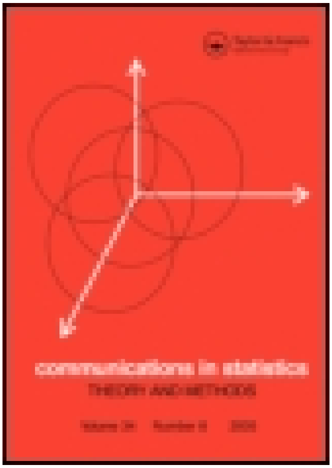

Communications in Statistics - Theory and Methods

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

Heteroscedasticity of residuals: a non-parametric alternative to the goldfeld-quandt peak test

Paul Horn [a]

[a] Department of Statistics , Princeton University , Princeton, New Jersey, 08544

Published online: 27 Jun 2007.

To cite this article: Paul Horn (1981) Heteroscedasticity of residuals: a non-parametric alternative to the goldfeld-quandt peak test, Communications in Statistics - Theory and Methods, 10:8, 795-808, DOI: 10.1080/03610928108828074

To link to this article: http://dx.doi.org/10.1080/03610928108828074

COMMUN. STATIST.-THEOR. METH., A10(8), 795-808 (1981)

HETEROSCEDASTICITY OF RESIDUALS: A NON-PARAMETRIC ALTERNATIVE TO THE GOLDFELD-QUANDT PEAK TEST

Paul Horn

Department of Statistics
Princeton University
Princeton, New Jersey 08544

Key Words and Phrases: regression; Theil residuals, least squares; homoscedasticity, trend test.

ABSTRACT

This paper discusses the problem of heteroscedasticity of residuals in a regression model. Tests that detect a monotonic relationship between mean and variance are presented. The theoretical shortcomings of these tests are discussed and the tests are compared by an empirical sampling computation of their powers.

1. INTRODUCTION

Consider the simple linear model $Y = X\beta + \varepsilon$ where Y is an n-dimensional random vector of observations, X is an $n \times m$ matrix of known elements, β is an m-dimensional vector of scalars, and ε is an n-dimensional random vector. It is assumed that the elements of ε have 0 expectation, the same variance, σ^2, and are uncorrelated. Then it is known that the residuals $\hat{\varepsilon}_i = y_i - X_i\hat{\beta}$, where $X_i = i^{th}$ row of X and $\hat{\beta}$ and is the least squares estimator of β, are in general correlated with unequal variances. Thus, tests that use the $\hat{\varepsilon}_i$'s to make inferences about the ε_i's should not be based on the assumption that the $\hat{\varepsilon}_i$'s are uncorrelated though the ε_i's are. Unfortunately, this assumption is needed to justify some of the tests to be discussed. Thus, the only way to assess the performance of such tests is by the computation of their sizes and powers by simulation.

2. THE TESTS OF GOLDFELD AND QUANDT

Goldfeld and Quandt (1965) present two tests for determining whether residuals from a least squares regression are homoscedastic. The first is a parametric F-test which assumes further that the ε_i's are normally distributed. The second is a non-parametric test that counts the number of records among the $\hat{\varepsilon}_i$'s and is appropriately called the "peak" test.

The assumptions in Goldfeld and Quandt's model are the same as above plus the additional assumption that if the variances increase with the means then they increase with one of X-variables. This variable, x_m, say, is called the deflator. Two models are considered. The first is $Y = X\beta + \varepsilon$ where ε has a diagonal constant covariance matrix. This model will be called the linear model. The second model is $Y = X\beta + \varepsilon$ where now the ε_i's are still uncorrelated but $\mathrm{var}(\varepsilon_i) = \sigma^2 x_{mi}^2$. The solution in the first case is least squares regression of Y on X, while the solution in the second is least squares regresion of $[Y/x_m]$ on $[X/x_m]$ where $[Y/x_m]$ is the vector Y divided, element-wise, by the appropriate element of the m^{th} column of X and $[X/x_m]$ is similar where each row of X is being divided. The model in the first case will be referred to as the linear model while that in the second case will be referred to as the ratio model.

The F-test consists of the following procedure:

(i) order the observations by the increasing values of the x_{mi};

(ii) delete k middle observations;

(iii) fit separate regressions to the first $(n-k)/2$ ($> m$) observations and the last $(n-k)/2$; and

(iv) define $R = S_2/S_1$, where S_1 and S_2 are the residual sums of squares due to the smaller and larger x_{mi}-values, respectively.

Under the null hypothesis of homoscedasticity, R has an F-distribution with $\{(n-k)/2\} - m$ and $\{(n-k)/2\} - m$ degrees of freedom since S_1 and S_2 have $(n-k)/2$ observations and m parameters to be estimated. Now, if the ratio model is the true one, but the linear model is fit then under the alternative R will tend to be large. A converse result holds if the reverse is true (see Goldfeld and Quandt, 1965).

Goldfeld and Quandt point out that the power of this test will depend on the deflator x_m. If the standard deviation of x_m is small relative to its mean then the power will be small. It is clear that even in the heteroscedastic situation, no test will reject with any regularity if the x_m's are tightly centered about a relatively large value.

The power also depends on k, the number of omitted observations. If $k = 0$ then the first $(n/2) - k$ observations are not very different from the second $(n/2) - k$ observations, and the test may not reject as often as it should. Yet if k is too large, then too many observations are discarded and there may not be enough degrees of freedom even though only extreme values (in the x_m-sense) of the observations are examined. So, it is unclear what is a good value of k. Monte Carlo results of Goldfeld and Quandt show that a k equal to about 25% of the observations does well for their simulation.

The second test, the non-parametric peak test, is based on the fact that if the ratio model is true but the linear model is fit, then the residuals will increase with the x_{mi}. The test proceeds as follows:

(i) examine the residuals $\hat{\varepsilon}_i$, where i corresponds to x_{mi}, and define a peak at observation j to occur when $|\hat{\varepsilon}_j| \geq |\hat{\varepsilon}_i|$ $\forall\ i < j$; and

(ii) reject the linear model as the true one if the number of peaks is large.

The power simulations of Goldfeld and Quandt show that the F-test does well with samples of 30 and 60. They also show that the peak test is a good competitor, attaining powers of between 2/3 and 3/4 those of the F-test when the errors are normally distributed.

3. ANOTHER NON-PARAMETRIC TEST

The type of heteroscedasticity considered in this paper is an upward trend in the error variances associated with a specific independent variable. We propose that a plausible non-parametric test statistic for testing heteroscedasticity is the trend statistic, D, as described by Lehmann (1967). The trend statistic D is defined as $\Sigma(R_i - i)^2$, where R_i is the rank of the i^{th} absolute residual (corresponding to x_{mi}, which has rank i since the x_{mi}'s are already ordered). It is clear that small values of D indicate an upward trend and a rejection of the null hypothesis.

The trend statistic is mathematically convenient since for large samples it is approximately normally distributed. It is also appealing since it is a function of Spearman's rank correlation coefficient, r_s, a well known non-parametric statistic;
$D = \left((n^3 - n)/6\right)(1 - r_s)$ where n is the sample size.

4. USE OF THEIL'S THEOREMS TO OBTAIN UNCORRELATED RESIDUALS

As was pointed out in the introduction, use of tests based on the least squares residuals should not depend on the assumption that they are uncorrelated. Unfortunately, the null distributions of the two non-parametric tests in the previous sections have been determined only for uncorrelated residuals. This condition does not hold since the least squares residuals are, in general, correlated.

In a series of theorems, Theil (1971) shows that $n - m$ *un-*correlated residuals can be gotten from n observations. These residuals are defined as

$$\tilde{\varepsilon}^* = \hat{\tilde{\varepsilon}}_1 - X_1X_0^{-1}\left[\sum_{h=1}^{H} \frac{d_h}{1+d_h} \tilde{q}_h\tilde{q}_h'\right]\hat{\tilde{\varepsilon}}_0$$

where

$$Y = \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} X_0 \\ X_1 \end{bmatrix}\beta + \begin{bmatrix} \tilde{\varepsilon}_0 \\ \tilde{\varepsilon}_1 \end{bmatrix} = \begin{bmatrix} X_0 \\ X_1 \end{bmatrix}\hat{\beta} + \begin{bmatrix} \hat{\tilde{\varepsilon}}_0 \\ \hat{\tilde{\varepsilon}}_1 \end{bmatrix} ,$$

X_0 is an $m \times m$ matrix of m rows of X, $\hat{\beta}$ is the least squares slope of Y on X, $\hat{\tilde{\varepsilon}}_0$, $\hat{\tilde{\varepsilon}}_1$ are least squares residuals of the observations corresponding to X_0 and X_1, respectively, the d_h's are the positive square roots of the eigenvalues (that are < 1) of $X_0(X'X)^{-1}X_0'$, H is the number of these, and $\tilde{q}_h$ is the associated right eigenvector of d_h^2, normalized to have unit length.

Hedayat, Raktoe and Talwar (1977) propose that in the special case of simple linear regression through the origin, the peak test can be used on Theil's T-residuals, since these residuals are uncorrelated. They show that if the error variances are ordered as $\sigma_i^2 < f(x)\sigma_j^2$ where $f(x)$ (which is larger than 1) is a function of the independent variables then the corresponding T-residuals have ordered variances: var $\varepsilon_i^* <$ var ε_j^* where ε_i^* (ε_j^*) is the T-residual corresponding to the i^{th} (j^{th}) observation, $x_i < x_j$, and $i, j \neq k$ where the k^{th} observation (usually the central observation) was used to generate the $n - 1$ T-residuals. Thus, they require more than a simple monotonic increase of variance. Also, if the implications of the authors' claims are examined, it is seen that statements concerning the T-residuals do not directly imply anything about the original error variances. However, there is an unbiasedness result in that for a subset of the alternative the power is greater than the size. Thus, a plausible association does exist.

Since regression through the origin is not used as often as simple linear regression, we have attempted to extend the authors' result to simple two-parameter regression. In this case two points $(x_{(1)},y_{(1)})$ and $(x_{(2)},y_{(2)})$, say, are removed to calculate $n - 2$ T-residuals. Even in the simplest case, where $x_{(2)} - \bar{x} = \bar{x} - x_{(1)}$,

$\bar{x} = (1/n)\Sigma X_i$, a condition on the independent variables is required in addition to a condition on the variances similar to that in the paper by Hedayat et al. Though results concerning the association between heteroscedasticity of errors and that of T-residuals appear weak in the one-parameter regression, they are much weaker in the two-parameter case.

5. EXPERIMENTAL SAMPLING RESULTS

It appears that whatever theory exists for this problem will not be of much use in assessing which of the proposed tests is preferable. The simulations to be discussed are similar to those given by Goldfeld and Quandt, and Hedayat, Raktoe and Talwar. Here the linear model is fit to data generated from the ratio model in order to obtain empirical estimates of the powers of five tests:

i) the F-test,
ii) the peak test based on least squares residuals,
iii) the peak test based on T-residuals,
iv) the trend test based on least squares residuals, and
v) the trend test based on T-residuals.

Following Hedayat et al., 31 independent x's were generated from the uniform distribution with mean = 30, 40, 50 and standard deviation = 10, 20, 25 (only the absolute values of the x's were used). For each x-sample 10,000 samples of 31 ε-values were generated from an N(0,1) distribution.

The first simulation (Table I) is regression through the origin. As expected, the F-test does reasonably well throughout. Also, Hedayat et al.'s claim seems to hold here: the peak test, when based on the T-residuals, has more power than when based on least squares residuals. However, the power of the trend test, when based on least squares residuals, appears only slightly less than when based on T-residuals. In none of these situations does the power based on T-residuals exceed that based on least squares (for the same test) by more than two (maximum) standard errors. In both cases, least squares and T-residuals, it appears that the trend test is more powerful than the peak test.

TABLE I

Empirical Power Based on 10,000 Simulations (α = .05)

	μ_x = 30			μ_x = 40			μ_x = 50		
	σ_x = 10,	20,	25	σ_x = 10,	20,	25	σ_x = 10,	20,	25
F-Test K=1	.6687	.9695	.9849	.4796	.9012	.9606	.3576	.7934	.8966
Peak Test on L.S. Resid.	.3463	.9093	.9183	.2326	.6736	.9036	.1691	.4648	.6645
Peak Test on T-Resid.	.3509	.9152	.9189	.2373	.6813	.9098	.1736	.4662	.6715
Trend Test on L.S. Resid.	.5926	.9883	.9950	.3997	.9098	.9839	.2993	.7465	.9055
Trend Test on T-Resid.	.5973	.9895	.9952	.4038	.9109	.9858	.2982	.7543	.9112

True model: $y_i = x_i(\beta + \varepsilon_i)$ Sample size = 31

Tested model: $y_i = x_i\beta + \varepsilon_i$ Standard error $\leq$ .005

In the second simulation (Table II), an unrestricted two-parameter linear model is generated. Since to compute the T-residuals we now need to remove two points, the sample size has been increased to 32. Again, the ratio model is the true one and the linear model is tested. As in the first simulation, the F-test does well in most of the situations. However, now the T-residuals are no longer superior to the least squares residuals. In every situation the powers of the peak test on T-residuals are less than those of the peak test on least squares residuals by more than two (maximum) standard errors. The same holds true for the trend test.

It seems that a result similar to Hedayat et al. for unrestricted linear regression, if it exists, is much weaker than in the restricted case. As was mentioned earlier, when we tried to extend the result to the unrestricted case we got a much weaker result, one that required restrictions on the independent variables. Hence, it is not surprising that the simulations show a decrease in the power of the non-parametric tests based on T-residuals. As in

TABLE II

Empirical Power Based on 10,000 Simulations (α = .05)

	$\mu_x = 30$			$\mu_x = 40$			$\mu_x = 50$		
	σ_x = 10,	20,	25	σ_x = 10,	20,	25	σ_x = 10,	20,	25
F-Test K=2	.6712	.9624	.9832	.4697	.8969	.9545	.3464	.7832	.8886
Peak Test on L.S. Resid.	.3504	.7919	.8403	.2340	.5956	.7662	.1776	.4310	.5987
Peak Test on T-Resid.	.2924	.7127	.7689	.2034	.5192	.6801	.1586	.3669	.5156
Trend Test on L.S. Resid.	.6181	.9660	.9826	.4295	.8880	.9553	.3188	.7520	.8880
Trend Test on T-Resid.	.5407	.8685	.9144	.3723	.7833	.8554	.2777	.6557	.7862

True model: $y_i = \beta_0 + X_i(\beta_1 + \varepsilon_i)$ Sample size = 32

Tested model: $y_i = \beta_0 + X_i\beta_1 + \varepsilon_i$ Standard error $\leq$ .005

the restricted regression case, the trend test shows itself to have greater power than the peak test using the same set of residuals.

To be sure that the empirical powers presented thus far are due to the sensitivity of the tests and not to the possibility that they have a propensity toward rejection, two more simulations were run. Table III presents the percentage of times that a correct hypothesis is rejected, i.e., the empirical size, in the restricted regression case. As one would expect, the F-test has empirical size approximately equal to the nominal .05. While both non-parametric tests have approximate size = .05 with T-residuals, they appear to be a bit low with the least squares residuals. However, in none of these cases does the empirical size deviate from the nominal level by more than two (maximum) standard errors.

The empirical sizes in the unrestricted regression case are presented in Table IV. As in the restricted regression case, the only large deviations from the nominal .05 occur with the non-parametric tests based on least squares residuals. Unlike the one-parameter model, in this model the empirical sizes are larger than

TABLE III

Empirical Size Based on 10,000 Simulations (α = .05)

	$\mu_x = 30$			$\mu_x = 40$			$\mu_x = 50$		
	σ_x = 10,	20,	25	σ_x = 10,	20,	25	σ_x = 10,	20,	25
F-Test K=1	.0493	.0512	.0497	.0517	.0541	.0481	.0487	.0499	.0473
Peak Test on L.S. Resid.	.0474	.0413	.0459	.0502	.0447	.0463	.0459	.0469	.0471
Peak Test on T-Resid.	.0504	.0486	.0505	.0498	.0497	.0492	.0502	.0526	.0494
Trend Test on L.S. Resid.	.0433	.0425	.0400	.0491	.0468	.0409	.0483	.0461	.0423
Trend Test on T-Resid.	.0514	.0509	.0499	.0483	.0533	.0461	.0520	.0502	.0499

Model: $y_i = X_i\beta + \varepsilon_i$ Standard error $\leq$ .005 for least squares cases
Sample size = 31 $\leq$ .003 other cases

TABLE IV

Empirical Size Based on 10,000 Simulations (α = .05)

	$\mu_x = 30$			$\mu_x = 40$			$\mu_x = 50$		
	σ_x = 10,	20,	25	σ_x = 10,	20,	25	σ_x = 10,	20,	25
F-Test K=2	.0509	.0502	.0487	.0536	.0516	.0507	.0525	.0502	.0482
Peak Test on L.S. Resid.	.0559	.0530	.0511	.0564	.0523	.0573	.0518	.0586	.0536
Peak Test on T-Resid.	.0515	.0501	.0493	.0474	.0490	.0479	.0511	.0513	.0500
Trend Test on L.S. Resid.	.0519	.0523	.0526	.0569	.0552	.0524	.0521	.0567	.0508
Trend Test on T-Resid.	.0477	.0485	.0514	.0533	.0518	.0497	.0503	.0518	.0461

Model: $y_i = \beta_0 + X_i\beta_1 + \varepsilon_i$ Standard error $\leq$ .005 for least squares cases
Sample size = 32 $\leq$ .003 other cases

the nominal level. However, as in the former model, the empirical sizes do not deviate from the nominal level by more than two (maximum) standard errors.

The empirical powers given here are reasonably consistent with those given by Goldfeld and Quandt. However, they are quite different from those presented in the paper of Hedayat, Raktoe and Talwar. In the latter paper, which deals only with regression through the origin, the power of the peak test T-residuals exceeds 50% in only 3 out of 9 cases with normally distributed errors. Even more curiously, under these same conditions the empirical power of the F-test is less than the nominal size in 7 out of 9 cases. In the other two cases the empirical power exceeds the nominal size by a mere .2%. It is because of these questionable figures that one finds the empirical results of Hedayat, Raktoe and Talwar difficult to believe.

6. NON-NORMAL ERRORS

To see how the tests performed with non-normal errors, we simulated the two-parameter regression situation twice, each time with a different type of non-normal error. The first type of error distribution was skewed, a chi-square with 4 degrees of freedom (adjusted to have zero mean). The second type of error distribution had heavy tails, a t with 2 degrees of freedom. Tables V and

TABLE V

Empirical Power Based on 10,000 Simulations (α = .05)
Errors Distributed as Chi-Square with 4 Degrees of Freedom (Adjusted to Have Zero Mean)

	μ_x = 30			μ_x = 40			μ_x = 50		
	σ_x = 10,	20,	25	σ_x = 10,	20,	25	σ_x = 10,	20,	25
F-Test K=2	.6250	.9255	.9624	.4796	.8294	.9195	.3963	.7303	.8351
Peak Test on L.S. Resid.	.2526	.7202	.7563	.1741	.4978	.6883	.1339	.3390	.4875
Peak Test on T-Resid.	.2359	.6396	.7002	.1600	.4292	.5945	.1226	.3051	.4261
Trend Test on L.S. Resid.	.6295	.9592	.9809	.4718	.8858	.9493	.3734	.7575	.8856
Trend Test on T-Resid.	.5753	.8827	.9182	.4338	.8065	.8731	.3465	.6805	.8066

True model: $y_i = \beta_0 + X_i(\beta_1 + \varepsilon_i)$ Sample size = 32
Tested model: $y_i = \beta_0 + X_i\beta_1 + \varepsilon_i$ Standard error $\leq$ .005

TABLE VI

Empirical Power Based on 10,000 Simulations (α = .05)
Errors Distributed as t with 2 Degrees of Freedom

	$\mu_x = 30$			$\mu_x = 40$			$\mu_x = 50$		
	$\sigma_x = 10$,	20,	25	$\sigma_x = 10$,	20,	25	$\sigma_x = 10$,	20,	25
F-Test K=2	.5825	.8604	.8953	.4908	.7500	.8325	.4415	.6598	.7454
Peak Test on L.S. Resid.	.1750	.5110	.5706	.1291	.3107	.4662	.1102	.2120	.3006
Peak Test on T-Resid.	.1527	.4545	.5360	.1152	.2672	.4026	.1004	.1893	.2649
Trend Test on L.S. Resid.	.4737	.8709	.9174	.3449	.7348	.8509	.2700	.5885	.7320
Trend Test on T-Resid.	.4232	.7529	.8172	.3120	.6319	.7268	.2586	.5139	.6259

True model: $y_i = \beta_0 + X_i(\beta_1 + \varepsilon_i)$ Sample size = 32
Tested model: $y_i = \beta_0 + X_i\beta_1 + \varepsilon_i$ Standard error $\leq$ .005

VI present the empirical powers for the two respective error distributions. As with normal errors, the empirical power of the F-test is quite respectable, as is that of the trend test for both sets of residuals. The power of the peak test using either set of residuals is well below that of the other tests. Again, as with the normal errors, the powers based on the T-residuals for either non-parametric test are below those of the least squares residuals.

The situation does not look so good for the trend and F-tests when we examine the empirical sizes, Tables VII and VIII. These tables show that the trend test is testing at about 9.5% on least squares residuals and about 10.5% on T-residuals. The F-test looks even worse; testing at about 12% in the chi-square situation, and about 23.5% in the t situation. The peak test looks good here, maintaining the nominal 5% in all situations.

7. CONCLUSION

The Monte Carlo results show that the trend test is more powerful than the peak test for moderate sample sizes (approxi-

TABLE VII

Empirical Size Based on 10,000 Simulations (α = .05)
Errors Distributed as Chi-Square with 4 Degrees of Freedom (Adjusted to Have Zero Mean)

	$\mu_x = 30$			$\mu_x = 40$			$\mu_x = 50$		
	σ_x = 10,	20,	25	σ_x = 10,	20,	25	σ_x = 10,	20,	25
F-Test K=2	.1146	.1159	.1204	.1200	.1153	.1174	.1199	.1265	.1193
Peak Test on L.S. Resid.	.0517	.0500	.0479	.0506	.0533	.0528	.0515	.0538	.0541
Peak Test on T-Resid.	.0527	.0492	.0470	.0470	.0496	.0485	.0458	.0514	.0550
Trend Test on L.S. Resid.	.0974	.0963	.0951	.0991	.0990	.1005	.0992	.1061	.1034
Trend Test on T-Resid.	.1080	.1036	.1088	.1097	.1047	.1088	.1091	.1099	.1096

Model: $y_i = \beta_0 + X_i\beta_1 + \varepsilon_i$ Sample size = 32
Standard error $\leq$ .005

TABLE VIII

Empirical Size Based on 10,000 Simulations (α = .05)
Errors Distributed as t with 2 Degrees of Freedom

	$\mu_x = 30$			$\mu_x = 40$			$\mu_x = 50$		
	σ_x = 10,	20,	25	σ_x = 10,	20,	25	σ_x = 10,	20,	25
F-Test K=2	.2422	.2363	.2356	.2357	.2340	.2361	.2437	.2340	.2301
Peak Test on L.S. Resid.	.0533	.0509	.0518	.0562	.0514	.0516	.0562	.0496	.0542
Peak Test on T-Resid.	.0498	.0503	.0539	.0495	.0526	.0491	.0515	.0473	.0488
Trend Test on L.S. Resid.	.0954	.0899	.0912	.0977	.0899	.0862	.0914	.0937	.0871
Trend Test on T-Resid.	.1070	.0980	.1083	.1028	.1012	.0980	.1032	.1037	.0957

Model: $y_i = \beta_0 + X_i\beta_1 + \varepsilon_i$ Sample size = 32
Standard error $\leq$ .005

mately 30). Given that the errors were normally distributed, the trend test shows itself to be a good competitor to the parametric F-test and even outperforms it in many situations.

In the usual unrestricted linear regression it is probably best to use the trend test with least squares residuals. Although the empirical sizes suggest that we may reject more often than we should in some cases, unless we are very conservative, the differences in power are strong enough to warrant the risk involved. Even in those rare instances when regression through the origin is called for, the little extra power gained by using the T-residuals may not be worth the computational effort.

If we believe we have non-normal errors, use of the trend test would be risky while use of the F-test would be foolhardy. The safe way out is to use the peak test on least squares residuals. Though, if the error distribution is heavy-tailed heteroscedasticity may be difficult to detect.

ACKNOWLEDGEMENTS

The author wishes to thank Professor Peter Bloomfield for his helpful suggestions and comments. Support for this research was provided in part by the U. S. Department of Energy under Contract No. EY-76-S-02-2310-A002, awarded to the Department of Statistics, Princeton University, Princeton, N. J.

BIBLIOGRAPHY

[1] Goldfeld, S. and Quandt, R. (1965). Some tests for homoscedasticity. *J. Amer. Statist. Assoc.*, *60*, 539.

[2] Hedayat, A., Raktoe, B., and Talwar, P. (1977). Examination and analysis of residuals: A test for detecting a monotonic relation between mean and variance in regression through origin. *Commun. Statist.-Theor. Meth.*, *A6* (6), 497.

[3] Lehmann, E. L. (1975). *Non-Parametrics: Statistical Methods Based on Ranks*. New York: Holden-Day.

[4] Theil, H. (1971). *Principles of Econometrics*. New York: John Wiley & Sons.

Received September, 1979, Revised June, 1980.

Recommended and Refereed Anonymously.