ECONOMETRICA

VOLUME 48 MAY, 1980 NUMBER 4

A HETEROSKEDASTICITY-CONSISTENT COVARIANCE MATRIX ESTIMATOR AND A DIRECT TEST FOR HETEROSKEDASTICITY

BY HALBERT WHITE[1]

This paper presents a parameter covariance matrix estimator which is consistent even when the disturbances of a linear regression model are heteroskedastic. This estimator does not depend on a formal model of the structure of the heteroskedasticity. By comparing the elements of the new estimator to those of the usual covariance estimator, one obtains a direct test for heteroskedasticity, since in the absence of heteroskedasticity, the two estimators will be approximately equal, but will generally diverge otherwise. The test has an appealing least squares interpretation.

1. INTRODUCTION

IT IS WELL KNOWN that the presence of heteroskedasticity in the disturbances of an otherwise properly specified linear model leads to consistent but inefficient parameter estimates and inconsistent covariance matrix estimates. As a result, faulty inferences will be drawn when testing statistical hypotheses in the presence of heteroskedasticity.

If the investigator has a formal model of the process generating the differing variances, these difficulties are easily eliminated by performing an appropriate linear transformation on the data, based on this model. However, even when such a model is available, it may be incorrect. Often, several models are considered (e.g., Griliches [**10**]), but still without the certain knowledge that any of them is correct. In this situation one can test each of the alternative transformed models for remaining heteroskedasticity (using any of several available tests), and eliminate those which fail. But what is one to do if all fail the heteroskedasticity test? Although the investigator will have a fairly good idea of the parameter values of the linear model, there remains a considerable difficulty in assessing the precision of the parameter estimates and testing hypotheses due to the possible inconsistency of the usual covariance matrix estimator.

In this paper I resolve this difficulty by presenting a covariance matrix estimator which is consistent in the presence of heteroskedasticity, but does not rely on a (possibly incorrect) specific formal model of the structure of the heteroskedasticity. Thus, even when heteroskedasticity cannot be completely eliminated, proper inferences can be drawn. Under appropriate conditions, a natural test for heteroskedasticity can be obtained by comparing the consistent estimator to the usual covariance matrix estimator; in the absence of heteroskedasticity, both estimators will be about the same—otherwise, they will generally diverge. The test shares the advantage of the covariance estimator, in that no formal structure on the nature of the heteroskedasticity is imposed, in contrast to the tests suggested

[1] I am greatly indebted to David Tanny, Jon Wellner, Dennis Carlton, the referees, and the participants of the University of Western Ontario and Harvard/MIT econometrics workshops for helpful comments and suggestions. Any errors are mine alone.

by Goldfeld and Quandt [**8**], Rutemiller and Bowers [**20**], Glejser [**6**], or Harvey [**12**].

2. THE HETEROSKEDASTICITY-CONSISTENT COVARIANCE ESTIMATOR

To begin, assume that the model has the following structure:

ASSUMPTION 1: The model is known to be

$$Y_i = X_i\beta_0 + \varepsilon_i \qquad (i = 1, \ldots, n)$$

where (X_i, ε_i) is a sequence of independent not (necessarily) identically distributed (i.n.i.d.) random vectors, such that X_i (a $1 \times K$ vector) and ε_i (a scalar) satisfy $E(X_i'\varepsilon_i) = 0$. ε_i is unobservable while Y_i and X_i are observable. β_0 is a finite unknown $K \times 1$ parameter vector to be estimated.

By assuming that the elements of the sequence (X_i, ε_i) are i.n.i.d., the case of fixed regressors with (possibly) heteroskedastic errors is automatically covered. Also covered by this assumption is the case in which observations are obtained not from a controlled experiment (as the fixed regressor assumption requires) but rather from a (possibly) stratified cross section, a case frequently encountered in applied microeconomics. Note that by assuming only that X_i and ε_i are uncorrelated, we automatically cover the less general but frequently encountered cases in which $E(\varepsilon_i|X_i) = 0$ or X_i and ε_i are independent, with $E(\varepsilon_i) = 0$. Thus, we allow heteroskedasticity of the form $E(\varepsilon_i^2|X_i) = g(X_i)$, where g is a known (possibly parametric) function. Such a situation arises, for example, in the random coefficients model of Hildreth and Houck [**14**].

Next, we make the following assumption.

ASSUMPTION 2: (a) There exist positive finite constants δ and Δ such that, for all i, $E(|\varepsilon_i^2|^{1+\delta}) < \Delta$ and $E(|X_{ij}X_{ik}|^{1+\delta}) < \Delta$ $(j, k = 1, \ldots, K)$; (b) $\bar{M}_n = n^{-1}\sum_{i=1}^{n} E(X_i'X_i)$ is nonsingular for (all) n sufficiently large,[2] such that $\det \bar{M}_n > \delta > 0$.

The first part of this assumption ensures that the error variances are uniformly bounded (the condition being slightly stronger than this) and that the elements of the average covariance matrix of the regressors are also uniformly bounded. In the fixed regressor case, the condition requires that the regressors themselves be uniformly bounded. The second part ensures the eventual nonsingularity of the average covariance matrix of the regressors and the uniform boundedness of the elements of the inverse. Note that $\bar{M}_n$ is not required to converge to any particular limit.

[2] In what follows the qualifier "all" will be implicitly understood.

Define the ordinary least squares (OLS) estimator $\hat{\beta}_n = (X'X)^{-1}X'Y$, where X is the $n \times K$ matrix with rows X_i and Y is the $n \times 1$ vector with elements Y_i. It is simple to prove the following lemma.

LEMMA 1: *Given Assumptions 1 and 2, $\hat{\beta}_n$ exists almost surely for n sufficiently large and $\hat{\beta}_n \xrightarrow{a.s.} \beta_0$.*

Proofs are given in the Mathematical Appendix. Thus, OLS provides a strongly consistent estimator for β_0 in the i.n.i.d. regressor and error case defined by Assumptions 1 and 2.

With the next assumption, an asymptotic normality result can be obtained.

ASSUMPTION 3: (a) There exist positive finite constants δ and Δ such that for all i $E(|\varepsilon_i^2 X_{ij} X_{ik}|^{1+\delta}) < \Delta$ $(j, k = 1, \dots, K)$; (b) The average covariance matrix $\bar{V}_n = n^{-1} \sum_{i=1}^n E(\varepsilon_i^2 X_i' X_i)$ is nonsingular for n sufficiently large, such that $\det \bar{V}_n > \delta > 0$.

Note that with fixed regressors or stochastic regressors independent of ε_i, Assumption 3(a) is implied by Assumption 2 (a).

The uniform boundedness of the elements of $\bar{V}_n$ in Assumption 3 is guaranteed by Assumption 3 (a). Assumption 3 (a) and (b) ensure the uniform boundedness of $\bar{V}_n^{-1}$.

Together, Assumptions 1–3 allow the multivariate Liapounov central limit theorem given by White [**23**] to be applied.

The asymptotic normality result is as follows.

LEMMA 2. *Under Assumptions* 1–3,

$$\sqrt{n}\,\bar{V}_n^{-1/2}\bar{M}_n(\hat{\beta}_n - \beta_0) \overset{A}{\sim} N(0, I_K).$$

This result is slightly more general than the asymptotic normality results usually given since the "asymptotic covariance matrix" $\bar{M}_n^{-1}\bar{V}_n\bar{M}_n^{-1}$ is not required to converge to any particular limit. In the fixed regressor case, $\bar{M}_n^{-1}\bar{V}_n\bar{M}_n^{-1}$ has the familiar form $(X'X/n)^{-1}(X'\Omega X/n)(X'X/n)^{-1}$, where Ω is the $n \times n$ diagonal matrix with diagonal elements $\sigma_i^2 = E(\varepsilon_i^2)$.

Now consider the problem of testing hypotheses. In particular, consider testing the linear hypotheses

$$H_0\colon R\beta_0 = r \quad \text{vs.} \quad H_1\colon R\beta_0 \neq r,$$

where R is a finite $q \times K$ matrix of full row rank and r is a finite $q \times 1$ vector. It can be shown that given H_0

$$n(R\hat{\beta}_n - r)'[R\bar{M}_n^{-1}\bar{V}_n\bar{M}_n^{-1}R']^{-1}(R\hat{\beta}_n - r) \overset{A}{\sim} \chi_q^2$$

under Assumptions 1–3. This statistic is not computable however, since $\bar{M}_n^{-1}\bar{V}_n\bar{M}_n^{-1}$ is not known. If it were possible to replace $\bar{M}_n^{-1}\bar{V}_n\bar{M}_n^{-1}$ with a consistent estimator, the usual asymptotic tests (the normal and χ^2 tests analogous to the familiar finite sample t and F tests) could be performed.

The difficulty evidently arises in estimating $\bar{V}_n$. If we consider the fixed regressor case in which $\bar{V}_n = (X'\Omega X/n)$, it might appear that we must successfully estimate each diagonal element of Ω. As the model has been set out, this could require estimating n different variances σ_i^2, obviously an impossible task when only n observations are available (cf. Goldfeld and Quandt [**9**, p. 86]). But this way of looking at the problem is misleading. What is actually required is to estimate $n^{-1}\sum_{i=1}^n E(\varepsilon_i^2 X_i'X_i)$, an average of expectations. To do this it is not necessary to estimate each expectation separately. Under the conditions given above, a consistent estimator is $n^{-1}\sum_{i=1}^n \varepsilon_i^2 X_i'X_1$. Unfortunately, ε_i is not observable; however, ε_i can be estimated by $\hat{\varepsilon}_{in} = Y_i - X_i\hat{\beta}_n$, which leads us to consider the estimator

$$\hat{V}_n = n^{-1}\sum_{i=1}^n \hat{\varepsilon}_{in}^2 X_i'X_i.$$

In the fixed regressor case, this amounts to replacing the ith diagonal element of Ω, σ_i^2, with $\hat{\varepsilon}_{in}^2$, the ith squared residual.

With the next condition, the estimator $\hat{V}_n$ becomes the key to solving the problem of obtaining a heteroskedasticity-consistent covariance estimator.

ASSUMPTION 4: There exist positive constants δ and Δ such that for all i

$$E(|X_{ij}^2 X_{ik} X_{il}|^{1+\delta}) < \Delta \qquad (j, k, l = 1, \ldots, K).$$

Uniformly bounded fixed or stochastic regressors are sufficient for Assumption 4 to hold.[3]

We now present the first main result.

THEOREM 1: (i) $|\hat{V}_n - \bar{V}_n| \xrightarrow{a.s.} 0$ *under Assumptions* 1, 2, 3(*a*) *and* 4; (ii)

$$|(X'X/n)^{-1}\hat{V}_n(X'X/n)^{-1} - \bar{M}_n^{-1}\bar{V}_n\bar{M}_n^{-1}| \xrightarrow{a.s.} 0$$

under Assumptions 1, 2, 3(*a*), *and* 4; (iii)

$$n(R\hat{\beta}_n - r)'[R(X'X/n)^{-1}\hat{V}_n(X'X/n)^{-1}R']^{-1}(R\hat{\beta}_n - r) \overset{A}{\sim} \chi_q^2,$$

given H_0 *and Assumptions* 1–4.

The first part of Theorem 1 states that $\hat{V}_n$ consistently estimates $\bar{V}_n$, resolving the difficulty discussed in the preceding paragraphs. The second part

[3] Note also that Assumption 4 is sufficient for the second condition in Assumption 2(a), since

$$E(|X_{ij}X_{ik}|^{1+\delta}) \leq E(|X_{ij}^2 X_{ik}^2|^{1+\delta}) + 1.$$

provides the heteroskedasticity-consistent covariance matrix estimator, $(X'X/n)^{-1}\hat{V}_n(X'X/n)^{-1}$. Using this estimator to test linear hypotheses in the usual way gives correct results asymptotically, as the third part of the theorem demonstrates. This result fills a substantial gap in the econometrics literature, and should be useful in a wide variety of applications. Since the convergence to normality of $\sqrt{n}(\hat{\beta}_n-\beta_0)$ can be shown to be uniform, the heteroskedasticity-consistent covariance estimator can also be shown to be appropriate for use in constructing asymptotic confidence intervals.

In fact, results similar to propositions (i) and (ii) of Theorem 1 were stated over a decade ago by Eicker [**5**], although Eicker considers only fixed and not stochastic regressors. Also, if $r=0$ and if R is the $1\times K$ vector with ith element equal to unity and the rest zero, then the χ_1^2 test statistic of (iii) is precisely the square of the asymptotic normal statistic (analogous to the t test) proposed by Eicker [**4**] for the heteroskedastic case in an even earlier classic paper. It is somewhat surprising that these very useful facts have remained unfamiliar to practicing econometricians for so long.

The estimator $\hat{V}_n$ is also similar to an estimator proposed by Hartley, J. N. K. Rao, and Kiefer [**11**] and C. R. Rao [**18**], namely

$$\tilde{V}_n=n^{-1}\sum_{i=1}^{n}\hat{\sigma}_{in}^2X_i'X_i$$

where $\hat{\sigma}_{in}$ is the minimum norm quadratic unbiased estimator (MINQUE) for σ_i^2, discussed by C. R. Rao [**18**] and Chew [**2**]. $\hat{\varepsilon}_{in}^2$ is not a MINQUE for σ_i^2; nevertheless, it is straightforward to show that $\hat{V}_n$ and $\tilde{V}_n$ are asymptotically equivalent, so that Theorem 1 also holds for $\tilde{V}_n$. Note, however, that due to the complexity of the formula for $\hat{\sigma}_{in}^2$, $\tilde{V}_n$ is always more difficult to compute than $\hat{V}_n$ and becomes more so as n increases.

The results of Theorem 1 can be extended to the nonlinear case $Y_i=f(X_i,\theta)+\varepsilon_i$ by replacing X_{ij} with $\partial f(X_i,\theta)/\partial\theta_j$ in all computations, under conditions given by White [**23**]. In this case, it resembles the covariance matrix estimator given by Berndt, Hall, Hall, and Hausman [**1**]. An expression analogous to $\hat{V}_n$ can also be easily obtained for instrumental variables estimators. (In the framework of two-stage least squares, one simply replaces X_i with $\hat{X}_i$, the projection of X_i on the space spanned by the instruments; however, as usual, $\hat{\varepsilon}_{in}^2$ is still computed using X_i.)

3. A DIRECT TEST FOR HETEROSKEDASTICITY

Goldfeld and Quandt [**9**, Ch. 3] provide several examples to show that heteroskedasticity can result in serious inconsistencies in the usual least squares covariance matrix estimator[4] $\hat{\sigma}_n^2(X'X/n)^{-1}$ where $\hat{\sigma}_n^2=n^{-1}\sum_{i=1}^n(Y_i-X_i\hat{\beta}_n)^2$.

[4] The estimator $\hat{\sigma}_n^2$ is used instead of $s_n^2=(n-K)^{-1}\sum_{i=1}^n(Y_i-X_i\hat{\beta}_n)^2$ to avoid unnecessarily cluttering the proofs. Obviously, all asymptotic results are valid if s_n^2 replaces $\hat{\sigma}_n^2$.

Similar problems can be caused by a lack of independence of ε_i and X_i, even in the absence of heteroskedasticity. Of course, the heteroskedasticity consistent covariance matrix given in Theorem 1 allows one to make proper inferences and construct appropriate confidence intervals regardless of these complications. However, this estimator is not as simple to compute as $\hat{\sigma}_n^2(X'X/n)^{-1}$, so it is particularly useful to have a simple indicator of whether or not $\hat{\sigma}_n^2(X'X/n)^{-1}$ is in fact inconsistent which does not rely on direct computation of $(X'X/n)^{-1}\hat{V}_n(X'X/n)^{-1}$. Further, if this inconsistency can be attributed to heteroskedasticity (as will often be possible) one also has an indication of whether there is a potential efficiency payoff to a more careful modeling of the variance structure.

In this section, we derive a simple test for conditions which ensure the consistency of $\hat{\sigma}_n^2(X'X/n)^{-1}$. The test is based on the fact that in the presence of homoskedasticity ($\sigma_i^2=\sigma_0^2$, for all i) and with ε_i independent of X_i, $\bar{V}_n = \sigma_0^2 n^{-1}\sum_{i=1}^n E(X_i'X_i)$. In this case $\bar{V}_n$ can be consistently estimated either by $\hat{V}_n$ or by $\hat{\sigma}_n^2(X'X/n)$. Comparing the elements of $\hat{V}_n$ and $\hat{\sigma}_n^2(X'X/n)$ thus provides an indication of whether or not $\hat{\sigma}_n^2(X'X/n)^{-1}$ is a consistent covariance matrix estimator. Theorem 2 below makes precise the sense in which $\hat{V}_n$ and $\hat{\sigma}_n^2(X'X/n)$ must be sufficiently far apart to indicate inconsistency, and Corollary 1 provides a simple procedure for detecting possible inconsistency. To obtain these results, we must add additional structure.

First, we make the following assumption.

Assumption 5: There exist positive constants δ and Δ such that, for all i, $E(|\varepsilon_i^4|^{1+\delta})<\Delta$, and $E(|X_{ij}X_{ik}X_{il}X_{im}|^{1+\delta})<\Delta$ ($j, k, l, m = 1, \ldots, K$).

Note that Assumption 5 is sufficient for Assumption 4 and Assumption 2(a). Next, define the products

$$\Psi_{is} \equiv X_{ik}X_{il} \qquad (s=1,\ldots,K(K+1)/2;\ k=1,\ldots,K;\ l=1,\ldots,k)$$

and let Ψ_i be the $1\times K(K+1)/2$ vector with elements Ψ_{is}. Thus, Ψ_i is the vector containing the elements of the lower triangle of the matrix $X_i'X_i$. Also, define

$$\bar{\Psi}_{ns} \equiv n^{-1}\sum_{i=1}^n E(\Psi_{is}) \qquad (s=1,\ldots,K(K+1)/2).$$

The vector $\bar{\Psi}_n$ with elements $\bar{\Psi}_{ns}$ contains the elements of the lower triangle of $\bar{M}_n$. Similarly, we define the $1\times K(K+1)/2$ vector $\hat{\Psi}_n = n^{-1}\sum_{i=1}^n \Psi_i$. $\hat{\Psi}_n$ contains the elements of the lower triangle of $X'X/n$. Now assume the following:

Assumption 6: The average covariance matrix

$$\bar{B}_n = n^{-1}\sum_{i=1}^n E([\varepsilon_i^2-\sigma_i^2]^2(\Psi_i-\bar{\Psi}_n)'(\Psi_i-\bar{\Psi}_n))$$

is nonsingular for n sufficiently large such that $\det \bar{B}_n > \delta > 0$; also, for n sufficiently large $n^{-1} \sum_{i=1}^{n} E((\varepsilon_i^2 - \sigma_i^2)^2) > \delta > 0$.

In the theorem which follows, we are particularly concerned with the behavior of the random variables $(\Psi_i - \bar{\Psi}_n)'[\varepsilon_i^2 - \sigma_i^2]$. Assumptions 5 and 6 help to ensure that the appropriate conditions are satisfied to invoke asymptotic normality. $\bar{B}_n$ is the average covariance matrix of these random variables; with X_i independent of ε_i, its elements are uniformly bounded by Assumption 5. The inverse $\bar{B}_n^{-1}$ exists for n sufficiently large, and with X_i independent of ε_i has uniformly bounded elements by Assumptions 5 and 6.

Finally, we make the following assumption:

ASSUMPTION 7: There exist positive constants δ and Δ such that for all i

$$E(|X_{ij}^4 \Psi_{is} \Psi_{it}|^{1+\delta}) < \Delta \qquad (j = 1, \ldots, K; s, t, = 1, \ldots, K(K+1)/2).$$

In vector notation, we can express the lower triangle of the matrix difference $\hat{V}_n - \hat{\sigma}_n^2(X'X/n)$ as the difference vector

$$D_n(\hat{\beta}_n, \hat{\sigma}_n^2) = n^{-1} \sum_{i=1}^{n} \Psi'_{in}[\hat{\varepsilon}_{in}^2 - \hat{\sigma}_n^2].$$

We also use an estimator of $\bar{B}_n$, defined as

$$\hat{B}_n = n^{-1} \sum_{i=1}^{n} [\hat{\varepsilon}_{in}^2 - \hat{\sigma}_n^2]^2 (\Psi_i - \hat{\Psi}_n)'(\Psi_i - \hat{\Psi}_n).$$

With this structure, we can now give a direct test for inconsistency of the usual least squares covariance matrix estimator, $\hat{\sigma}_n^2(X'X/n)^{-1}$.

THEOREM 2: *Given Assumptions* 1, 2(*b*), 3(*b*), *and* 5–7, *if* ε_i *is independent of* X_i *and* $E(\varepsilon_i^2) = \sigma_0^2$ *for all i, then*

$$(1) \qquad nD_n(\hat{\beta}_n, \hat{\sigma}_n^2)'\hat{B}_n^{-1}D_n(\hat{\beta}_n, \hat{\sigma}_n^2) \overset{A}{\sim} \chi^2_{K(K+1)/2}.$$

Note that the null hypothesis maintains not only that the errors are homoskedastic, but also that they are independent of the regressors, *and* that the model is correctly specified in the sense that Assumptions 1, 2(b), 3(b), and 5–7 hold. Failure of any of these conditions can lead to a statistically significant test statistic. Essentially, the statistic (1) is testing a joint hyoothesis that the model's specification of the first *and* second moments of the dependent variable is correct. To discover that (1) is not statistically significant for a given model is particularly good news since it implies that not only is the variance specification of the model correct (subject to caveats below) but also that the linear specification $Y_i = X_i\beta_0 + \varepsilon_i$ is correct. In this sense, the test based on (1) is a general test for model misspecification.

In fact, the test statistic of Theorem 2 is computationally identical to the linear version of a test for model misspecification proposed in White [**22**] for the i.i.d. regressor nonlinear model case. In that situation, the power of this statistic with respect to the null hypothesis that the explanatory model is correct up to an independent additive disturbance derives from the fact that $(X'X/n)^{-1}\hat{V}_n(X'X/n)^{-1}$ is a consistent estimator of the parameter covariance matrix even when the model is misspecified, while $\hat{\sigma}_n^2(X'X/n)^{-1}$ is inconsistent. This inconsistency arises from the fact that when the model is misspecified, the disturbance of the incorrect specification contains the specification error, so that it is impossible for the regressors to be independent of the disturbance term, as the consistency of $\hat{\sigma}_n^2(X'X/n)^{-1}$ requires. (See White [**22** and **24**] for further discussion.)

When the null hypothesis of Theorem 2 is rejected, the appropriate conclusion to be drawn depends upon whether one is willing to maintain the correctness of the model's specification (most importantly, that the true relation is indeed $Y_i = X_i\beta_0 + \varepsilon_i$). If so, the most plausible reason for rejection is heteroskedasticity. In the regressor case, this is the *only* reason. In the stochastic regressor case, lack of independence of regressors and errors alone can lead to rejection; however, in the most commonly encountered models, it is heteroskedasticity which leads to lack of independence, through a dependence of the error variance on the regressors. Thus, when the model is maintained to be correct, rejection may reasonably be attributed to heteroskedasticity, so that there is a potential efficiency gain to be realized from a more careful modeling of the variance structure.

If the investigator is less confident about the correctness of the linear model, the test indicates only that something is wrong, but not what. A more thorough investigation of the model's specification is indicated.[5] In what follows, we assume the model is correct and treat (1) as a heteroskedasticity test.

As given, the statistic (1) is even more cumbersome to compute than $\hat{V}_n$. However, by slightly modifying the conditions of Theorem 2, we can obtain a comparable statistic which is very easy to obtain. In particular, replace Assumption 7 in Theorem 2 with the assumption that the ε_i are homokurtic—that is, $E(\varepsilon_i^4) = \mu_4$ for all i. Now consider the artificial regression

$$\text{(2)} \qquad \begin{aligned} \hat{\varepsilon}_{in}^2 &= \alpha_0 + \alpha_1\Psi_1 + \alpha_2\Psi_2 + \ldots + \alpha_{K(K+1)/2}\Psi_{K(K+1)/2} \\ &= \alpha_0 + \sum_{j=1}^{K}\sum_{k=j}^{K} \alpha_s X_{ij}X_{ik} \qquad (i = 1, \ldots, n) \end{aligned}$$

[5] If the investigator is unsure about the correctness of the model's specification, a practical way of proceeding would be to use the statistic (1), and if rejection occurs, apply a specification test of the type proposed by Hausman [**13**] or White [**24**]. These latter tests are sensitive to model misspecification, but not heteroskedasticity. Hence, accepting the null hypothesis of no model misspecification would indicate that rejection of the null hypothesis of Theorem 2 is indeed due to heteroskedasticity, whereas rejection of the null hypothesis of no misspecification indicates that rejection of the null hypothesis of Theorem 2 is due to model misspecification. Since these tests are in general dependent, the formal size of this sequential procedure will be difficult to determine. Note also that both specification tests mentioned above may detect only a lack of independence between errors and regressors, instead of misspecification.

where the α's are parameters to be estimated by OLS. Equation (2) directs us to regress the squared estimated residuals on all second order products and cross-products of the original regressors. The next result uses the fact that (1) is asymptotically equivalent to testing the joint hypothesis $\alpha_1 = \alpha_2 = \ldots = \alpha_{K(K+1)/2} = 0$ using the standard R^2 statistic from the regression (2), given the homokurtosis assumption. Formally, we have the following corollary.

COROLLARY 1: *Given Assumptions* 1, 2(*b*), 3(*b*), 4–6, *if* ε_i *is independent of* X_i, *and* $E(\varepsilon_i^2) = \sigma_0^2$, $E(\varepsilon_i^4) = \mu_4$ *for all i, then*

$$nR^2 \overset{A}{\sim} \chi^2_{K(K+1)/2} \tag{3}$$

where R^2 *is the* (*constant-adjusted*) *squared multiple correlation coefficient from the regression* (2).

All the remarks made following Theorem 2 apply here as well. It is useful to note at this point, however, that Assumption 6 can fail identically. In particular, if the estimating equation contains a constant term, say $X_{i1} \equiv 1$, then equation (2) contains a redundant constant; equivalently, the corresponding (first) element of $D_n(\hat{\beta}_n, \hat{\sigma}_n^2)$ and the row and column of $\bar{B}_n$ corresponding to the constant will be identically zero, implying a singular $\bar{B}_n$. In this case, it is entirely appropriate to delete the redundant constant from (2) (equivalently, the corresponding element of $D_n(\hat{\beta}_n, \hat{\sigma}_n^2)$ and row and column of $\hat{B}_n$) and proceed, reducing the χ^2 degrees of freedom by one. Redundancies will also occur if the X_i contains a constant and polynomial terms (e.g., the translog production function). For example if $X_{i1} \equiv 1$ and $X_{i3} = X_{i2}^2$, then $X_{i1}X_{i3} = X_{i2}X_{i2}$. Again, the redundant term is simply dropped and degrees of freedom are reduced by one.

Note that when the homokurtosis assumption fails (so that (1) is appropriate), the only effect is that the nominal size of the test associated with (3) becomes incorrect. The test statistic (3) will still have unit power asymptotically.

This result allows one to perform the test without first computing the matrix $\hat{V}_n$. If the test is passed, it indicates the adequacy of $\hat{\sigma}_n^2(X'X/n)^{-1}$ for hypothesis testing, and one could stop at this point (although in some cases further efficiency gains might be possible—see below). If the test is failed, one can proceed to calculate the heteroskedasticity-consistent covariance matrix estimator using the identity

$$D_n(\hat{\beta}_n, \hat{\sigma}_n^2) \equiv \left(n^{-1} \sum_{i=1}^{n} (\Psi_i - \hat{\Psi}_n)'(\Psi_i - \hat{\Psi}_n)\right)\hat{\alpha}_n$$

where $\hat{\alpha}_n$ is the $K(K+1)/2 \times 1$ vector containing the OLS estimates of $\alpha_1, \ldots, \alpha_{K(K+1)/2}$. By adding the appropriate element of $D_n(\hat{\beta}_n, \hat{\sigma}_n^2)$ to each element of $\hat{\sigma}_n^2(X'X/n)$, one obtains $\hat{V}_n$.

In the form (3), the test resembles a Lagrange multiplier test for a specific class of normal heteroskedastic alternatives considered by Godfrey [**7**]. Note that normality is not assumed here. From the discussion above, it is clear that the power of Godfrey's test must derive, as does that of the statistic (3), from the inconsistency of $\hat{\sigma}_n^2(X'X/n)$ for $\bar{V}_n$.

The present testing procedure is also similar to the modified Glejser procedure proposed by Goldfeld and Quandt [**9**, p. 93]; however, while Goldfeld and Quandt accepted or rejected the homoskedasticity hypothesis on the basis of unidirectional t tests associated with the estimated α's, the test proposed here considers all α's jointly. Its performance as a heteroskedasticity test should compare to the modified Gleijser test in a manner roughly analogous to the performance of Goldfeld and Quandt's FIML χ^2 test relative to their FIML t test [**9**, pp. 94–100].

Although most previous tests for heteroskedasticity have relied upon imposing some more or less formal structure on the nature of the heteroskedasticity and then testing to see if this structure is found (e.g., Goldfeld and Quandt [**8**], Glejser [**6**], Rutemiller and Bowers [**20**], Harvey [**12**], and Godfrey [**7**]), the present test does not require specifying the heteroskedastic structure. Heuristically, one expects the statistics (1) or (3) to have good power against all heteroskedastic alternatives which result in inconsistency for the usual covariance matrix estimator; for n sufficiently large, the power of the test in these cases will approach unity. Tests which correctly formalize the heteroskedasticity should have some power advantage in finite samples, but (1) or (3) should equal or dominate tests which incorrectly specify the variance structure. If theory suggests that some of the α's in (2) should be zero (as in the original Hildreth-Houck [**14**] model) or negligible, power (and degrees of freedom) may be gained by omitting some terms in (2). A finite sample investigation of the power of (3) absolutely and relative to existing tests will be undertaken in future work.

Are there cases in which heteroskedasticity does *not* result in inconsistency for $\hat{\sigma}_n^2(X'X/n)^{-1}$? If so, one should expect the power of (1) and (3) to be low in absolute terms for these cases. Such circumstances are easily characterized, although it might sometimes be difficult to detect these cases independently of the statistics given. This characterization is given by the next simple result.

THEOREM 3: *Given Assumptions* 1, 2(*a*) *and* ε_i *independent of* X_i, $|\hat{\sigma}_n^2(X'X/n)-\bar{V}_n|\xrightarrow{\text{a.s.}}0$ *if and only if*

$$\left|n^{-1}\sum_{i=1}^{n}(\sigma_i^2-\bar{\sigma}_n^2)(E(X_i'X_i)-\bar{M}_n)\right|\to 0 \quad \text{as} \quad n\to\infty.$$

The necessary and sufficient condition for the implied consistency of $\hat{\sigma}_n^2(X'X/n)^{-1}$ provided above says essentially that the second moments of the errors, ε_i, must be "uncorrelated" with (asymptotically orthogonal to) *all* second moments and cross-moments of the regressors, X_i. Obviously, this occurs when $\sigma_i^2\equiv\sigma_0^2$, the case of homoskedasticity; but it can also occur when $E(X_i'X_i)\equiv M_0$ (a fixed finite matrix), e.g. when the X_i are obtained from a truly random sample, so that the regressors are i.i.d. In this last mentioned case, it can be shown that (1) has precisely the same $\chi^2_{K(K+1)/2}$ distribution, so that in the presence of heteroskedasticity the power of the test is precisely the apparent size of the test. Unfortunately, these cases are not the only possible ones for which the test will have low power, since one can construct an unlimited number of them simply by letting σ_i^2 and $E(X_i'X_i)$ be independent random variables with finite means (taking these moments as fixed *a posteriori*). A small χ^2 statistic will indicate either homoskedasticity or any of these other possibilities. Note, however, that if a variance structure satisfies $E(\varepsilon_i^2|Q_i)=g(Q_i)$ for some variables Q_i, then Q_i will have to be appropriately orthogonal to X_i for the conditions of Theorem 3 to hold; usually $Q_i = X_i$, so this cannot occur.

Theorem 3 generalizes a simple result given by Malinvaud [**16**, pp. 303–04] for the single regressor case. There, Malinvaud shows sufficiency; the present result shows that necessity is also valid. This result indicates that heteroskedasticity will be easily detected *only* when it makes a difference for drawing inferences using the usual least squares covariance matrix estimator. However, it does not necessarily follow that in such cases efficiency cannot be improved. The effects of heteroskedasticity on the usual covariance matrix estimator and on the efficiency of the least squares estimator are quite separate. In the case of a diagonal error covariance matrix, the necessary and sufficient condition for the efficiency of least squares is precisely homoskedasticity (cf. McElroy [**17**]). Thus, taking proper account even of variance structures which satisfy the conditions of Theorem 3 can improve estimator efficiency. The extent of the improvement can be directly assessed (at least asymptotically) by comparing the heteroskedasticity-consistent covariance matrix of the appropriate weighted least squares estimator to that of the OLS estimator.

In fact, since the covariance matrix is being consistently estimated, any number of alternative variance structures may be evaluated in this way, and the structure with the greatest resulting efficiency can be chosen. One particular such structure suggested by equation (2) involves choosing weights

$$\hat{\omega}_{in} = (\max\,[\hat{\alpha}_{on} + \Psi_i\hat{\alpha}_n, \delta > 0])$$

(where δ is arbitrarily chosen) and forming the weighted least squares estimator $\tilde{\beta}_n = (X'\hat{\Omega}_n^{-1}X)^{-1}X'\hat{\Omega}_n^{-1}Y$, where $\hat{\Omega}_n$ is the diagonal matrix with diagonal elements $\hat{\omega}_{in}$. Both Glejser [**6**] and Goldfeld and Quandt [**9**] have found in Monte Carlo studies that even incorrect heteroskedasticity corrections can improve estimator efficiency over OLS. In future work, possible efficiency improvements from the above choice of weights will be investigated using the heteroskedasticity-consistent covariance matrix.

4. SUMMARY AND CONCLUDING REMARKS

This paper has presented general conditions under which a consistent estimator of the OLS parameter covariance matrix can be obtained, regardless of the presence of heteroskedasticity in the disturbances of a properly specified linear model. Since this estimator does not require a formal modeling of the structure of the heteroskedasticity and since it requires only the regressors and the estimated least squares residuals for its computation, the estimator of Theorem 1 should have wide applicability. Additional conditions are given which allow the investigator to test directly for the presence of heteroskedasticity. If found, elimination of the heteroskedasticity by a more careful modeling of the stochastic structure of the model can yield improved estimator efficiency.

Until now, one had either to model heteroskedasticity correctly or suffer the consequences. The fact that the covariance matrix estimator and heteroskedasticity test given here do not require formal modeling of the heteroskedastic structure is a great convenience, but it does not relieve the investigator of the burden of carefully specifying his models. Instead, it is hoped that the statistics presented here will enable researchers to be even more careful in specifying and estimating econometric models. Thus, when a formal model for heteroskedasticity is available, application of the tools presented here will allow one to check the validity of this model, and undertake further modeling if indicated. But even when heteroskedasticity cannot be completely eliminated, the heteroskedasticity-consistent covariance matrix of Theorem 1 allows correct inferences and confidence intervals to be obtained.

University of Rochester

Manuscript received August, 1978; revision received March, 1979.

MATHEMATICAL APPENDIX

All symbols, definitions, and assumptions are as given in the text.

LEMMA 1: *Given Assumptions* 1 *and* 2, $\hat{\beta}_n$ *exists almost surely for* (*all*) *n sufficiently large and* $\hat{\beta}_n \xrightarrow{\text{a.s.}} \beta_0$.

PROOF: $\hat{\beta}_n$ exists almost surely for (all) n sufficiently large provided $(X'X/n)$ is nonsingular almost surely for n sufficiently large. When this is true, Assumption 1 allows us to write

$$\hat{\beta}_n = (X'X/n)^{-1}(X'Y/n) = \beta_0 + (X'X/n)^{-1}(X'\varepsilon/n).$$

Now $E(|X_{ij}X_{ik}|^{1+\delta}) < \Delta$ for all i and $j, k = 1, \ldots, K$ by Assumption 2(a); since the X_i are mutually independent, it follows by Markov's strong law of large numbers (e.g. Chung [**3**, p. 125]) that

$$|(X'X/n) - \bar{M}_n| \xrightarrow{\text{a.s.}} 0$$

where the notation is understood to indicate convergence of the matrices element by element. Now by Assumption 2(b) $\bar{M}_n$ is nonsingular for n sufficiently large. By the continuity of the matrix inverse, it follows that $(X'X/n)$ is nonsingular almost surely for n sufficiently large (so that $\hat{\beta}_n$ exists); the elements of $\bar{M}_n^{-1}$ are uniformly bounded for n sufficiently large, so that

$$|(X'X/n)^{-1} - \bar{M}_n^{-1}| \xrightarrow{\text{a.s.}} 0.$$

Next, $E(|X_{ij}\varepsilon_i|^{1+\delta}) < \Delta$ for all i and $j = 1, \ldots, K$ as a result of the Hölder inequality and Assumption 2(a) which ensures $E(|X_{ij}^2|^{1+\delta}) < \Delta$, $E(|\varepsilon_i^2|^{1+\delta}) < \Delta$. Since the $X_i'\varepsilon_i$ are independent, Markov's strong law implies

$$\left|(X'\varepsilon/n)-n^{-1}\sum_{i=1}^{n}E(X_i'\varepsilon_i)\right|\xrightarrow{\text{a.s.}}0.$$

Since $\bar{M}_n^{-1}$ has uniformly bounded elements for n sufficiently large by Assumption 2(b), and since $E(X_i'\varepsilon_i)$ has uniformly bounded elements, uniform continuity implies

$$\left|(X'X/n)^{-1}(X'\varepsilon/n)-\bar{M}_n^{-1}n^{-1}\sum_{i=1}^{n}E(X_i'\varepsilon_i)\right|\xrightarrow{\text{a.s.}}0.$$

But $E(X_i'\varepsilon_i)=0$ under Assumption 1. Thus

$$(X'X/n)^{-1}(X'\varepsilon/n)\xrightarrow{\text{a.s.}}0$$

given Assumptions 1 and 2, implying $\hat{\beta}_n\xrightarrow{\text{a.s.}}\beta_0$. *Q.E.D.*

LEMMA 2: *Under Assumptions* 1–3,

$$\sqrt{n}\bar{V}_n^{-\frac{1}{2}}\bar{M}_n(\hat{\beta}_n-\beta_0)\overset{A}{\sim}N(0,I_K).$$

PROOF: Consider the quantity $n^{-\frac{1}{2}}\sum_{i=1}^{n}X_i'\varepsilon_i$. Under Assumption 1 the random vectors $X_i'\varepsilon_i$ are independent with $E(X_i'\varepsilon_i)=0$ and average covariance matrix

$$\bar{V}_n=n^{-1}\sum_{i=1}^{n}E(\varepsilon_i^2X_i'X_i)$$

which is positive definite for n sufficiently large by Assumption 3(a). Thus, we can define the symmetric positive definite matrix $\bar{V}_n^{-\frac{1}{2}}$ such that $(\bar{V}_n^{-\frac{1}{2}})^2=\bar{V}_n^{-1}$. The elements of $\bar{V}_n^{-\frac{1}{2}}$ are uniformly bounded under Assumptions 2 and 3, as are $E(|X_{ij}\varepsilon_i|^{2+\delta})$ for some $\delta>0$ and all i,j given Assumption 3(b). Thus, by the Minkowski inequality for some $\delta>0$, $\sum_{i=1}^{n}E|\lambda'\bar{V}_n^{-\frac{1}{2}}X_i'\varepsilon_i|^{2+\delta}/n^{(2+\delta)}/2\to 0$ for all λ in R^K. Hence, the multivariate Liapounov central limit theorem (White [**23**, Theorem 3.1]) implies that

$$\bar{V}_n^{-\frac{1}{2}}n^{-\frac{1}{2}}\sum_{i=1}^{n}X_i'\varepsilon_i\overset{A}{\sim}N(0,I_K).$$

Now

$$\sqrt{n}\bar{V}_n^{-\frac{1}{2}}\bar{M}_n(\hat{\beta}_n-\beta_0)=\bar{V}_n^{-\frac{1}{2}}\bar{M}_n(X'X/n)^{-1}\bar{V}_n^{\frac{1}{2}}\bar{V}_n^{-\frac{1}{2}}n^{-\frac{1}{2}}\sum_{i=1}^{n}X_i'\varepsilon_i$$

almost surely for n sufficiently large. Next,

$$|\bar{V}_n^{-\frac{1}{2}}\bar{M}_n(X'X/n)^{-1}\bar{V}_n^{\frac{1}{2}}-I_K|\xrightarrow{P}0$$

by Lemma 3.2 of White [**23**], since $|(X'X/n)^{-1}-\bar{M}_n^{-1}|\xrightarrow{\text{a.s.}}0$ under Assumption 2 as argued in Lemma 1, so that

$$\left|\sqrt{n}\bar{V}_n^{-\frac{1}{2}}\bar{M}_n(\hat{\beta}_n-\beta_0)-\bar{V}_n^{-\frac{1}{2}}n^{-\frac{1}{2}}\sum_{i=1}^{n}X_i'\varepsilon_i\right|\xrightarrow{P}0. \tag{a.1}$$

It then follows from Lemma 3.3 of White [**23**] that

$$\sqrt{n}\bar{V}_n^{-\frac{1}{2}}\bar{M}_n(\hat{\beta}_n-\beta_0)\overset{A}{\sim}N(0,I_K).$$ *Q.E.D.*

THEOREM 1: (i) $|\hat{V}_n - \bar{V}_n| \xrightarrow{a.s.} 0$ *under Assumptions* 1, 2, 3(*a*) *and* 4; (ii)

$$|(X'X/n)^{-1}\hat{V}_n(X'X/n)^{-1} - \bar{M}_n^{-1}\bar{V}_n\bar{M}_n^{-1}| \xrightarrow{a.s.} 0$$

under Assumptions 1, 2, 3(*a*) *and* 4; (iii)

$$n(R\hat{\beta}_n - r)'[R(X'X/n)^{-1}\hat{V}_n(X'X/n)^{-1}R']^{-1}(R\hat{\beta}_n - r) \overset{A}{\sim} \chi^2_q,$$

given H_0 *and Assumptions* 1–4.

PROOF: (i) Since β_0 is finite, there exists a compact neighborhood ν of β_0 such that $(\beta_j - \beta_{0j})$ is finite, $j = 1, \ldots, \kappa$, where β_j and β_{0j} are the jth elements of β and β_0 in ν, respectively. There also exists a finite vector $\tilde{\beta}$ (not necessarily in ν) with elements $\tilde{\beta}_j$ such that $(\beta_j - \beta_{0j})^2 \leq (\tilde{\beta}_j - \beta_{0j})^2$ for all β in ν, so that for all β in ν

$$|X_{ij}^2(\beta_j - \beta_{0j})^2 X_{ik}X_{il}| \leq |X_{ij}^2(\tilde{\beta}_j - \beta_{0j})^2 X_{ik}X_{il}|.$$

Now

$$|(Y_i - X_i\beta)^2 X_{ik}X_{il}| = |(\varepsilon_i - X_i(\beta - \beta_0))^2 X_{ik}X_{il}|$$

under Assumption 1, and it is a direct consequence of the elementary inequality

$$|a+b|^r \leq 2^{r-1}|a|^r + 2^{r-1}|b|^r \; (r \geq 1)$$

that there exist finite positive constants $\lambda_0, \ldots, \lambda_K$ such that for all β in ν

$$\text{(a.2)} \qquad |(Y_i - X_i\beta)^2 X_{ik}X_{il}| \leq \lambda_0|\varepsilon_i^2 X_{ik}X_{il}| + \sum_{j=1}^{K} \lambda_j |X_{ij}^2(\beta_j - \beta_{0j})^2 X_{ik}X_{il}|$$

$$\leq \lambda_0|\varepsilon_i^2 X_{ik}X_{il}| + \sum_{j=1}^{K} \lambda_j |X_{ij}^2 X_{ik}X_{il}|(\tilde{\beta}_j - \beta_{oj})^2 \equiv m_{kl}(X_i, \varepsilon_i).$$

Thus $|(Y_i - X_i\beta)^2 X_{ik}X_{il}|$ is dominated by $m_{kl}(X_i, \varepsilon_i)$.

Now it also follows from the above inequality that there exist finite positive constants $\mu_0, \ldots, \mu_K$ such that, for $\delta > 0$,

$$E(|m_{kl}(X_i, \varepsilon_i)|^{1+\delta}) \leq \mu_0\lambda_0^{1+\delta}E(|\varepsilon_i^2 X_{ik}X_{il}|^{1+\delta})$$
$$+ \sum_{j=1}^{K} \mu_j\lambda_j^{1+\delta}E(|X_{ij}^2 X_{ik}X_{il}|^{1+\delta})(\tilde{\beta}_j - \beta_{0j})^{2+2\delta}.$$

Since $E(|\varepsilon_i^2 X_{ik}X_{il}|^{1+\delta})$ is uniformly bounded by Assumption 3(a), since $E(|X_{ij}^2 X_{ik}X_{il}|^{1+\delta})$ is uniformly bounded by Assumption 4, $j = 1, \ldots, K$, and since $\mu_0, \ldots, \mu_K$, $\lambda_0, \ldots, \lambda_K$, and $(\tilde{\beta}_1 - \beta_{01})^{2+2\delta}, \ldots, (\tilde{\beta}_K - \beta_{0K})^{2+2\delta}$ are finite positive constants, $E(|m_{kl}(X_i, \varepsilon_i)|^{1+\delta})$ is uniformly bounded on ν. It then follows from White [**23**, Lemma 2.3] that

$$\sup_{\beta \in \nu} \left|n^{-1}\sum_{i=1}^{n}(Y_i - X_i\beta)^2 X_{ik}X_{il} - n^{-1}\sum_{i=1}^{n} E((Y_i - X_i\beta)^2 X_{ik}X_{il})\right| \xrightarrow{a.s.} 0.$$

Since $\hat{\beta}_n \xrightarrow{a.s.} \beta_0$ under Assumptions 1 and 2 as established by Lemma 1, it follows from White [**23**, Lemma 2.6] that

$$\left|n^{-1}\sum_{i=1}^{n}(Y_i - X_i\hat{\beta}_n)^2 X_{ik}X_{il} - n^{-1}\sum_{i=1}^{n} E(\varepsilon_i^2 X_{ik}X_{il})\right| \xrightarrow{a.s.} 0 \qquad (k, l = 1, \ldots, K)$$

or $|\hat{V}_n - \bar{V}_n| \xrightarrow{a.s.} 0$, the desired result.

(ii) In Lemma 1 it was established that $|(X'X/n)^{-1} - \bar{M}_n^{-1}| \xrightarrow{a.s.} 0$, and $\bar{M}_n^{-1}$ has uniformly bounded elements for n sufficiently large under Assumption 2. Combining this result with (i) above,

$$|(X'X/n)^{-1}\hat{V}_n(X'X/n)^{-1} - \bar{M}_n^{-1}\bar{V}_n\bar{M}_n^{-1}| \xrightarrow{\text{a.s.}} 0$$

follows by uniform continuity, since $\bar{V}_n$ has uniformly bounded elements under Assumption 3(a).

(iii) We can write

$$R\hat{\beta}_n - r = R\beta_0 - r + R(\hat{\beta}_n - \beta_0).$$

Applying H_0: $R\beta_0 = r$ and multiplying by $\sqrt{n}$ yields

$$\sqrt{n}(R\hat{\beta}_n - r) = \sqrt{n}R(\hat{\beta}_n - \beta_0).$$

For convenience let $\bar{\Gamma}_n \equiv R\bar{M}_n^{-1}\bar{V}_n\bar{M}_n^{-1}R'$ (which exists, has uniformly bounded elements, and is nonsingular for n sufficiently large given Assumptions 1–3) and define the symmetric positive definite matrix $\bar{\Gamma}_n^{-\frac{1}{2}}$ such that $(\bar{\Gamma}_n^{-\frac{1}{2}})^2 = \bar{\Gamma}_n^{-1}$. The elements of $\bar{\Gamma}_n^{-\frac{1}{2}}$ are also uniformly bounded by Assumption 2 and 3. Consider the quantity

$$\sqrt{n}\bar{\Gamma}_n^{-\frac{1}{2}}(R\hat{\beta}_n - r) = \sqrt{n}\bar{\Gamma}_n^{-\frac{1}{2}}R(\hat{\beta}_n - \beta_0).$$

Since $\bar{\Gamma}_n^{-\frac{1}{2}}R$ is uniformly bounded, it follows from (a.1) that

$$\left|\sqrt{n}\bar{\Gamma}_n^{-\frac{1}{2}}R\bar{M}_n^{-1}\bar{V}_n^{\frac{1}{2}}\bar{V}_n^{-\frac{1}{2}}\bar{M}_n(\hat{\beta}_n - \beta_0) - \bar{\Gamma}_n^{-\frac{1}{2}}R\bar{M}_n^{-1}n^{-\frac{1}{2}}\sum_{i=1}^{n} X_i'\varepsilon_i\right| \xrightarrow{P} 0$$

and by Lemma 3.3 of White [**24**] that $\sqrt{n}\bar{\Gamma}_n^{-\frac{1}{2}}R(\hat{\beta}_n - \beta_0)$ has the same asymptotic distribution as

$$n^{-\frac{1}{2}}\sum_{i=1}^{n}\bar{\Gamma}_n^{-\frac{1}{2}}R\bar{M}_n^{-1}X_i'\varepsilon_i,$$

provided this is multivariate normal. The random variables $\bar{\Gamma}_n^{-\frac{1}{2}}R\bar{M}_n^{-1}X_i'\varepsilon_i$ are independent with $E(\bar{\Gamma}_n^{-\frac{1}{2}}R\bar{M}_n^{-1}X_i'\varepsilon_i) = 0$ and covariance matrix

$$\begin{aligned} n^{-1}\sum_{i=1}^{n}&\bar{\Gamma}_n^{-\frac{1}{2}}R\bar{M}_n^{-1}E(\varepsilon_i^2X_i'X_i)\bar{M}_n^{-1}R'\bar{\Gamma}_n^{-\frac{1}{2}} \\ &= \bar{\Gamma}_n^{-\frac{1}{2}}R\bar{M}_n^{-1}\bar{V}_n\bar{M}_n^{-1}R'\bar{\Gamma}_n^{-\frac{1}{2}} = \bar{\Gamma}_n^{-\frac{1}{2}}\bar{\Gamma}_n\bar{\Gamma}_n^{-\frac{1}{2}} \\ &= I_K. \end{aligned}$$

By the multivariate Liapounov central limit theorem,

$$n^{-\frac{1}{2}}\sum_{i=1}^{n}\bar{\Gamma}_n^{-\frac{1}{2}}R\bar{M}_n^{-1}X_i'\varepsilon_i \overset{A}{\sim} N(0, I_K)$$

provided that for some $\delta > 0$,

$$\sum_{i=1}^{n} E|\lambda'\bar{\Gamma}_n^{\frac{1}{2}}R\bar{M}_n^{-1}X_i'\varepsilon_i|^{2+\delta}/n^{(2+\delta)/2} \to 0 \tag{a.3}$$

for all λ in $\mathbb{R}^q$.

Equation (a.3) holds by the Minkowski inequality since $\bar{\Gamma}_n^{-\frac{1}{2}}R\bar{M}_n^{-1}$ has uniformly bounded elements given 2 and 3 and since for some δ, $E(|X_{ij}\varepsilon_i|^{2+\delta})$ is uniformly bounded for all i, j by Assumptions 3(a). Thus, we may infer

$$\sqrt{n}\bar{\Gamma}_n^{-\frac{1}{2}}(R\hat{\beta}_n - r) \overset{A}{\sim} N(0, I_K). \tag{a.4}$$

Since $\bar{\Gamma}_n$ is not computable, consider the quantity

$$\sqrt{n}\hat{\Gamma}_n^{-\frac{1}{2}}(R\hat{\beta}_n - r)$$

where $\hat{\Gamma}_n^{-\frac{1}{2}}$ is the symmetric positive definite matrix such that $(\hat{\Gamma}_n^{-\frac{1}{2}})^2 = \hat{\Gamma}_n^{-1}$, and $\hat{\Gamma}_n \equiv R(X'X/n)^{-1}\hat{V}_n(X'X/n)^{-1}R'$ exists and is nonsingular almost surely for n sufficiently large given

Assumptions 1–4. Since R is finite it follows from (ii) above that

$$|\hat{\Gamma}_n - \bar{\Gamma}_n| \xrightarrow{P} 0.$$

It then follows from Lemma 3.3 of White [23] that

$$\sqrt{n}\hat{\Gamma}_n^{-\frac{1}{2}}(R\hat{\beta}_n - r) \overset{A}{\sim} N(0, I_K)$$

and that

$$n(R\hat{\beta}_n - r)'\hat{\Gamma}_n^{-1}(R\hat{\beta}_n - r) \overset{A}{\sim} \chi_q^2$$

or, substituting for $\hat{\Gamma}_n$,

$$n(R\hat{\beta}_n - r)'[R(X'X/n)^{-1}\hat{V}_n(X'X/n)^{-1}R']^{-1}(R\hat{\beta}_n - r) \overset{A}{\sim} \chi_q^2. \qquad Q.E.D.$$

THEOREM 2: *Given Assumptions* 1, 2(*b*), 3(*b*), 5–7, *if* ε_i *is independent of* X_i *and* $E(\varepsilon_i^2) = \sigma_0^2$ *for all* i, *then*

$$nD_n(\hat{\beta}_n, \hat{\sigma}_n^2)'\hat{B}_n^{-1}D_n(\hat{\beta}_n, \hat{\sigma}_n^2) \overset{A}{\sim} \chi^2_{K(K+1)/2}.$$

PROOF: Let $\theta = (\beta, \sigma)$ and consider the quantities

$$D_n^s(\theta) = n^{-1}\sum_{i=1}^n [(Y_i - X_i\beta)^2 - \sigma^2](\Psi_{is} - \bar{\Psi}_{ns}) \qquad (s = 1, \dots, K(K+1)/2)$$

where

$$\bar{\Psi}_{ns} = n^{-1}\sum_{i=1}^n E(\Psi_{is}), \quad \Psi_{is} = X_{ik}X_{il}$$

$$(s = 1, \dots, K(K+1)/2; k = 1, \dots, K; l = 1, \dots, k).$$

With homoskedastic disturbances independent of X_i, the random variables $[\varepsilon_i^2 - \sigma_0^2](\Psi_{is} - \bar{\Psi}_{ns})$ have expectation zero. Define the average covariance matrix $\bar{B}_n(\theta)$ with elements

$$\bar{B}_n^{st}(\theta) = n^{-1}\sum_{i=1}^n E([(Y_i - X_i\beta)^2 - \sigma^2]^2(\Psi_{is} - \bar{\Psi}_{ns})(\Psi_{it} - \bar{\Psi}_{nt})) \quad (s, t = 1, \dots, K(K+1)/2).$$

By Assumption 6, $\bar{B}_n(\theta_0)$ is nonsingular for n sufficiently large, with uniformly bounded elements given Assumption 5 and independence of X_i and ε_i. For n sufficiently large, we can define the symmetric positive definite matrix $\bar{B}_n(\theta_0)^{-\frac{1}{2}}$ such that $[\bar{B}_n(\theta_0)^{-\frac{1}{2}}]^2 = \bar{B}_n(\theta_0)^{-1}$.

Assumptions 5, 6, and the independence of X_i and ε_i ensure the eventual uniform boundedness of the elements of $\bar{B}_n(\theta_0)^{-1}$. Let $D_n(\theta_0)$ be the $K(K+1)/2 \times 1$ vector with elements $D_n^s(\theta_0)$. Then

$$\text{(a.5)} \qquad \sqrt{n}\bar{B}_n(\theta_0)^{-\frac{1}{2}}D_n(\theta_0) \overset{A}{\sim} N(0, I_{K(K+1)/2})$$

by the multivariate Liapounov central limit theorem, provided that for some $\delta > 0$

$$\sum_{i=1}^n E|\lambda'\bar{B}_n(\theta_0)^{-\frac{1}{2}}(\Psi_i - \bar{\Psi}_n)'(\varepsilon_i^2 - \sigma_0^2)|^{2+\delta}/n^{(2+\delta)/2} \to 0$$

for all λ in $\mathbb{R}^{K(K+1)/2}$. But this holds by the Minkowski inequality since $\bar{B}_n(\theta_0)^{-\frac{1}{2}}$ has uniformly bounded elements and since $E|(\Psi_{is} - \bar{\Psi}_{ns})(\varepsilon_i^2 - \sigma_0^2)|^{2+\delta}$ is uniformly bounded for some $\delta > 0$ given Assumption 5 and the independence of X_i and ε_i.

To obtain the desired result, we first show that

$$|\sqrt{n}D_n^s(\hat{\theta}_n)-\sqrt{n}D_n^s(\theta_0)|\xrightarrow{P}0,$$

where

$$D_n^s(\hat{\theta}_n)=n^{-1}\sum_{i=1}^{n}[(Y_i-X_i\hat{\beta}_n)^2-\hat{\sigma}_n^2]\Psi_{is}$$

and

$$\hat{\sigma}_n^2=n^{-1}\sum_{i=1}^{n}(Y_i-X_i\hat{\beta}_n)^2.$$

Since

$$n^{-1}\sum_{i=1}^{n}[(Y_i-X_i\hat{\beta}_n)^2-\hat{\sigma}_n^2]\bar{\Psi}_{ns}$$

is identically zero, we have

$$D_n^s(\hat{\theta}_n)=n^{-1}\sum_{i=1}^{n}[(Y_i-X_i\hat{\beta}_n)^2-\hat{\sigma}_n^2](\Psi_{is}-\bar{\Psi}_{ns}).$$

Next, $\hat{\theta}_n\xrightarrow{\text{a.s.}}\theta_0$ under Assumption 1, 2, and the homoskedasticity assumption, so that there exists a sequence $(\tilde{\theta}_n)$ which is tail equivalent to $(\hat{\theta}_n)$ such that each $\tilde{\theta}_n$ takes its values in a convex compact neighborhood ν of θ_0, such that all θ in ν are finite. Let $\tilde{\beta}_n'=(\tilde{\beta}_{n1},\dots,\tilde{\beta}_{nK})$ and let $\beta_0'=(\beta_{01},\dots,\beta_{0K})$. Since $D_n^s(\theta)$ satisfies the conditions of Jennrich [**15**, Lemma 3], there exists a measurable function $\bar{\theta}_n$ lying between $\tilde{\theta}_n$ and θ_0 such that

$$\begin{aligned}D_n^s(\tilde{\theta}_n)=D_n^s(\theta_0)&+\sum_{j=1}^{K}(\tilde{\beta}_{nj}-\beta_{0j})\partial D_n^s(\bar{\theta}_n)/\partial\beta_j\\&+(\tilde{\sigma}_n^2-\sigma_0^2)\partial D_n^s(\bar{\theta}_n)/\partial\sigma^2.\end{aligned}$$

Rearranging and multiplying by $\sqrt{n}$ gives

$$|\sqrt{n}D_n^s(\tilde{\theta}_n)-\sqrt{n}D_n^s(\theta_0)|=\left|\sum_{j=1}^{K}\sqrt{n}(\tilde{\beta}_{nj}-\beta_{0j})\partial D_n^s(\bar{\theta}_n)/\partial\beta_j+\sqrt{n}(\tilde{\sigma}^2-\sigma_0^2)\partial D_n^s(\bar{\theta}_n)/\partial\sigma^2\right|.$$

Using the triangle inequality and factoring yields

$$\begin{aligned}\text{(a.5)}\qquad |\sqrt{n}D_n^s(\tilde{\theta}_n)-\sqrt{n}D_n^s(\theta_0)|\leq&\sum_{j=1}^{K}|\sqrt{n}(\tilde{\beta}_{nj}-\beta_{0j})|\cdot|\partial D_n^s(\bar{\theta}_n)/\partial\beta_j|\\&+|\sqrt{n}(\tilde{\sigma}_n^2-\sigma_0^2)|\cdot|\partial D_n^s(\bar{\theta}_n)/\partial\sigma^2|.\end{aligned}$$

Now each term $|\sqrt{n}(\tilde{\beta}_{nj}-\beta_{0j})|$ is $O_p(1)$ by the tail equivalence of $\tilde{\beta}_n$ and $\hat{\beta}_n$ and Lemma 2, given Assumptions 1–3. Also, $|\sqrt{n}(\tilde{\sigma}_n^2-\sigma_0^2)|$ is easily shown to be $O_p(1)$ given Assumptions 1–3, provided the ε_i^2 are uniformly asymptotically negligible, as ensured by Assumption 5. Next,

$$\partial D_n^s(\bar{\theta}_n)/\partial\beta_j=-2n^{-1}\sum_{i=1}^{n}(Y_i-X_i\bar{\beta}_n)X_{ij}(\Psi_{is}-\bar{\Psi}_{ns})$$

and

$$\partial D_n^s(\bar{\theta}_n)/\partial\sigma^2=-n^{-1}\sum_{i=1}^{n}\Psi_{is}-\bar{\Psi}_{ns}.$$

Assumptions 1, 2, and 5 ensure that $|(Y_i-X_i\beta)X_{ij}\Psi_{is}|$ and $|(Y_i-X_i\beta)X_{ij}|$ are appropriately dominated (arguing as in part (i) of Theorem 1 above) so that Lemma 2.3 of White [**23**] can be applied to obtain

$$\sup_{\theta\in\nu}\left|n^{-1}\sum_{i=1}^{n}(Y_i-X_i\beta)X_{ij}(\Psi_{is}-\bar{\Psi}_{ns})-n^{-1}\sum_{i=1}^{n}E((Y_i-X_i\beta)X_{ij}(\Psi_{is}-\bar{\Psi}_{ns}))\right|\xrightarrow{\text{a.s.}}0.$$

Since $\bar{\beta}_n \xrightarrow{\text{a.s.}} \beta_0$, it follows from White [23, Lemma 2.6] that

$$\left| n^{-1} \sum_{i=1}^{n} (Y_i - X_i\bar{\beta}_n) X_{ij}(\Psi_{is} - \bar{\Psi}_{ns}) - n^{-1} \sum_{i=1}^{n} E(\varepsilon_i) E(X_{ij}(\Psi_{is} - \bar{\Psi}_{ns}))\right| \xrightarrow{\text{a.s.}} 0.$$

Under Assumption 1, and independence of X_i and ε_i, $E(\varepsilon_i) = 0$, so $|\partial D_n^s(\bar{\theta}_n)/\partial\beta_j| \xrightarrow{\text{a.s.}} 0$. In Lemma 1, it was shown that $|(X'X/n) - \bar{M}_n| \xrightarrow{\text{a.s.}} 0$ under Assumption 2, so $|\partial D_n^s(\bar{\theta}_n)/\partial\sigma^2| \xrightarrow{\text{a.s.}} 0$. Hence, (a.6) and the immediately preceding arguments imply

$$|\sqrt{n} D_n^s(\tilde{\theta}_n) - \sqrt{n} D_n^s(\theta_0)| \xrightarrow{P} 0 \qquad (s = 1, \ldots, K(K+1)/2)$$

by 2c.4x(a) of Rao [19]. The uniform boundedness of the elements of $\bar{B}_n(\theta_0)^{-\frac{1}{2}}$ then ensures that

$$|\sqrt{n}\bar{B}_n(\theta_0)^{-\frac{1}{2}} D_n(\tilde{\theta}_n) - \sqrt{n}\bar{B}_n(\theta_0)^{-\frac{1}{2}} D_n(\theta_0)| \xrightarrow{P} 0.$$

The desired result will follow in a straightforward way from Lemma 3.3 of White [24] if we can find a consistent estimator for $\bar{B}_n(\theta_0)$. Accordingly, consider the estimator $\hat{B}_n$ with elements

$$\hat{B}_n^{st} = n^{-1} \sum_{i=1}^{n} [(Y_i - X_i\hat{\beta}_n)^2 - \hat{\sigma}_n^2]^2 (\Psi_{is} - \hat{\Psi}_{ns})(\Psi_{it} - \hat{\Psi}_{nt})$$

where $\hat{\Psi}_{ns} = n^{-1} \sum_{i=1}^{n} \Psi_{is}$. Expanding $\hat{B}_n^{st}$ gives

$$\begin{aligned}
n^{-1} &\sum_{i=1}^{n} [(Y_i - X_i\hat{\beta}_n)^2 - \hat{\sigma}_n^2]^2 (\Psi_{is} - \hat{\Psi}_{ns})(\Psi_{it} - \hat{\Psi}_{nt}) \\
&= n^{-1} \sum_{i=1}^{n} (Y_i - X_i\hat{\beta}_n)^4 \Psi_{is}\Psi_{it} - 2\hat{\sigma}_n^2 n^{-1} \sum_{i=1}^{n} (Y_i - X_i\hat{\beta}_n)^2 \Psi_{is}\Psi_{it} \\
&\quad + \hat{\sigma}_n^4 n^{-1} \sum_{i=1}^{n} \Psi_{is}\Psi_{it} - \hat{\Psi}_{ns} n^{-1} \sum_{i=1}^{n} (Y_i - X_i\hat{\beta}_n)^4 \Psi_{it} \\
&\quad + 2\hat{\sigma}_n^2 \hat{\Psi}_{ns} n^{-1} \sum_{i=1}^{n} (Y_i - X_i\hat{\beta}_n)^2 \Psi_{it} - \hat{\sigma}_n^4 \hat{\Psi}_{ns} n^{-1} \sum_{i=1}^{n} \Psi_{it} \\
&\quad - \hat{\Psi}_{nt} n^{-1} \sum_{i=1}^{n} (Y_i - X_i\hat{\beta}_n)^4 \Psi_{is} + 2\hat{\sigma}_n^2 \hat{\Psi}_{nt} n^{-1} \sum_{i=1}^{n} (Y_i - X_i\hat{\beta}_n)^2 \Psi_{is} \\
&\quad - \hat{\sigma}_n^4 \hat{\Psi}_{nt} n^{-1} \sum_{i=1}^{n} \Psi_{is} + \hat{\Psi}_{ns}\hat{\Psi}_{nt} n^{-1} \sum_{i=1}^{n} (Y_i - X_i\hat{\beta}_n)^4 \\
&\quad - 2\hat{\Psi}_{ns}\hat{\Psi}_{nt} n^{-1} \sum_{i=1}^{n} (Y_i - X_i\hat{\beta}_n)^2 + \hat{\Psi}_{ns}\hat{\Psi}_{nt}\hat{\sigma}_n^4.
\end{aligned}$$

As in Lemma 1, $|\hat{\Psi}_{ns} - \bar{\Psi}_{ns}| \xrightarrow{\text{a.s.}} 0$ under Assumption 2, and $\bar{\Psi}_{ns}$ is uniformly bounded by Assumption 2. Also $|\hat{\sigma}_n^4 - \sigma_0^4| \xrightarrow{\text{a.s.}} 0$ under Assumptions 1, 2, and the homoskedasticity assumption. Consider a convex compact neighborhood ν of β_0 such that all β in ν are finite. Using the same kind of argument as that at the outset of the proof of Theorem 1, part (i), there exists β^* finite with elements β_j^* such that $(\beta_j - \beta_{0j})^4 \leq (\beta_j^* - \beta_{0j})^4$ for all β in ν, so that for all β in ν

$$|(Y_i - X_i\beta)^4 \Psi_{is}\Psi_{it}| \leq \lambda_0 |\varepsilon_i^4 \Psi_{is}\Psi_{it}| + \sum_{j=1}^{K} \lambda_j |X_{ij}^4 \Psi_{is}\Psi_{it}| (\beta_j^* - \beta_{0j})^4 \equiv \zeta_{st}(X_i, \varepsilon_i)$$

where $\lambda_0, \ldots, \lambda_K$ are finite positive constants. Similarly,

$$|(Y_i - X_i\beta)^4| \leq \lambda_0|\varepsilon_i^4| + \sum_{j=1}^{K} \lambda_j|X_{ij}^4|(\beta_j^* - \beta_{0j})^4 \equiv \eta(X_i, \varepsilon_i)$$

for all β in ν. Since $|\Psi_{is}| \leq \Psi_{is}^2 + 1$, it follows that

$$|(Y_i - X_i\beta)^4\Psi_{is}| \leq |(Y_i - X_i\beta)^4\Psi_{is}^2| + |(Y_i - X_i\beta)^4| \leq \zeta_{ss}(X_i, \varepsilon_i) + \eta(X_i, \varepsilon_i)$$

for all β in ν. Similarly, for all β in ν

$$|(Y_i - X_i\beta)^2\Psi_{is}\Psi_{it}| \leq \zeta_{st}(X_i, \varepsilon_i) + |\Psi_{is}\Psi_{it}|$$

and

$$|(Y_i - X_i\beta)^2\Psi_{is}| \leq \zeta_{ss}(X_i, \varepsilon_i) + 1.$$

Continuing the same kind of argument as in part (i) of Theorem 1, it is easily shown that $E(|\zeta_{st}(X_i, \varepsilon_i)|^{1+\delta})$ is uniformly bounded under Assumptions 5, 7 and independence of X_i and ε_i, and that $E(|\eta(X_i, \varepsilon_i)|^{1+\delta})$ is uniformly bounded under Assumption 5. It is also easily shown that

$$\|E(|\zeta_{ss}(X_i, \varepsilon_i) + \eta(X_i, \varepsilon_i)|^{1+\delta}), \quad E(|\zeta_{st}(X_i, \varepsilon_i) + |\Psi_{is}\Psi_{it}\|^{1+\delta}), \quad \text{and} \quad E(|\zeta_{ss}(X_i, \varepsilon_i) + 1|^{1+\delta})$$

are uniformly bounded given Assumptions 5, 7 and independence of X_i and ε_i. Thus by Lemma 2.3 of White [**23**]

$$\sup_{\beta \in \nu} \left|n^{-1} \sum_{i=1}^{n} (Y_i - X_i\beta)^4\Psi_{is}\Psi_{it} - n^{-1} \sum_{i=1}^{n} E((Y_i - X_i\beta)^4\Psi_{is}\Psi_{it})\right| \xrightarrow{\text{a.s.}} 0,$$

$$\sup_{\beta \in \nu} \left|n^{-1} \sum_{i=1}^{n} (Y_i - X_i\beta)^2\Psi_{is}\Psi_{it} - n^{-1} \sum_{i=1}^{n} E((Y_i - X_i\beta)^2\Psi_{is}\Psi_{it})\right| \xrightarrow{\text{a.s.}} 0,$$

$$\sup_{\beta \in \nu} \left|n^{-1} \sum_{i=1}^{n} (Y_i - X_i\beta)^4\Psi_{it} - n^{-1} \sum_{i=1}^{n} E((Y_i - X_i\beta)^4\Psi_{it})\right| \xrightarrow{\text{a.s.}} 0,$$

$$\sup_{\beta \in \nu} \left|n^{-1} \sum_{i=1}^{n} (Y_i - X_i\beta)^2\Psi_{it} - n^{-1} \sum_{i=1}^{n} E((Y_i - X_i\beta)^2\Psi_{it})\right| \xrightarrow{\text{a.s.}} 0,$$

and

$$\sup_{\beta \in \nu} \left|n^{-1} \sum_{i=1}^{n} (Y_i - X_i\beta)^4 - n^{-1} \sum_{i=1}^{n} E((Y_i - X_i\beta)^4)\right| \xrightarrow{\text{a.s.}} 0 \qquad (s, t = 1, \ldots, K(K+1)/2).$$

Since $\hat{\beta}_n \xrightarrow{\text{a.s.}} \beta_0$ under Assumptions 1 and 2 as established by Lemma 1, it follows from White [**23**, Lemma 2.6] that

$$\left|n^{-1} \sum_{i=1}^{n} (Y_i - X_i\hat{\beta}_n)^4\Psi_{is}\Psi_{it} - n^{-1} \sum_{i=1}^{n} E(\varepsilon_i^4\Psi_{is}\Psi_{it})\right| \xrightarrow{\text{a.s.}} 0,$$

$$\left|n^{-1} \sum_{i=1}^{n} (Y_i - X_i\hat{\beta}_n)^2\Psi_{is}\Psi_{it} - n^{-1} \sum_{i=1}^{n} E(\varepsilon_i^2\Psi_{is}\Psi_{it})\right| \xrightarrow{\text{a.s.}} 0,$$

$$\left|n^{-1} \sum_{i=1}^{n} (Y_i - X_i\hat{\beta}_n)^4\Psi_{it} - n^{-1} \sum_{i=1}^{n} E(\varepsilon_i^4\Psi_{it})\right| \xrightarrow{\text{a.s.}} 0,$$

$$\left|n^{-1} \sum_{i=1}^{n} (Y_i - X_i\hat{\beta}_n)^2\Psi_{it} - n^{-1} \sum_{i=1}^{n} E(\varepsilon_i^2\Psi_{it})\right| \xrightarrow{\text{a.s.}} 0,$$

and

$$\left|n^{-1} \sum_{i=1}^{n} (Y_i - X_i\hat{\beta}_n)^4 - n^{-1} \sum_{i=1}^{n} E(\varepsilon_i^4)\right| \xrightarrow{\text{a.s.}} 0.$$

By Markov's strong law of large numbers

$$\left| n^{-1} \sum_{i=1}^{n} \Psi_{is}\Psi_{it} - n^{-1} \sum_{i=1}^{n} E(\Psi_{is}\Psi_{it}) \right| \xrightarrow{\text{a.s.}} 0,$$

given Assumption 5. Since all relevant expectations are contained in a compact subset of a Euclidean space, uniform continuity and the foregoing facts imply $|\hat{B}_n^{st} - \bar{B}_n^{st}(\theta_0)| \xrightarrow{\text{a.s.}} 0$, $s, t = 1, \ldots, K(K+1)/2$, as straightforward (but tedious) algebra will verify. It then follows from White [**23**, Lemma 3.3] that

$$nD_n(\hat{\theta}_n)'\hat{B}_n^{-1}D_n(\hat{\theta}_n) \overset{A}{\sim} \chi^2_{K(K+1)/2}. \qquad \textit{Q.E.D.}$$

COROLLARY 1: *Given Assumptions* 1, 2(*b*), 3(*b*), 4–6, *if* ε_i *is independent of* X_i, *and* $E(\varepsilon_i^2) = \sigma_0^2$, $E(\varepsilon_i^4) = \mu_4$ *for all i, then*

$$nR^2 \overset{A}{\sim} \chi^2_{K(K+1)/2}$$

where R^2 *is the (constant-adjusted) squared multiple correlation coefficient of the regression* (2).

PROOF: If X_i and ε_i are independent, $E(\varepsilon_i^2) = \sigma_0^2$ and $E(\varepsilon_i^4) = \mu_4$ for all i, then

$$\bar{B}_n^{st}(\theta_0) = (\mu_4 - \sigma_0^4) \cdot n^{-1} \sum_{i=1}^{n} E((\Psi_{is} - \bar{\Psi}_{ns})(\Psi_{it} - \bar{\Psi}_{nt})) \qquad (s, t = 1, \ldots, K(K+1)/2)$$

for which a strongly consistent estimator is

$$\tilde{B}_n^{st} = n^{-1} \sum_{i=1}^{n} (\hat{\varepsilon}_{in}^2 - \hat{\sigma}_n^2)^2 \cdot n^{-1} \sum_{i=1}^{n} (\Psi_{is} - \hat{\Psi}_{ns})(\Psi_{it} - \hat{\Psi}_{nt}) \qquad (s, t = 1, \ldots, K(K+1)/2)$$

given Assumption 5. Arguing exactly as in Theorem 2 above, it follows that

$$nD_n(\hat{\beta}_n, \hat{\sigma}_n^2)'\tilde{B}_n^{-1}D_n(\hat{\beta}_n, \hat{\sigma}_n^2) \overset{A}{\sim} \chi^2_{K(K+1)/2} \tag{a.7}$$

where $\tilde{B}_n$ is the matrix with elements $\tilde{B}_n^{st}$. The left-hand side of (a.7) can be written more explicitly as

$$\begin{aligned} &nD_n(\hat{\beta}_n, \hat{\sigma}_n)'\tilde{B}_n^{-1}D_n(\hat{\beta}_n, \hat{\sigma}_n^2) \\ &= n\left(n^{-1} \sum_{i=1}^{n} \Psi_i'[\hat{\varepsilon}_{in}^2 - \hat{\sigma}_n^2]\right)'\left(n^{-1} \sum_{i=1}^{n} (\Psi_i - \hat{\Psi}_n)'(\Psi_i - \hat{\Psi}_n)\right)^{-1} \\ &\quad \cdot \left(n^{-1} \sum_{i=1}^{n} \Psi_i'[\hat{\varepsilon}_{in}^2 - \hat{\sigma}_n^2]\right) \Big/ \left(n^{-1} \sum_{i=1}^{n} (\hat{\varepsilon}_{in}^2 - \hat{\sigma}_n^2)^2\right). \end{aligned} \tag{a.8}$$

The estimates of $\alpha_1, \ldots, \alpha_{K(K+1)/2}$ from the regression (2) may be written in mean-deviation form as

$$\hat{\alpha}_n = \left(n^{-1} \sum_{i=1}^{n} (\Psi_i - \hat{\Psi}_n)'(\Psi_i - \hat{\Psi}_n)\right)^{-1}\left(n^{-1} \sum_{i=1}^{n} (\Psi_i - \hat{\Psi}_n)'[\hat{\varepsilon}_{in}^2 - \hat{\sigma}_n^2]\right).$$

Since

$$n^{-1} \sum_{i=1}^{n} \Psi_i'[\hat{\varepsilon}_{in}^2 - \hat{\sigma}_n^2] = n^{-1} \sum_{i=1}^{n} (\Psi_i - \hat{\Psi}_n)'[\hat{\varepsilon}_{in}^2 - \hat{\sigma}_n^2],$$

(a.8) becomes

$$\begin{aligned} &nD_n(\hat{\beta}_n, \hat{\sigma}_n^2)'\tilde{B}_n^{-1}D_n(\hat{\beta}_n, \hat{\sigma}_n^2) \\ &= n\hat{\alpha}_n'\left(n^{-1} \sum_{i=1}^{n} (\Psi_i - \hat{\Psi}_n)'(\Psi_i - \hat{\Psi}_n)\right)\hat{\alpha}_n \Big/ \left(n^{-1} \sum_{i=1}^{n} (\hat{\varepsilon}_{in}^2 - \hat{\sigma}_n^2)^2\right). \end{aligned} \tag{a.9}$$

Except for the proportionality factor n, the right-hand side of (a.9) is easily recognized to be in the form for the constant-adjusted R^2 given by Theil [**21**, p. 176, equation (4.4)]. Hence

$$nD_n(\hat{\beta}_n, \hat{\sigma}_n^2)'\tilde{B}_n^{-1}D_n(\hat{\beta}_n, \hat{\sigma}_n^2) = nR^2 \overset{A}{\sim} \chi^2_{K(K+1)/2}$$

where R^2 is the constant-adjusted squared multiple correlation coefficient for the regression (2). *Q.E.D.*

THEOREM 3: *Given Assumptions* 1, 2(*a*), *and* ε_i *independent of* X_i,

$$|\hat{\sigma}_n^2(X'X/n) - \bar{V}_n| \xrightarrow{\text{a.s.}} 0$$

if and only if

$$\left| n^{-1} \sum_{i=1}^{n} (\sigma_i^2 - \bar{\sigma}_n^2)(E(X_i'X_i) - \bar{M}_n) \right| \to 0.$$

PROOF: (i) By the triangle inequality

$$|\hat{\sigma}_n^2(X'X/n) - \bar{V}_n| \leqslant |\hat{\sigma}_n^2(X'X/n) - \bar{\sigma}_n^2\bar{M}_n| + |\bar{\sigma}_n^2\bar{M}_n - \bar{V}_n|.$$

Since $\bar{\sigma}_n^2$ and $\bar{M}_n$ are uniformly bounded by Assumption 2(a) and since $|\hat{\sigma}_n^2 - \bar{\sigma}_n^2| \xrightarrow{\text{a.s.}} 0$ and $|(X'X/n) - \bar{M}_n| \xrightarrow{\text{a.s.}} 0$ by Markov's strong law of large numbers, Lemma 3.2 of White [**23**] implies $|\hat{\sigma}_n^2(X'X/n) - \bar{\sigma}_n^2\bar{M}_n| \xrightarrow{\text{a.s.}} 0$. With X_i and ε_i independent,

$$\left| n^{-1} \sum_{i=1}^{n} (\sigma_i^2 - \bar{\sigma}_n^2)(E(X_i'X_i) - \bar{M}_n) \right| = |\bar{V}_n - \bar{\sigma}_n^2\bar{M}_n| \to 0$$

by hypothesis. Hence, $|\hat{\sigma}_n^2(X'X/n) - \bar{V}_n| \xrightarrow{\text{a.s.}} 0$.

(ii) Suppose

$$\left| n^{-1} \sum_{i=1}^{n} (\sigma_i^2 - \bar{\sigma}_n^2)(E(X_i'X_i) - \bar{M}_n) \right| \to 0$$

does not hold. Then there exists $\delta > 0$ such that if n is any natural number there is a natural number $m(n) \geqslant n$ such that

$$|\bar{V}_{mjk} - \bar{\sigma}_m^2\bar{M}_{mjk}| \geqslant \delta$$

for some indices $j, k \in \{1, \dots, K\}$. Given Assumptions 1 and 2(a), there exists $n_0(\delta/2)$ such that

$$\left| \hat{\sigma}_n^2 n^{-1} \sum_{i=1}^{n} X_{ij}X_{ik} - \bar{\sigma}_n^2\bar{M}_{njk} \right| < \delta/2$$

almost surely for all $n \geqslant n_0(\delta/2)$. The triangle inequality implies

$$\left| \hat{\sigma}_n^2 n^{-1} \sum_{i=1}^{n} X_{ij}X_{ik} - \bar{V}_{njk} \right| \geqslant \left| \bar{V}_{njk} - \bar{\sigma}_n^2\bar{M}_{njk} \right| - \left| \hat{\sigma}_n^2 n^{-1} \sum_{i=1}^{n} X_{ij}X_{ik} - \bar{\sigma}_n^2\bar{M}_{njk} \right|.$$

In view of the above, there exist $m(n) \geqslant n \geqslant n_0(\delta/2)$ and $\delta > 0$ such that, almost surely,

$$\left| \hat{\sigma}_m^2 m^{-1} \sum_{i=1}^{m} X_{ij}X_{ik} - \bar{V}_{mjk} \right| \geqslant \delta/2$$

so that $|\hat{\sigma}_n^2(X'X/n) - \bar{V}_n| \xrightarrow{\text{a.s.}} 0$ cannot hold, a contradiction. *Q.E.D.*

REFERENCES

[**1**] BERNDT, E. K., B. H. HALL, R. E. HALL, AND J. A. HAUSMAN: "Estimation and Inference in Nonlinear Structural Models," *Annals of Economic and Social Measurement*, 3/4 (1974), 653–665.

[2] CHEW, VICTOR: "Covariance Matrix Estimation in Linear Models," *Journal of the American Statistical Association*, 65 (1970), 173–181.

[3] CHUNG, K. L.: *A Course in Probability Theory*. New York: Harcourt, Brace and World, 1968.

[4] EICKER, F.: "Asymptotic Normality and Consistency of the Least Squares Estimators for Families of Linear Regressions," *Annals of Mathematical Statistics*, 34 (1963), 447–456.

[5] ———: "Limit Theorems for Regressions with Unequal and Dependent Errors," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1. Berkeley: University of California Press, 1967, pp. 59–82.

[6] GLEJSER, H.: "A New Test for Heteroskedasticity," *Journal of the American Statistical Association*, 64(1969), 316–323.

[7] GODREY, L. G.: "Some Tests for Specification Error," unpublished paper, University of York, February, 1978.

[8] GOLDFELD, S. M., AND R. E. QUANDT: "Some Tests for Homoskedasticity," *Journal of the American Statistical Association*, 60 (1965), 539–559.

[9] ———: *Nonlinear Methods in Econometrics*. Amsterdam: North Holland Publishing Co., 1972.

[10] GRILICHES, ZVI: "Cost Allocation in Railroad Regulation," *The Bell Journal of Economics and Management Science*, 3 (1972), 26–41.

[11] HARTLEY, H. O., J. N. K. RAO, AND G. KIEFER: "Variance Estimation with One Unit Per Stratum," *Journal of the American Statistical Association*, 64(1969), 841–851.

[12] HARVEY, A. C.: "Estimating Regression Models with Multiplicative Heteroskedasticity," *Econometrica*, 44(1976), 461–466.

[13] HAUSMAN, J. A.: "Specification Tests in Econometrics," *Econometrica*, 46 (1978), 1251–1272.

[14] HILDRETH, C., AND J. P. HOUCK: "Some Estimators for a Linear Model with Random Coefficients," *Journal of the American Statistical Association*, 68(1968), 584–595.

[15] JENNRICH, ROBERT I.: "Asymptotic Properties of Non-linear Least Squares Estimators," *Annals of Mathematical Statistics*, 40 (1969), 633–643.

[16] MALINVAUD, E.: *Statistical Methods of Econometrics*. Amsterdam: North Holland Publishing Co., 1970.

[17] MCELROY, F. W.: "A Necessary and Sufficient Condition that Ordinary Least-Squares Estimators Be Best Linear Unbiased," *Journal of the American Statistical Association*, 62(1967), 1302–1304.

[18] RAO, C. R.: "Estimation of Heteroskedastic Variances in Linear Models," *Journal of the American Statistical Association*, 65(1970), 161–172.

[19] ———: *Linear Statistical Inference and Its Applications*. New York: John Wiley and Sons, 1973.

[20] RUTEMILLER, H. C., AND D. A. BOWERS: "Estimation in a Heteroskedastic Regression Model," *Journal of the American Statistical Association*, 63(1968), 552–557.

[21] THEIL, H.: *Principles of Econometrics*. New York: John Wiley and Sons, 1971.

[22] WHITE, HALBERT: "Using Nonlinear Least Squares to Approximate Unknown Regression Functions," University of Rochester Department of Economics Discussion Paper, Rochester, April, 1978.

[23] ———: "Nonlinear Regression on Cross-Section Data," *Econometrica*, 48 (1980), 721–746.

[24] ———: "Using Least Squares to Approximate Unknown Regression Functions," *International Economic Review*, forthcoming.