Biometrical Letters
Vol. 45(2008), No. 2, 43-55

Notes on the use of the Goldfeld-Quandt test for heteroscedasticity in environment research

Anna Budka, Danuta Kachlicka, Maria Kozłowska

Department of Mathematical and Statistical Methods, Poznan University of Life Sciences, Wojska Polskiego 28, 60-637 Poznań, Poland, e-mail: budka@up.poznan.pl

SUMMARY

Some methods of using the Goldfeld-Quandt test are described. The use of classical statistics (mean, standard deviation) or position statistics (median, average deviation, median absolute deviation) is proposed. The results can be used to test the hypothesis that the residuals from a linear regression are homoscedastic, when the experimental distribution of the considered variable is symmetrical or asymmetrical (right-sided skew or left-sided skew). Using Monte Carlo method, properties of these modifications of the Goldfeld-Quandt procedure are explored. A comparison of the effectiveness of the described methods for environment research is presented.

Key words: Goldfeld-Quandt test, heteroscedasticity, linear regression, position statistics

1. Introduction

The pure sciences have wide application to experimentation in the field of environmental engineering. Statistics and econometrics offer quantitative methods which are an important tool in the analysis of results of environment research. These include research on the content of the air, air quality being an important issue. High air quality means that the level of pollution of the air is low. Pollution may come from sources which are either natural or anthropogenic (caused by human activity). Natural pollution is produced mainly by forest fires or the decomposition of living organisms, while anthropogenic sources include transport, fuel burning, industrial processes and others. The quality composition of the air is variable, and what is worse there arise secondary pollutants, which are often more toxic and harmful than the primary ones. The course of these processes depends on many factors, including the spatial distribution of concentrations. Another important aspect is the “asphalt effect”. When we cover the soil surface with an impermeable layer of concrete or asphalt, the damp from the ground cannot pass to the air. The air becomes dry and susceptible to pollution.

Therefore a fundamental criterion to consider when buying a plot of building land is its position. Research into the effect on land prices of, among other things, the distance of the plot from the nearest highway and airport requires certain conditions to be checked, which we do using the Goldfeld-Quandt test. The object of this paper is to present certain modifications of the use of the Goldfeld-Quandt test for investigating the presence of heteroscedasticity.

2. Framework

We consider thelinear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \qquad (1)$$

where $\mathbf{y}$ is an n x 1 vector of observations on the dependent variable, $\mathbf{X}=[\mathbf{1}, \mathbf{x}_1, \ldots, \mathbf{x}_{k-1}]$ is an n x k matrix of observations on k-1 independent variables, $\boldsymbol{\beta}$ is a k x 1 vector of unknown parameters, and $\boldsymbol{\varepsilon}$ is an n x 1 random vector of errors, where $\mathbf{1}$ is the vector of ones. We assume that errors are normally distributed and the value of the error term corresponding to an observed value of the dependent variable is statistically independent of the value of the error term corresponding to any other observed value of the variable. The error is the distance between the observed value of the dependent variable and its expected value. This expectation is labelled E($\mathbf{y}$) and E($\mathbf{y}$) = $\mathbf{X}\boldsymbol{\beta}$, because we assume that E($\boldsymbol{\varepsilon}$) = $\mathbf{0}$, where $\mathbf{0}$ is the null vector. In order for the estimates of the parameters not to be biased (or best linear unbiased estimates) one more assumption should hold: the variance should be constant. The variance matrix is labelled D($\boldsymbol{\varepsilon}$) and D($\boldsymbol{\varepsilon}$) = $\sigma^2\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. If we observe departures from these assumptions then the estimators of the parameters are biased and consequently they are not best or other drawbacks may to be. Hence control of the model is a very important part of the regression analysis.

It is well known that in the presence of heteroscedasticity of error variances D($\boldsymbol{\varepsilon}$)=diag($\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$), the least squares method has two major drawbacks:

inefficient parameter estimates and biased variance estimates which make standard hypothesis tests inappropriate.

Known tests for heteroscedasticity based on error analysis include, for example, the most popular Goldfeld-Quandt test, the Breusch-Pagan test and the White test, described respectively by Goldfeld and Quandt (1965), Breusch and Pagan (1979), White (1980). The literature on testing for heteroscedasticity includes many more tests (see Dufour et al., 2004). The test for heteroscedasticity in regression models based on the Goldfeld–Quandt methodology defined by Carapeto and Holt (2003) also deserves attention.

Most of these tests are what Goldfeld and Quandt call non-constructive tests, in that they can be used to determine the presence or absence of heteroscedasticity, but reveal nothing about the form of the variance structure (Buse, 1984).

We are interested in testing for heteroscedasticity in situations with known deflator. Our null hypothesis is H_0: $\sigma_i^2 = \sigma^2$ for i=1,2,…,n, and the alternative hypothesis is H_1: ~H_0 (the symbol ~ denotes negation). The Goldfeld-Quandt test is proposed by, for example, Goldfeld and Quandt (1965), Buse (1984) or Maddala (2006) for verification of the null hypothesis.

The test proposed by Goldfeld-Quandt (1965) is carried out in four steps. In the first step we must sort the multivariate variable with respect to the choice of one independent variable, for example x_d, $d \in \{l: l=1,2,\dots,k-1\}$. This variable is a potential deflator. Next, in the second step, we can discard one or more central observations. In the third step, we fit separate regression analyses to each of two remaining sets of observations. In the last step, we form the statistic, which has F-distribution. The statistic is the quotient of the mean squares for error from the separate regressions.

Let $x_{d1} \leq x_{d2} \leq \dots \leq x_{dn}$, where the second subscript denotes the number of observations of the variable x_d. We obtain a set of observations of the multivariate variable in the new ordering. Now we consider model (1) in the form

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \mathbf{X}_3 \end{bmatrix} \cdot \boldsymbol{\beta} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \boldsymbol{\varepsilon}_3 \end{bmatrix}, \tag{2}$$

where for j=1, 2, 3, $\mathbf{y}_j$ and $\boldsymbol{\varepsilon}_j$ are n_j x 1 vectors, $\mathbf{X}_j$ is an n_j x k dimension matrix. Goldfeld and Quandt (1965) state that the vector $\mathbf{y}_2$ includes some central observations after they are sorted with respect to the chosen potential deflator. The choice of the n_2 x 1 dimension vector $\mathbf{y}_2$ is the subject of consideration in the next section. The choice of the vector $\mathbf{y}_2$ follows as a consequence of the choice of central observations of the variable x_d. Goldfeld and Quandt (1965) state that the dimension of vectors $\mathbf{y}_1$ and $\mathbf{y}_3$ is the same, but other authors, for example Thursby (1982), Dufour et al. (2004), take into account varying dimensions of these vectors.

Goldfeld and Quandt (1965) proposed using the test for heteroscedasticity after the removal of some number of central observations. They did not specify how many observations should be removed. They gave the relative frequency of cases in which the false hypothesis is rejected for samples of dimension n=30 and n=60 after omitting 0, 4, 8, 12 or 16 central observations, and they estimated the power of the test. They obtained the largest frequency for n=30 and n=60 after the omission, respectively, of 8 and 16 central observations (equal to 26.7%). Buse (1984) analysed the problem for n=20, 40, 80 removing 20% of central observations. The same dimension of the removed set of observations was used by Dufour et al. (2004) for n=50 and n=100. Maddala (2006) suggests the removal of central observations to increase the power of the test, but he does not answer the question of how many observations to remove.

3. Results

We must quote here two sentences formulated by Goldfeld and Quandt (1965): (i) “The power of this test will clearly depend upon the value of n_2, the number of omitted observations; for every large value of n_2 the power will be small but it is not obvious that the power increases monotonically as n_2 tends to 0”, (ii) “The power of the test will clearly depend on the nature of the sample of values for the variable which is the deflator. Thus, if the variance of x_d is small relative to the mean of x_d the power can be expected to be small and conversely”.

We believe that the number of omitted observations depends on the precision of the measurements carried out and on their distribution.

The distribution of the independent variable called the deflator may be defined by a cumulative probability function labelled $F(x_d)$. If for a value $x_{d(q)}$, where $q \in (0;1)$, the cumulative probability function $F(x_{d(q)})$ is equal to q, then $x_{d(q)}$ is called a quantile of order q. Lower quartile, median and upper quartile are measures of location and are labelled $x_{d(1/4)}$, $x_{d(1/2)}$ and $x_{d(3/4)}$ respectively. Let us recall that many authors suggest removing 20% of observations; these are values of the variable x_d within the interval $(x_{d(2/5)}; x_{d(3/5)})$.

When for a set of several measurements the mean has been calculated, the standard deviation can be calculated for the set too. The standard deviation tells us how repeatable the placement of the measure is, and how much this contributes to the uncertainty of the mean value. From these, the estimated standard deviation of the mean (the standard error) may be calculated. The standard deviation of the mean has also been called a standard uncertainty. The standard uncertainty is a margin whose size can be thought of as ‘plus or minus one standard error’. We propose using classical statistics or position statistics with the aim of defining the set of removed observations.

3.1. Method based on classical statistics

Proposition 1. When the variable x_d has symmetrical distribution, the mean is the central point of the distribution. We assume that standard error is the measure of deviation from the mean. From the set of observations of variable x_d, such that $x_{d1} \le x_{d2} \le \ldots \le x_{dn}$, we propose to remove all observations belonging to the interval defined by the mean of the variable plus or minus the standard error, that is

$$x_{di} \in \left(\hat{\mu} - \frac{\hat{\sigma}}{\sqrt{n}}; \hat{\mu} + \frac{\hat{\sigma}}{\sqrt{n}}\right), \tag{3}$$

where $\hat{\mu}$ and $\hat{\sigma}$ are estimates of the mean and standard deviation of the variable x_d respectively.

Let us notice that, for the variable $x_d \sim N(\mu;\sigma)$, the orders of the quantile $x_{d(q1)} = \mu - \sigma/\sqrt{n}$ and $x_{d(q2)} = \mu + \sigma/\sqrt{n}$ have the forms:

$$q_1 = F(\mu - \frac{\sigma}{\sqrt{n}}) = P(x_d < \mu - \frac{\sigma}{\sqrt{n}}) = P(\frac{x_d - \mu}{\sigma} < -\frac{1}{\sqrt{n}}) = \\ = F_{0;1}(-\frac{1}{\sqrt{n}}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-1/\sqrt{n}} e^{-\frac{1}{2}x_d^2} dx,$$

$$q_2 = F(\mu + \frac{\sigma}{\sqrt{n}}) = P(\frac{x_d - \mu}{\sigma} < \frac{1}{\sqrt{n}}) = \\ = F_{0;1}(\frac{1}{\sqrt{n}}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{1/\sqrt{n}} e^{-\frac{1}{2}x_d^2} dx$$

Hence, we have $q_2 - q_1 = 2F_{0;1}(1/\sqrt{n}) - 1$. In the case when n=16 the difference of the orders is equal to 1/5, that is $q_2 - q_1 = 2F_{0;1}(0.25) - 1 = 1/5$. Let us notice that for increasing sample size, we have

$$\lim_{n\to\infty}(q_2 - q_1) = \lim_{n\to\infty}(2F_{0;1}(\frac{1}{\sqrt{n}}) - 1) = 0.$$

When x_d has the normal distribution, the length of the interval (3) depends on the size of the sample. In the case we will remove for example 20% of observations when n=16 or 10% when n=81.

3.2. Methods based on position statistics

When the variable x_d has asymmetrical distribution then the median med(x_d) is the central point of the distribution. The median is in the middle of the sorted data ($x_{d1} \le x_{d2} \le \ldots \le x_{dn}$). In a similar way we can define the first quartile to be 1/4 of the way through the sorted data ($x_{d(1/4)}$), and the third quartile to be 3/4 of the way through the sorted data ($x_{d(3/4)}$).

For investigation of the variability of the median, a position statistic called median absolute deviation is used. The median absolute deviation (MAD) is defined as the median of the absolute value of the difference between a variable and its median, that is MAD(x_d) = med|x_d – med(x_d)|.

For normally distributed $x_d \sim N(\mu;\sigma)$, the MAD is given by:

$$MAD(x_d) = \sigma \cdot med|(x_d - \mu)/\sigma| = 0.67\sigma$$

Let us notice that μ–MAD(x_d)=med(x_d)–MAD(x_d)=$x_{d(1/4)}$, μ+MAD(x_d)= =med(x_d)+MAD(x_d)= $x_{d(3/4)}$ and F(μ+0.5·MAD(x_d)) − F(μ–0.5·MAD(x_d))=0.25 because

$$F(\mu + 0.67\sigma/2) - F(\mu - 0.67\sigma/2) = P(|\frac{x_d - \mu}{\sigma}| < 0.67/2) = \\ = 2F_{0;1}(0.67/2) - 1 = 0.25 .$$

We propose taking MAD(x_d) as the upper limitation of the length of the interval including omitted observations of the variable x_d.

We may investigate the symmetry of distribution of the variable x_d by calculating quartiles. When the median is equal to half of the inter-quartile range, that is med(x_d)=½ ($x_{d(3/4)}$ −$x_{d(1/4)}$), then the distribution is symmetrical. For right-sided skew distribution we have med(x_d)<½ ($x_{d(3/4)}$–$x_{d(1/4)}$), and for left-sided skew distribution med(x_d)>½ ($x_{d(3/4)}$ −$x_{d(1/4)}$). For variable x_d having asymmetrical distribution we propose some methods for defining the set of omitted observations.

Proposition 2. Adopting the asymptotic approach for building the confidence interval for the median applying standard error (see Dawid, 1981) we propose to define the set of removed observations by the interval "median plus or minus standard error". Hence we propose removing the following observations :

$$x_{di} \in (\text{med}(x_d) - \frac{\hat{\sigma}}{\sqrt{n}}; \text{med}(x_d) + \frac{\hat{\sigma}}{\sqrt{n}}) . \quad (4)$$

Proposition 3. Another measure of variation is mean absolute deviation. The mean absolute deviation (MeanAD) is defined as follows:

$$\text{MeanAD}(x_d) = \frac{1}{n}\sum_i | x_{di} - \text{med}(x_d) | . \quad (5)$$

The mean absolute deviation takes account of values from zero to half of the range of observations. Hence we believe that we may remove observations belonging to the interval "median plus or minus the mean absolute deviation divided by the square root of the number of observations n", namely

$$x_{di} \in (\text{med}(x_d) - \frac{\text{MeanAD}(x_d)}{\sqrt{n}}; \text{med}(x_d) + \frac{\text{MeanAD}(x_d)}{\sqrt{n}}) . \quad (6)$$

Proposition 4. In experiments there is no such thing as a perfect measurement. Each measurement contains a degree of uncertainty due to the limits of instruments and the people using them. The uncertainty index for sample size n has the form 1.859 MAD(x_d) /sqrt(n), where sqrt(n) denotes the square root of n. The range of omitted values of x_d can be defined as the median plus or minus the value of the uncertainty index. We propose removing the following central observations:

$$x_{di} \in \left(\text{med}(x_d) - \frac{1.859 \cdot \text{MAD}(x_d)}{\sqrt{n}};\text{med}(x_d) + \frac{1.859 \cdot \text{MAD}(x_d)}{\sqrt{n}}\right) \qquad (7)$$

Above we have given four propositions for defining the set of central values of observations of the independent variable x_d called a deflator. These propositions concern the use of measures of variation such as the standard error, the mean absolute deviation and the median absolute deviation. The range of removed observations is defined as an interval whose central point is the mean or median of the variable x_d. We will check the usefulness of the above propositions on generated data.

4. Monte Carlo simulation

In order to compare the effectiveness of the described modifications of the Goldfeld-Quandt procedure we performed a Monte Carlo study using SEPATH from Statistica 7.1. We considered sample size n=100 and we used the single regressor model as applied for example in Goldfeld and Quandt (1965), Griffiths and Surekha (1986), Carapeto and Holt (2003). However we performed our study in a different way than these authors. In the first step, values y_i and x_i for i=1,2,…,100, were generated from a normal distribution N(0;1), such that these values (y_i;x_i) had highly significant correlation. In the second step we sorted (y_i;x_i) with respect to x_i. In the third step we omitted some central observations. We used one of six methods of defining the set. In the next step we calculated the Goldfeld-Quandt statistics. Nine hundred replications were generated. In the last step we calculated the p-value for the Goldfeld-Quandt test and the cumulative frequency of cases in which the null hypothesis is not rejected. The power of the Goldfeld-Quandt test was calculated using Statistica 7.1 for first five hundred replications.

We concentrate now on comparing different methods for omitting central observations. Two standard methods are considered here; the first method was described by Goldfeld and Quandt (1965) and the second by, for example, Buse (1984). The first method involves use of the test without removing central observations, and in the second method twenty percent of central observations are omitted. These two methods are compared with our four propositions.

Table 1 displays the experimentally calculated mean p-value, cumulative frequency of cases in which the null hypothesis is not rejected on significance level 0.05 or 0.01 or 0.001 and mean power of the Goldfeld-Quandt test tabulated by six methods.

Table 1. Mean p-value, cumulative frequency of cases in which null hypothesis is not rejected (α=0.05, 0.01, 0.001), mean power of the Goldfield-Quandt test for Monte-Carlo experiments and number of removed observations (nro)

Set of omitted observations	p-value	frequency			power	Nro
		0.05	0.01	0.001		
Null	0.2518	0.891	0.978	0.997	0.1267	0.0
20% central observations	0.2488	0.904	0.983	0.998	0.1257	20.0
mean ± standard error	0.2486	0.900	0.981	0.998	0.1245	7.9
median ± standard error	0.2521	0.896	0.983	0.998	0.1233	8.8
median ± MeanAD/sqrt(n)	0.2540	0.898	0.982	0.998	0.1225	7.1
median ± 1.859·MAD/sqrt(n)	0.2511	0.897	0.984	0.998	0.1245	10.6

The magnitude of the differences in the p-values of our propositions compared with the two methods from earlier papers is small (<1%). We obtained the smallest p-value for the interval “mean plus or minus standard error”. The power given in Table 1 clearly indicates that the power of the Goldfeld-Quandt test depend upon the number of omitted observations but it is not true that power increases monotonically as this number tends to null.

This consideration shows that the standard method recommended by many authors (20% omitted central observations) gives the same result as our methods. When we make measurements, we have no way of knowing how accurate the values are. The use of variation measures such as standard error, median absolute deviation or standard uncertainty make it possible to avoid taking an overoptimistic view of reality. By performing 'truthful' analysis, we can refine the Goldfeld-Quandt procedure. We uncover and take advantage the major sources of the uncertainty. In doing so, we use information about measurements.

5. Research problem

Regression models in environment research, and particularly the estimated standard errors, rely upon the assumption that the residuals are independently and identically distributed. Two common violations of this assumption are serial correlation and heteroscedasticity. We assume that serial correlation is not a significant issue. However heteroscedasticity, which refers to non-constant error variance, is a common problem in environment regression models.

The present analysis is based on the results of research described in example 4.7 by Maddala (2006). In this example six variables were described. The price of land per acre is here called the dependent variable labelled y, the percentage afforested area is labelled x_1, the distance of the land parcel from the airport is labelled x_2, the distance from the highway is labelled x_3, the area of the land parcel is labelled x_4 and the month in which it was sold is labelled x_5. The dependent variable y and two independent variables x_2 and x_3 were transformed and denoted respectively lny, $\ln x_2$ and $\ln x_3$. The independence variable $\ln x_3$ is called the deflator. Hence we can write $x_d = \ln x_3$. Table 2 displays classical and position statistics tabulated by the six variables.

The analysis began with a study of homoscedasticity for each independent variable separately and the dependent variable (Table 3). We performed the Goldfeld-Quandt test using six methods of defining the set of omitted observations, separately. The same analyses were made for two groups of independent variables, in which we sorted these data with respect to $x_d = \ln x_3$. We performed the Goldfeld-Quandt test for two groups to compare their results. We calculated the median absolute deviation of the deflator $MAD(x_d) = MAD(\ln x_3) = 0.397$ and inter-quantile range $(x_{d(3/5)} - x_{d(2/5)}) = 0.2809$. Table 3 examines the p-value for these analyses and the number of omitted observations of the deflator $\ln x_3$ for two standard methods and our four propositions for defining the set of omitted observations. For the last two analyses we also calculated the power of the test.

Table 2. Classical statistics and position statistics for the variables considered in example 4.7 by Maddala (2006)

	lny	x_1	$\ln x_2$	$\ln x_3$	x_4	x_5
Mean	8.393	0.317	2.657	1.431	61.93	23.8
Standard deviation	0.523	0.452	0.510	0.827	116.78	11.5
Standard error	0.064	0.055	0.062	0.101	14.27	1.4
Median	8.466	0.000	2.741	1.411	20.00	24.0
Minimum observation	6.770	0.000	1.253	-2.303	3.50	2.0
Maksimum observation	9.700	1.000	3.757	3.418	656.00	46.0

Before we consider the influence of a defined of set of omitted observations on the p-value and power of the Goldfeld-Quandt test, it is necessary to make a short digression on the relation between the magnitude of the sets considered. In the example the greatest number of observations belongs to the second standard set ($20\%\cdot n=13$), but the smallest p-value and the highest power is for the propositions “mean plus or minus standard deviation” and “median plus or minus standard deviation”. In the respective sets we have 9 observations in each. For the next two propositions, based on mean absolute deviation or median absolute deviation, we obtained worse results.

6. Conclusion

To conclude our considerations we offer four methods for building the set of omitted observations. On the basis of the Monte Carlo simulation we generally

prefer first proposition. Proposition 1 is preferred when the deflator has symmetrical distribution, proposition two when the distribution is asymmetrical.

For the considered research problem the results of each application of the Goldfeld-Quandt test are different. In many cases our propositions are better than the second standard method considered here. Methods based on measurement of the deviation of the deflator are better than the removal of 20% central observations without regard to their distribution.

We should reiterate that the use of variation measures such as standard error, median absolute deviation or standard uncertainty makes it possible to avoid taking an overoptimistic view of reality.

Table 3. The p-value for the Goldfeld-Quandt test for simple linear regression and two multivariate linear regression with $x_d=\ln x_3$ (nro – number of removed observations of variable $\ln x_3$)

Set of omitted observations	$\ln x_2$	$\ln x_3$	x_4	x_5	nro	x_1; $\ln x_2$; $\ln x_3$; x_4; x_5		$\ln x_2$; $\ln x_3$	
	p-value	p-value	p-value	p-value		p-value	power	p-value	power
null	0.0405	0.0049	0.3406	0.1952	0	0.0019	0.8357	0.0011	0.8770
20% central observations	0.0196	0.0013	0.1984	0.3279	13	0.0025	0.8159	0.0016	0.8543
mean ± standard error	0.0578	0.0005	0.2847	0.4206	9	0.0016	0.8485	0.0006	0.9062
median ± standard error	0.1608	0.0005	0.2639	0.4206	9	0.0016	0.8485	0.0006	0.9062
median ± MeanAD/sqrt(n)	0.0147	0.0042	0.2474	0.4206	6	0.0053	0.7273	0.0010	0.8768
median ± 1.859·MAD/sqrt(n)	0.0147	0.0038	0.2636	0.3980	7	0.0042	0.7523	0.0010	0.8768

REFERENCE

Breusch T., Pagan A. (1979): A Simple Test for Heteroscedasticity and Random Coefficient Variation. Econometrica 47: 1287-1294.

Buse A. (1984): Tests for additive heteroskedasticity Goldfeld and Quandt revisited. Empirical Economics 9: 199-216.

Carapeto M., Holt W. (2003): Testing for heteroscedasticity in regression models. Journal of Applied Statistics 30(1): 13-20.

Dawid H.A. (1981): Order Statistics. New York, Wiley.

Dufour J.-M., Khalaf L., Bernard J.-T., Genest I. (2004): Simulation-based finite-sample tests for heteroskedasticity and ARCH effects. Journal of Econometrics 122: 317-347.

Griffiths W.E., Surekha K. (1986): A Monte Carlo evaluation of the power of some tests for heteroscedasticity. Journal of Econometrics 31: 219-231.

Goldfeld S.M., Quandt R.E. (1965): Some tests for homoscedasticity. American Statistical Association Journal 60: 539-547.

Maddala G.S. (2006): Ekonometria. PWN, Warszawa

Thursby G. (1982): Misspecification, heteroscedasticity, and the chow and Goldfeld-Quandt tests. The Review of Economics and Statistics 64(2): 314-321.

White H. (1980): A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity. Econometrica 48: 817-838.