Journal of Econometrics 17 (1981) 107–112. North-Holland Publishing Company

A NOTE ON STUDENTIZING A TEST FOR HETEROSCEDASTICITY

Roger KOENKER

Bell Laboratories, Murray Hill, NJ 07974, USA

Received February 1981, final version received June 1981

Breusch and Pagan (1979) have recently proposed a convenient test for heteroscedasticity in general linear models. This note derives the asymptotic distribution of their test under sequences of contiguous alternatives to the null hypothesis of homoscedasticity. The test is shown to possess asymptotically incorrect size (nominal significance level) except in the case of strictly Gaussian disturbances. A slight modification of the test is proposed which corrects this defect.

1. Introduction

Several authors including Goldfeld and Quandt (1972), Amemiya (1977) and most recently Breusch and Pagan (1979) have studied versions of the following 'linear scale' model of heteroscedasticity,

$$y_i = \boldsymbol{x}_i\beta + u_i, \tag{1.1}$$

$$u_i = \sigma_i \varepsilon_i, \tag{1.2}$$

$$\sigma_i = 1 + h(z_i\gamma), \tag{1.3}$$

where $\boldsymbol{x}_i$ and z_i are k and p dimensional row vectors of known constants, β and γ are k and p dimensional vectors of unknown parameters, h is a known, smooth positive function, $h(0)=0$, and the subscript i indexes n observations.

Breusch and Pagan (1979) have proposed a Lagrange multiplier test of the null hypothesis

$$H_0 : \gamma = 0, \tag{1.4}$$

and they derive the asymptotic distribution of their test statistic under the null hypothesis and Gaussian assumptions on ε.

In this note we relax the Gaussian error assumption and study the asymptotic behavior of the test statistic proposed by Breusch and Pagan under a class of contiguous alternatives to the null hypothesis. Two conclusions emerge from this analysis:

(1) The asymptotic power of the Breusch and Pagan test is extremely sensitive to the kurtosis of the distribution of ε, and

(2) the asymptotic size of the test is correct only in special case of Gaussian kurtosis.

The former conclusion is expanded upon in Koenker and Bassett (1981) where alternative, robust tests for heteroscedasticity are suggested. The latter conclusion implies that the significance levels suggested by Breusch and Pagan will be correct only under Gaussian conditions on ε. Since such conditions are generally assumed on blind faith and are notoriously difficult to verify, a modification of the Breusch and Pagan test is suggested which correctly 'studentizes' the test statistic and leads to asymptotically correct significance levels for a reasonably large class of distributions for ε. This modification has also been independently suggested by Payen (1980).

2. The result

The following regularity conditions are employed below. The Euclidean norm of $\boldsymbol{x}$ is denoted by $\|\boldsymbol{x}\|$.

(A.1) *Design.* The sequences of design points $\{\boldsymbol{x}_i\}$ and $\{\boldsymbol{z}_i\}$ satisfy the conditions

(i) $$\lim_{n\to\infty} 1/n \sum \boldsymbol{x}_i \boldsymbol{x}_i' = \boldsymbol{Q},$$

(ii) $$\lim_{n\to\infty} 1/n \sum \boldsymbol{z}_i \boldsymbol{z}_i' = \boldsymbol{D},$$

(iii) $$\max_{i=1,\ldots,n} \|\boldsymbol{z}_i\| < M \quad \text{for all } n,$$

for positive definite matrices $\boldsymbol{Q}$ and $\boldsymbol{D}$ and positive constant M.

(A.2) *Errors.* The errors $\{\varepsilon_i\}$ are independent and identically distributed (iid) with distribution function F, with mean zero, variance σ^2 and finite fourth moment.

(A.3) *Sequence of alternatives.* There exists a fixed $\gamma_0 \in \boldsymbol{R}^p$ such that $\gamma = \gamma_0/\sqrt{n}$ in (1.3) for samples of size n, further, $h(0)=0$ and $h'(0)<\infty$.

The design assumptions embodied in (A.1) are quite standard, but may be weakened slightly. The moment conditions on F are obviously essential. The specification of a sequence of alternatives which converges to the null at rate $1/\sqrt{n}$ is a classical device employed to study the asymptotic efficiency (power) of test statistics. In the terminology of Hajek and Sidek (1967) such a sequence of alternatives is *contiguous to the null hypothesis*. Its role will become apparent below.

Breusch and Pagan base their test on the vector of least-squares residuals

$$\hat{\boldsymbol{u}}=[\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']\boldsymbol{u}, \tag{2.1}$$

and the variance estimate

$$\hat{\sigma}^2=\hat{\boldsymbol{u}}'\hat{\boldsymbol{u}}/n. \tag{2.2}$$

Let $\hat{w}_i=\hat{u}_i^2-\hat{\sigma}^2$. Their test statistic is

$$\hat{\xi}=\tfrac{1}{2}\hat{\boldsymbol{w}}'\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\hat{\boldsymbol{w}}/\hat{\sigma}^4.$$

They note that with Gaussian F and under the null hypothesis $H_0:\gamma=0$, $\hat{\xi}$ converges in distribution to a central chi-squared random variable with p degrees of freedom.

This result is extended to contiguous sequences of alternatives and error distributions satisfying the moment conditions of (A.2) in the following:

Theorem. Given the model (1.1)–(1.3) and assumptions (A.1)–(A.3) let $\phi=V(\varepsilon^2)$, $\lambda=\frac{1}{2}\phi/\sigma^4$, $\alpha_0=2h'(0)\gamma_0$, and $\eta=\alpha_0'\boldsymbol{D}\alpha_0/\phi$. Then $\xi^*=\hat{\xi}/\lambda\to\chi_p^2(\eta)$, i.e., ξ^* converges in distribution to a non-central chi-square law with p degrees of freedom and non-centrality η.

Proof. We adopt some of the notation and techniques of Amemiya (1977). Decompose $\hat{\boldsymbol{w}}$ into four parts

$$\hat{\boldsymbol{w}}=\boldsymbol{v}_1-2\boldsymbol{v}_2+\boldsymbol{v}_3+\boldsymbol{v}_4, \tag{2.3}$$

where $\boldsymbol{v}_j$, $j=1,\ldots,4$, denote n-vectors with ith components,

$$v_{1i}=u_i^2-\sigma^2, \tag{2.4}$$

$$v_{2i}=u_i\boldsymbol{x}_i(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}, \tag{2.5}$$

$$v_{3i}=(\boldsymbol{x}_i(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u})^2, \tag{2.6}$$

$$v_{4i}=\sigma^2-\hat{\sigma}^2. \tag{2.7}$$

Let $\hat{\alpha}$ denote the least-squares estimate

$$\hat{\alpha}=(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\hat{\boldsymbol{w}}. \tag{2.8}$$

We begin by establishing the following asymptotic linearity result:

$$\sqrt{n}\,\hat{\alpha}=\alpha_0+(\boldsymbol{Z}'\boldsymbol{Z}/n)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\varepsilon_i^2-\sigma^2)z_i+o_p(1). \tag{2.9}$$

The four $\boldsymbol{v}$ terms are treated sequentially. The interesting term involves $\boldsymbol{v}_1$, the others will be seen to be asymptotically negligible. First,

$$\sqrt{n}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'v_1=(\boldsymbol{Z}'\boldsymbol{Z}/n)^{-1}\frac{1}{\sqrt{n}}\sum z_i[[1+2h(z_i\gamma)+h^2(z_i\gamma)]\varepsilon_i^2-\sigma^2].$$

Using (A.1.-iii) and (A.3) we expand h around 0 to obtain the right-hand side of (2.9). Next

$$\sqrt{n}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{v}_2=(\boldsymbol{Z}'\boldsymbol{Z}/n)^{-1}\left[\frac{1}{n}\sum z_iu_i\boldsymbol{x}_i\right][\sqrt{n}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}]$$
$$=o_p(1),$$

since the first term converges to D^{-1}, the last term converges in distribution to a Gaussian law and therefore is $O_p(1)$ and the middle term converges in quadratic mean, and therefore in probability, to zero using (A.1) and (A.2). Next,

$$\sqrt{n}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{v}_3=(\boldsymbol{Z}'\boldsymbol{Z}/n)^{-1}\frac{1}{\sqrt{n}}\sum z_i[\boldsymbol{x}_i(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}]^2$$
$$\leqq\frac{1}{\sqrt{n}}(\boldsymbol{Z}'\boldsymbol{Z}/n)^{-1}[\max\|z_i\|][\boldsymbol{u}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}]$$
$$=o_p(1),$$

since the first two terms are bounded by hypothesis (A.1) and the last term is $O_p(1)$ since it converges to a χ^2 law. Finally, it is easily shown that

$$\hat{\sigma}^2=\sigma^2+o_p(1),$$

so (2.9) is verified.

From (A.1) and (A.2) a Lindeberg condition on the independent summands in (2.9) is easily checked and we conclude that $\sqrt{n}\,\hat{\alpha}$ converges to a p-variate Gaussian law with mean α_0 and covariance matrix $\phi \boldsymbol{D}^{-1}$. Thus,

$$\xi^* = n\hat{\alpha}'\boldsymbol{D}\hat{\alpha}/\phi \rightarrow \chi_p^2(\eta).$$

But Breusch and Pagan's test statistic is

$$\begin{aligned}\hat{\xi} &= \tfrac{1}{2}\hat{\alpha}'(\boldsymbol{Z}'\boldsymbol{Z})\hat{\alpha}/\hat{\sigma}^4 \\ &= \frac{n}{2}\hat{\alpha}'\boldsymbol{D}\hat{\alpha}/\sigma^4 + \mathrm{o}_p(1) \\ &= \lambda\xi^* + \mathrm{o}_p(1).\end{aligned}$$

When the distribution of ε is Gaussian then $\phi = 2\sigma^4$ so $\lambda = 1$, and the asymptotic size of the Breusch–Pagan statistic is correct.

The obvious 'solution' to this problem is to studentize the quadratic form suggested by Breusch and Pagan by dividing by $\hat{\phi} \equiv (1/n)\sum \hat{w}_i^2$ instead of by $2\hat{\sigma}^4$ as they suggest. Since $\hat{\phi} \rightarrow \phi$ in probability it is clear that the modified test statistic $\hat{\alpha}'Z'Z\hat{\alpha}/\hat{\phi} \rightarrow \chi_p^2(\eta)$. A referee has observed that this modified test statistic is simply nR^2, where R^2 is the uncentered coefficient of determination in the regression of $\hat{w}$ on Z. This 'solution' is not entirely satisfactory, as is noted in Koenker and Bassett (1981), since the *power* of the resulting test may be quite poor except under idealized Gaussian conditions. This is easily seen in the appearance of ϕ in the denominator of the noncentrality parameter of the limiting distributions. Nevertheless, the modified test is easy to compute and provides a simple means of checking for a rather general class of heteroscedastic phenomena.[1]

3. Conclusion

The point raised by this note might plausibly be considered common knowledge, except for the fact that papers continue to appear proposing tests for heteroscedasticity based on Gaussian assumptions. The seminal papers of Box (1953) and Tukey (1960) both note very graphically the dangers inherent in such an approach. More recently, Bickel (1978) and Ruppert and Carroll (1979) make similar points to those raised in the present note concerning

[1]Perhaps it should also be noted that the power of the test is also low when $h'(0)$ is small. In the case of contiguous sequences of alternatives this is evident from the form of α_0 in the statement of the theorem. However, even for fixed alternatives similar problems appear, e.g. let $\sigma_i = 1 + z_i^2$ and suppose the scalar z's are uniformly distributed on the interval $[-1, 1]$, then the test will have asymptotic power equal to its size.

tests for and models of heteroscedasticity originally suggested by Anscombe (1961).

References

Amemiya, T., 1977, A note on a heteroscedastic model, Journal of Econometrics 6, 365–370.

Anscombe, F.J., 1961, Examination of residuals, Proceedings of the 5th Berkeley Symposium 1, 1–36.

Bickel, P.J., 1978, Using residuals robustly I: Tests for heteroscedasticity, nonlinearity, Annals of Statistics 6, 266–291.

Box, G.E.P., 1953, Non-normality and tests on variance, Biometrika 40, 318–335.

Breusch, T.S. and A.R. Pagan, 1979, A simple test for heteroscedasticity and random coefficient variation, Econometrica 47, 1287–1294.

Goldfeld, S.M. and R. Quandt, 1972, Nonlinear methods in econometrics (North-Holland, Amsterdam).

Hajek, J. and Z. Sidak, 1967, Theory of rank tests (Academic Press, New York).

Koenker, R.W. and G.W. Bassett, 1980, Robust tests for heteroscedasticity based on regression quantiles, Econometrica, forthcoming.

Payen, J.-F., 1980, A simple test for homoscedasticity in the case of nonnormality, INSEE working paper no. 8007.

Ruppert, D. and R.J. Carroll, 1981, On robust tests for heteroscedasticity, Annals of Statistics 9, 206–210.

Tukey, J.W., 1960, A survey of sampling from contaminated distributions. in: I. Olkin, ed., Contributions to probability and statistics (Stanford University Press, Stanford, CA).