working paper
department
of economics

AN ASYMMETRIC LEAST SQUARES TEST
OF HETEROSCEDASTICITY

by
Whitney K. Newey
Princeton University
and
James L. Powell
M.I.T.

March 1984

massachusetts
institute of
technology
50 memorial drive
cambridge, mass. 02139

AN ASYMMETRIC LEAST SQUARES TEST OF HETEROSCEDASTICITY

by
Whitney K. Newey
Princeton University
and
James L. Powell
M.I.T.

March 1984
Working Paper #337

AN ASYMMETRIC LEAST SQUARES TEST OF HETEROSCEDASTICITY[1/]

by

Whitney K. Newey and James L. Powell

1. Introduction

For the linear regression model, several tests of the null hypothesis of homoscedastic disturbances against the alternative of heteroscedasticity of linear form have recently been investigated. One category of tests use the residuals from a preliminary fit of the regression equation of interest; this group includes the tests proposed and studied by Anscombe [1961], Glejser [1969], Goldfeld and Quandt [1972], Harvey [1976], Godfrey [1978], Bickel [1978], Breusch and Pagan [1979], and White [1980]. In their simplest form, these tests of homoscedasticity are tests that the coefficients of a second-stage regression of the squared values of the residuals (or monotonic transformations of them, e.g., the absolute residuals) on transformations of the regressors are zero.

An alternative approach has been studied by Koenker and Bassett [1982a]; the test they propose is based upon the regression analogues of order statistics, termed "regression quantiles," introduced in Koenker and Bassett [1978]. For this test, the null hypothesis of homoscedasticity is rejected if the slope coefficients of the regression equation, estimated at different quantiles of the distribution of the dependent variable, are significantly different from each other. Comparing the asymptotic efficiency of this test relative to a corresponding "squared residual regression" test, the authors found some inefficiency of the "regression quantiles" test when the error distribution is Gaussian, but this conclusion was reversed for contaminated Gaussian error distributions, and the efficiency gains of the latter test appeared to be substantial even for low levels of contamination.

Koenker and Bassett's approach to heteroscedasticity testing differs from the "squared residual regression" methods in two important respects. First, estimation of regression parameters and testing for heteroscedasticity are unified in the regression quantile approach; by comparing regression coefficients estimated at different quantiles, the question of heteroscedasticity is recast as a question concerning differences in alternative measures of "location" of the conditional distribution of the dependent variable. Also, the regression quantile test involves "robust" estimation of these location measures, so the precision of this statistical procedure depends on the ability to identify the percentiles of the error distribution, rather than on moments of the error terms. It is not immediately apparent which of of these two characteristics of their test -- the "specification test" form (as in Hausman [1978]) or the robustness of quantile estimation -- is primarily responsible for the superior performance of the regression quantiles test for nonnormal errors, and the object of the present study is to determine which of these factors predominates. To this end, a least squares analogue of the regression quantile test is investigated, in order to focus on the relationship of heteroscedasticity to the divergence of location measures.

In the following section, these least squares analogues of regression quantiles, termed "asymmetric least squares" estimators, are defined, and their asymptotic behavior under local heteroscedasticity and a general error distribution is investigated in Section 3 below. The general results are specialized in Section 4 to the contaminated normal error distributions considered by Koenker and Bassett, and the efficiency of the proposed test is determined, relative to the regression quantile test and tests using magnitudes of least squares residuals, for various degrees of contamination. A comparison of these tests for a specific empirical example follows, and the paper concludes with some qualitative observations about the applicability of the procedures considered.

2. Definition of the Asymmetric Least Squares Estimators

The observable data $\{(y_i, x_i'): i = 1, \ldots, n\}$, are assumed to be generated by the linear model

$$(2.1) \qquad y_i = x_i'\beta_o + u_i,$$

where $\{x_i\}$ is a set of fixed regression vectors of dimension p with first component $x_{i1} \equiv 1$, β_o is a conformable vector of unknown error terms, and $\{u_i\}$ is a set of mutually independent scalar error terms. A convenient specification of heteroscedasticity in this context is of the form

$$(2.2) \qquad u_i = \sigma_i \varepsilon_i,$$

where the $\{\varepsilon_i\}$ are independent and identically distributed error terms with distribution function $F(\lambda)$ and $\{\sigma_i\}$ is a sequence of unknown scale parameters; the null hypothesis of homoscedasticity thus specifies the $\{\sigma_i\}$ to be constant across i.

The "regression quantile" estimators of the parameter vector β_o, proposed by Koenker and Bassett [1978a], are defined as those vectors $\tilde{\beta}(\theta)$ which minimize the function

$$(2.3) \qquad Q_n(\beta;\ \theta) \equiv \sum_{i=1}^{n} \rho_\theta(y_i - x_i'\beta)$$

over β in R^p for fixed values of θ in $(0, 1)$, where $\rho_\theta(\cdot)$ is a convex loss function of the form

$$(2.4) \qquad \rho_\theta(\lambda) \equiv |\theta - 1(\lambda < 0)|\cdot|\lambda|,$$

with "$1(A)$" denoting the indicator function of the event "A". Under the null hypothesis of homoscedasticity (and with additional regularity conditions cited below), the probability limits of the regression quantile estimators $\{\tilde{\beta}(\theta)\}$ for different choices of θ deviate from β_o only in their intercept terms; that is, when $\sigma_i \equiv 1$,

$$(2.5) \qquad \operatorname*{plim}_{n\to\infty} \tilde{\beta}(\theta) = \beta_o + \eta(\theta)e^1,$$

where e^1 is the first standard basis vector for R^p and $\eta(\theta) \equiv F^{-1}(\theta)$, the quantile function for the error terms $\{\varepsilon_i\}$. Under the heteroscedastic alternative, the probability limits for the slope coefficients (when they exist) will in general also vary with θ, with differences depending on the relationship between the scale parameters $\{\sigma_i\}$ and the regressors $\{x_i\}$.

The regression quantile estimators are thus a class of empirical "location" measures for the dependent variable whose sampling behavior involves the true regression coefficients and the stochastic behavior of the error terms; their "robustness" follows from the absolute error loss component of the criterion function (2.4). To obtain a similar class of location measures which do not involve robustness considerations, we consider replacing the "check function" criterion of (2.4) with the following "asymmetric least squares" loss function:

$$(2.6) \qquad \rho_\tau(\lambda) \equiv |\tau - 1(\lambda < 0)| \cdot \lambda^2, \quad \text{for } \tau \text{ in } (0, 1).$$

The corresponding class of asymmetric least squares estimators $\{\hat{\beta}(\tau)\}$ are defined to minimize

$$(2.7) \qquad R_n(\beta; \tau) \equiv \sum_{i=1}^{n} \rho_\tau(y_i - x_i'\beta)$$

over β, for $\rho_\tau(\cdot)$ given in (2.6).2/

To determine the class of location parameters which are estimated by $\{\hat{\beta}(\tau)\}$, consider the scalar parameter $\mu(\tau)$ which minimizes the function $E[\rho_\tau(\varepsilon_1 - m) - \rho_\tau(\varepsilon_1)]$ over m, where the expectation is taken with respect to the distribution of the i.i.d. residuals $\{\varepsilon_i\}$, which is assumed to have finite mean and to be absolutely continuous with respect to Lebesgue measure. The parameter $\mu(\tau)$ is easily shown to be a solution of the equation

$$(2.8) \qquad \mu(\tau) = \frac{(1-\tau)E[\varepsilon_1 \cdot 1(\varepsilon_1 < \mu(\tau))] + \tau E[\varepsilon_1 \cdot 1(\varepsilon_1 \geqq \mu(\tau))]}{(1-\tau)F(\mu(\tau)) + \tau[1 - F(\mu(\tau))]}$$

$$= \alpha_\tau E[\varepsilon_1 | \varepsilon_1 < \mu(\tau)] + (1 - \alpha_\tau)E[\varepsilon_1 | \varepsilon_1 \geqq \mu(\tau)],$$

where

$$(2.9) \qquad \alpha_\tau \equiv (1-\tau)F(\mu(\tau))[(1-\tau)F(\mu(\tau)) + \tau[1 - F(\mu(\tau))]]^{-1};$$

thus the parameter $\mu(\tau)$, hereafter referred to as the "τ^{th} weighted mean," summarizes the distribution of the error terms in much the same way that the quantile function $\eta(\theta) \equiv F^{-1}(\theta)$ does. When the error terms $\{u_i\}$ of the linear model (2.1) are homoscedastic, then, we will have

$$\text{(2.10)} \qquad \operatorname*{plim}_{n\to\infty} \hat{\beta}(\tau) = \beta_0 + \mu(\tau)e^1$$

under suitable regularity conditions, while for heteroscedastic disturbances the slope coefficients in large samples can be expected to vary with the index τ as well; nonzero differences in the slope coefficients of $\hat{\beta}(\tau)$ across τ can thus be taken as evidence of heteroscedasticity.

Of course, the loss function given in (2.6) is not the only one which could be used to construct a "specification test" of heteroscedasticity; if, for example, the median and the mean of the error distribution differed, a similar test could be constructed using the loss function

$$\text{(2.11)} \qquad \rho_\theta \equiv |\lambda|^\theta,$$

where slope coefficient estimates would be compared when a function like $Q_n(\cdot)$ of (2.3) is minimized for θ set equal to one and two. One reason to focus on the loss function $\rho_\tau(\lambda)$ of (2.6) is that the corresponding "location" function $\mu(\tau)$ will be an invertible function of τ even if the error terms are symmetrically distributed, due to the asymmetry of the loss function itself.[3] Hence the test which uses the "asymmetric least squares" loss function can be expected to have reasonable power regardless of the shape of the error density.[4]

An additional advantage of the asymmetric least squares estimators relative to regression quantiles is that the loss function $\rho_\tau(\lambda)$ is continuously differentiable in λ, so the estimators $\hat{\beta}(\tau)$ can be computed as iterative weighted least squares estimators, i.e., as solutions to the equations

$$(2.12) \qquad \hat{\beta}(\tau) = [\sum_{i=1}^{n} |\tau - 1(y_i < x_i'\hat{\beta}(\tau))| x_i x_i']^{-1} \cdot \sum_{i=1}^{n} |\tau - 1(y_i < x_i'\hat{\beta}(\tau))| x_i y_i .$$

(Note that the classical least squares estimator is a special case, corresponding to $\tau = 1/2$). Furthermore, and perhaps more importantly, consistent estimation of the asymptotic covariance matrix of the $\{\hat{\beta}(\tau)\}$ under the null hypothesis does not require estimation of the density function of the error terms, as shown below; unlike the regression quantile test of homoscedasticity, then, the test statistic using the asymmetric least squares estimators will involve no "smoothing" of the empirical distribution or quantile function of the estimated residuals.[5/] These convenient properties of the asymmetric least squares test, along with its relatively favorable performance in the efficiency comparisons of Section 4, suggest that it merits consideration for use in practice, and should not simply be regarded as a standard of comparison for the other tests of heteroscedasticity considered below.

3. Large Sample Properties of the Estimators and Test Statistic

The asymptotic theory for the asymmetric least squares estimators and test for heteroscedasticity will be developed under the following assumptions:

Assumption E: The error terms $\{\varepsilon_i\}$ of (2.2) are i.i.d. with $E(\varepsilon_1)^{2+\zeta} < \infty$ for some $\zeta > 0$, and have distribution function $F(\lambda)$ which is continuous with density $f(\lambda)$.

Assumption R: The regression vectors $\{x_i\}$ have $x_{1i} \equiv 1$, and $n^{-1} \sum_i x_i x_i' \to D$, where D is positive definite.

Assumption S: The scale terms $\{\sigma_i\}$ of (2.2) are of the form $\sigma_i = 1 + x_i'\gamma_n$, where $\gamma_n = \gamma_o/\sqrt{n}$ for some fixed p-dimensional vector γ_o.

These conditions are identical to Assumptions A1 to A3 of Koenker and Bassett [1982a], except that our condition $E(\varepsilon_1)^{2+\eta} < \infty$ replaces their assumption of uniform positivity of the error density $f(\lambda)$ in λ. Following Koenker and Bassett, then, we consider only local heteroscedasticity which is linear in the regression vector x_i, and restrict attention to the case of fixed dimension for the unknown parameter vectors β_o and γ_o, assumed functionally independent (note, though, that Assumption S does not restrict the exogenous variables upon which σ_i depends, because of the possibility of redefining the original regression vector x_i).

For a vector $(\tau_1, \tau_2, \ldots, \tau_m)'$ of weights ordered so that $0 < \tau_1 < \ldots < \tau_m < 1$, let $\mu_o \equiv (\mu(\tau_1), \ldots, \mu(\tau_m))'$ be the corresponding vector of τ_i^{th} weighted means (as defined in (2.8)), and define the $(m{\times}p)$-dimensional vectors $\hat{\xi} \equiv \text{vec}[\hat{\beta}(\tau_1), \ldots, \hat{\beta}(\tau_m)]$ and $\xi_o \equiv (\iota_m \otimes \beta_o) + (\mu_o \otimes e^1)$, where ι_m is an m-vector of ones. Then we have the following result, analogous to Theorem 3.1 of Koenker and

Bassett [1982a]:

Theorem 1: Under the linear heteroscedastic model (2.1) and (2.2) with Assumptions E, R, and S, the random vector $\hat{\xi}$ of asymmetric least squares estimators is asymptotically normal,

$$\sqrt{n}\,(\hat{\xi} - \xi_0) \overset{d}{\to} N(\mu_0 \otimes \gamma_0, \Sigma \otimes D^{-1}),$$

where the matrix Σ has elements $\sigma_{ij} \equiv c(\tau_i, \tau_j)/d(\tau_i)\cdot d(\tau_j)$, for $c(\tau, \theta) \equiv E\{|\tau - 1(\varepsilon_1 < \mu(\tau))|\cdot|\theta - 1(\varepsilon_1 < \mu(\theta))|(\varepsilon_1 - \mu(\tau))(\varepsilon_1 - \mu(r))\}$ and $d(\tau) \equiv \tau[1 - F(\mu(\tau))] + (1 - \tau)F(\mu(\tau))$.

Proof: Our argument follows the approach taken by Ruppert and Carroll [1980], Koenker and Bassett [1982a], and Powell [1983], among others. Consider the random function

$$(3.1) \qquad M_n(\delta, \tau) \equiv \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i \psi_\tau((1 + x_i'\gamma_n)\varepsilon_i - x_i'\delta - \mu(\tau)),$$

where

$$(3.2) \qquad \psi_\tau(\lambda) \equiv |\tau - 1(\lambda < 0)|\cdot\lambda \ .$$

The behavior of this function is of interest because, letting $\hat{\delta}(\tau) \equiv \hat{\beta}(\tau) - \beta_0 - \mu(\tau)\cdot e^1$, we can write

$$(3.3) \qquad M_n(\hat{\delta}(\tau), \tau) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i \psi_\tau(y_i - x_i'\hat{\beta}(\tau))$$

$$= 0,$$

the latter equality holding because $\hat{\beta}(\tau)$ minimizes $R_\tau(\beta)$ of (2.7).

Assumption R ensures that $\max_i \|x_i\|/\sqrt{n} \to 0$; thus, it can be shown that, for any $L > 0$ (and $\gamma_n = \gamma_0/\sqrt{n}$),

$$(3.4) \qquad \sup_{\|\delta\| \leq L/\sqrt{n}} \| M_n(\delta, \tau) - M_n(0, \tau) - E[M_n(\delta, \tau) - M_n(0, \tau)] \| = o_p(1)$$

by a straightforward modification of the proof of Lemma 4.1 of Bickel [1975]. Since

$$(3.5) \qquad E[\psi_\tau((1 + x_i'\gamma_n)\varepsilon_i - x_i'\delta - \mu(\tau))] = \mu(\tau)x_i'\gamma_n - d(\tau)x_i'\delta$$

$$+ (1 - 2\tau)\int_{\mu(\tau)}^{(x_i'\delta+\mu(\tau))(1+x_i'\gamma_n)^{-1}} [(1 + x_i'\gamma_n)\lambda - x_i'\delta - \mu(\tau)]dF(g)$$

(where $d(\tau)$ is defined in the statement of Theorem 1), it is evident that $E[M_n(0, \tau)] = O(x_i'\gamma_n) = o(1)$ and that

$$(3.6) \qquad E[M_n(\delta, \tau)] = D[\mu(\tau)\gamma_o - d(\tau)\cdot(\sqrt{n}\delta)] + o(1)$$

when $\delta = O(n^{-1/2})$. Equations (3.3), (3.4), (3.6), and the monotonicity of $\psi_\tau(\lambda)$ imply

$$(3.7) \qquad \sqrt{n}(\hat{\delta}(\tau) \equiv \sqrt{n}(\hat{\beta}(\tau) - \beta_o - \mu(\tau)e^1) = O_p(1)$$

by Lemma 5.1 of Jureckova [1977], so $\sqrt{n}\hat{\delta}(\tau)$ satisfies the asymptotic linearity relationship

$$(3.8) \qquad \sqrt{n}\hat{\delta}(\tau) = \mu(\tau)\gamma_o - [d(\tau)D]^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i\psi_\tau(\varepsilon_i - \mu(\tau)) + o_p(1).$$

Finally, writing $\sqrt{n}(\hat{\xi} - \xi_o) = \sqrt{n}\ \mathrm{vec}[\hat{\delta}(\tau_i), \ldots, \hat{\delta}(\tau_m)]$, the result of Theorem 1 follows from application of the Lindeberg-Feller central limit theorem to the right-hand side of (3.8).

In order for the result of Theorem 1 to be useful in constructing a large sample test of the null hypothesis of homoscedasticity, a consistent estimator of the asymptotic covariance matrix $\Sigma \otimes D^{-1}$ must be provided. Unlike the regression quantile estimators, natural estimators of the components of Σ can be constructed using the asymmetric least squares residuals

$$(3.9) \qquad \hat{u}_i(\tau) \equiv y_i - x_i'\hat{\beta}(\tau) .$$

Define

$$(3.10) \qquad \hat{c}(\tau, \theta) \equiv \frac{1}{n} \sum_{i=1}^{n} \psi_\tau(\hat{u}_i(\tau)) \cdot \psi_\theta(\hat{u}_i(\theta)) ,$$

for $\psi_\tau(\lambda)$ defined in (3.2) and

$$(3.11) \qquad \hat{d}(\tau) \equiv \frac{1}{n} \sum_{i=1}^{n} |\tau - 1(\hat{u}_i(\tau) < 0)| ;$$

these are just the sample analogues of the population moments defining $c(\tau, \theta)$ and $d(\tau)$ of Theorem 1.

Theorem 2: Under the conditions of Theorem 1, the estimators $\hat{c}(\tau_i, \tau_j)$ and $\hat{d}(\tau_i)$ are consistent for $i, j = 1, \ldots, m$; that is, $\hat{c}(\tau_i, \tau_j) - c(\tau_i, \tau_j) = o_p(1)$ and $\hat{d}(\tau_i) - d(\tau_i) = o_p(1)$.

Proof: Only consistency of $\hat{d}(\tau)$ will be shown here; consistency of $\hat{c}(\tau, \theta)$ can be shown analogously. Writing

$$(3.12) \qquad \hat{d}(\tau) - d(\tau) = \frac{1}{n} \sum_{i=1}^{n} [|\tau - 1(\varepsilon_i < \mu(\tau))| - E|\tau - 1(\varepsilon_i < \mu(\tau)|]$$
$$+ \frac{1}{n} \sum_{i=1}^{n} [|\tau - 1(\hat{u}_i(\tau) < 0)| - |\tau - 1(\varepsilon_i < \mu(\tau)|],$$

the first term in this sum converges to zero in probability by Tchebyshev's inequality. Thus

(3.13) $$|\hat{d}(\tau) - d(\tau)| \leqq$$

$$|2\tau - 1| \frac{1}{n} \sum_{i=1}^{n} 1(|\varepsilon_i - \mu(\tau)| \leqq |x_i'(\hat{\delta}(\tau) - \mu(\tau)\gamma_n)/(1 + x_i'\gamma_n)|)$$

$$+ o_p(1),$$

for $\hat{\delta}(\tau)$ defined in (3.7). By Theorem 1 and Assumptions R and S,

(3.14) $$(1 + x_i'\gamma_n)^{-1}(x_i'\hat{\delta}(\tau) + \mu(\tau)x_i'\gamma_n) = o_p(1),$$

and since all moments of the indicator function $1(|\varepsilon_1 - \mu(\tau)| \leqq d)$ are $O(d)$ by the continuity of the error distribution, the sum in (3.13) converges to zero in probability, again by Tchebyshev's inequality.

With the results of Theorems 1 and 2, a large sample χ^2 test of the null hypothesis of homoscedasticity using the vector $\hat{\xi}$ of asymmetric least squares estimators can be constructed. The definition of ξ_0 implies a set of linear restrictions $H\xi_0 = 0$, where the transformation H yields a vector of slope coefficients corresponding to pairwise differences of the vectors $\beta_0 + \mu(\tau_i)e^1$ and $\beta_0 + \mu(\tau_{i+1})e^1$. As Koenker and Bassett [1982a] point out, the matrix H can be written in the form

(3.15) $$H \equiv \Delta \otimes \Psi,$$

where Δ is an $(m - 1)\times m$ differencing matrix with i,j^{th} element equal to $\delta_{i,j} - \delta_{i,j-1}$, and Ψ is a $(p - 1)\times p$ selection matrix with k,l^{th} element $\delta_{k,l-1}$, for $\delta_{i,j}$ the Kronecker delta. The test of heteroscedasticity can thus be based upon the statistic

$$(3.16) \qquad \sqrt{n}H\hat{\xi} \equiv \Psi[n(\hat{\beta}(\tau_2) - \hat{\beta}(\tau_1), \ldots, \hat{\beta}(\tau_m) - \hat{\beta}(\tau_{m-1})) ;$$

this statistic is asymptotically normal under the conditions of Theorem 1, with zero mean only when $\gamma_0 = 0$. The asymmetric least squares test of homoscedasticity thus uses the test statistic

$$(3.17) \qquad T_{LS} \equiv n(H\hat{\xi})'[H(\hat{\Sigma} \otimes \hat{D}^{-1})H']^{-1}(H\hat{\xi}),$$

where $\hat{\Sigma}$ is computed using (3.10) and (3.11) and

$$(3.18) \qquad \hat{D} \equiv \frac{1}{n}\sum_{i=1}^{n} x_i x_i' .$$

Corollary 1: Under the conditions of Theorem 1, the test statistic T_{LS} defined in (3.17) has a limiting noncentral χ^2 distribution with $(m - 1)\cdot(p - 1)$ degrees of freedom and noncentrality parameter

$$\delta_{LS} \equiv (\Delta\mu_0)'(\Delta\Sigma\Delta')^{-1}(\Delta\mu_0)[(\Psi\gamma_0)'(\Psi D^{-1}\Psi')(\Psi\gamma_0)]$$

$$\equiv \kappa_{LS}\,[(\Psi\gamma_0)'(\Psi D^{-1}\Psi')(\Psi\gamma_0)] .$$

Again, this result parallels that of Theorem 4.1 of Koenker and Bassett [1978a]; this test and the corresponding regression quantile test are both consistent (i.e., for fixed significance level their power tends to one as $\|\Psi\gamma_0\| \to \infty$), and the regression quantile test statistic has a limiting noncentrality parameter of the form

$$(3.19) \qquad \delta_{RQ} \equiv \kappa_{RQ}\,[(\Psi\gamma_0)'(\Psi D^{-1}\Psi')(\Psi\gamma_0)] ,$$

where the scalar κ_{RQ} involves the differences $[\eta(\theta_i) - \eta(\theta_{i-1})]$ of quantiles and the precision with which these differences are estimated. Thus the relative efficiency of the two tests is governed by the relative accuracy of estimation of weighted means versus quantiles, a property which is exploited in the following section.

4. Asymptotic Relative Efficiencies of Alternative Tests

In this section, the efficiency comparisons of the regression quantiles test to a squared residual regression test made by Koenker and Bassett [1982a] are extended to the asymmetric least squares test. In this context, Koenker and Bassett's original calculations are revised; due to an algebraic error (described below), their Figures 1 and 2 give a misleading depiction of the relative performance of the tests for the class of nonnormal error distributions they considered.

Following Koenker and Bassett's setup, we consider the two-parameter class of contaminated Gaussian distributions, with cumulative distributions of the form

$$(4.1) \qquad F(\lambda|\alpha, \sigma) = (1 - \alpha)\Phi(\lambda) + \alpha\Phi(\lambda/\sigma),$$

for $\Phi(\cdot)$ denoting the standard normal cumulative and for α in the interval $(0, 1)$. For this class of distributions, the τ^{th} weighted mean satisfies

$$(4.2) \qquad \mu(\tau) = \frac{(2\tau - 1)[(1 - \alpha)\phi(\mu(\tau)) + (\alpha/\sigma)\phi(\mu(\tau)/\sigma)]}{\tau + (1 - 2\tau)[(1 - \alpha)\Phi(\mu(\tau)) + \alpha\Phi(\mu(\tau)/\sigma)]},$$

for $\phi(\cdot)$ the standard normal density function. To conform to Koenker and Bassett's framework, we consider only the efficiency of the asymmetric least squares test using a single difference of symmetrically chosen weights, i.e., a test based upon $\hat{\beta}(\tau) - \hat{\beta}(1 - \tau)$, for

$\frac{1}{2} < \tau < 1$. For this test, the scalar κ_{LS} of Corollary 1 above can be shown (after some tedious algebra) to be

$$(4.3) \qquad \kappa_{LS} = 2[\mu(\tau)]^2[d(\tau)]^2 \cdot$$

$$[(1-2\tau)\{1 - (1-\alpha)\Phi(\mu(\tau)) - \alpha\sigma\Phi(\mu(\tau)/\sigma)\} - (1-\tau)[\mu(\tau)]^2]^{-1}.$$

The corresponding regression quantile test uses $\tilde{\beta}(\theta) - \tilde{\beta}(1-\theta)$, the difference in symmetric regression quantiles; for θ in $(\frac{1}{2}, 1)$, the corresponding scalar κ_{RQ} governing the power of the regression quantile test is given by Koenker and Bassett to be

$$(4.4) \qquad \kappa_{RQ} = 2[\eta(\theta)]^2[(1-\alpha)\phi(\eta(\theta)) + (\alpha/\sigma)\phi(\eta(\theta)/\sigma)]^2/(1-\theta)(2\theta-1) .$$

Koenker and Bassett compared the scalar κ_{RQ} with the corresponding term κ_{SR} for a heteroscedasticity test using squared residuals from a preliminary least squares fit of equation (2.1), a test closely related to those investigated by Breusch and Pagan [1979] and White [1980]. More generally, tests for heteroscedasticity can be based on the sample correlation of $\rho(\hat{u}_i)$ with the regressors x_i, where $\hat{u}_i h = y_i - x_i'\hat{\beta}(.50)$ is the least squares residual and $\rho(\cdot)$ is an even function. To obtain a test with more asymptotic power than the squared residual regression test for (heavy-tailed) nonnormal disturbances, we might choose, say, $\rho(u) = |u|^\alpha$ for $1 \leq \alpha < 2$ rather than $\alpha = 2$.

The test statistic for this type of test is

$$(4.5) \qquad T_\rho = nR_\rho^2 ,$$

the sample size n times the constant-adjusted R^2 of the regression of $\rho(\hat{u}_i)$ on x_i. Bickel [1978] has obtained the asymptotic properties of such tests when it is assumed that $\gamma_0 = \beta_0$, for γ_0 defined in Assumption S, but his results can be extended to the more general linear scale model of Assumption S. With some additional regularity conditions (such as the boundedness of $E[\rho(\varepsilon_1)]^2$) which can be verified for the cases we consider here, the test statistic T_ρ of (4.5) can be shown to have a limiting noncentral chi-square distribution with $(p - 1)$ degrees of freedom and noncentrality parameter

$$\text{(4.6)} \qquad \delta_\rho \equiv [E(\rho'(\varepsilon_1)\varepsilon_1]^2[\mathrm{Var}(\rho(\varepsilon_1))]^{-1}(\Psi\gamma_0)'(\Psi D^{-1}\Psi')(\Psi\gamma_0)$$

$$\equiv \kappa_\rho \, (\Psi\gamma_0)'(\Psi D^{-1}\Psi')(\Psi\gamma_0)$$

under the conditions of Theorem 1.

In our application we focus attention on the squared residual regression test $(\rho(u) = u^2$, versions of which have been considered by White [1980] and Breusch and Pagan [1979]) and the more "robust" test which uses absolute residuals (i.e., $\rho(u) = |u|$, as in Glejser [1969] and Bickel [1978]). For the former test, the scalar $\kappa_\rho \equiv \kappa_{SR}$ is

$$\text{(4.7)} \qquad \kappa_{SR} = 4[3(1 + \alpha(\sigma^4 - 1))/(1 + \alpha(\sigma^2 - 1))^2 - 1]^{-1}$$

when the errors are contaminated Gaussian, while for the latter test, $\kappa_\rho \equiv \kappa_{AR}$ for this distribution is

$$\text{(4.8)} \qquad \kappa_{AR} = [\pi(1 + \alpha(\sigma^2 - 1))/2(1 + \alpha(\sigma - 1))^2 - 1]^{-1} .$$

The local power of the squared residual regression, absolute residual regression, regression quantile, and asymmetric least squares tests may be compared by computing their Pitman asymptotic relative efficiencies (AREs); since the limiting degrees of freedom for all of these test statistics are equal, these AREs are just the ratios of the respective noncentrality parameters, which in turn reduce to the ratios of the respective κ coefficients. However, the noncentrality parameters of the regression quantiles and asymmetric least squares tests depend upon the particular weights (θ and τ, respectively) chosen. Rather than considering the AREs for these tests for a range of weights, we consider only the weights $[1 - \theta, \theta] = [.14, .86]$ for the regression quantiles test and $[1 - \tau, \tau] = [.42, .58]$ for the asymmetric least squares test. These values of θ and τ were selected after a preliminary calculation of the weights which maximized the respective noncentrality parameters in a grid search for each α and σ considered; the results of this optimization are given in Table 1. As the table shows, the optimal θ values are typically between .75 and .90, and decrease as α and σ increase (although there is a sharp reversal in this pattern for values of α near .50). The optimal values of τ for the asymmetric least squares test are usually between .51 and .75, and also typically decrease with increasing α and σ. The average of the optimal θ values is .86 and the average of the optimal τ values is .58.

It is important to note that the value of the noncentrality parameter is usually quite insensitive to moderate perturbation of the weights from their optimal values. For example, for the regression quantiles test, when $\alpha = .05$ and $\sigma = 5$, use of $\theta = .86$ rather than the optimal $\theta = .89$ results in an efficiency loss of only 3 percent (though for $\alpha = 0$, the efficiency loss rises to 18 percent, with optimal $\theta = .93$).

Table 2 gives the AREs of the regression quantile, asymmetric least squares, and absolute residual regression tests, all relative to the squared residual regression test. One striking feature of this table is the nearly identical performance of the absolute residual regression test and the asymmetric least squares test. The ARE of the asymmetric least squares test never differs from the ARE of the absolute residual regression test by more than two percent. Also, both of these tests are more efficient than the squared residuals regression test except when α and σ are large (or when $\alpha = 0$). The ARE of the asymmetric least squares test is small for $\sigma = 2$, but increases substantially as σ increases.

Another interesting feature of Table 2 is the behavior of the AREs of the regression quantile test. For $\sigma = 2$ the squared residual regression test is always more efficient than the regression quantile test, and for $\sigma = 3$ the asymmetric least squares (or absolute residual regression) test is efficient relative to the regression quantile test. For $\sigma = 4$ and $\sigma = 5$, the regression quantiles test is the most efficient of all tests considered when α is between 5 and 20 percent; for $\sigma = 4$, however, its efficiency gain over the asymmetric least squares test is not particularly large, amounting, for example, to 28 percent at $\alpha = .10$.

These results on the ARE of the regression quantile test relative to the squared residual regression test are quite different from those reported in Koenker and Bassett [1982a]. For example, when $\theta = .75$, the relative scale σ is five, and there is 20 percent contamination, we find the ARE of the regression quantile test to be 1.64, rather than the "40+" figure reported previously. This difference is explained by an error in equations (4.12) and (4.14) of Koenker and Bassett

[1982];[6/] the term corresponding to κ_{SR} in these expressions is "$4[\mathrm{Var}(\varepsilon_1{}^2)]^{-1}$" instead of the correct $\kappa_{SR} = 4[E(\varepsilon_1)^2]^2[\mathrm{Var}(\varepsilon_1{}^2)]^{-1}$. This omission overstates the ARE of the regression quantile test for $\sigma > 1$, particularly when the contamination percentage α is large; hence the "iso-efficiency" contours of Figures 1 and 2 of Koenker and Bassett [1982a] should actually be shifted upward and "U" shaped, with the ARE of the regression quantile test sharply declining as the distinction between the "contaminating" and "contaminated" distributions of the error vanishes.

It should be noted, however, that for sufficiently large σ and sufficiently small α, dramatic efficiency gains of the regression quantiles test to the other procedures are attainable. For example, for $\alpha = .0125$ and $\sigma = 10$, the ARE of the regression quantile test is 20.80, over twice as large as that for the asymmetric least squares and absolute residual regression tests. This value rises to 62.42 when σ increases to 50, representing a twenty-fold improvement over the other procedures; the improvement, though, drops off quite rapidly as α increases. Thus the regression quantile test should perform very well for large data sets which contain a few sizable outliers (perhaps due to keypunching errors).

5. A Numerical Example Revisited

Turning now to a more practical comparison of the performance of the tests considered above, we consider the food expenditure/income example of Koenker and Bassett [1982a]. A surprising feature of their analysis of this example was that, while the scatter diagram of the observations and fitted regression quartile lines suggested heteroscedastic disturbances (with scale increasing as income increased), the regression quantile test they preformed could not reject the null hypothesis of homoscedasticity at an (asymptotic) 5% level. As the results below demonstrate, this inability to reject homoscedasticity is due to the lack of precision of the quartile estimation for these data; the asymmetric least squares slope coefficients, while exhibiting less movement across the τ weights considered, do provide stronger evidence of the heteroscedasticity suggested by casual inspection of the data.

Of the 235 budget surveys considered by Koenker and Bassett, only 224 of the observations were readily available for our calculations; on the (reasonable) presumption that this subsample is representative of the entire sample, we have computed the asymmetric least squares estimators for weights $\tau = .42$, $.50$, and $.58$. The classical least squares estimate of the slope coefficient $(\tau = .50)$ is .847, which approximates the regression quantile slope for the lower quartile. The asymmetric least squares slope estimates for $\tau = .42$ and $\tau = .58$ are .841 and .854, respectively; the iterative procedure described in Section 2 above was used to calculate these estimates, and converged in four iterations for both weighting factors. Calculation of $\hat{d}(\tau)$ and $\hat{c}(\tau, \theta)$ of section 4 for this example produced the following estimate of the Σ matrix for $(\hat{\beta}(.42), \hat{\beta}(.50), \hat{\beta}(.58))$:

$$\hat{\Sigma} = \begin{bmatrix} 2.00 & 1.92 & 1.85 \\ 1.92 & 1.87 & 1.81 \\ 1.85 & 1.81 & 1.77 \end{bmatrix} \times 10^{-2} . \tag{5.1}$$

Thus, while the differences in the asymmetric least squares slope estimates are an order of magnitude smaller than the corresponding quartile estimates, the asymmetric least squares estimates display higher correlation across weights, so that their differences (used to construct the heteroscedasticity test statistic) are very precisely estimated.

The dependent variable x_i (log income) has

$$(5.2) \qquad \sum_{i=1}^{n} (x_i - \bar{x})^2 = 42.76$$

for this example, so setting

$$(5.3) \qquad \Delta = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} , \quad \hat{\xi}' = (.841, .847, .854) ,$$

the test statistic of (3.17) is calculated to be $T_{LS} = 10.517$ here; since the upper 1% critical value of a χ_2^2 random variable is 9.21, the null hypothesis of homoscedasticity is rejected at standard significance levels using the asymmetric least squares test.

In order to include the absolute and squared residual regression tests in this comparison, we also consider the regression quantile and asymmetric least squares tests which use

$$(5.4) \qquad \Delta = [1 \quad 0 \quad -1]$$

as the difference matrix, i.e., the regression quantile test based upon $\tilde{\beta}(.25) - \tilde{\beta}(.75)$ and the asymmetric least squares test using $\hat{\beta}(.42) - \hat{\beta}(.58)$. With the covariance matrix estimate for "moderate" smoothing reported by Koenker and Bassett, the regression quantile test statistic is computed to be 3.644, which again is not significant (relative to a χ_1^2 distribution) at the 5% level. On the other hand, the corresponding asymmetric least squares test statistic is 9.898, with marginal significance level less than 1%. The absolute residual regression and squared residual regression test statistics -- n times the constant-adjusted R^2 of the regression of the absolute or squared residuals on x_i and a constant -- are 10.494 and 13.304, respectively. Thus application of these more common procedures provides even stronger evidence of heteroscedasticity for this example.

6. Conclusions

From the results of Section 4, we conclude that, on grounds of asymptotic relative efficiency, the asymmetric least squares or absolute residual regression tests are preferred to the regression quantile test for contaminated Gaussian distributions with small to moderate relative scale, while the regression quantile test would be preferred if the percent contamination is not large but relative scale is large. The similar performance of the absolute residual regression test -- a "robust" version of the squared residual regression test -- and the asymmetric least squares test -- a "less robust" version of the regression quantile test -- indicates that the relative performance of "residuals" tests to "location" tests of heteroscedasticity is entirely explained by the robustness of the criterion function involved in the estimation method, and is not a result of the "specification test" form of the regression quantile and asymmetric least squares tests of homoscedasticity. Given the computational convenience of the absolute residuals regression test, and the similar AREs of the absolute residual and asymmetric least squares tests, the former appears to be the preferred test for heteroscedasticity in the presence of contaminated disturbances, except when the contaminating distribution has very long tails, when the regression quantiles test would be preferred.

FOOTNOTES

1/ This research was supported by National Science Foundation Grant SES-8309292 at the Massachusetts Institute of Technology. We are grateful to Anil K. Bera, Jerry Hausman, Roger Koenker, Richard Quandt, and participants at workshops at MIT and the 1983 Econometric Society Winter Meetings for their helpful comments. We also thank Roger Koenker for providing the data used in Section 5 below.

2/ Equivalently, we could define $\rho_{\tau}(\lambda) = [\rho_{\theta}(\lambda)]^2$, for $\rho_{\theta}(\cdot)$ defined in (2.4) and $\tau = \theta^2/[\theta^2 + (1 - \theta)^2]$.

3/ To see this, note that the function $\mu(\tau)$ is differentiable in τ with $d\mu(\tau)/d\tau = E|\varepsilon_1 - \mu(\tau)|\cdot[(1 - \tau)F(\mu(\tau)) + \tau(1 - F(\mu(\tau))] > 0$.

4/ It is also possible that a heteroscedasticity test based upon differences of location measures would reject due to misspecification of the regression function. An interesting avenue of research would be a comparison of the power of such tests to the tests of nonlinearity studied by Bickel [1978], which involve regressions of odd functions of the residuals on transformations of the regressors. It is not clear whether a location measure test would have reasonable power against this latter type of alternative.

5/ However, Koenker and Bassett [1982b] have recently shown how estimation of the density function of the errors can be avoided when testing linear hypotheses concerning $\tilde{\beta}(.50)$ through use of the Lagrange multiplier form of the test statistic, and their approach can apparently be adapted to the present case.

6/ Roger Koenker has pointed out that this error was originally brought to his attention by Alistair Hall at Warwick University.

Table 1

Optimal Values of Regression Quantile (Asymmetric Least Squares)

Weights for Various Contaminated Gaussian Distributions

Contamination Percentage α \ Relative Scale σ	2	3	4	5
.05	.92 (.78)	.90 (.61)	.90 (.51)	.89 (.51)
.10	.91 (.73)	.88 (.52)	.87 (.51)	.87 (.51)
.15	.90 (.70)	.87 (.51)	.85 (.51)	.84 (.51)
.20	.89 (.67)	.85 (.51)	.83 (.51)	.82 (.51)
.25	.88 (.66)	.84 (.51)	.82 (.51)	.80 (.51)
.30	.88 (.66)	.83 (.51)	.80 (.51)	.79 (.51)
.35	.87 (.66)	.82 (.51)	.79 (.51)	.77 (.51)
.40	.87 (.66)	.81 (.51)	.77 (.51)	.76 (.51)
.45	.87 (.66)	.80 (.51)	.97 (.51)	.97 (.51)
.50	.87 (.66)	.97 (.52)	.97 (.51)	.97 (.97)

Table 2

Local Efficiencies of Tests Relative to Squared Residual Regression Test

Relative Scale σ	Contamination Proportion α	Regression Quantile	Asymmetric Least Squares	Absolute Residual Regression
1	-	.57	.88	.88
2	.0125	.66	.99	.98
	.025	.72	1.07	1.06
	.05	.82	1.17	1.16
	.10	.92	1.26	1.25
	.15	.94	1.26	1.26
	.20	.92	1.24	1.24
	.25	.88	1.21	1.20
	.30	.83	1.17	1.17
	.40	.74	1.10	1.09
	.50	.66	1.04	1.03
3	.0125	1.12	1.55	1.54
	.025	1.46	1.86	1.85
	.05	1.79	2.04	2.04
	.10	1.84	1.90	1.90
	.15	1.64	1.68	1.69
	.20	1.40	1.50	1.51
	.25	1.17	1.37	1.37
	.30	.98	1.26	1.26
	.40	.68	1.11	1.11
	.50	.51	1.02	1.02
4	.0125	2.23	2.71	2.70
	.025	2.95	3.08	3.08
	.05	3.30	2.88	2.89
	.10	2.83	2.20	2.21
	.15	2.19	1.76	1.78
	.20	1.67	1.49	1.50
	.25	1.25	1.31	1.32
	.30	.92	1.19	1.20
	.40	.50	1.04	1.04
	.50	.34	.95	.95
5	.0125	4.10	4.29	4.29
	.025	5.07	4.27	4.28
	.05	4.96	3.36	3.39
	.10	3.62	2.22	2.24
	.15	2.54	1.68	1.70
	.20	1.78	1.39	1.41
	.25	1.22	1.22	1.23
	.30	.81	1.10	1.11
	.40	.34	.96	.97
	.50	.26	.89	.90

REFERENCES

Anscombe, F.J. [1961], "Examination of Residuals," Proceedings of the Fifth Berkeley Symposium, 1: 1-36.

Bickel, P.J. [1978], "Using Residuals Robustly I: Tests for Heteroscedasticity and Nonlinearity," Annals of Statistics 6: 266-291.

Breusch, T.S. and A.R. Pagan [1980], "A Simple Test for Heteroscedasticity and Random Coefficient Variation," Econometrica, 47: 1287-1294.

Glejser, H. [1969], "A New Test for Heteroscedasticity," Journal of the American Statistical Association, 64: 316-323.

Godfrey, L.G. [1978], "Testing for Multiplicative Heteroskedasticity," Journal of Econometrics, 8: 227-236.

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

Harvey, A.C. [1976], "Estimating Regression Models with Multiplicative Heteroscedasticity," Econometrica, 44: 461-466.

Hausman, J.A. [1978], "Specification Tests in Econometrics," Econometrica, 46: 1251-1271.

Jureckova, J. [1977], "Asymptotic Relations of M-Estimates and R-Estimates in the Linear Model," Annals of Statistics, 5: 464-472.

Koenker, R. and G. Bassett [1978], "Regression Quantiles," Econometrica, 46: 33-50.

Koenker, R. and G. Bassett [1982a], "Robust Tests for Heteroscedasticity Based on Regression Quantiles," Econometrica, 50: 43-61.

Koenker, R. and G. Bassett [1982b], "Tests of Linear Hypotheses and L_1 Estimation," Econometrica, 50: 1577-1583.

Powell, J.L. [1983], "The Asymptotic Normality of Two-Stage Least Absolute Deviations Estimators," Econometrica, 51: 1569-1575.

Ruppert, D. and R.J. Carroll [1980], "Trimmed Least Squares Estimation in the Linear Model," Journal of the American Statistical Association, 75: 828-838.

White, H. [1980], "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," Econometrica, 48: 817-838.