Statistics

1981

UCD Department of Agricultural Economics

WORKING PAPER SERIES

University of California, Davis
Department of Agricultural Economics

ESTIMATION OF HETEROSCEDASTIC REGRESSION MODELS WHOSE VARIANCES ARE FUNCTIONS OF EXOGENOUS VARIABLES

by

John M. Antle

Working Paper No. 81-5

Estimation of Heteroscedastic Regression Models Whose Variances are Functions of Exogenous Variables

Heteroscedasticity of a regression model's disturbances is a standard estimation problem. The assumption that the regression disturbances' variances are functions of exogenous variables is often a feasible maintained hypothesis in applied research, and numerous techniques have been proposed for estimation of such models (see [8], Ch. 4, for a survey of recent research). In this paper I propose a general heteroscedastic regression model for which one can obtain consistent estimates of the error variances in the case where the variances are linear functions of exogenous variables. The proposed estimation technique is a generalization of Goldfeld and Quandt's [6] "Modified Glejser Method" and of Amemiya's [1] generalized least squares method, and is based on the principle that, not only the mean and variance but, generally, all of the moments of the dependent variable may be functions of exogenous variables. Using this principle, a general model is proposed for which one can obtain single equation estimates of the parameters of both the regression model and the variance function which are asymptotically equivalent to Aitken estimators, and this is achieved without making explicit assumptions about the distribution of the regression model's disturbance term (such as normality). Since the Aitken estimator can be shown to be asymptotically normal under standard conditions, this method provides an asymptotic test for the hypothesis of heteroscedasticity without imposing an explicit distribution on the regression model. In addition, I show that if the third moment of the regression distrubance is nonzero (and also possibly a function of exogenous variables) one can employ a joint Aitken procedure, which is a heteroscedastic version of Zellner's [13] "seemingly unrelated regression technique," to increase estimation efficiency. Asymptotically valid tests for symmetry of the disturbance distribution are also devised. One problem with the proposed estimation method is that negative estimates of variances may be obtained. Using standard programming techniques, I also show how the model may be estimated under the restriction that the estimated variances are non-negative.

Several other model specifications and estimation methods have been proposed for heteroscedastic regression models which are similar to the one described in this paper (Glejser [4], Harvey [7]). While these models exhibit some desirable properties, I show that they require explicit distributional assumptions for the regression error term. Therefore, the model described in this paper is more general in the important sense that one need not assume the error term follows a particular distribution to obtain estimates of the model's parameters and standard errors with desirable asymmptotic properties.

The paper begins with a description of the heterscedastic model and the single equation Aitken estimators. The following sections discuss the joint Aitken estimator, the estimation procedure with the variances constrained to be positive, and the alternative estimation methods.

1. The Model and Single Equation Estimation Method.

The linear regression model is given by

$$y_t = x_t\beta + u_t,\ E(u_t) = 0,\ E(u_t u_{t'}) = 0 \text{ for } t \neq t' \qquad (1)$$

The (1xk) vector $x_t = (1, x_{t1}, \ldots, x_{tk})$. In addition let

$$u_t^2 = z_t\gamma_2 + v_t, \quad E(v_t) = 0 \qquad (2)$$

where z_t is a vector of variables exogenous to u_t.

Note that the $(1\mathrm{x}\ell)$ vector $z_t = (1, z_{t1}, \ldots, z_{t\ell})$ could include x_t as a subvector. Now defining

$$E(u_t^i) = \mu_{it},$$

we have

$$\begin{aligned} E(u_t^2) &= \mu_{2t} = z_t\gamma_2, \\ E(v_t^2) &= \mu_{4t} - \mu_{2t}^2, \qquad (3) \\ E(v_t\, v_{t'}) &= 0, \quad t \neq t'. \end{aligned}$$

Therefore the regression model (1) has a heterscedastic error structure, and consistent estimates of the regression variances μ_{2t} can be obtained if the parameter vector γ_2 can be consistently estimated. I will now show that under standard assumptions on the x_t and z_t one can obtain consistent estimates of γ_2 and, thus, implement the Aitken procedure for estimation of β. Let X be the $(\mathrm{T}\mathrm{x}\mathrm{k})$ matrix of the x_t, let Z be the $(\mathrm{T}\mathrm{x}\ell)$ matrix of the z_t, and let u, u^2 and v be the (Tx1) vectors of the u_t, u_t^2 and v_t.

<u>Assumptions</u>. <u>As</u> $T \to \infty$, $X'X/T$ <u>and</u> $Z'Z/T$ <u>converge</u> <u>to</u> <u>nonsingular</u> <u>matrices</u> <u>and</u> $\operatorname{plim} X'u/T = \operatorname{plim} Z'v/T = 0$.

Under these assumptions it follows that least squares estimation produces a consistent estimate $\hat{\beta}$ of β. Thus, we have the residuals

$$\hat{u}_t = y_t - \hat{y}_t = x_t\beta + u_t - x_t\hat{\beta} = u_t + x_t\,(\beta-\hat{\beta}). \qquad (4)$$

Using (4) we have

$$\hat{u}_t^2 = u_t^2 + [x_t(\beta-\hat{\beta})]^2 + u_t x_t(\beta-\hat{\beta}).$$

In general, $E(\hat{u}_t^2) \neq \mu_t$, but using a well-known limit theorem (Rao [11], p. 122) it follows that the bias is zero in the limit because $\operatorname{plim}\hat{\beta} = \beta$:

$$\operatorname{plim} \hat{u}_t^2 = [u_t + z_t\,(\beta - \operatorname{plim}\hat{\beta})]^2 = u_t^2. \qquad (5)$$

We now consider the following theorems:

Theorem 1: Let the least squares estimator of γ_2 in equation (2) be given by

$$\hat{\gamma}_2 = (Z'Z)^{-1} Z' u^2.$$

Also define the feasible least squares estimator of γ_2 as

$$\tilde{\gamma}_2 = (Z'Z)^{-1} Z' \hat{u}^2.$$

Then $\operatorname{plim} \tilde{\gamma}_2 = \operatorname{plim} \hat{\gamma}_2 = \gamma$ and $\tilde{\gamma}_2$ converges in distribution to $\hat{\gamma}_2$.

Proof: From (5) it follows that

$$\operatorname{plim} \tilde{\gamma}_2 = \operatorname{plim} \hat{\gamma}_2$$

and under the assumptions made above it also follows that $\operatorname{plim} \hat{\gamma}_2 = \gamma_2$. Since $\tilde{\gamma}_2$ converges to $\hat{\gamma}_2$ in probability, $\tilde{\gamma}_2$ also converges to $\hat{\gamma}_2$ in distribution (Rao [11], p. 122). Q.E.D.

Theorem 2: Let Σ be the diagonal matrix of the μ_{2t} and let $\hat{\Sigma}$ be the diagonal matrix of the $z_t\hat{\gamma}_2$, t=1, . . ., T. Also, assume $\lim_{T\to\infty} Z'\Sigma^{-1} Z/T$ is finite and nonsingular.

Then the feasible Aitken estimator

$$\hat{\beta}^A = (X'\hat{\Sigma}^{-1}X)^{-1} X'\hat{\Sigma}^{-1} y$$

converges in distribution to the Aitken estimator

$$\beta^A = (X'\Sigma^{-1}X)^{-1} X'\Sigma^{-1} y.$$

Proof: It is sufficient to show that

$$\operatorname{plim} \frac{X'\hat{\Sigma}^{-1}X}{T} = \lim \frac{X'\Sigma^{-1}X}{T} \tag{6}$$

$$\operatorname{plim} \frac{X'\hat{\Sigma}^{-1}u}{\sqrt{T}} = \operatorname{plim} \frac{X'\Sigma^{-1}u}{\sqrt{T}} \tag{7}$$

By Theorem 1 and (3) we have

$$\text{plim } z_t \tilde{\gamma}_2 = z_t \gamma_2 = \mu_{2t}, \tag{8}$$

and hence (6) is satisfied for each element of the matrix.

Equation (7) follows from (8) and the fact that each element of $\hat{\Sigma}^{-1}u$ therefore converges in distribution to the corresponding element of $\Sigma^{-1}u$ (Rao [11], p. 122). Hence, $\text{plim} \sqrt{T}(\hat{\beta}^A - \beta^A) = 0$ and therefore $\hat{\beta}^A$ converges in distribution to β^A. <u>Q.E.D.</u>

It is also possible to obtain Aitken estimates of the variance regressions by letting

$$E(u_t^4) = \mu_{4t} = z_t \gamma_4$$

$$u_t^4 = z_t \gamma_4 + w_t, \quad E(w_t) = 0.$$

By the above arguments it is clear that regression of $\hat{u}_t^4$ on z_t will produce consistent estimates of γ_4, that is, $\text{plim } \tilde{\gamma}_4 = \gamma_4$. Therefore, we can also prove the following result:

<u>Theorem 3</u>: <u>Let Ω be the diagonal matrix of</u> $[\mu_{4t} - \mu_{2t}^2] = E(v_t^2)$, <u>and let $\hat{\Omega}$ be the diagonal matrix of</u> $[z_t\hat{\gamma}_4 - (z_t\hat{\gamma}_2)^2]$, t=1, . . ., T. <u>Also assume $\lim_{T\to\infty} \frac{Z'\Omega Z}{T}$ is a finite nonsingular matrix. Then the feasible Aitken estimator</u>

$$\hat{\gamma}_2^A = (Z' \hat{\Omega}^{-1} Z)^{-1} Z' \hat{\Omega}^{-1} \hat{u}^2$$

<u>converges in distribution to the Aitken estimator</u>

$$\gamma_2^A = (Z' \Omega^{-1} Z)^{-1} Z' \Omega^{-1} u^2.$$

<u>Proof</u>: First note that

$$\text{plim } Z'\hat{\Omega}^{-1}Z/T = \lim Z' \Omega^{-1}Z/T$$

because the elements of $\hat{\Omega}$ converge in probability to the elements of Ω. In addition:

$$\text{plim } \frac{Z'\hat{\Omega}^{-1}\hat{u}^2}{\sqrt{T}} = \text{plim } \frac{Z'\hat{\Omega}^{-1}u^2}{\sqrt{T}}$$

by the fact that the elements of $\hat{\Omega}$ converge in probability to the elements of Ω, and by the fact that $\hat{u}^2$ converges in probability to u^2 (see 5). Therefore, each element of $\hat{\Omega}^{-1}\hat{u}^2$ converges in distribution to the corresponding element of $\Omega^{-1} u^2$ (Rao [11], p. 122). Hence, $\text{plim} \sqrt{T}\, (\hat{\gamma}_2^A - \gamma_2^A) = 0$. <u>Q.E.D.</u>

The reader should note that μ_{4t} need not be a function of z_t for the above results to hold, since any consistent estimate such as $\hat{\mu}_4 = \sum_{t=1}^{T} \hat{u}_t^4/T$ can be used in the event that μ_4 is a constant across all observations.

We have now established the asymmptotic equivlance of the feasible Aitken estimators $\hat{\beta}^A$ and $\hat{\gamma}_2^A$ to their respective true Aitken estimators β^A and γ_2^A. The standard proof of asymmptotic normality of the Aitken estimator can be applied to β^A and γ_2^A (Theil [12], Ch. 8) since the weighted distrubances $u_t/\sqrt{\mu_{2t}}$ and $u_t^2/(\mu_{4t} - \mu_{2t}^2)^{1/2}$ are independently distributed with zero mean and unit variance for all t.[1] Therefore, a large sample estimate $\hat{\gamma}_2^A$ is approximately normal, and one can employ standard test procedures to test the hypothesis of heteroscedasticity (some $\gamma_{2\ell} \neq 0$ for some $\ell > 2$) against the hypothesis of homoscedasticity ($\gamma_{2\ell} = 0$ for $\ell > 2$).

A large sample estimation algorithm for β and γ_2, based on the above Theorems, can be defined as follows: First, run the regression

$$y_t = x_t\beta + u_t \tag{9}$$

to obtain a consistent estimate $\hat{\beta}$ of β. Second, use $\hat{\beta}$ to compute $\hat{u}_t$ and run the regressions

$$\hat{u}_t^2 = z_t\gamma_2 + v_t \tag{10}$$

$$\hat{u}_t^4 = z_t\gamma_4 + w_t$$

to obtain consistent estimates $\hat{\gamma}_2$ and $\hat{\gamma}_4$ of γ_2 and γ_4. Third, use the $\hat{\gamma}_2$ and $\hat{\gamma}_4$ to estimate the variances of (9) and (10) and compute the feasible Aitken estimators for these equations. These latter regressions can be accomplished by weighted least squares with weights given by

$$(z_t\hat{\gamma}_2)^{-1/2}, \quad [z_t\hat{\gamma}_4 - (z_t\hat{\gamma}_2)^2]^{-1/2}$$

respectively for (9) and (10).

In concluding this section, we emphasize that estimators for β and γ_2 which are asymmptotically equivalent to Aitken estimators have been devised without imposing any explicit distributional assumptions on the regression model. Only the condition of independence of the u_t across observations (see equation 1) and the standard assumptions concerning asymmptotic behavior of the data matrices are required for these results.

2. The Joint Aitken Estimator

I now show that the disturbances u_t and v_t of the regression model (1) and the "variance equation" (3) may be contemporaneously correlated. Therefore, we can increase estimation efficiency by taking these cross-equation correlations into account. Since the disturbances are heteroscedastic both within and across equations, the estimator is similar to Zellner's "seemingly unrelated regression" estimator with the addition of heteroscedasticity.

The cross-equation correlations are computed in a straightforward fashion, as follows:

$$E(u_t\ v_t') = E(u_t u_t^{2\prime} - u_t z_t'\gamma_2) = \mu_{3t} \quad , \quad t = t'$$
$$= 0 \quad , \quad t \neq t'.$$

Note that a consistent estimate of μ_{3t} can be obtained by regressing $\hat{u}_t^3$ on z_t to obtain the estimate $\hat{\gamma}_3$. Now we let Γ be the diagonal matrix of the $z_t\gamma_3$ and we let $\hat{\Gamma}$ be the diagonal matrix of the $z_t\hat{\gamma}_3$. Also let:

$$R = \begin{bmatrix} X & 0 \\ 0 & Z \end{bmatrix}, \quad q = \begin{bmatrix} y \\ u^2 \end{bmatrix}, \quad \hat{q} = \begin{bmatrix} y \\ \hat{u}^2 \end{bmatrix}, \quad \delta = \begin{bmatrix} \beta \\ \gamma_2 \end{bmatrix}$$

$$\Theta = \begin{bmatrix} \Sigma & \Gamma \\ \Gamma & \Omega \end{bmatrix}, \quad \hat{\Theta} = \begin{bmatrix} \hat{\Sigma} & \hat{\Gamma} \\ \hat{\Gamma} & \hat{\Omega} \end{bmatrix}.$$

The following Theorem is a straightforward generalization of Theorems 2 and 3:

<u>Theorem 4</u>: <u>Let</u> $\Theta^{-1} = [\Theta^{ij}]$, $i,j = 1,2$, <u>and assume</u> $\lim R'\Theta^{ij}R'/T$ <u>is finite and nonsingular. Then the feasible joint Aitken estimator</u>

$$\hat{\delta}^A = (R'\hat{\Theta}^{-1}R)^{-1}R'\hat{\Theta}^{-1}\hat{q}$$

<u>converges in distribution to the estimator</u>

$$\delta^A = (R'\Theta^{-1}R)^{-1}R'\Theta^{-1}q.$$

The joint estimation procedure is, therefore, an extension of the single equation approach, and requires the estimation of the parameter vector γ_3 of the third moment in addition to γ_2 and γ_4. The reader may note that if the distribution of the original dependent variable, y, is symmetric, then $\gamma_3 = 0$ and the single equation Aitken estimators are as efficient as the joint estimator. Consequently, it is of interest to test the hypothesis $\gamma_3 = 0$ against $\gamma_3 \neq 0$. To do so we may formulate the models

$$u_t^3 = z_t\gamma_3 + s_t = \mu_{3t} + s_t \ , \ E(s_t) = 0$$

$$u_t^6 = z_t\gamma_6 + r_t = \mu_{6t} + r_t \ , \ E(r_t) = 0$$

Noting that

$$E(u_t^3 - z_t\gamma_3)^2 = \mu_{6t} - \mu_{3t}^2 = z_t\gamma_6 - (z_t\gamma_3)^2$$

we can obtain a feasible Aitken estimate of γ_3 by regressing $\hat{u}_t^3$ and $\hat{u}_t^6$ on z_t to obtain consistent estimates of γ_3 and γ_6 and then running a weighted least squares regression with weights $[z_t\hat{\gamma}_6 - (z_t\hat{\gamma}_3)^2]^{-1/2}$. Since the resulting estimate $\hat{\gamma}_3^A$ of γ_3 is asymmptotically normal one can then test the hypothesis $\gamma_3 = 0$ using standard test statistics.[2]

3. Constraining Estimates of Even Moments to be Non-negative

Although even moments are non-negative by definition, in utilizing the estimation techniques described in the previous two sections it is possible that some estimates of even moments may be negative. If the model is correctly specified, the problem may be due to either small sample bias or sampling error in parameter estimates (recall from section 1, least squares estimates of the γ_i are biased but consistent). Therefore, it may be necessary to constrain the estimates to be positive in finite samples; in the limit as T approaches infinity consistency holds and the problem disappears. Since the non-negativity restriction holds with certainty, it can be demonstrated that the restricted estimator is superior to the unrestricted estimator in terms of mean-squared error. Hence, use of a restricted estimator can improve small sample properties of even-moment estimates and will not affect their asymmptotic properties.

Restricted Aitken estimators can be implemented as follows. First, estimate γ_2 by solving the nonlinear minimization problem:

$$\min_{\gamma_2} \sum_{t=1}^{T} [\hat{u}_t^2 - z_t\gamma_2]^2 \text{ subject to } z_t\gamma_2 > 0.$$

Second, estimate γ_4 by solving

$$\min_{\gamma_4} \sum_{t=1}^{T} [\hat{u}_t^4 - z_t\gamma_4]^2 \text{ subject to } [z_t\gamma_4 - (z_t\hat{\gamma}_2)^2] > 0.$$

Note that the inequality constraint in this latter problem simultaneously forces $\hat{\gamma}^4$ to satisfy the restriction that $\mu_4 > 0$ as well as the restriction that the variance of u_t^2 be non-negative. These inequality constrained minimization problems can be solved with appropriate software (see [10] for an example), and the resulting estimates can then be employed in the weighted regressions to obtain Aitken estimates. Similar procedures could be implemented for the even moment regressions discussed in section 2.

4. Choice Among Alternative Estimation Methods

Two other estimation methods similar to the one described in section 1 have been proposed by Glejser [4] and Harvey [7]. The existence of several methods poses the problem to the researcher of choosing among them. In this section I show that the heteroscedastic structure they both use requires the assumption that the regression disturbance follows a particular probability distribution such as the normal. The method proposed in section 1 above is, therefore, more general because it is valid for whatever distribution u_t may follow. However, the Glejser and Harvey methods do always provide positive estimates of the regression variances, whereas the model in section 1 may produce negative estimates and, thus, require the use of the constrained estimation method described in section 3. We may conclude, therefore, that the Glejser or Harvey methods are attractive if the researcher is willing to impose a specific distribution, such as the normal, on the regression model. However, if the researcher is not willing to impose a normal distribution on the regression model, or if the researcher would like to test for asymmetry of the distribution, the method described in section 1 is preferable.

Both the Glejser and Harvey methods are based on the "multiplicative heteroscedasticity" structure of the general form

$$u_t = g(z_t)\varepsilon_t, \quad E(\varepsilon_t) = 0 \tag{11}$$

where ε_t is an independently distributed random variable. Glejser assumes g is linear

$$g(z_t) = z_t\gamma_2$$

whereas Harvey assumes g is log-linear:

$$g(z_t) = \exp\ \{z_t\gamma_2\}$$

The Glejser model implies μ_{2t} is proportional to $(z_t r_2)^2$ and the Harvey model implies μ_{2t} is proportional to $\exp\ \{2z_t\ \gamma_2\}$. But in general it also follows from (11) that the ith moment is

$$\mu_{it} = E(u_t^i) = g(z_t)^i\ E(\varepsilon_t^i) \tag{12}$$

The "multiplicative heteroscedasticity" model (11) therefore implies that the exogenous variables z_t affect all nonzero moments of u_t through the powers of $g(z_t)$. Clearly, this condition holds only under special circumstances; in particular, when the u_t are normal variates it can be shown that

$$\mu_{2r,t} = \frac{\mu_{2t}^r\ (2r!)}{2^r\ r!} = \frac{g(z_t)2r\ (2r!)}{2^r\ r!}$$

so that (12) is satisfied. However, the moments of u_t are not necessarily related to z_t in this manner, and in general need not be functions of z_t at all. Therefore, as we asserted above, the Glejser and Harvey models are valid only if u_t follows a distribution, such as the normal, which satisfies (12).

5. Conclusions

In this paper I have demonstrated that under the hypothesis that a regression model's moments are linear functions of exogenous variables, estimators of the parameters of the regression model and other moment functions can be obtained which are asymmptotically equivalent to Aitken estimators. These results require only the assumptions of independent regression distrubances and the usual convergence conditions for the data matrices; therefore, this model is more general than other models which require specific distributional assumptions. Under the conditions of the Lindberg-Feller central limit theorem the Aitken estimators are asymmptotically normal and, thus, provide a means of testing for heteroscedasticity and skewness of the regression distribution. When the distribution has a nonzero third moment, I have shown that a joint Aitken estimation procedure is possible. To overcome the problem of negative estimates of even moments an inequality-constrained estimation procedure was outlined.

The model discussed in this paper is based on the assumption that moments are linear functions of exogenous variables. It would appear that the results obtained here could be extended to the case of nonlinear functions by utilizing the nonlinear consistency results of Malinvaud [7]. This generalization could prove to be a useful extension of the methods discussed here.

jw 8/24/81 P11

FOOTNOTES

[1]Note, however, that the proof of asymmptotic normality of the Aitken estimator does not require the assumption that the weighted disturbances are identically distributed; rather, it can be seen from an examination of the standard proofs (e.g., Theil [12], Ch. 8) based on the Lindberg-Feller theorem that independence of the u_t is sufficient. This result is fortunate since in general the weighted disturbances are not necessarily identically distributed even though they have the same mean and variance (for example, if $\mu_{3t} = z_t\gamma_3 \neq 0$, the weighted disturbances are not identically distributed). See Gnedenko and Kolmogorov [5] or Dhrymes [3], Ch. 3, for further discussion of this point.

[2]In general, the joint estimation procedure may include any number of moment functions. For the general model see Antle [2].

References

[1] Amemiya, T. "A Note on a Heteroscedastic Model." Journal of Econometrics 6(1977): 365-370, "Corrigenda," Journal of Econometrics 8(1977): 265.

[2] Antle, J. M. "Theory and Measurement of Output Distributions." University of California, Davis, Department of Agricultural Economics, Working Paper #81-3, 1981.

[3] Dhrymes, P. J. Econometrics: Statistical Foundations and Applications. Springer-Verlay, 1974.

[4] Glejser, H. "A New Test for Heteroscedasticity." Journal of the American Statistical Association 64.

[5] Gnedenko, B. V. and A. N. Kolmogorov. Limit Distributions for Some of Independent Random Variables. Addison Wesley, 1954.

[6] Goldfeld, S. M. and R. E. Quandt. Nonlinear Methods in Econometrics. Amsterdam, North-Holland, 1972.

[7] Harvey, A. C. "Estimating Regression Models with Multiplicative Heteroscedasticity." Econometrica 44(1976): 461-465.

[8] Judge, G. G., W. E. Griffiths, R. C. Hill, and T-C. Lee. The Theory and Practice of Econometrics. New York: Wiley, 1980.

[9] Malinvaud, E. "The Consistency of Nonlinear Regressions." Annals of Mathematical Statistics 41(1970): 956-969.

[10] Murtagh, B. A. and M. A. Saunders. MINOS User's Guide. Technical Report 77-9 (February 1977), Dept. of Operations Research, Stanford University.

[11] Rao, C. R. Linear Statistical Inference and Its Application. New York: Wiley, 1973.

[12] Theil, H. Principles of Econometrics. New York: Wiley, 1971.

[13] Zellner, A. "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias." Journal of the American Statistical Association 57(1962): 348-368.