Economics Letters 7 (1981) 313–318
North-Holland Publishing Company

EFFICIENT TESTS FOR NORMALITY, HOMOSCEDASTICITY AND SERIAL INDEPENDENCE OF REGRESSION RESIDUALS
Monte Carlo Evidence

Anil K. BERA and Carlos M. JARQUE

Australian National University, Canberra ACT 2600, Australia

Received 23 April 1981

In this paper we study the performance of various tests for normality (N), homoscedasticity (H) and serial independence (I) of regression residuals (u) under one, two and three directional departures from $H_0: u \sim NHI$.

1. Introduction and description of Monte Carlo study

In an earlier paper [see Jarque and Bera (1980)] we used the Lagrange multiplier procedure to derive an efficient test, LM_{NHI}, for the hypothesis that regression residuals (u) are 'well behaved', i.e., $H_0: u \sim NHI$. (For notation used here we refer the reader to that paper.) We also noted the one and two directional tests that arise as particular cases of the procedure used, i.e., LM_N, LM_H, LM_I, LM_{NH}, LM_{HI} and LM_{NI} which are asymptotically efficient for testing H_0 when u is, respectively, HI, NI, NH, I, N and H (e.g., LM_{NH} is asymptotically efficient for testing H_0 when $u \sim I$). In this paper we report the results of extensive simulation experiments, designed to study the power of these tests under various residual distributional assumptions.

Departures from H_0 may arise because u is serially correlated ($\bar{I}$) and/or heteroscedastic ($\bar{H}$) and/or non-normal ($\bar{N}$). Regarding *serial correlation*, we generate residuals from an autoregressive process $u_t = \rho u_{t-1} + \epsilon_t$ where, to study the effect of 'weak' and 'strong' autocorrelation, we set $\rho = \rho 1 = 0.3$, and $\rho = \rho 2 = 0.7$. We consider *heteroscedasticity* of the form $E[\epsilon_t^2] = \sigma_t^2 = 25 + \alpha z_t$, where $\sqrt{z_t}$ is generated from a Normal (10,25). We use $\alpha = \alpha 1 = 0.25$, and $\alpha = \alpha 2 = 1.25$ (in our case these values represent 'weak' and 'strong' heteroscedasticity). Regarding *non-*

0165-1765/81/0000–0000/$02.75

Table 1
Estimated power using 1000 replications (10 percent significance level).

		LM_N	LM_H	LM_I	LM_{NH}	LM_{HI}	LM_{NI}	LM_{NHI}
One directional								
$\bar{N}HI$:	t	*0.496*	0.183*	0.108*	0.466	0.167*	0.466	0.443
	B	*0.174*	0.061*	0.115*	0.076	0.082*	0.157	0.112
	Log	*1.000*	0.586*	0.050*	0.999	0.531*	1.000	1.000
$N\bar{H}I$:	$\alpha 1$	0.178*	*0.466*	0.101*	0.418	0.403	0.156*	0.386
	$\alpha 2$	0.329*	*0.809*	0.090*	0.757	0.737	0.290*	0.708
$NH\bar{I}$:	$\rho 1$	0.125*	0.105*	*0.581*	0.102*	0.506	0.500	0.424
	$\rho 2$	0.166*	0.075*	*0.990*	0.108*	0.987	0.985	0.979
Two directional								
$\bar{N}\bar{H}I$:	$t,\alpha 1$	0.540	0.481	0.101*	*0.653*	0.433	0.503	0.622
	$B,\alpha 1$	0.161	0.577	0.110*	*0.519*	0.511	0.149	0.470
	$Log,\alpha 1$	0.999	0.591	0.047*	*1.000*	0.521	1.000	1.000
	$t,\alpha 2$	0.662	0.757	0.100*	*0.839*	0.698	0.613	0.818
	$B,\alpha 2$	0.234	0.860	0.107*	*0.806*	0.809	0.225	0.760
	$Log,\alpha 2$	1.000	0.587	0.043*	*1.000*	0.522	1.000	1.000
$N\bar{H}\bar{I}$:	$\alpha 1,\rho 1$	0.165*	0.413	0.573	0.359	*0.666*	0.524	0.590
	$\alpha 2,\rho 1$	0.287*	0.724	0.568	0.683	*0.851*	0.568	0.801
	$\alpha 1,\rho 2$	0.179*	0.188	0.993	0.204	*0.994*	0.989	0.980
	$\alpha 2,\rho 2$	0.208*	0.352	0.994	0.347	*0.997*	0.992	0.994
$\bar{N}H\bar{I}$:	$t,\rho 1$	0.455	0.153*	0.582	0.417	0.540	*0.680*	0.648
	$B,\rho 1$	0.135	0.076*	0.571	0.075	0.485	*0.527*	0.457
	$Log,\rho 1$	1.000	0.550*	0.568	1.000	0.734	*1.000*	1.000
	$t,\rho 2$	0.303	0.093*	0.997	0.236	0.991	*0.991*	0.987
	$B,\rho 2$	0.137	0.063*	0.996	0.087	0.991	*0.991*	0.986
	$Log,\rho 2$	0.989	0.296*	0.996	0.980	0.995	*1.000*	1.000
Three directional								
$\bar{N}\bar{H}\bar{I}$:	$t,\alpha 1,\rho 1$	0.502	0.421	0.573	0.584	0.690	0.712	*0.754*
	$B,\alpha 1,\rho 1$	0.135	0.500	0.562	0.434	0.694	0.518	*0.662*
	$Log,\alpha 1,\rho 1$	1.000	0.542	0.567	0.999	0.715	1.000	*1.000*
	$t,\alpha 2,\rho 1$	0.584	0.703	0.577	0.790	0.838	0.761	*0.874*
	$B,\alpha 2,\rho 1$	0.205	0.788	0.557	0.718	0.866	0.526	*0.828*
	$Log,\alpha 2,\rho 1$	0.999	0.558	0.565	0.999	0.721	1.000	*1.000*
	$t,\alpha 1,\rho 2$	0.327	0.221	0.995	0.342	0.993	0.993	*0.989*
	$B,\alpha 1,\rho 2$	0.148	0.217	0.995	0.200	0.994	0.989	*0.990*
	$Log,\alpha 1,\rho 2$	0.987	0.295	0.994	0.977	0.993	1.000	*1.000*
	$t,\alpha 2,\rho 2$	0.356	0.351	0.994	0.452	0.997	0.992	*0.995*
	$B,\alpha 2,\rho 2$	0.165	0.394	0.995	0.353	0.991	0.989	*0.989*
	$Log,\alpha 2,\rho 2$	0.986	0.296	0.994	0.969	0.993	1.000	*1.000*
NHI:		0.100	0.100	0.100	0.100	0.100	0.100	0.100

normal alternatives we consider three distributions $g(\epsilon_t)$: Students t_5, Beta (3,2) and the Lognormal (say t, B and Log), which cover a wide range of skewness and kurtosis measures. In all, we generate residuals from 35 alternative distributions; consisting of 7 one directional, 16 two directional and 12 three directional departures from H_0. Residuals under H_0 were also generated, in order to obtain the empirical 10 percent significance point for each test.

We consider a linear model with $K = 4$ regressors. We set $X_{t1} = 1$ and generate X_2 from a Normal, X_3 from a Uniform and X_4 from a χ^2_2. For every experiment, we generate $T = 50$ pseudorandom variates (residuals) from a given distribution; obtain the Ordinary Least Squares (OLS) estimated residuals; compute the test statistics considered and see whether H_0 is rejected by each test. We carried out 1000 replications. The estimated power of each test (obtained by counting the number of times H_0 was rejected and dividing this by 1000) for every one of the 35 distributions is given in table 1.

2. Analysis of results

In table 1 the power of the optimal test is underlined, which in most cases coincides with the maximum power. For example, if u is generated with heteroscedasticity parameter $\alpha 1$ and autocorrelation $\rho 2$, say $u \sim N\bar{H}\bar{I}(\alpha 1, \rho 2)$, LM_{HI} (which is the optimal test when $u \sim N\bar{H}\bar{I}$) gives the maximum power equal to 0.994. In the table, *quantities with a star* (*) should be approximately equal to the significance level 0.10. We find that in terms of significance level LM_I and LM_{NH} are robust. This is not the case for all other one and two directional tests, e.g., when $u \sim \bar{N}HI(\text{Log})$ the power of LM_H is 0.586, which is undesirably high and may lead one to infer (with high probability) that the residuals are $\bar{H}$ when in fact they are H but $\bar{N}$.

For the *non-starred* quantities we observe the following interesting results.

We first look at a given *column*, i.e. a given test, and compare its power for various distributions. Considering the column for LM_I, we observe that LM_I is robust in the presence of $\bar{H}$ and/or $\bar{N}$. To see this, note that the power of LM_I is invariant to the occurrance of $\bar{H}$ and/or $\bar{N}$. For example, the powers of LM_I wherever $\rho 1$ appears are approximately the same: these are 0.581, 0.573, 0.568, 0.582 0.571, 0.568, 0.573, 0.562, 0.567, 0.577, 0.557 and 0.565 respectively for $u \sim NH\bar{I}(\rho 1)$, $N\bar{H}\bar{I}(\alpha 1, \rho 1)$,

$N\overline{H}\overline{I}(\alpha 2,\rho 1)$, $\overline{N}H\overline{I}(t,\rho 1)$, $\overline{N}H\overline{I}(B,\rho 1)$, $\overline{N}H\overline{I}(\text{Log},\rho 1)$, $\overline{N}H\overline{I}(t,\alpha 1,\rho 1)$, $\overline{N}\overline{H}\overline{I}(B,\alpha 1,\rho 1)$, $\overline{N}\overline{H}\overline{I}(\text{Log},\alpha 1,\rho 1)$, $\overline{N}\overline{H}\overline{I}(t,\alpha 2,\rho 1)$, $\overline{N}\overline{H}\overline{I}(B,\alpha 2,\rho 1)$ and $\overline{N}H\overline{I}(\text{Log},\alpha 2,\rho 1)$. For LM_H, however, power may decrease in the presence of $\overline{I}$, e.g., for $u \sim N\overline{H}I(\alpha 2)$ power is 0.809 and for $N\overline{H}\overline{I}(\alpha 2,\rho 2)$ power is 0.352. Hence, using LM_H one may wrongly infer, with high probability (1–0.352), that u is H when in fact it is $\overline{H}$ and $\overline{I}$. Similarly for LM_N in the presence of $\overline{H}$ and $\overline{I}$, e.g., for $u \sim \overline{N}HI(t)$ power is 0.496 and for $\overline{N}\overline{H}\overline{I}(t,\alpha 1,\rho 2)$ power decreases to 0.327. These results show that $\overline{I}$ may seriously affect the performance of H and N tests and highlight the possible consequences of using the one directional tests LM_H and LM_N in the presence of two and three directional departures from H_0. Similar evidence is found when using a two directional test in three directional departures from H_0.

We now look at a given *row* (i.e., a given distribution) and compare the relative power of the tests. Here we observe that the use of one directional tests, when there is more than one directional departures from H_0, may lead to a substantial loss in power with respect to the use of a two or three directional test. For example, when $u \sim \overline{N}\overline{H}I(B,\alpha 2)$ the power of LM_N is 0.234, the power of LM_{NH} is 0.806 and that of LM_{NHI} is 0.760. Similarly, when $u \sim N\overline{H}\overline{I}(\alpha 1,\rho 2)$ the powers are 0.188, 0.994 and 0.980 respectively for LM_H, LM_{HI} and LM_{NHI}; and finally, when $u \sim \overline{N}H\overline{I}(\text{Log},\rho 1)$ the power of LM_I is 0.568 and that of LM_{NI} and LM_{NHI} is 1.000. A similar result holds for two directional tests (with respect to three directional tests) in three directional departures from H_0. For example, when $u \sim \overline{N}\overline{H}\overline{I}(t,\alpha 1,\rho 2)$ power of LM_{NH} is 0.342 and that of LM_{NHI} is 0.989. It is interesting to note that, for each distribution, when comparing LM_{NHI} with the other tests, its power is insignificantly less than the corresponding maximum power. Therefore, we can conclude that, when using LM_{NHI}, there may be, firstly, a considerable gain in power with respect to two (one) directional tests, in three (two and three) directional departures from H_0; and secondly, little loss in power (with respect to the optimal test) in one and two directional departures from H_0. This evidence should motivate the use of LM_{NHI} in all time-series studies where there is no prior knowledge on the distribution of the residuals. In cross sectional studies, where $u \sim I$, the results in table 1, which show that not a great loss in power occurs by using LM_{NH} when $u \sim \overline{N}HI$ or $N\overline{H}I$, should encourage the use of LM_{NH}.

If H_0 is accepted on the basis of LM_{NHI}, 'classical' regression analysis would follow. If H_0 is rejected then, to adjust the model, one would like to know in which direction(s) the departure from H_0 is. This problem requires further attention. In the light of our simulation results, our recommendation would be to proceed by testing for autocorrelation first, using LM_I. This has been found to be robust in the presence of $\bar{N}$ and/or $\bar{H}$. *If serial independence (I) is accepted*, we would proceed to test for NH using LM_{NH}. If NH is rejected, we would suggest the use of a multiple comparison procedure [see Savin (1980)]. In this we would adjust for $\bar{H}$; or $\bar{N}$; or $\overline{NH}$ respectively if $LM_N < a$ and $LM_H \geqslant b$; or $LM_N \geqslant a$ and $LM_H < b$; or $LM_N \geqslant a$ and $LM_H \geqslant b$, where a and b are appropriate significance points for LM_N and LM_H. If $LM_N < a$ and $LM_H < b$, no adjustment would be made, although LM_{NH} has rejected NH. Such contradictory results are possible when using multiple comparison procedures [see Savin (1980, p. 261)]. *If I is rejected*, we would proceed—as above—to test for NH, but computing LM_{NH} (i.e., $LM_N + LM_H$) with estimated residuals of the Cochrane–Orcutt transformed model, rather than the original OLS estimated residuals. Finally we note that, since we are carrying out a sequence of tests, significance levels would have to be adjusted [see Savin (1980, p. 257)].

3. Concluding remarks

We also included in our simulation study the White and MacDonald (1980) modified W' test for $\bar{N}$ (derived under HI); the Payen (1980) P test for H (derived under $\bar{N}I$) and the Durbin–Watson D–W test for I (derived under NH). The performance of W' was similar to that of LM_N. P detected $\bar{H}$ quite well for $N\bar{H}I$ but its performance was rather bad, compared to LM_{NH}, for $\bar{N}\bar{H}I$. Lastly, D–W (as LM_I) was robust in the presence of $\bar{N}$ and/or $\bar{H}$.

All our results depend on the choice of T, K and the way the regressors are generated. Hence we repeated our study, changing T, K and the regressors. We found that, in all cases, our conclusions did not vary substantially from those stated here. Numerical results not reported are available from the authors.

References

Jarque, C.M. and A.K. Bera, 1980, Efficient tests for normality, homoscedasticity and serial independence of regression residuals, Economics Letters 6, 255–259.

Payen, J.-F., 1980, A simple test for homoscedasticity in case of non-normality, Paper presented at the Fourth Econometric Society World Congress 1980 (Aix-en-Provence, France).

Savin, E.N., 1980, The Bonferroni and the Scheffé multiple comparison procedures, Review of Economic Studies XLVII, 255–273.

White, H. and G.M. MacDonald, 1980, Some large-sample tests for non-normality in the linear regression model, Journal of the American Statistical Association 75, 16–28.