Journal of Economic and Social Research 2 (1) 2000, 1-11

The Inconsistency of the Breusch-Pagan Test

Asad Zaman[1]

Abstract. In regression models, Breusch-Pagan (BP) is a widely used test for heteroskedasticity. Breusch and Pagan (1979) derived this as a Lagrange Multiplier test for the null of homoskedasticity versus the general alternative that error variances depend in a (possibly nonlinear) way on some regressors $Z_2,\ldots,Z_m$. Shown here is that the apparent generality of the BP test is illusory. The test is consistent if and only if there is some *linear* relationship between the error variances and the regressors. This characterization of the asymptotic properties of the Breusch-Pagan test is based on a new necessary and sufficient condition for consistency of the F-test in linear regression models. This general result is also of interest in its own right.

JEL Classification Codes: C10, C12.

Key Words: F-test; Breusch-Pagan Test; Heteroskedasticity Tests; Test Consistency; Test Inconsistency.

1. Introduction

In a classic article, Breusch and Pagan (1979) introduced a Lagrange Multiplier test for heteroskedasticy which appears to allow for very general types of alternatives; see also Bickel (1978). Specifically, in a regression model $y_t = x'_t \beta + \varepsilon_t$, where $Var(\varepsilon_t) = \sigma^2_t = f(\gamma_0 + \gamma' z_t)$, Breusch and Pagan give a test of the null hypothesis $H_0: \gamma = 0$ for arbitrary smooth functions f. The object of this paper is to show that this apparent generality is an illusion and that the test is consistent only for $f(x)=x$, the identity function. Nonlinear functions f are tested for as alternatives *only* to the extent that they are correlated with the regressors z. In particular, for any

[1] Lahore University of Management Science, Lahore, PAKISTAN

non-zero value of γ such that $Cov(f(\gamma_0+ \gamma' z_t), z_t) = 0$, the Breusch-Pagan test has no power asymptotically (i.e. it is inconsistent).

In order to obtain this result, first a general necessary and sufficient condition for the consistency of the *F*-test in regression models is developed. This result is new and provides a powerful tool for examining test consistency for a large class of tests.

2. The Consistency of the F Test

In this section, the consistency of the *F*-test in linear regression models allowing for substantial misspecification is presented. Almost every symbol to follow will depend on the sample size *T*, but it will be notationally convenient to suppress this dependence.

Suppose $y = y(T)$ is a $T \times 1$ vector of observations on a dependent variable and that we wish to "explain" *y* by means of the regressors $1 = 1(T)$, a $T \times 1$ vector of *1*'s, and a $T \times K$ matrix $X = X(T)$ of observations on the independent variables. It will be convenient, and entail no loss of generality, to assume that the regressors are in the form of differences from means, so that $X' 1 = 0$. Define $P = X (X'X)^{-1} X'$ and $Q = \boldsymbol{I} - \{P+(1/T)11'\}$ and let $d_P = K$ be the rank of *P* and $d_Q = T - (K+1)$ be the rank of *Q*. Denoting by $\boldsymbol{X}$ the vector space spanned by the columns of *X* and by $\boldsymbol{Q}$ the vector space orthogonal to $\boldsymbol{X}$ and *1*, note that *Py* is the projection of *y* onto $\boldsymbol{X}$ and *Qy* is the projection of *y* onto $\boldsymbol{Q}$.

In the linear regression model $y = \beta_0 1 + X\beta + \varepsilon$, the standard *F* statistic for testing the null hypothesis $\beta = 0$ can be written as

$$F = \frac{||Py||^2 / d_p}{||Qy||^2 / d_q} = \frac{y' P^* y}{y' Q^* y}$$

where $P^* = P / d_P$ and $Q^* = Q / d_Q$. Effectively, the *F* statistic compares the average projection of *y* on $\boldsymbol{X}$ to the average projection of *y* on $\boldsymbol{Q}$. Here the word "average" indicates that the squared length of the projection is divided by the dimension of the space on which the projection is made.

In order to allow for misspecification, we assume that $y = \mu + \varepsilon$, where $\mu = \mu(T)$ is the $T \times 1$ (nonstochastic) mean vector of the dependent variable and that the $T \times 1$ vector of errors $\varepsilon = \varepsilon(T)$ satisfies the condition that follows. There exists a constant B such that for all T and all $T \times T$ projection matrices M, the following inequality holds:

$$Var(\varepsilon' M\varepsilon) \leq B \ tr(M) \qquad (1)$$

Hypothesis (1) is a mild assumption that will typically be satisfied by most error sequences. Lemma 1 below gives one set of sufficient conditions which ensures (1). All proofs are given in the appendix.

Lemma 1 *If a sequence of errors $\varepsilon_1, \varepsilon_2, \ldots$ (A) forms a martingale, and (B) for all i,j, $Var(\varepsilon_i, \varepsilon_j) \leq B < \infty$, then it also satisfies (1).*

Intuitively speaking, the F test is designed to assess whether the regressors X have a significant relationship with μ, the mean of y. The theorem below gives necessary and sufficient conditions for the consistency of F test.

Theorem 1 *Suppose $y = \mu + \varepsilon$ where ε satisfies the condition 1 given above. Then, the F test rejects the null with probability one if*

$$\lim_{T\to\infty} \frac{\mu' P^* \mu}{1 + \mu' Q^* \mu} = \infty \qquad (2)$$

For the converse, let $\Sigma = \Sigma(T)$ be the covariance matrix of the errors ε, and suppose that $0 < m < \lambda_{min}(\Sigma)$; that is, the smallest eigenvalue of the covariance matrix of the errors is bounded away from zero for all T. If the F test rejects the null with probability one then condition (2) must hold.

The theorem states that, under mild assumptions, the F test for the significance of a set of regressors X will reject the null with probability one if and only if the average projection of the mean vector μ of y on the space $\boldsymbol{X}$ is substantially larger (infinitely larger asymptotically) than its average projection on $\boldsymbol{Q}$, the space orthogonal to X and 1. Because the assumptions made on the error term are very weak, the theorem demonstrates a very strong robustness property of the F-test. Evaluated in terms of test consistency, the test detects the presence of a linear relationship between X and μ if and only if this relationship is much (asymptotically infinitely) stronger than the linear relationship between μ and any omitted variables. Since the assumptions on the error term are very weak (and allow for substantial autocorrelation and heteroskedasticity), this theorem shows that the F-test is robust to such misspecifications with respect to the property of test consistency. This is quite different from *size robustness* properties -- it is well known that the size of the F test is sensitive to such misspecifications. See Zaman (1996, Section 8.10) for some discussion and references regarding the size robustness of F tests.

An important implication of this theorem is that asymptotic local power (as measured by Pitman Efficiency, for example), may not imply test consistency. Suppose, for example, that $y_t = f(\alpha + \beta' x_t) + \varepsilon_t$ for $t=1,2,\ldots,T$, where ε_t is an i.i.d. sequence of errors with common distribution $N(0,\sigma^2)$. If f is any smooth function, it is easily seen that the Lagrange Multiplier test of the null hypothesis H_0: $\beta = 0$ is equivalent to the overall F statistic for the linear regression $y_t = \alpha + \beta x_t + \varepsilon_t$; see Section 11.4.2 of Zaman (1996) for a derivation. Thus the Lagrange Multiplier principle suggests that the overall F statistic for a linear regression tests for the presence of *any* smooth relationship between the regressors and the dependent variable. However, our characterization of the consistency of the F-test shows that the F-test will “detect” (i.e. be consistent for) nonlinear relationships f only to the extent that $f(\alpha + \beta' x_t)$ is linearly correlated with x_t. In particular, if a non-zero value of β is such that $Cov\,(f(\alpha + \beta' x_t), x_t) = 0$, then the usual F test will be unable to reject the null hypothesis that $\beta = 0$ even asymptotically.

3. The Inconsistency of the Breusch-Pagan Test

Derived here is the inconsistency of the Breusch-Pagan test as a consequence of our Theorem 1. Suppose $y_t = \beta' x_t + \varepsilon_t$, where ε_t are independent $N(0, \sigma_t^2)$. It will be convenient to adopt the following notational conventions:

• $[a_t]$ refers to a $T \times 1$ column vector with t-th element a_t.

• $[b_{ij}]$ refers to a $T \times T$ matrix with (i,j) entry b_{ij}.

Let e_t be the residuals from an OLS regression. Breusch and Pagan (1979) derived the Lagrange Multiplier (LM) test of the null hypothesis $H_0\!: \gamma = 0$ given that $\sigma_t^2 = f(\alpha + \gamma' z_t)$ for some smooth function f; here z_t is the t-th row of the matrix of regressors Z being tested as explanatory variables for fluctuations in the variance σ_t^2. Koenker (1981) showed that the original LM statistic is very sensitive to the assumption of normality, while the asymptotically equivalent statistic, based on the TR^2 of the (auxiliary) regression of $[e_t^2]$ on a constant and Z, remains robust to non-normality. Since the overall F statistic for a regression is a monotonic transform of the TR^2, it is clear that using the overall F for the auxiliary regression will be asymptotically equivalent to the Breusch-Pagan test. Assume, without loss of generality, that the regressors $Z = Z(T)$ have been differenced from their means so that $Z'1 = 0$. Let $P = Z(Z'Z)^{-1}Z'$ be the matrix of the projection onto the column space of Z and let Q be the matrix of the projection onto the vector space orthogonal to 1 and the columns of Z.

Theorem 2 *Assume that the variances σ_t^2 are bounded above: for all t, $\sigma_t^2 \le M < \infty$. Then the Breusch-Pagan test rejects the null hypothesis of homoskedasticity with probability one if*

$$\lim_{T\to\infty} \frac{[\sigma_t^2]' P * [\sigma_t^2]}{1 + [\sigma_t^2]' Q * [\sigma_t^2]} = \infty \qquad (3)$$

The converse also holds if the variances are bounded away from zero for all t, $\sigma_t^2 > c > 0$.

Thus, consistency of the Breusch-Pagan test requires that the average projection of the vector $[\sigma_t^2]$ of variances on the column space of the regressors Z to be large relative to the average projection on the orthogonal complement of this space. This shows that the Breusch-Pagan test only detects linear relationships between the variables Z and the vector of variances $[\sigma_t^2]$.

A Random Coefficients Example: Suppose that x_t is i.i.d. $N(0,1)$ and $y_t = a_t x_t$ where a_t is *i.i.d.* $N(a,1)$. This random coefficient model can be rewritten as $y_t = a x_t + \varepsilon_t$, where $\sigma_t^2 = Var(\varepsilon_t | x_t) = x_t^2$. By letting

$f(x) = x^2$, it is clear that $\sigma_t^2 = f(a+b\, x_t)$ when $a = 0$ and $b = 1$. Then, let $[e_t^2]$ be the vector of squared residuals from an OLS regression of y on x, and use the F statistic for the regression of $[e_t^2]$ on 1 and x to test for heteroskedasticity. In this case, our earlier results imply that this test will not reject the null, since x^2 and x are uncorrelated (because $x \sim N(0,1)$). This test is asymptotically equivalent to the Breusch-Pagan test, and hence the Breusch-Pagan will also not reject the null asymptotically. Thus, the Breusch-Pagan test fails to detect this nonlinear relationship. For some finite sample properties of the Breusch-Pagan test for random parameter variation and comparisons with a modified likelihood ratio test, see Zaman (1997).

Appendix

Here Theorems 1 and 2 and also Lemma 1, which is useful in verifying condition (1) for error sequences, are proven.

Proof of Lemma 1: Let $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_T)$ be a sequence of random variables satisfying properties (A) and (B) of Lemma 1, and let Σ be the $T \times T$ covariance matrix of the vector ε. We will use Σ_{ij} for the (i,j) entry of the matrix Σ. The martingale property ensures that $\Sigma_{ij} = 0$ if $i \neq j$.

Let P be any idempotent matrix. We aim to show that $Var(\varepsilon' P \varepsilon) \leq 2\, B\, tr\, P$. To prove this, note that

$$
\begin{aligned}
Var(\varepsilon' P \varepsilon) &= E\, \{tr(\varepsilon\, \varepsilon' - \Sigma)P\}^2 \\
&= \boldsymbol{\Sigma}_{\mathbf{i,j}}\, \boldsymbol{\Sigma}_{\mathbf{k,l}}\, E([\varepsilon_i\varepsilon_j - \sigma_{i,j}]\ [\varepsilon_k\varepsilon_l - \sigma_{k,l}]\ P_{i,j}\, P_{k,l}) \\
&= \boldsymbol{\Sigma}_{\mathbf{i,j}}\, E(\varepsilon_i\varepsilon_j - \sigma_{i,j})^2\, P_{i,j}^{\ 2} \\
&\leq 2B\, \boldsymbol{\Sigma}_{\mathbf{i,j}}\, P_{i,j}^{\ 2} = 2\, B\, tr(P'P) = 2\, B\, tr\, P.
\end{aligned}
$$

In this derivation, we have used the fact that if ε_t forms a martingale, then the term $E\,(\varepsilon_i\varepsilon_j - \sigma_{i,j})(\varepsilon_k\varepsilon_l - \sigma_{k,l}) = 0$ unless $i = j$, $k = l$ or else $i = k$, $j = l$.

Proof of Theorem 1: Define $\Delta = 1 + \mu' Q^* \mu$, $N = \Delta^{-1} (y' P^* y)$, and $D = \Delta^{-1} (y' Q^* y)$. It is immediately seen that $F = N / D$. We will show that ED converges to a strictly positive quantity and $Var(D)$ goes to 0. From this it follows that the convergence of F to $+\infty$ is equivalent to the convergence of N to $+\infty$. Then, we will show that N goes to infinity if and only if the hypothesis of the theorem holds.

Step 1: $E D = \Delta^{-1} (\mu' Q^* \mu + tr\, \Sigma Q^*)$. Now $tr\, \Sigma Q^* \leq M\, tr\, Q^* = M$. It is easily deduced that ED is bounded away from $+\infty$. If the variances Σ_{tt} are greater than $m > 0$, then ED is also bounded away from 0, since $tr\, \Sigma Q^* \geq m\, tr\, Q^* = m$. Next we will show that $Var(D) \rightarrow 0$.

To prove this, first note that $Var(X+Y) \leq 2\, Var(X) + 2\, Var(Y)$. Now, $Var(\Delta^{-1} (y'Q^*y) = \Delta^{-2}\, Var(2\varepsilon'Q^*\mu + \varepsilon'Q^*\varepsilon) \leq 2\Delta^{-2}(2\, Var(\varepsilon'v) + Var(\varepsilon'\, Q^*\, \varepsilon))$, where $v = Q^* \mu$. It is clear that $Var\, (\varepsilon'v) \leq B\, ||v||^2$ where B_2 is the upper bound on the error variances. The assumption (1) permits us to conclude that $Var(\varepsilon'Q^*\varepsilon) = Var(\varepsilon'Q\varepsilon) / d_Q^2 \leq B(tr\, Q) / d_Q^2 = B_4 / d_Q$. We thus conclude that

$$Var(D) \leq 2\, \Delta^{-2} (2\, B_2\, ||Q^*\mu||^2 + B_4 / d_Q).$$

Now $||Q^*\mu||^2 = \mu'Q\mu / d_Q^2 = \mu'Q^*\mu / d_Q$ so that $Var(D) \leq (4B/d_Q)\, (\mu'Q^*\mu + (1/2))/(1+ \mu'Q^*\mu)^2$. This will converge to 0 as d_Q goes to infinity.

Step 2: We will now show that $EN \rightarrow +\infty$. Also, if we define $S_N^2 = Var(N)$, then both $S_N \rightarrow +\infty$ and $EN/S_N \rightarrow +\infty$. From these facts we can conclude that $N \rightarrow +\infty$ with probability one as follows below. First, we need to show that for any (large, positive) constant k, $P(N > k)$ converges to one. Note that $P(N > k) = P(((N\text{-}EN) / S_N) > (k\text{-}EN) / S_N)$. Then, let $X = (N\text{-}EN) / S_N$. If S_N and EN/S_N both go to $+\infty$, then $(k\text{-}EN) / S_N \rightarrow -\infty$ so that the probability in question converges to $P(X > -\infty)$. Since X has mean 0 and variance 1, this probability converges to unity by, for example, Chebyshev's Inequality. Remaining to be shown is that EN and EN / S_N converge to $+\infty$.

First, consider that $EN = \Delta^{-1} (\mu'P^*\mu + tr\, \Sigma P^*)$. Ignoring the term $tr\, \Sigma P^*$ which is positive, the remaining term is assumed to converge to $+\infty$ as the main hypothesis of the theorem we are proving. Next, consider that $Var(N) \leq \Delta^{-2} (4B / d_P)(\mu'P^*\mu + (1/2))$ following the same logic as for $Var(D)$. With $S_N^2 = Var(N)$ it follows that

$$EN / S_N \geq 2\sqrt{B / d_p}\,\frac{\mu' P^* \mu + tr\ \Sigma P^*}{(\mu' P^* \mu + (1/2))^{1/2}}$$

It is clear that this goes to $+\infty$ provided that $\mu' P^* \mu$ does too, which is entailed by the hypothesis of the theorem.

Conversely, suppose that the hypothesis of the theorem does not hold. It is immediately seen that EN fails to go to $+\infty$. Since D is bounded away from 0 and $N = 0$, it is also seen that F cannot converge to $+\infty$ with probability one.

Proof of Theorem 2:

Step 1: Define $M_T = M = \{\boldsymbol{\Sigma_t}\, x_t x_t'\}^{-1}$. A key quantity which occurs in the proof is $a_t = x'_t\, M_T\, x_t$. We will need to bound this as seen below. The largest possible projection of the vector $a = [a_t] = (a_1, \ldots, a_T)'$ is onto itself, so that $||a'Pa|| \leq \boldsymbol{\Sigma_t}\, a_t^2$. To bound this, note that

$$\boldsymbol{\Sigma_t}\, a_t = \boldsymbol{\Sigma_t}\, tr\, x_t'\, M_T\, x_t = tr\, (\boldsymbol{\Sigma_t}\, x_t x_t')\, M_T = tr\, M_T^{-1}\, M_T = K$$

Since $a_t \geq 0$, it follows that $a_t \leq K$ so that $1 = \boldsymbol{\Sigma_t}\, (a_t / K) \geq \boldsymbol{\Sigma_t}\, a_t^2 / K^2$. This implies that $||a||^2 \leq K^2$. From this it follows that $0 \leq a' P^*\, a \leq K^2 / d_P$ and that $0 \leq a'\, Q^*\, a \leq K^2 / d_Q$.

Step 2: To prove the theorem, showing that $[e_t^2]\, W [e_t^2]$ behaves asymptotically similar to $[\varepsilon_t^2]\, W [\varepsilon_t^2]$ for the matrices $W = P^*$ and $W = Q^*$ suffices, since applying Theorem 1 to the second form yields the result immediately. We will therefore analyze the difference between the two quadratic forms and show that they remain bounded asymptotically. From this the result will follow. Define $z_t = X'\, W\, x_t$ and note that:

$$e_t^2 = (\varepsilon_t - x_t'\, W\, X\, \varepsilon)^2 = (\varepsilon_t - z_t'\varepsilon)^2 = \varepsilon_t^2 - 2\, \varepsilon_t\, (z_t'\varepsilon) + (z_t'\varepsilon)^2.$$

Thus, the difference D that we wish to show is asymptotically negligible can be written as

$$\begin{aligned} D &= [e_t^2]\, W [e_t^2] - [\varepsilon_t^2]\, W [\varepsilon_t^2] \\ &= [(z_t'\varepsilon)^2]\, W [(z_t'\, \varepsilon)^2] - 4 [(z_t'\, \varepsilon)^2]\, W [\varepsilon_t\, (z_t'\varepsilon)] + 2\, [(z_t'\varepsilon)^2]\, W [\varepsilon_t^2] \end{aligned}$$

$$+4\ [\varepsilon_t(z_t'\varepsilon)]\,W[\varepsilon_t(z_t'\varepsilon)] - 4\ [\varepsilon_t(z_t'\varepsilon)]\,W[\varepsilon_t^2].$$

Now, we will show that each of the five terms in the difference converge in quadratic mean to zero asymptotically. This will prove the result.

Consider that the first term $T_1 = [(z_t'\varepsilon)^2]\ W\,[(z_t'\varepsilon)^2]$. Note that

$$ET_1 = tr\ E[(z_t'\varepsilon)^2][(z_t'\varepsilon)^2]'W = tr\ [E\ (z_i'\varepsilon)^2\ (z_j'\varepsilon)^2]\ W \le \boldsymbol{\Sigma_t}\ E\ (z_t'\ \varepsilon)^4$$

The last inequality follows from Amemiya's Lemma, according to which $tr\ AB \le (tr\ A)\ \lambda_{max}(B)$ when A and B are positive semi definite matrices. Since $W = P,\ Q$ are projection matrices, the largest eigenvalue is 1. Since $z_i'\varepsilon$ is normal, its fourth moment is just 3 times its variance so that

$$ET_1 = 3\ \boldsymbol{\Sigma_t}\ z_t'\Sigma z_t\,.$$

Now Σ is a diagonal matrix with elements bounded above by $M < \infty$ and below by $m > 0$. It follows that the difference between $z_t'\Sigma z_t$ and $a_t = z_t'z_t$ is bounded, and since $\boldsymbol{\Sigma_t}\ a_t = K$ as established in Step 1, we conclude that $|ET_1 - K| \le C$, and it follows that T_1 cannot go to infinity. We make a similar, but more complex calculation, to show that the variance of T_1 is similarly bounded:

$$Var\ (T_1) = E\{\ tr\ (\ [(z_t'\varepsilon)^2]\ [(z_t'\varepsilon)^2]' - E[(z_t'\varepsilon)^2]\ [(z_t'\varepsilon)^2]')W\}^2$$

$$= \boldsymbol{\Sigma_{i,j=1}\Sigma_{k,l=1}}\ Cov((z_i'\varepsilon)^2(z_j'\varepsilon)^2,\ (z_k'\varepsilon)^2\ (z_l'\varepsilon)^2)W)_{ij}W_{kl}\,.$$

To calculate the required covariance, we need the following formula: if X_1, X_2, X_3, X_4 are jointly normal, and $\sigma_{i,j} = Cov(X_i,X_j)$, then

$$\begin{aligned} EX_1^2X_2^2X_3^2X_4^2 \quad &= \sigma_{11}\ \sigma_{22}\ \sigma_{33}\ \sigma_{44} \\ &+2\sigma_{11}\sigma_{22}\sigma_{34}^2 + 2\sigma_{11}\sigma_{23}^2\sigma_{44} + 2\sigma_{12}^2\sigma_{33}\sigma_{44} + 2\sigma_{13}^2\sigma_{22}\sigma_{44} + 2\sigma_{11}\sigma_{24}^2\sigma_{33} \\ &+4\sigma_{12}^2\sigma_{34}^2 + 4\sigma_{13}^2\sigma_{24}^2 + 4\sigma_{14}^2\sigma_{23}^2 \\ &+8\sigma_{11}\sigma_{23}\sigma_{24}\sigma_{34} + 8\sigma_{22}\sigma_{13}\sigma_{14}\sigma_{34} + 8\sigma_{33}\sigma_{12}\sigma_{14}\sigma_{34} + 8\sigma_{44}\sigma_{12}\sigma_{13}\sigma_{23} \\ &+16\sigma_{12}\sigma_{34}\sigma_{13}\sigma_{24} + 16\sigma_{12}\sigma_{34}\sigma_{14}\sigma_{23} + 16\sigma_{13}\sigma_{24}\sigma_{14}\sigma_{23}\,. \end{aligned}$$

By Cauchy-Schwartz, we have $\sigma_{ij}^2 \le \sigma_{ii}\sigma_{jj}$. It follows from the formula that $E\ X_1^2X_2^2X_3^2\ X_4^2\ \le 100\sigma_{11}\sigma_{22}\sigma_{33}\sigma_{44}$ so that

$$Var(T_1) \le \boldsymbol{\Sigma_{i,j}\Sigma_{k,l}}\ (z_i'\Sigma z_i)\ (z_j'\Sigma z_j)\ (z_k'\Sigma z_k)\ (z_l'\Sigma z_l)W_{ij}W_{kl}$$

$$=(\boldsymbol{\Sigma}_{\mathbf{i,j=1}}\, (z_i'\Sigma z_i)(z_i'\Sigma z_j) W_{ij})^2 .$$

This is bounded by the square of the mean.

Completing the proof requires showing that the other four terms in the difference of $[\varepsilon_t^2]W[\varepsilon_t^2]$ and $[e_t^2]W[e_t^2]$ remain bounded by a finite quantity with probability 1. These follow exactly the same procedure outlined above, and in fact are slightly easier. Thus, these proofs are omitted for brevity.

Now consider the effect of replacing W by the matrix P^*. The difference between $[\varepsilon_t^2]W[\varepsilon_t^2]$ and $[e_t^2]W[e_t^2]$ for $W = P^*$ remains bounded by a finite quantity with probability 1 asymptotically. Thus, convergence to infinity of one of the terms is equivalent to convergence to infinity of the other. For $W = Q^* = Q / d_Q$ in the denominator, the difference is again bounded by a finite quantity. Dividing by d_q which goes to infinity, makes the difference go to zero asymptotically. This means that the hypothesis of the theorem applied to $[e_t^2]$ gives the same results as the hypothesis applied to $[\varepsilon_t^2]$. This is what we desire to prove.

References

Bickel, P. J. (1978) "Using Residuals Robustly {I}: Tests for Heteroscedasticity, Nonlinearity." *The Annals of Statistics.* 6:266-291.

Breusch, T. and Pagan, A. (1979) "A Simple Test of Heteroskedasticity and Rrandom Coefficient Variation." *Econometrica* 47:1287-1294.

Koenker, R. (1981) "A Note on Studentizing a Test for Heteroskedasticity." *Journal of Econometrics* 17:107-112.

Zaman, A. (1996) *Statistical Foundations for Econometric Techniques*. New York: Academic Press.

Zaman, A. (1997) "Random Coefficients and Heteroskedasticity." Technical Report 98-4, Economics Department, Bilkent University, 06533 Ankara, Turkey.