_1 is the orthogonal projection on the span of Z, and then testing $\alpha=0$, provide a vlid test for H_0. This is called the J test.

The process of replacing b, in (10) by $\hat{b}_1$ is arbitrary, since b_1 may be replaced by any estimate that is asymptotically uncorrelated with the disturbances under H_0. Fisher and McAleer (1981) drawing upon Atkinson's (1970) suggsetion that quantities be evaluated under the null, proposed an alternative to the J procedure replacing b_1 with $\hat{b}_{10}$, a consistent estimate of the asymptotic expectation of b_1 under H_0. They also argued that this will improve the properties of the test not only does it entail re-evaluating the entire statistic under H_0 but because it leans in the direction of H_0 if the model under H_1 is performing better than expected.

In practice they replace b_1 with an estimat eof the expected value of $\hat{b}_1$ under H_0, namely P_1P_0Y where P_0 is the orthogonal projection on the span of X. The equation (11) then becomes.

$$Y = Xb_0 + \alpha\, P_1P_0Y + U \qquad (12)$$

It is straightforward to show that under H_0, the t ratio for $\hat{\alpha}$ from (12) is asymptotically equivalent to the J statistic from (11). This test is called the JA test.

Unlike the J test which enjoys only large sample validity, the JA test is exact when the regressors X and Z are fixed in repeated sampling with JA being distributed as t with $(n-k_0-1)$ degrees of freedom. This is because the additional regressor for the JA test namely P_1P_0Y, is obtained by regressing P_0Y on the columns of Z, whence P_1P_0Y is a function of P_0Y and is independent of e_0; i.e. the residuals from the hull model.

As we pointed out earlier, since the process of replacing b_1 in (10) using $\hat{b}_1$ or $\hat{b}_{10}$ is purely arbitrary we can use any estimate for b_1 without violating the asymptotic properties of the test. Immediately one thinks of a weighted average of the above two estimates for b_1 namely,

$(\lambda_1\hat{b}_1 + \lambda_2\hat{b}_{10})/(\lambda_1 + \lambda_2)$, where λ_1 and λ_2 are appropriate weights for $\hat{b}_1$ and $\hat{b}_{10}$. This will lead to the artificial regression

$$Y = Xb_0 + \alpha^1 Z(\lambda_1 b_1 + \lambda_2 b_{10}) + U$$

or

$$Y = Xb_0 + \alpha^1(\lambda_1 P_1 + \lambda_2 P_1 P_0)Y + U \tag{13}$$

and a test of $\alpha^1 = 0$ can be used as a valid test for the null hypothesis as in the case of the J and the JA tests. This test we call the JJA test.

Since we are incorporating more information regarding H_0 and H_1 through the wighted average of the two estimates we may expect more power relative to the case of a single estimate. On the other hand better small sample performance can be expected as the second estimate leans the test statistic in the direction of H_0. However, these points should be regarded as indicative rather than definitive.

5. Relationship between tests based on centered likelihood ratios and artificial regressions

The important point to note in the discussion in Section 3 is that while both T_0 and TA_0 modify a log-likelihood ratio by subtracting its asymptotic expectation under the null model, the log-likelihood ratios differ in that only the Cox statistic uses the maximised log-likelihood of the specific alternative. Another possibility in place of $\hat{l}_1$ or $\hat{l}_{10}$ is to use $\hat{l}_1^*$ the calculated value of the log-likelihood function under H_1, evaluated at $\theta_1^* = (\hat{b}_1^*, \hat{\sigma}_{10}^2)$, where $\hat{b}_0^*$ is an alternative estimate of b_1. which gives

$$T_0^* = (\hat{l}_0 - \hat{l}_1^*) - n[\text{plim}_0\, n^{-1} (\hat{l}_0 - \hat{l}_1^*)]\,(\theta_0 - \hat{\theta}_0) \tag{14}$$

Following Pesaran (1982b) and McAleer (1987) and restricting ourselves to a class of estimators of b_1 which are linear in Y, we substitute RY for $\hat{b}_1^*$ in the artificial regression.

$$Y = X\hat{b}_0 + \alpha \hat{Z}b_1^* + U \tag{15}$$

which is the regression that Davidson and MacKinnon (1981) and Fisher and McAleer (1981) used in their precedures for testing non-nested models. The k_1 x n matrix R is also assumed, for our purposes, to be non-stochastic, or at least uncorrelated with the error in large samples.

It is straightforward to relate tests based upon different modified likelihood ratios through $\hat{l}_1^*$ and the test of $H_0' : \alpha = 0$ in the artificial regression (15). If the tests are related to the Cox test through the artificial regression of the form given in (15), we can restrict the possibilities for R to a specific set of matrices. The likelihood function for H_1 evaluated at $(\hat{b}_1^*, \hat{\sigma}_{10}^2) = (Y^1R^1, \hat{\sigma}_{10}^2)$ is

$$\hat{l}_1^* = -\tfrac{n}{2}\log 2\pi - \tfrac{n}{2}\log\hat{\sigma}_{10}^2 - \tfrac{n}{2}]n^{-1}Y^1(I-ZR)^1(I-ZR)Y]/\hat{\sigma}_{10}^2$$

so that T_0^* may be written as

$$T_0^* = \tfrac{n}{2}[n^{-1}Y' (I-ZR)'(I-ZR)Y - \hat{\sigma}_{10}^2]/\hat{\sigma}_{10}^2 \tag{16}$$

In large samples the mean of T_0^* should be equal to zero since

$$Y' (I-ZR)' (I-ZR) - n\hat{\sigma}_{10}^2 = Y' [-ZR-R'Z'+R'Z'ZR+P_0P_1P_0]Y$$

upon substituting $Xb_0 + U$ for Y in (16), the following condition must hold for T_0^* to have zero mean under H_0 :

$$X'[ZR + R'Z - R'Z'ZR]X = X'P_1X \tag{17}$$

The condition in (17) indicates that ZR must be of the form $(\lambda_1P_1 + \lambda_2P_1P_0)/(\lambda_1 + \lambda_2)$, where λ_1 and λ_2 are fixed constants. Hence

$$T_0^* = \left[\frac{\lambda_1(\lambda_1+2\lambda_2)}{(\lambda_1+\lambda_2)^2} TL_0 + \frac{\lambda_2^2}{(y_1+\lambda_2)^2} TA_0 \right] \tag{18}$$

When $\lambda_2 = 0$, $ZR = P_1$, then $T_0^* = TL_0$. An asymptotically equivalent test statistic is obtained as the t-ratio of the OLS estimate of α in the artificial regression (11), which is obtained b} replacing $Z\hat{b}_1^*$ in (15) by P_1Y. The test $\alpha = 0$ in (11) is the J test of Davidson and Mac Kinnon (1981).

When $\lambda_1 = 0$, $ZR = P_1P_0$, then $T_0^* = TA_0$. An asymptotically equivalent test to TA_0 is the t ratio of the OLS estimate of α in the artificial regression (12), which is obtained by replacing $Z\hat{b}_1^*$ in (15) by P_1P_0Y.

The test $\alpha = 0$ in (12) is the JA test of Fisher and McAleer (1981).

When $\lambda_1, \lambda_2 = 0$, $ZR = (\lambda_1P_1 + \lambda_2P_1P_0) / (\lambda_1 + \lambda_2)$, and T_0^* is given as in (18). As previously explained, an asymptotically eqrivalent test statistic to TL_0 or TA_0 is therefore obtained as the t-ratio of the OLS estimate of α' in the artificial regression (13), which is obtained by replacing $Z\hat{b}_1^*$ in (15) by $(\lambda_1P_1 + \lambda_2P_1P_0)/(\lambda_1 + \lambda_2)$. The test $\alpha^1 = 0$ in (13) is the JJA test.

6. Small sample corrections and adjusted J and JJA tests

Apart from the JA test two other tests (J and JJA) are valid, only asymptotically. Therefore in theory we cannot expect better performance of these tests in small samples. Godfrey and Pesaran (1983) pointed out that the small sample significance levels of the J test are not so well behaved and are often too high. They also indicated that, as in the case of the Cox test, this may be due to the non-zero expectation of the statistic under H_0 in small samples. If this is the case, even for the JJA test, one cannot expect better performance of significance levels in small samples because it also suffers from the same problem. Along the lines suggested by Godfrey and Pesaran 1983), we can also derive mean and the variance adjusted J and JJA statistics to eliminate the size deficiency in small samples.

6.1 Adjusted J test

Consider the numerator of the statistic of the J test given by

$$J = e'_0 \hat{Y}_1$$

where e_0 is the OLS residual vector of model H_0 and $Y_1 = Z\hat{b}_1$ is the predicted value of Y from H_1.

Using E_0 to denote the expectation under H_0, we have

$$E_0\ (J) = \sigma_0^2 \operatorname{tr} M_0P_1 = \sigma_0^2\, [k_1 - \operatorname{tr}\ P_1P_0]$$

Define

$$\tilde{J} = e'_0\hat{Y}_1 - E_0\,(J) = U'_0BU_0 + q'U_0 \qquad (19)$$

where

$$B = M_0P_1 - \operatorname{tr}\ \frac{(M_0P_1)}{n-k_0}\ M_0$$

and

$$q = M_0P_1\, X\, b_0 = -M_0M_1\, X\, b_0$$

Assuming U_0 is independent $N\,(0, \sigma_0^2)$ from (19) we have

$$V_0\ (J\,) = 2\sigma_0^4 \operatorname{tr}\,(B^2) + \sigma_0^2 q'q$$

We are now in a position to define the following statistic for testing H_0. Define

$$MJ = \frac{e'_0\hat{Y}_1 - \overset{z}{\sigma_0^2} \operatorname{tr} M_0P_1}{[2\tilde{\sigma}_0^2 \operatorname{tr}\,(B^2) + \tilde{\sigma}_0^2\, e'_{100}\, e_{100}]^{\frac{1}{2}}} \qquad (20)$$

where $\tilde{\sigma}_0^2 = e_0'e_0/(n\text{-}k_0)$ and e_{100} stands for the OLS estimator of $q(= -M_0M_1 X b_0)$ and can be computed consistently by the OLS residual vector from the regression of e_{01} on X (e_{01} is the OLS residual vector of the regression of $\hat{Y}_0$ $(=X\hat{b}_0)$ on the explanatory variable Z of the model H_1. Following Godfrey and Pesaran (1983) we can show that under H_0, MJ is approximately distributed as N (0, 1).

6.2 Adjusted JJA test

It can esily be shown that the statistic for the adjusted JJAI test statistic (assuming $\lambda_1 = \lambda_2 = 1$) can be expresseed as

$$\text{MJJAI} = \frac{Y'P_1Y - Y'P_0P_1P_0Y_0 - \tilde{\sigma}_0^2 \,\text{tr}\, M_0P_1}{[2\tilde{\sigma}_0^4 \,\text{tr}\,(B^2) + 4\tilde{\sigma}_0^2 \, e'_{100}e_{100}]^{\frac{1}{2}}} \tag{21}$$

where $B = P_1 - P_0P_1P_0 - \text{tr}\ \dfrac{(M_0P_1)}{n\text{-}k_0} M_0$

As in the case of the MJ test it can easily be established that, under H_0, MJJAI is also approximately distributed as N(0, 1).

It is noteworthy to mention here that the mean of the JA test is zero even in small samples and also its variance is already in adjusted form. Therefore no adjustment can be made as in the other similar tests.

The existence of asymptotically equivalent tests with substantially different computational costs raises the issue of whether there are significant finite sample size and power differences among these alternative procedures. Indeed, in most practical applications it will be difficult if not impossible, to tell whether a large number of observations guarantees that the finite sample distributions at the different test statistics are close to their asympotois distributions. Therefore it is necessary to evaluate the finite sample properties of these tests.

In the rest of the paper we shall employ Monte Carlo experiments to evaluate the small sample performances of three closely related and asymptotically equivalent test procedures namely the J test, the JA test and the JJA test. We shall consider three variants of the JJA test by assigning different values to λ_1 and λ_2. Those are the JJAI test ($\lambda_1 = 1_2$, $\lambda_2 = 1$), the JJA2 test ($\lambda_1 = 1, \lambda_2 = 2$) and the JJA3 test ($\lambda_1 = 1$, $\lambda_2 = 3$). Adjusted J and JJAI tests will also be included in our experiments to see whether any significant improvement will occur be keeping small sample means under H_0 equal to zero and adjusting the variance accordingly.

7. Description of the Monte Carlo experiments

Previous Monte Carlo studies have reported that the J test has poor size when the false model has more parameters than the true model (i. e. $k_1 > k_0$), while the JA test has poor power when $k_0 > k_1$. This indicates that the number of parameters of the model will play a crucial role in determining the small sample performance of these tests. Also the correlation among the regressor variables across the two models may affect the properties of the tests. Therefore we will pay attention to these points when we designing our experiments.

The model used to generate the n observations was

$$Y_t = \sum_{i=I}^{k_0} \times_{ti} + U_t, \; t=1,2\ldots\ldots n \qquad (22)$$

The values of X_{ti} were generated according to N(0, 1) using TSP Random procedure and kept fixed in repeated sampling. The values of U_t were generated according to $N(0, \sigma_u^2)$ by setting the value of σ_u^2 so as to ensure the population multiple correlation coefficient of (22), the true data generating scheme, is equal to R^2. That is

$$\sigma_U^2 = k_0(1 - R^2)/R^2$$

The values of the explanatory variables of the alternative false model were generated in the following manner :

$$z_{ti} = \gamma_i x_{ti} + v_{ti} \quad i=1,2,\ldots\min(k_0,k_1)$$

and, if $k_i > k_0$

$$z_{ti} = v_{ti}, \; i=k_0+1,\ldots,k_1$$

with $v_{ti} \sim N(o, 1)$, for $t=1, 2,\ldots,n$.

The values of γ were set according to

$$\gamma_i = \rho_i/(1 - \rho_i^2)^{\frac{1}{2}}, \quad i = 1,2,\ldots,\min(k_0, k_1 l$$

This ensures that the simple correlation between x_{ti} and z_{ti} is equal to ρ_i, for $i=1, 2,\ldots\min(k_0, k_1)$ and that other x_{ti} and z_{ti} are uncorrelated. In all experiments we considered fixed alternatives to (22) and set $\rho_i = \rho$. Note that k_0 and k_1, here denote the number of variables specific to the two models, i. e. the number of non-overlapping variables in H_0 and H_1 respectively. Thus we restrict our experimental design parameters to (a) the population multiple correlation coefficient of the data generating process, (b) the correlation coefficient between the regressers of the true and the false models and (c) the number of non-overlapping variables in the two models.

Since the aim of this study is to compare the small sample performance of various tests we selected n = 20 as our sample size. We conducted several experiments with various combinations of values of R^2, ρ^2 and the number of non-overlappng variables. Each experiment was replicated 500 times. Since our interest is in testing the adequacy of the models rather then in discriminating between them we used two sided tests in our study.

For each of the 500 replications in each experiment we computed the test statistics J^2, JA^2, $JJAI^2$, $JJA2^2$, $JJA3^2$, MJ^2 and $MJJAI^2$. We then computed estimates of the significance levels of these tests by calculating the proportion of times that their corresponding test statistics exceeded : the 5 percent critical value of the F_1,n-k_0-1 distribtion in the case of J^2, JA^2, $JJAI^2$, $JJA2^2$ and $JJA3^2$, and the 5 percent critical value of the x^2_1 distribution in the case of MJ^2 and $MJJAI^2$.

Since the hypotheses considered here are non-nested, a null is taken to be so only temporarily. Thus, the definition of the power of the test as the probability of rejecting the false null hypothesis is not directly applicable to the non-nested hypotheses. Therefore, we defined power as the probability of making the correct decision as used in Pesaran (1974) and Godfrey and Pesaran (1983) ; that is the probability of accepting the true model and, rejecting the false one. We therefore, computed the proportion of times that each of the tests resulted in the correct inference.

8. Results of the Monte Carlo experiments

8.1 Case of unequal number of regressors

Several interesting features emerge from these experiments. The small sample significance levels of the J test are not well behaved and are often too high. When the false model has fewer regressors than the true model, i. e. $k_1 < k_0$, ($k_i = 2$, $k_1 = 4$), this problem is not as serious as when the false model contains more regressors than the true model, i. e. $k_1 > k_0$, ($k_1 = 4$, $k_2 = 2$), very high values in the range 6—25 being obtained. Although the estimated significance levels of the JJA tests are higher than the nominal value the gradual reduction of this difference can be observed when λ_2 assumes higher values. Especially for $k_1 > k_0$ this feature can be clearly seen. When kl $< k_0$ there is not much difference among the significance levels of the JJA tests and except for a very few cases they assume values very close to the nominal level. The significance levels of the MJ test are well behaved irrespectively of the number of regressors in the models but in the case of the MJJAI test, under rejection of H_0 cna be observed, more often.

TABLE 1

Power and Size Estimates for Cases with n=20, $k_o=2$, $k_1=4$ and $\rho^2=R^2$

	Estimated Power							Estimated Size						
$\rho^2=R^2$	J	JA	JJAI	JJA2	JJA3	MJ	MJJA1	J	JA	JJA1	JJA2	JJA3	MJ	MJJA1
0.30	49.8	33.0	55.0	52.8	51.4	46.2	43.4	24.8	5.6	15.2	11.8	9.4	6.6	3.2
0.50	67.2	61.0	71.6	68.0	66.6	72.0	66.6	17.0	5.8	10.0	9.0	7.8	5.8	4.6
0.80	80.0	78.2	80.4	80.4	80.0	81.0	80.8	11.8	5.2	9.0	7.4	7.4	6.4	5.4
0.85	80.6	80.2	81.8	81.6	80.6	81.7	80.2	11.6	5.4	8.8	7.4	7.2	6.2	7.0
0.90	80.6	81.0	82.0	81.8	82.2	81.8	80.8	11.0	5.4	8 4	7.6	7.4	6.5	7.0

TABLE 2

Power and Size Estimates for Cases with n=20, $k_o=4$, $K_1=2$ and $\rho^2=R^2$

$\rho^2=R^2$	J	JA	JJA1	JJA2	JJA3	MJ	MJJAI	J	JA	JJA1	JJA2	JJA3	MJ	MJJA1
0.30	71.8	6.2	73.0	65.8	60.6	31.2	26.8	14.0	4.4	9.0	7.2	5.4	5.8	4.2
0.50	81.6	10.6	85.8	84.0	80.6	64.0	56.2	15.6	5.2	10.2	7.8	6.2	7.6	3.6
0.80	93.8	6.8	94.4	94.6	94.0	95.8	96.6	6.2	4.8	5.6	5.4	5.4	4.2	3.4
0.85	94.2	5.4	95.4	96.4	96.2	95.0	97.2	5.8	4.4	4.6	3.6	3.8	5.0	2.8
0.90	95.0	2.6	95.6	95.6	95.6	96.0	96.6	5.0	4.4	4.4	4.4	4.0	4.0	3.4

TABLE 3

Power and Size Estimates for Cases with n=20 and R^2=0.50

			Estimated Power							Estimated Size						
k_o	k_1	ρ^2	J	JA	JJA1	JJA2	JJA3	MJ	MJJA	J	JA	JJA1	JA2	JJA3	MJ	MJJA1
2	4	0.25	76.4	66.6	81.6	84.0	83.6	79.2	78.6	18.2	4.6	11.4	7.0	6.2	5.2	2.8
2	4	0.30	75.8	65.4	80.2	81.0	81.0	78.4	77.4	17.4	5.8	11.8	7.2	6.8	5.0	2.8
2	4	0.50	67.2	61.0	71.6	68.0	66.6	72.0	66.6	17.0	5.8	10.0	9.0	7.8	5.8	4.6
2	4	0.80	35.6	34.2	37.2	36.2	36.0	41.8	34.0	19.0	5.6	11.4	10.2	8.8	6.0	5.0
2	4	0.85	30.4	27.6	30.8	30.8	30.2	34.8	24.8	19.2	5.8	14.0	10.4	9.2	5.4	5.2
2	4	0.90	19.0	16.8	19.8	19.6	18.2	14.4	16.6	18.6	5.8	15.8	11.0	10.4	6.0	5.6
2	4	0.95	9.0	7.0	9.8	9.8	9.8	10.4	5.0	18.8	5.8	16.8	15.0	11.8	7.0	5.2
4	2	0.25	84.2	16.4	90.0	88.4	85.6	60.4	59.0	12.2	5.4	5.8	6.2	6.6	5.6	4.6
4	2	0.30	83.6	11.2	88.4	88.4	86.2	60.8	58.4	13.0	4.8	7.2	5.8	5.0	5.2	4.6
4	2	0.50	81.6	10.6	85.8	84.0	80.6	64.0	56.2	15.6	5.2	10.2	7.8	6.2	7.6	3.6
4	2	0.80	86.2	6.8	81.4	71.2	64.6	61.8	54.0	7.4	5.4	7.0	5.4	7.0	5.8	2.6
4	2	0.85	84.8	6.2	77.8	69.8	62.2	60.6	53.6	6.0	5.4	6.4	6.2	6.2	5.0	1.0
4	2	0.90	82.0	5.0	74.8	65.8	56.8	57.4	51.6	6.4	5.2	5.8	6.4	6.6	4.8	1.6
4	2	0.95	76.4	2.4	70.2	65.6	56.4	52.6	49.8	5.2	5.4	6.2	6.0	5.4	5.0	1.8

TABLE 4

Power and Size Estimates for Cases with n = 20 and $R^2 = 0.50$

			Estimated Power							Estimated Size						
k_0	k_1	ρ^2	J	JA	JJA1	JJA2	JJA3	MJ	MJJAI	J	JA	JJA1	JJA2	JJA3	MJ	MJJA1
2	2	0.25	84.2	63.4	80.8	79.6	79.2	80.0	76.6	7.6	5.8	8.6	8.4	8.4	5.0	4.6
2	2	0.30	81.6	62.2	80.4	78.6	77.6	78.4	76.4	9.0	5.4	7.8	8.0	8.0	5.4	4.8
2	2	0.50	73.2	59.0	69.4	67.0	65.6	72.2	66.8	7.4	5.2	6.6	7.0	7.2	5.6	4.4
2	2	0.80	36.6	32.8	34.4	34.4	33.4	40.4	31.4	5.4	5.0	5.6	5.0	5.0	4.6	4.2
2	2	0.85	28.0	25.8	27.4	27.2	26.4	32.4	24.8	5.6	4.6	5.0	5.0	5.0	5.4	3.8
2	2	0.90	20.2	18.0	18.6	18.6	18.6	21.6	16.0	5.6	4.6	5.0	5.0	5.0	5.6	4.4
2	2	0.95	9.8	9.0	9.6	9.4	9.4	11.2	6.2	5.0	5.0	5.0	5.0	5.0	5.6	4.4
4	4	0.25	71.6	32.2	82.6	84.8	83.0	60.8	60.0	24.6	4.2	12.6	9.0	6.6	5.0	3.2
4	4	0.30	72.6	34.4	83.0	85.0	81.2	60.6	61.2	23.6	4.2	13.0	7.6	7.2	5.8	3.5
4	4	0.50	80.8	52.4	84.2	79.2	73.6	59.4	55.0	14.2	4.8	8.2	5.4	4.2	4.4	2.0
4	4	0.80	57.6	30.8	43.4	38.0	36.0	31.0	25.2	8.6	5.0	5.0	5.6	5.6	4.6	2.0
4	4	0.85	42.6	21.4	30.2	26.8	24.4	26.0	16.0	8.0	5.0	5.0	5.6	5.6	4.0	1.8
4	4	0.90	29.8	16.2	20.2	18.6	18.2	19.6	11.0	7.0	5.0	5.2	5.2	5.2	4.0	1.8
4	4	0.95	16.0	10.2	11.4	10.2	10.2	12.4	6.2	7.0	4.8	5.0	5.0	5.0	4.0	1.8

Although we cannot make meaningful comparisons of power of the tests with different values of significance levels, we can observe that, for $k_1 > k_0$, the powers of the JJA tests are always greater than that of the J test. A similar pattern can be seen even for $k_1 < k_0$ but in the lower range of p^2. It can be observed that the behaviour of the JA test is quite different from that of the other tests. The power of the JA test is very sensitive to the number of regressors in the models. It has considerable power when $k_1 > k_0$, but the performance is very poor when $k_1 < k_0$. For tests with acceptable size MJ has the highest power in almost all cases. When $k_1 > k_0$ the powers of all the tests decrease as ρ^2 increases but in the case of $k_1 < k_0$ some irregularities can be seen.

8.2 Case of equal number of regressors

The results for the case of an equal number of regressors, namely $(k_0, k_1) = (2, 2)$ and $(4, 4)$, are interesting in that the size of the J test is generally acceptable for $k_0 = k_1 = 2$, but not for $k_0 = k_1 = 4$, especially in the lower range of p^2. Except for a very few cases with very low values of p^2, the size of the JJA tests are acceptable even for $k_0 = k_1 = 4$. As in the previous case the downward trend of the significance levels of the JJA tests can be seen especially in the lower range of $p2$ when $k_0 = k_1 = 4$. The size of the MJ test is well behaved in both cases but for $k_0 = k_1 = 4$ the significance levels of the MJJAI tests are very low.

As far as power is concerned, when $k_0 = k_1 = 2$, the power of the J test is slightly higher than that of the JJA tests while when $k_0 = k_1 = 4$, especially in the lower range of p^2. the power of the JJA tests are assured higher value than that of the J test. Although the power of the JA test is considerable in both cases that will not reflect any superiority over the other comparable tests. In both cases the power of the MJ test is slightly higher than that of the MJJAI test.

9. Conclusions

What are we to make of these results ? According to previous Monte Carlo studies it has been revealed that the J test has very high significance levels when the false model contains more explanatory variables than the true model (i. e. $k_1 > k_0$) while the JA test has poor power for $k_1 < k_0$. Our findings also generally confirm these results. In practical situations it will not be known whether the true model has more or fewer number of regressors than the false one the selection of the appropriate testing procedure is hardly, possible. Although the behaviour of the J and the JA tests are highly sensitive to the number of regressors in the two models, our findings reveal that the behaviour of the JJA tests (especially JJA2 and JJA3) are not so sensitive to the number of regressors in the models. Also we can observe that their significance levels are always closer to the nomiral level than that of the J test and their power, generally, is in line with that of the J test. These features with their simplicity and ease of implementation (using existing computer packages) reflect the attractiveness of these tests in applied econometric studies. Although the behaviour of the MJ test is quite acceptable in both respects (power and size) computational complexity hampers the wide application of this test.

If one is prepared to accept a slightly higher size at less computational cost then JJA tests (especially JJA2 and JJA3) would be a better choice. On the other hand, if one is more concerned about the optimum properties of the tests rather than their computational complexity then the $\tilde{N}$ test suggested by Godfrey and Pesaran (1983) would be the best choice because the $\tilde{N}$ test is slightly superior to the MJ test but both involve roughly the same amount of computation. The JA test can also be recommended if the models being tested have the equal number of (relatively few) non-overlapping variables. Finally as to the choice of non-nested tests, the mean and variance adjusted tests seem to have no advantages over the JJA tests to balance their computational complexity.

References

Atkinson, A C (1970), "A Method for Discriminating Between Models", *Journal of the Royal Statistical Society B*, 32, 323-53.

Cox, D. R. (1967), ' Tests of Separate Families of Hypotheses", *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probabil iy*, 1, Berkeley : University of California Press, 105-23.

Cox, D. R. (1962), "Further Results on Test of Separate Families of Hypotheses", *Journal of the Royal Statist cal Society B*, 24, 406-24.

Davidson, R. and MacKinnon, J. G' (1981), "Several Tests for Model Specification in the Presence of Alternative Hypotheses , *Econometrica*, 49, 781-93.

Davidson, R. and Mackinnon J. G. (1982), "Some Non-nested Hypotheses Tests and the Relations among them", *Review of Economic Studies*, 49, 551-65.

Fisher, G. R. and McAleer, M. (1981), "Alternative Procedures and Associated Tests of Significance for Non-nested Hypotheses", *Journal of Econometrices*, 16, 103-19.

Friedman, M. and Heiselman, D (1963), "The Relative Stability of Monetary Velocity and the Investment Multiplier in the United States 1987-1958, in the Commission on Monetary and Credit Research Study *Stabilistion policies*, Prentice-Hall, Engelwood Cliffs, New Jersey.

Godfrey, L. G. and Pesaran, M. H. (1983), "Tests of Non-nested Regression Models : Small Sample Adjustments and Monte Carlo Evidence", *Journal of Econometrics*, 21, 133-54.

Jorgenson, D. W. and Siebert, C. D. (1968), "A Comparison of Alternative Theories of Corporate Investment Behaviour", *American Economic Review*, 58, 681-712.

McAleer, M. (1987), "Specification Tests for Separate Models: A Survey", in *Specification Analysis in the Linear Models* (in honour of Donald Cochrane), eds : King, M. L. and Giles, D. E. A. Routledge & Kedan Paul Ltd., London.

Pesaran, M. H. (1974), "On the General Problem of Model Selection", *Review of Econometric Studies*, 41, 153-71.

Pesaran, M. H. (1982a), "Comparison of Local Power of Alternative Tests of Non-nested Regression Models", *Econometrica* 50, 1287-1305.

Pesaran, M. H. (1982b), "On the Comprehensive Method of Testing Non-nested Regression Models", *Journal of Econometrics*, 18, 263-74.

Pesaran, M. H. and Deaton, A. S. (1978), ' Testing Non-nested Nonlinear Regression Models' *Econometrica*, 46, 677-94.