INFERENCE IN THE PRESENCE OF HETEROSKEDASTICITY

by

Walter Sudmant

B.Sc, Simon Fraser University, 1977

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in the Department

of

Economics

SIMON FRASER UNIVERSITY

July, 1988

ISBN 0-315-48832-8

APPROVAL

Name: Walter Sudmant

Degree: Master of Arts

Title of Thesis: Inference in the Presence of Heteroskedasticity.

Examining Committee:

Chairman: Lawrence A. Boland

Dr. P. Kennedy
Professor
Senior Supervisor

Dr. T. Heaps
Associate Professor

Dr. R. Holmes
Professor
Business Administration
Simon Fraser University
External Examiner

Date Approved: July 19, 1988

Title of Thesis/Project/Extended Essay

Inference in the Presence of Heteroskedasticity.

Author:

(signature)

Walter Sudmant

(name)

20 July 1988

(date)

Abstract

The size and power of small sample inferences using estimated regression coefficients in the presence of heteroskedastic errors are examined by means of Monte Carlo simulations. Six possible procedures are examined under a variety of heteroskedasticity generating processes.

Inferences using ordinary least squares (OLS) coefficients are computed using two possible estimators for the variance-covariance matrix: the jackknife estimate, and an estimate based on a functional form estimate of the error variance matrix. Estimated generalized least squares (EGLS) inferences are made by transforming the data using a variety of conventional functional forms to estimate the standard deviations of the errors. Simulations of EGLS inferences are initially based on the assumption that the heteroskedasticity generating function is known. Further simulations examine cases where the heteroskedasticity generating function is assumed to be unknown. Finally, the inferences arising from two pre-test estimators are examined: the Goldfeld-Quandt test is used to choose between OLS and EGLS estimates of the regression coefficients and the White test is used as a pre-test to select either the standard OLS variance-covariance matrix estimator or the jackknife estimator of the variance.

The results show no major differences in empirical size among the methods examined, and empirical sizes of tests were usually found to be not significantly different from nominal sizes. EGLS techniques were found to yield inferences of higher power than all of the other techniques when the functional form of the heteroskedasticity was assumed known. When an incorrect functional was used to obtain EGLS estimates the power of inferences based on EGLS techniques was reduced substantially, but in many cases EGLS inferences still yielded the inferences of highest power. Inferences based on the jackknife estimate were of superior power only in cases where the functional form chosen for EGLS was very different from the heteroskedasticity generating function.

The major recommendation to researchers indicated by these simulations is to make inferences using EGLS estimates after some preliminary diagnostic work on the OLS residuals whenever there is empirical evidence or a theoretical basis to relate the variances of the errors to an exogenous variable.

TABLE OF CONTENTS

LIST OF TABLES

Introduction

Much of the research concerning heteroskedasticity has been directed toward the detection of heteroskedasticity and subsequent estimation of regression coefficients with improved efficiency. Relatively little research has been done on inference with respect to the regression coefficients in the presence of heteroskedasticity. The consequences of heteroskedasticity for ordinary least squares (OLS) estimation are well known: the estimates of the regression coefficients are inefficient but unbiased, while the estimates of the variance-covariance matrix are biased. The conventional treatment for heteroskedasticity, where the variances of the error terms are not known to a constant of proportionality, is to estimate the regression coefficients using estimated generalized least squares (EGLS). While the improvement in efficiency and mean squared error for the EGLS estimates of the regression coefficents has been well established over a wide range of cases, the researcher wishing to make inferences concerning the regression coefficients is forced to rely upon the asymptotic properties of the EGLS estimates. The usual t-test statistic will only be asymptotically distributed as a t distribution, therefore the nominal size of the test may be incorrect in small samples. Furthermore, the EGLS estimates as well as the small sample properties of their distributions will depend upon the assumptions the researcher makes about the functional form of the heteroskedasticity. If an incorrect functional form is chosen to estimate the variances of the errors, the estimates of the variances as well as the estimate of the variance-covariance matrix may be biased asymptotically. The existing literature offers little reassurance that the inferences arising from EGLS do in fact have the properties of size and power suggested by the asmptotic theory, and there has been no research at all on the properties of inferences based on incorrect functional forms.

Recent research by White (1980), and MacKinnon and White (1985) has suggested that inferences can be obtained using the OLS estimates along with a consistent estimator of the OLS variance-covariance estimator. This technique is attractive in that no assumptions concerning the functional form of the heteroskedasticity are required, hence eliminating the potential for choosing the wrong functional form. Although the inferences still depend upon asymptotic properties, the Monte Carlo studies by MacKinnon and White show that for sample sizes as small as 30 the size of the statistic approximates the nominal critical value rather well.

A disadvantage of MacKinnon and White's technique of inference based on OLS estimates lies in the potential for reduced power of the tests. MacKinnon and White do not examine the power of their tests, nor do they compare their results with EGLS techniques; however, since power increases as the variance of a statistic decreases, and EGLS estimates are known to have lower variances than OLS estimates over a wide range of heteroskedasticity generating processes, it is reasonable to suspect that the cost of eliminating assumptions concerning the functional form of the heteroskedasticity is a reduction in power.

This study will investigate statistical inference concerning regression coefficients in the presence of heteroskedastic errors using Monte Carlo methods. Several possible procedures for model estimation are examined:

1. EGLS

Estimates of the coefficients are obtained by assuming a functional form for the variance of the error. The data are transformed by dividing each observation through by an estimate of the standard deviation of the error, and OLS is computed using the transformed data. Inferences are obtained by computing a t-statistic from the results of OLS computed on the transformed data. Following the terminology of MacKinnon and White, this t-statistic is referred to as a 'quasi t-statistic'.

2. OLS

The problem in obtaining inferences from the OLS estimates in the presence of heteroskedasticity is in estimating the variance-covariance matrix. Two approaches are examined. First, a variation on White's heteroskedasticity-consistent variance-covariance estimator known as the 'jackknife' is used. Second, the matrix of the variances of the errors is estimated using a functional form. This estimate is then substituted into the general formula for the variance-covariance matrix of the coefficient estimates when errors are non-spherical.

3. Pre-test Estimators

The procedures in 1. and 2. above assume that heteroskedasticity is known to exist a priori. In practice, the choice of the estimation technique for either the regression coefficients or the variance-covariance matrix may depend upon a pre-test for homoskedasticity. Two pre-test estimators are examined:

The Goldfeld-Quandt test detects heteroskedasticity related to some exogenous variable. Coefficients are then estimated by EGLS or OLS depending upon the outcome of the test, and inferences are made using the EGLS estimate of the variance-covariance matrix, or the simple OLS variance-covariance estimator as derived from the assumption of homoskedasticity. In the context of inference, both the numerator and denominator of the quasi t-statistic are pre-test estimators.

The White test detects heteroskedasticity only if it is of a form which biases the simple OLS variance-covariance estimator. Therefore when inferences are to be made using OLS coefficient estimates as in 2. above, a natural pre-test estimator for the variance-covariance estimator is the estimator which selects either the heteroskedasticity-consistent variance-covariance estimator when the White test rejects the null hypothesis of homoskedasticity, or the simple OLS variance-covariance matrix estimate when the White test does not reject the hypothesis of homoskedasticity.

Inferences arising from each of the above EGLS, OLS and pre-test procedures are made using a variety of heteroskedasticity generating processes, and a range of sample sizes. Results are presented to suggest conclusions on a number of questions of practical interest to researchers:

1. How closely do the sizes of inferences approximate their nominal sizes using several different critical values, and is there a tendency for the empirical sizes of tests to underestimate or overestimate the nominal size the researcher has selected?

2. At what sample size do the asymptotic properties take effect with regard to the inferences? For example, although Monte Carlo studies (Goldfeld and Quandt, 1974) have found EGLS coefficient estimates to be nearly unbiased in samples as small as T=30, this is no guarantee that valid inferences are possible at such a sample size.

3. Which inference procedures have higher power? In particular, it is hypothesised that EGLS techniques will yield test statistics with higher power due to the increased efficiency of the coefficient estimates.

The conclusions arising from the above questions are compared for different heteroskedasticity generating processes, and finally, the conclusions are tested for their robustness with respect to the choice of functional form of the heteroskedasticity. While the researcher will sometimes have theoretical justification for the choice of a functional form for EGLS, as in the case of the random coefficients model, most often economic theory will only suggest that the variance of the error is related to some exogenous variable, as for example, in the case of the variation in expenditure being positively related to income. Textbooks often provide several alternative functional forms for the heteroskedasticity but little advice on choosing between them in the absence of theory. Therefore inferences arising from EGLS estimates are examined for a number cases in which the functional form chosen is different from the heteroskedasticity generating function.

A Review of the Literature

The research literature on small sample inference in the presence of heteroskedasticity using OLS and EGLS techniques consists mainly of Monte Carlo studies. No exact small sample distribution theory exists for the general case of heteroskedasticity in the linear regression model. Before discussing the Monte Carlo studies, a brief summary of the asymptotic theory is presented.

Assume the linear regression model

$$y = X\beta + \mu,$$

where y is a (T x 1) vector of observations on a dependent variable, X is a (T x k) matrix of observations on independent variables, μ is a (T x k) vector of error terms where

$$E(\mu) = 0,$$

and

$$E(\mu\mu') = \Omega \quad \text{with} \quad \Omega = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_T^2),$$

and not all σ_i^2 are identical.

Then a set of J linear restrictions on the regression coefficients β can be described as

$$R\beta = r,$$

where R is a (J x k) known matrix, and r is a (J x 1) known vector.

Judge et al (1985 p 177) give an asymptotic statistic for testing the linear restrictions $R\beta = r$ when the covariance matrix of the errors is non-spherical and coefficients are estimated by EGLS. For the EGLS estimate β^{egls}, the statistic:

$$L = (\beta^{egls}-\beta)'R'[R(X'\hat{\Omega}^{-1}X)^{-1}R']^{-1}R(\beta^{egls}-\beta)$$

is asymptotically distributed as χ^2 with degrees of freedom equal to the number of restrictions or equivalently, L/J is asymptotically distributed as the $F_{(J,T-k)}$ distribution. For a single restriction, $(\sqrt{L})/(\sqrt{J})$ is then asymptotically distributed as the t distribution with t-k degrees of freedom and is equivalent to the conventional t-statistic usually computed by running OLS on transformed data. The transformation consists of multiplying y and X by P where $(P'P)^{-1} = \hat{\Omega}$.

For the OLS estimate β^{ols}, Judge et al (p 426) present the statistic:

$$(\beta-\beta^{ols})R'[R(X'X)^{-1}(X'\hat{\Omega}X)(X'X)^{-1}R']^{-1}R(\beta-\beta^{ols})$$

asymptotically distributed as χ^2 with degrees of freedom equal to the number of restrictions, where $(X'\hat{\Omega}X)$ is a consistent estimator of $X'\Omega X$. Dividing by the number of restrictions and taking the square root, this can also be expressed as an F-distributed statistic. It is the asymptotic distribution of these two statistics which provides all of the justification for inferences using regression coefficients in the presence of heteroskedasticity. For small samples, the distributions of these statistics are not known, hence they are investigated by means of Monte Carlo studies.

Some Monte Carlo studies have been conducted in an experimental design context, where replicated observations are available for each value of the independent variable. While not directly applicable to most econometric problems, some conclusions of interest can be drawn from these studies.

Deaton (1983) compares the power of EGLS inferences to OLS inferences via a Monte Carlo study where for each value of x_t, the independent variable, there are replications of observations on the dependent variable. Replications vary from 3 to 25. EGLS in this study is defined as:

$$\beta^{egls} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y,$$

where each diagonal element of Ω is estimated by computing the sample variance of the OLS residuals within a group of replicated observations. OLS inferences are made using $s^2(X'X)^{-1}$ as the estimate of the variance-covariance matrix, where s^2 is computed as the sum of squared residuals over all observations, divided by T-k. Comparing the power of EGLS inferences to OLS inferences, Deaton finds OLS to be 'surprisingly robust to departures from the assumptions of homoskedasticity', however, these results on the power of OLS are highly suspect in light of his use of a biased estimator for $V(\beta^{ols})$.

Nozari (1984) examines the size and power of test statistics for a single coefficient from the model:

$$y = a + bx + e,$$

where x = 1....10, and replications of x vary from 1 to 10 (ie the sample size varies from 10 to 100 as the number of replications is varied from 1 to 10).

The variances of the errors are generated as powers of x, however the estimates of Ω are never based on the assumption of any functional form. Instead, Ω is estimated either by computing simple sample variances within a group of replications for a given x as by Deaton, or by one of several variations of the MINQUE (minimum norm quadratic unbiased estimator) procedure. EGLS estimates are computed by substituting the various estimates $\hat{\Omega}$ into the EGLS formula:

$$\hat{\beta}^{egls} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y$$

The study then compares size and power for three groups of estimates: EGLS, OLS using $V(\beta^{ols}) = (X'X)X'\hat{\Omega}X(X'X)^{-1}$, and OLS ignoring heteroskedasticity and computing $V(\beta^{ols}) = s^2(X'X)^{-1}$, as in Deaton.

The major recommendation from Nozari is that the problem of unequal variances is best dealt with by obtaining more replications; a suggestion of little help to econometricians. Of more interest is the finding that OLS inferences using the correct form of the variance-covariance estimator are superior to inference arising from the EGLS procedure, in that empirical sizes approximate nominal sizes more closely. In the case with no replications and sample size of 10, using MINQUE estimators for Ω, the size of the test for the OLS coefficient with nominal size of .10 varies from .12 to .15, whereas the size of the test for the EGLS coefficients are unacceptably large, ranging from .58 to .81. The results for inference employing OLS estimates and ignoring heteroskedasticity contradict Deaton's result, showing low power and poor approximations to nominal size. No substantial differences in power were found between EGLS inferences and OLS inferences using the correct formula for the variance-covariance estimate.

Kleijnen et al (1985), in a Monte Carlo study similar to that of Nozari compare the size and power of EGLS and OLS inferences. EGLS estimates are computed as in the study by Deaton; estimates of the error variances within groups of replicated observations are computed as the sample variance of the OLS residuals. The estimates of the error variances become the diagonal elements of $\hat{\Omega}$ which is used in the usual EGLS formula. The same estimate, $\hat{\Omega}$, is used to estimate the variance of the OLS estimates:

$$V(\beta^{ols}) = (X'X)^{-1}X'\hat{\Omega}X(X'X)^{-1}.$$

Kleijnen et al confirm Nozari's conclusion that OLS with the correct form of the variance-covariance matrix leads to superior inferences in terms of approximating the nominal size of the test statistic when the number of replications is small. But, they find EGLS inferences to have higher power in all cases. They recommend the use of EGLS when the number of replications exceeds 25. It appears from these findings that better estimates of the variance of the errors, as might result from the use of specific functional forms for the variance, may lead to the superiority of EGLS over OLS in making inferences in terms of both size and power. Kleijnen et al also recommend the use of higher levels of significance in testing hypotheses on regression coefficients, presumably because the empirical sizes of inferences tend to overestimate the nominal sizes, and researchers might wish to avoid these unknown but high probabilities of type I error. They offer no specific guidelines on this point.

The case of non-replicated observations is by far the more common case in the analysis of economic data. However, only one major study of inference in the presence of heteroskedasticity with non-replicated observations exists in the literature. MacKinnon and White (1985) conduct a series of Monte Carlo experiments to study the size (but not power) of tests on OLS regression coefficients in the presence of heteroskedasticity of unknown form. White (1980) has shown that the covariance matrix of the OLS estimates in the presence of heteroskedasticity can be consistently estimated by:

$$(X'X)^{-1}X'\hat{\Omega}X(X'X)^{-1},$$

where $\hat{\Omega} = \mathrm{diag}(e_1^2, e_2^2, \ldots\ldots, e_T^2)$, and e_i is the OLS residual.

MacKinnon and White use this heteroskedasticity consistent variance-covariance matrix estimator and three variations on it (degree of freedom corrections and the 'jackknife estimator') to compute test statistics for regression coefficients where the errors are generated as follows:

1. No heteroskedasticity: errors are independent, $N(0, \sigma^2)$. The purpose of this case is to determine how inference is affected by use of heteroskedasticity consistent variance-covariance matrices when the variances are homoskedastic. Note that the heteroskedasticity consistent variance-covariance matrix will still be an unbiased estimator of the variance-covariance matrix when there is no heteroskedasticity.

2. Structural change. There are two sets of observations: those with error variances of σ^2, and those with error variances of $k\sigma^2$. It is assumed that it is known to which set an observation belongs, but that k is unknown.

3. Random coefficients: assuming each of the coefficients in the regression model are random variables, the variance of the error term is easily derived as a linear function of the squares of the exogenous variables.

Some general conclusions from the study are: that the jackknife estimator of the variance-covariance estimator yields superior inferences to all of the other estimators considered; that the use of a heteroskedasticity consistent estimator can lead to misleading inferences when there is in fact no heteroskedasticity; and, predictably, that the usual OLS variance-covariance estimator, $s^2(X'X)^{-1}$, can lead to seriously misleading inferences in the presence of heteroskedasticity. An important practical observation which can be made from the results is that the size of the tests for the smallest sample (T=50) are often higher than that associated with the nominal critical values. Rejection proportions at the nominal critical value of 1 percent are frequently observed to be as high as 5 percent. Moreover, in the vast majority of cases shown in the tables, rejection frequencies are too high relative to the nominal critical value, a tendency also observed in the studies of Nozari and Kleijnen.

Since all of the test statistics examined are asymptotically normal, we expect good results for large samples. MacKinnon and White's study shows steady improvement in inference as sample size increases, but at the maximum sample size of 200, several test statistics for the case of structural change in the variance have empirical critical values significantly different from the nominal critical values, though the jackknife inferences perform well. No general conclusions are possible concerning the question of how inferences depend upon the specific form of heteroskedasticity.

A major omission in the Monte Carlo literature is an examination of power and size under the procedures typically used to correct for heteroskedasticity in econometric work. For example, a conventional approach would be to pre-test for heteroskedasticity with the Goldfeld-Quandt test, then (assuming the test indicates heteroskedasticity) to correct the estimates using the Park or Glejser procedure and test significance based on these final estimates. Although the literature contains considerable evidence on the efficacy of such a procedure in improving estimates, research on small sample inference on the regression coefficients of ultimate interest to the researcher have been largely ignored.

Design of the Monte Carlo Study

Previous studies on inference in the presence of heteroskedasticity have suggested that hypothesis testing using the OLS regression coefficients might be superior to inference based on EGLS estimates. In addition, the study by MacKinnon and White (1985) provides convincing evidence that good inferences employing the OLS regression coefficients can be obtained by using a heteroskedasticity consistent variance-covariance estimator, where knowlege of the specific form of the heteroskedasticity is not required. Unfortunately none of the the existing research addresses the problem of inference when a specific functional form for the heteroskedasticity is assumed. The following series of Monte Carlo experiments are designed to compare the size and power of several approaches to inference in the presence of a variety of heteroskedasticity generating processes.

The general approach is to generate data based on a particular model of heteroskedasticity and then to produce inferences from the data based on several different estimation procedures. In the first series of simulations (Part I) it is assumed that the process generating the variances of the errors is known; that is, the EGLS estimates are based on the same functional form as that which generated the data. In the second series of simulations (Part II) the EGLS procedure is computed using a functional form different from the one which is generating the heteroskedasticity. In part I we are therefore examining the case where the estimates of Ω, the variance-covariance matrix , and the EGLS estimates, are asymptotically unbiased; this is the case covered by asymptotic theory. However, it would be fair to question the practical value of these results since it can be argued that the practitioner rarely has theoretical reasons for assuming a specific functional form for the heteroskedasticity. Therefore the opposite extreme, where we deliberately choose an incorrect functional form, is also tested. Note that even as the sample tends to infinity, it will not be possible for the incorrect functional form to provide unbiased estimates of Ω, therefore the functional form inferences in this case are not even justified by asymptotic theory. In practice, the functional form might be chosen on the basis of testing the fit of several standard models and selecting the functional form which best fits the OLS residuals. Therefore these results may be viewed as a worst possible case, in the sense that a heuristic examination of the residuals would be likely to result in the selection of a functional form which better fits the data or possibly even the correct functional form. In these studies the incorrect functional form has simply been imposed on the data.

All Monte Carlo experiments are based on the model:

$$y_t = \beta_0 + \beta_1 x_t + u_t,$$

where $u_t \sim N(0, \sigma_t^2)$,

$$E(u_i u_j) = 0 \text{ for } i \neq j,$$

and σ_t^2 is determined by one of four models used to generate the errors.

$\beta = (\beta_0, \beta_1)$ is fixed as $(1,1)$.

For the size experiments the null hypothesis is taken as $\beta_1 = 1$; for power experiments various false nulls concerning β_1 (given in the results section) are chosen such that the absolute size of the powers is meaningful over all sample sizes tested. For both size and power simulations the alternate hypothesis is that β_1 is greater than the null hypothesis value. Hypotheses are tested by computing the conventional t-test: the difference between the estimated value of β_1 and the null hypothesis value is divided by the estimated standard deviation of β_1. Only hypotheses concerning β_1 are tested in order to simplify the results, and since economic research is usually concerned with inferences with respect to slope coefficients as opposed to intercept coefficients.

Each simulation is based on 1000 samples, and repeated for sample sizes of 30, 60, and 120. The exogenous variable x_t is taken as 30 numbers from a uniform distribution, U(0,1), replicated 2 and 4 times as required for the sample size. The purpose of the replication of x_t, as explained by MacKinnon and White, is to hold the degree of heteroskedasticity constant as the sample size is increased.

Models of Heteroskedasticity

Model 1: No heteroskedasticity

$$\sigma_t^2 = \sigma^2$$

By applying each of the procedures for estimation and inference in the presence of heteroskedasticity to data generated by model 1, information is obtained on how costly the technique is when it is mistakenly applied to homoskedastic data.

Model 2: Multiplicative heteroskedasticity

$$\sigma_t^2 = Kx_t^{\partial}$$

Model 3: The Random Coefficients Model and Mixed Heteroskedasticity

Assuming each of the coefficients β_i in the model are normally distributed random variables with means equal to (B_0, B_1), the variance of the error term is easily derived as a linear function of the squares of the exogenous variables:

$$\beta_i = B_i + v_{it}, \quad \text{where } v_{it} \sim N(0,w_i)$$

implies $\quad y_t = B_0 + B_1 x_t + v_{0t} + v_{1t}x_t + u_t,$

and assuming all u_t and v_t are independent, the variance of the total error term can be written as:

$$\sigma_t^2 = a + bx_t^2$$

where a and b are functions of the w_i and σ^2.

Note that the above model is a special case of mixed heteroskedasticity, where the variance of the error is related to a constant and an exogenous variable as in the following model typically used to model heteroskedasticity:

$$\sigma_t^2 = a + bf(x_t) \ .$$

Model 4: Structural Change

For the sample size of 30 we have:

$$u_t \sim \begin{cases} N(0, \sigma^2) \text{ for } t = 1 \text{ to } 15 \\ N(0, \alpha\sigma^2) \quad t = 16 \text{ to } 30. \end{cases}$$

For larger sample sizes the variance of the errors alternates between σ^2 and $\alpha\sigma^2$ every 15 observations, keeping the relationship between u_t and x_t constant as sample size increases. Assume it is known at what points the structural change occurs, but α is unknown.

Estimation and Inference Procedures

1. OLS ignoring heteroskedasticity

Coefficients are estimated as $\beta^{ols} = (X'X)^{-1}X'y$ and the variance-covariance matrix is estimated as $V(\beta^{ols}) = s^2(X'X)^{-1}$. Inferences using the OLS coefficients are made using the usual t-ratio.

2. GLS

Coefficients are estimated as $\beta^{gls} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$ and the variance-covariance matrix is estimated as $V(\beta^{gls}) = s^2(X'\Omega^{-1}X)^{-1}$, where

$$s^2 = (e'\Omega^{-1}e)/(T-k),$$

e is the vector of residuals computed by applying the GLS coefficients to the original data, and Ω is the known matrix of the variance of the errors. Note that when Ω is fully known (as opposed to being known only up to a constant of proportionality) the variance-covariance matrix is given by:

$$(X'\Omega^{-1}X)^{-1},$$

the expectation of s^2 will be one, and the 't-ratio' will actually be distributed normally. Since the purpose of the GLS inferences is to serve a benchmark against which the others are to be compared, it is convenient to also estimate s^2, so that inferences using the GLS coefficients will be distributed as t-statistics.

3. OLS using functional form estimates of Ω

When errors are heteroskedastic, $V(\beta^{ols}) = (X'X)^{-1}X'\Omega X(X'X)^{-1}$.
An estimate of Ω is computed using an assumption about the functional form of the heteroskedasticity. For multiplicative or mixed heteroskedasticity Ω is estimated by:

i. regressing the OLS squared residuals on some function of X, as determined by the functional form chosen to represent the heteroskedasticity (in part I it will be the correct functional form, in part II, incorrect). For example, if we assume multiplicative heteroskedasticity, then the log of the squared OLS residuals are regressed on $\ln x_t$ and a constant.

ii. The predicted squared residuals from i. above are used as estimates for the diagonal elements of Ω. In the case of multiplicative heteroskedasticity, where $\sigma_t^2 = Kx_t^{\partial}$, Harvey (1976) has shown that the regression in i. above yields an inconsistent estimate of $\ln K$. However, when the errors are normally distributed, Harvey shows that $E(\ln K) = -1.2704$. This is of no consequence to EGLS estimation since variances of errors need only be estimated to a constant of proportionality, but the result is very useful in estimating $V(\beta^{ols})$ where an estimate of Ω is required. Harvey's result is used to adjust the estimates accordingly.

For the case of structural change, where the functional form is assumed known, the method outlined in Judge et al (1985 p 428) is followed. Ω is estimated by computing one OLS regression for the set of observations for which the variance of the errors is σ^2, and another independent OLS regression for the set of observations for which the variance of the errors is $\alpha^2\sigma^2$. The diagonal elements of Ω are estimated as the sum of squared residuals divided by ((T/2) - k) for each group.

Inferences are made by forming the t-ratio using $(\beta^{ols}-1)$ in the numerator and the square root of the appropriate element of the estimate of $V(\beta^{ols})$ in the denominator.

4. OLS and the Jackknife Estimator of the Variance-Covariance matrix

White's (1980) original heteroskedasticity consistent variance-covariance estimator, given by

$$(X'X)^{-1}X'\hat{\Omega}X(X'X)^{-1},$$

where $\hat{\Omega} = \mathrm{diag}(e_1^2, e_2^2, \ldots\ldots, e_T^2)$, and e_i is the OLS residual,

was examined in the study by MacKinnon and White along with a number of variations which have also been shown to be heteroskedasticity consistent.

The best of these estimators across all forms of heteroskedasticity examined by MacKinnon and White was found to be the jackknife estimator. Consequently, in these simulations, only the jackknife estimator is examined for the case of non-functional form estimation of $V(\beta^{ols})$.

The jackknife estimator is a very general estimator for the variance of a statistic based on recomputing the statistic with one observation deleted. Though rarely used in actual computations, the definition of the jackknife as suggested by Tukey (1958) provides a good intuitive rationale for this technique. For any statistic computed from T observations, we can calculate a series of T recomputed statistics by deleting each one of the T observations in turn, and recomputing the original statistic from T-1 observations. The variation among the T recomputed statistics can be used to obtain an estimate for the sample variance of the original statistic. The jackknife estimator for the variance of the statistic is defined as the sum of squared deviations from the mean of these T recomputed statistics, multiplied by a correction factor of (T-1)/T. Efron (1982 p13) shows that when the jackknife is applied to the sample mean, the jackknife estimate for the variance of the mean is identical to the usual estimate of the variance of the mean given by s^2/T. MacKinnon and White (after considerable algebraic manipulation) give the jackknife estimator for the variance-covariance estimator as:

$$((n-1)/n)\ (X'X)^{-1}\ [X'\hat{\Omega}X - (1/n)(X'ee'X)]\ (X'X)^{-1},$$

where $\hat{\Omega}$ is a TxT diagonal matrix with diagonal elements of e_t^2, and off diagonal elements of zero, and e is the vector of OLS residuals e_t. It is easily shown that the jackknife estimator is asymptotically equivalent to White's heteroskedasticity consistent variance-covariance matrix estimator.

5. EGLS estimation

Three common EGLS procedures are used:

i. The modified Glejser method is used in Part I where we have mixed heteroskedasticity of the random coefficients model variant. The modified Glejser method was chosen (as opposed to the Glejser method) because it more closely models the functional form of the heteroskedasticity and because according to Kennedy (1984) it is the form preferred by most practitioners in the case of mixed heteroskedasticity. The modified Glejser method consists of regressing the squared OLS residuals on a constant and on x_t^2, and obtaining predicted values for the squared OLS residuals. The data is then transformed by dividing through by the square root of these predicted squared residuals, and OLS is run on the transformed data. The quasi t-statistics obtained from this regression on transformed data are used to make inferences. A problem sometimes encountered with the modified Glejser method is that the predicted values for the squared residuals, and hence predicted variances of the errors, are negative. Negative values are replaced with the actual squared residual, as in the studies by Goldfeld and Quandt (1975).

ii. The method of Park, also very popular in textbooks and among practitioners, is used in Part I where we have multiplicative heteroskedasticity, and throughout Part II. In Part II the objective is to illustrate how well EGLS techniques perform when the wrong functional form is used to model the heteroskedasticity. The method of Park is chosen as the 'wrong' model here for two reasons. First, the functional form is very general. The method of Park consists of modelling the heteroskedasticity as:

$$\sigma_t^2 = Kx_t^{\partial},$$

To estimate σ_t^2 the log of the OLS squared residuals are regressed on a constant and $\ln x_t$. Therefore, the method of Park is a good choice when the researcher suspects heteroskedasticity related to the independent variable, but would like a model which could accomodate any power of x. Second, Kennedy (1985 p159) has found that in terms of the ratio of OLS variance to EGLS variance, the Park method performs very well over a variety of cases where the heteroskedasticity is not generated by the functional form employed by the Park method.

iii. The Constant Variance Within Subgroups EGLS procedure

This method, given by Judge et al consists of estimating the variances for the case of structural change in Part I as has already been described above (see OLS using the functional form estimates). The estimates of the standard deviations of the errors are then used to transform the data as in 1. and 2. above.

6. Pre-test estimators and inferences

i) The Goldfeld-Quandt test is used to detect heteroskedasticity and to choose between OLS and EGLS estimators and inferences. The Goldfeld-Quandt test, rather than some of the more general tests for heteroskedasticity (eg. Judge et al, p 450) is used here primarily because of the popularity of the test among practitioners, and the good results obtained for the test by Goldfeld and Quandt (1972) in their major study of heteroskedasticity. For sample sizes of 30, 60 and 120 the recommendations of Goldfeld and Quandt for eliminating 8, 16, and 32 observations are followed. In the case of the structural change model of part I, where we assume that it is known where the structural change takes place, no observations are eliminated, since elimination of observations in the case of a discontinuous change in the error variance would reduce the power of the test.

ii) The White test for heteroskedasticity is used to choose between the simple OLS variance-covariance matrix, $s^2(X'X)^{-1}$, when the hypothesis of homoskedasticity is not rejected, and the jackknife estimator when homoskedasticity is rejected. Inferences are made using the OLS coefficient estimates and this pre-test estimator of the variance. The White test consists of regressing the squared OLS residuals on a constant, the original regressors, the regressors squared, and all the cross products of regressors (In this case we regress on a constant, X, and X^2), and computing the product of T and the R^2 from this regression. TR^2 is shown by White (1980) to be asymptotically distributed as χ^2 with degrees of freedom equal to the number of regressors.

Measurement of Heteroskedasticity

Kennedy (1985) presents arguments for the use of an appropriate measure of heteroskedasticity. A measure of heteroskedasticity in Monte Carlo studies is necessary if: the results are to be applied in other contexts; the conclusions change as the degree of heteroskedasticity varies; or if conclusions are drawn from a number of different heteroskedasticity generating processes. Previous studies have used a number of measures, such as the ratio of largest to smallest variance, or the magnitude of some parameter in the function which generates the variance. Kennedy argues that all of these measures are inadequate, and suggests using the squared coefficient of variation of the variances themselves. If u_t is the variance of the error term for observation t, then :

$$CV = \frac{\Sigma(u_i - \bar{u})^2}{N(\bar{u})^2} .$$

In part I of the study all models of heteroskedasticity are chosen such that CV = .432. According to the examples from other studies provided by Kennedy, this would correspond to a moderately high degree of heteroskedasticity. To give some feeling for the severity of a CV of .432, notice that this would correspond to $\alpha = 4.84$ in the structural change model; that is, the variance for high variability observations is nearly five times as high as for low variability observations. In part II, CV varies over a range of values as shown in the tables. The reason for varying the CV in this second set of simulations was to test Park's EGLS method against some worst case alternatives. In particular, Park's method employs a model which assumes the variance of the errors decreases to zero when the value of the independent variable is zero (ie, there is no constant in $\sigma_t^2 = Kx^{\partial}$, and the inclusion of a constant term would make the function impossible to estimate using OLS). To test the efficacy of Park's method when there is mixed heteroskedasticity and a substantial constant term implicitly places an upper limit on CV, since the constant term has the effect of introducing a fixed component to the error variance across all values of x_t.

Results

The results of the Monte Carlo simulations are presented in tables 1 to 16. Tables 2 to 8 show results for cases where it is assumed the functional form is known and tables 9 to 16 show the results for cases where an incorrect functional form is used. Details of the coefficient estimates and variance-covariance matrix estimates are not presented, but in all cases except those specifically mentioned, the average variance-covariance estimates exhibited the relationship:

$$V(\beta^{gls}) < V(\beta^{egls}) < V(\beta^{ols}),$$

indicating that the results reflect cases of heteroskedasticity where the use of EGLS would be advisable under the usual justification of increased efficiency. The tables show the size and power of inferences using β_1, the slope coefficient, calculated 1000 times for each model under each of the seven techniques. Rejection proportions computed for cases of a true null hypothesis (ie the sizes of statistics) are reported for critical values from the t-distribution with nominal rejection proportions of .05, .025, and .01. Powers are calculated using empirically determined critical values derived from a ranking of the quasi t-statistics calculated for a true null hypothesis. In order to calculate power it is necessary to assume a false null hypothesis for β_1. For all models except the homoskedastic model the false null is $\beta_1 = 0.6$; for the homoskedastic models the false null was chosen as $\beta_1 = 0$. The choice of false nulls is somewhat arbitrary, however a very false null hypothesis leads to powers too close to one for comparison, and a nearly true null leads to very small powers which are also difficult to compare. The false nulls are chosen here such that for all sample sizes shown in a given table the powers range from approximately 0.20 to 0.80.

Rejection proportions from Monte Carlo simulations can be interpreted statistically by noting that the variance of a sample proportion is asymptotically normally distributed, and given by:

$$p(1-p)/n,$$

where p is the expected proportion of rejections, and n is the number of simulations. For expected rejection proportions of .05, .025 and .01, the values for two standard deviations of the sample proportion are .014, .010, .006. Therefore, for the .05 nominal size we can determined whether a particular empirical proportion is statistically different from .05 by noting whether it falls within the two standard deviation confidence interval ranging from .036 to .064.

Table 1 shows the effect of the various techniques on the size of inferences when the errors are homoskedastic. OLS and GLS are mathematically equivalent in this case. The EGLS procedure employed here is Park's method. The general conclusion arising from this table is that the procedures employed appear to have had no serious effect on the reliablity of the size of inferences. This agrees with the finding of MacKinnon and White that the jackknife estimator is only slightly less reliable than OLS in the case of no heteroskedasticity. The jackknife yielded the worst inferences in table 1, particularily for T=30, but the discrepancies are not large, and in accordance with the asymptotic theory, the problem disappears for T=120. One might expect that inferences using the jackknife could be improved by using the pre-test estimator of the variance-covariance matrix based on the White test, and the results confirm this. In fact either one of the pre-testing techniques results in inferences with empirical sizes almost identical to the OLS inferences.

Tables 2 through 4 show the size of inferences in the presence of the various forms of heteroskedasticity, where it is assumed that the functional form is known. Several conclusions apply to all three tables:

1. Most of the rejection proportions are not significantly different from the expected proportion of rejections; in terms of empirical size, the inferences obtained from all the procedures are quite reliable. This conclusion agrees with MacKinnon and White's findings for the jackknife, but differs quite dramatically from the Monte Carlo studies which use replicated observations and no functional forms to obtain EGLS estimates.

2. There is a tendency for the empirical rejection frequencies to be too high, a finding noted in all previous studies.

3. Comparing inferences resulting from the use of EGLS coefficients with those resulting from the use of OLS coefficients shows neither one to be consistently superior, contrary to Nozari's finding that inferences using OLS coefficients were superior.

4. Both pre-tests result in only very small improvement in accuracy of empirical sizes of inferences.

5. In agreement with Deaton, the inferences arising from the simple OLS procedure, ignoring heteroskedasticity, are quite good, and often superior to the EGLS inferences.

6. The most accurate empirical rejection frequencies are observed in the case of structural change. This is not surprising in the case of EGLS, since the estimate of one parameter is based on multiple observations. The jackknife inferences are also more accurate in this case.

7. Rejection frequencies for sample size 120 are no more accurate than the rejection frequencies for sample size 60, indicating that the asymptotic properties of the test statistic can be assumed to take effect at sample sizes of less than 60.

Turning now to a comparison of powers (tables 5 to 8), a consistent and sometimes dramatic advantage for EGLS is observed for all cases except the model with homoskedastic errors. The superiority of EGLS is most pronounced in the case of multiplicative heteroskedasticity (table 6), where the power of the inferences using EGLS coefficients sometimes exceeds the power of inferences using OLS coefficients by as much as 100 percent. The inferences based on the Goldfeld-Quandt pre-test are very similar to the EGLS inferences, particularily for large samples. In all three models, the advantage in power of the EGLS technique increases as sample size increases, a predictable result since EGLS approaches GLS asymptotically; ie, as sample size increases the variance-covariance matrix decreases not only due to the effect of increasing T as if GLS were used, but also due to the fact that there is less variation in the estimate of Ω. For the case of the homoskedastic error model (Table 5), EGLS yields inferences with the lowest power, though for sample sizes of 60 and 120, the power of EGLS inferences is only slightly lower than that of the other techniques. Use of the Goldfeld-Quandt pre-test to select between EGLS and OLS results in inferences with power nearly as high as simple OLS inference.

MacKinnon and White (1985) recommend against the use of the pre-test estimator based on the White test, and suggest simply using the jackknife whenever heteroskedasticity is suspected. This conclusion is based on the potential for lower power and biased variance-covariance matrix estimates in cases where existing heteroskedasticity is not detected. The results of this study confirm both these findings, but show the loss in power due to use of the pre-test based on the White test to be small, relative to the power of the jackknife inferences.

Tables 9 to 16 show the size and power of tests when Park's method of EGLS is employed for cases where the heteroskedasticity has not been generated by the functional form used in the Park method.

Since there was little difference between the results for sample sizes of 60 and 120, we show results only for T=30 and T=60. Moreover, as explained above, when the incorrect functional form is used, the asymptotic theory does not apply.

The functional forms chosen to generate the heteroskedasticity are shown on each table. Initially, forms were chosen with the same CV as in part I, but preliminary simulations revealed that the quality of inferences (as well as estimates) depended upon the absolute size of the constant term in the heteroskedasticity generating function. In following the approach of generating worst possible cases for EGLS, functional forms with relatively large constant terms were tested. The large constant term impairs the ability of the Park method to estimate the errors, since the functional form used in the Park method has no intercept. The observation regarding the importance of the constant term has also been made by Kennedy (1985) in the context of efficiency. Kennedy suggests correcting for the existence of a constant term when the constant exceeds 15 percent of the average variance. The results shown here are based on models where the constant term as a percentage of the average variance ranges from a low of 16 percent to a high of 57 percent. For any of these models with constant terms, the results of some preliminary simulations showed that EGLS inferences could be improved, both in terms of size and power, by following the advice of Kennedy and estimating the errors using the modified Glejser method. These preliminary simulations also showed that for constant terms of the relative size used here, the need for an intercept term in the functional form could be easily detected either by comparing the fit of the Park model with the modified Glejser model, or by applying Kennedy's 15 percent rule. Nevertheless, the intention of these tables is to show results for the effects of choosing the incorrect functional form, and to provide some indication of the robustness of EGLS inferences to incorrect choices of the functional form.

In terms of empirical size (tables 9 to 12), the use of an incorrect functional form has very little effect. All of the findings (points 1 to 6 above) which applied in the case of known functional form apply here as well. Even for the case of structural change, where we are estimating a step function with a simple curve, the inferences arising from EGLS techniques are good.

Comparing powers of inferences (tables 13 to 16), we see there are some consequences to using the incorrect functional form. Inferences arising from the use of EGLS estimates still have generally higher power than the inferences arising from OLS estimates in tables 13 to 15, but in the case of structural change (table 16), the inferences arising from the jackknife estimator of the variance are of similar power to those of EGLS. It should be noted however, that it is unlikely that a researcher examining data showing heteroskedasticity due to structural change would have chosen Park's method of EGLS. It is illustrated here as a worst possible case, and in fact, the differences do not rule out even this very unlikey form of EGLS.

Table 13 compares the power of inferences when heteroskedasticity is generated by the random coefficients model and estimated by the Park model. The results show a clear advantage of EGLS inferences despite the incorrect functional form, and as in part I, no significant advantage to the Goldfeld-Quandt pre-test. In table 14, the heteroskedasticity generating function is altered by increasing the intercept by a factor of 5. These results show only a very slight advantage in power to using EGLS inferences over jackknife inferences. Similar results obtain for the linear model of heteroskedasticity in table 15. Note that in table 14 the particular model of heteroskedasticity has had the effect of reducing the CV quite substantially from the levels used in Tables 13 and 16 (and all of Part I). The effect of this is that the Goldfeld-Quandt test is now less powerful, hence the power of the inferences using the pre-test coefficient estimates are also less powerful.

It is also interesting to note that all the cases where the EGLS inferences had less power than the OLS inferences corresponded to cases where the average variance of the EGLS estimates was higher than the average variances of the OLS estimates. That is, EGLS inferences are inferior in cases where EGLS ought not to have been applied on the basis of reduced efficiency.

Table 1
Empirical Size of Inferences

$\sigma_t^2 = 1$; CV = 0

Nominal Size	.050	.025	.010
T=30			
OLS	.064	.034	.011
GLS	.064	.034	.011
OLS/functional form Ω	.048	.025	.013
OLS/jackknife	.073	.047	.022
EGLS	.074	.045	.020
PRE/ols or egls	.065	.034	.011
PRE/var=ols or jackknife	.062	.031	.010
T=60			
OLS	.056	.028	.015
GLS	.056	.028	.015
OLS/functional form Ω	.052	.028	.014
OLS/jackknife	.064	.037	.017
EGLS	.060	.033	.020
PRE/ols or egls	.056	.028	.015
PRE/var=ols or jackknife	.056	.028	.015
T=120			
OLS	.047	.020	.007
GLS	.047	.020	.007
OLS/functional form Ω	.045	.023	.008
OLS/jackknife	.054	.027	.011
EGLS	.049	.020	.008
PRE/ols or egls	.047	.020	.007
PRE/var=ols or jackknife	.048	.021	.007

Table 2
Empirical Size of Inferences

$\sigma_t^2 = Kx_t^{\partial}$, K=1, ∂=1.38; CV = .432

Nominal Size	.050	.025	.010
T=30			
OLS	.047	.022	.011
GLS	.047	.022	.005
OLS/functional form Ω	.059	.036	.015
OLS/jackknife	.066	.033	.020
EGLS	.071	.028	.008
PRE/ols or egls	.068	.031	.010
PRE/var=ols or jackknife	.045	.023	.012
T=60			
OLS	.047	.026	.012
GLS	.055	.024	.008
OLS/functional form Ω	.053	.031	.016
OLS/jackknife	.060	.030	.014
EGLS	.061	.028	.010
PRE/ols or egls	.058	.027	.009
PRE/var=ols or jackknife	.055	.028	.014
T=120			
OLS	.052	.024	.002
GLS	.032	.013	.002
OLS/functional form Ω	.058	.023	.005
OLS/jackknife	.055	.027	.009
EGLS	.036	.013	.002
PRE/ols or egls	.036	.013	.002
PRE/var=ols or jackknife	.055	.027	.009

Table 3
Empirical Size of Inferences

$\sigma_t^2 = 1 + 10x_t^2$;   CV = .432

Nominal Size	.050	.025	.010
T=30			
OLS	.057	.027	.007
GLS	.048	.020	.010
OLS/functional form Ω	.056	.026	.007
OLS/jackknife	.060	.032	.013
EGLS	.058	.031	.021
PRE/ols or egls	.064	.031	.021
PRE/var=ols or jackknife	.051	.023	.007
T=60			
OLS	.055	.028	.012
GLS	.048	.015	.005
OLS/functional form Ω	.049	.021	.008
OLS/jackknife	.059	.029	.012
EGLS	.054	.029	.014
PRE/ols or egls	.052	.029	.015
PRE/var=ols or jackknife	.056	.024	.010
T=120			
OLS	.060	.037	.021
GLS	.056	.036	.013
OLS/functional form Ω	.057	.032	.020
OLS/jackknife	.056	.032	.020
EGLS	.063	.037	.017
PRE/ols or egls	.063	.037	.017
PRE/var=ols or jackknife	.056	.032	.020

Table 4

Empirical Size of Inferences

σ_t^2 = 1 or 4.84 (structural change model); CV = .432

Nominal Size	.050	.025	.010
T=30			
OLS	.038	.013	.004
GLS	.048	.019	.010
OLS/functional form Ω	.041	.014	.004
OLS/jackknife	.055	.027	.007
EGLS	.050	.022	.008
PRE/ols or egls	.049	.022	.008
PRE/var=ols or jackknife	.036	.014	.003
T=60			
OLS	.046	.026	.007
GLS	.047	.028	.012
OLS/functional form Ω	.051	.026	.007
OLS/jackknife	.065	.031	.010
EGLS	.054	.030	.014
PRE/ols or egls	.054	.030	.014
PRE/var=ols or jackknife	.053	.028	.008
T=120			
OLS	.051	.028	.010
GLS	.059	.031	.013
OLS/functional form Ω	.050	.027	.010
OLS/jackknife	.049	.025	.013
EGLS	.062	.032	.012
PRE/ols or egls	.062	.032	.012
PRE/var=ols or jackknife	.050	.028	.010

Table 5
Empirical Power of Inferences

$\sigma_t^2 = 1$ CV = 0

Empirical Size	.050	.025	.010
T=30			
OLS	.437	.340	.223
GLS	.437	.340	.223
OLS/functional form Ω	.454	.349	.192
OLS/jackknife	.428	.302	.192
EGLS	.396	.284	.169
PRE/ols or egls	.436	.325	.221
PRE/var=ols or jackknife	.437	.344	.231
T=60			
OLS	.659	.544	.311
GLS	.659	.544	.311
OLS/functional form Ω	.652	.528	.323
OLS/jackknife	.652	.537	.320
EGLS	.624	.507	.280
PRE/ols or egls	.651	.543	.312
PRE/var=ols or jackknife	.650	.543	.309
T=120			
OLS	.932	.876	.779
GLS	.932	.876	.779
OLS/functional form Ω	.935	.861	.781
OLS/jackknife	.930	.868	.761
EGLS	.929	.880	.783
PRE/ols or egls	.933	.878	.781
PRE/var=ols or jackknife	.932	.879	.776

Table 6
Empirical Power of Inferences

$\sigma_t^2 = Kx_t^{\partial}$, K=1, ∂=1.38; CV = .432

Empirical Size	.050	.025	.010
T=30			
OLS	.252	.167	.075
GLS	.385	.290	.209
OLS/functional form Ω	.203	.126	.082
OLS/jackknife	.242	.173	.071
EGLS	.311	.238	.154
PRE/ols or egls	.294	.208	.141
PRE/var=ols or jackknife	.256	.169	.073
T=60			
OLS	.371	.259	.147
GLS	.573	.474	.384
OLS/functional form Ω	.348	.204	.104
OLS/jackknife	.367	.249	.107
EGLS	.523	.443	.313
PRE/ols or egls	.521	.425	.326
PRE/var=ols or jackknife	.362	.262	.115
T=120			
OLS	.527	.448	.342
GLS	.877	.832	.739
OLS/functional form Ω	.514	.422	.333
OLS/jackknife	.525	.426	.309
EGLS	.867	.807	.699
PRE/ols or egls	.867	.807	.699
PRE/var=ols or jackknife	.525	.426	.309

Table 7
Empirical Power of Inferences

$\sigma_t^2 = 1 + 10x_t^2$; CV = .432

Empirical Size	.050	.025	.010
T=30			
OLS	.155	.100	.064
GLS	.191	.131	.062
OLS/functional form Ω	.150	.099	.067
OLS/jackknife	.169	.102	.044
EGLS	.175	.121	.023
PRE/ols or egls	.172	.120	.024
PRE/var=ols or jackknife	.157	.101	.065
T=60			
OLS	.247	.173	.100
GLS	.307	.219	.169
OLS/functional form Ω	.246	.173	.100
OLS/jackknife	.233	.159	.102
EGLS	.296	.196	.093
PRE/ols or egls	.301	.199	.104
PRE/var=ols or jackknife	.231	.179	.107
T=120			
OLS	.368	.237	.111
GLS	.495	.352	.222
OLS/functional form Ω	.373	.236	.104
OLS/jackknife	.367	.238	.098
EGLS	.476	.334	.212
PRE/ols or egls	.476	.335	.212
PRE/var=ols or jackknife	.367	.238	.098

Table 8
Empirical Power of Inferences

$\sigma_t^2 = 1$ or 4.84 (structural change model); CV = .432

Empirical Size	.050	.025	.010
T=30			
OLS	.229	.176	.117
GLS	.249	.178	.081
OLS/functional form Ω	.234	.181	.113
OLS/jackknife	.248	.159	.101
EGLS	.256	.166	.083
PRE/ols or egls	.262	.166	.083
PRE/var=ols or jackknife	.226	.175	.113
T=60			
OLS	.349	.240	.164
GLS	.415	.274	.148
OLS/functional form Ω	.340	.238	.164
OLS/jackknife	.316	.240	.178
EGLS	.377	.264	.150
PRE/ols or egls	.377	.264	.149
PRE/var=ols or jackknife	.336	.241	.172
T=120			
OLS	.592	.434	.311
GLS	.769	.657	.518
OLS/functional form Ω	.596	.424	.310
OLS/jackknife	.606	.482	.275
EGLS	.756	.667	.518
PRE/ols or egls	.756	.667	.518
PRE/var=ols or jackknife	.613	.434	.311

Table 9

Empirical Size Using Park's method of EGLS

$\sigma_t^2 = 1 + 10x_t^2$; CV = .432

Nominal Size	.050	.025	.010
T=30			
OLS	.055	.025	.012
GLS	.043	.017	.005
OLS/functional form Ω	.058	.030	.016
OLS/jackknife	.069	.038	.012
EGLS	.069	.033	.016
PRE/ols or egls	.071	.032	.016
PRE/var=ols or jackknife	.052	.022	.010
T=60			
OLS	.061	.031	.015
GLS	.052	.027	.015
OLS/functional form Ω	.064	.036	.022
OLS/jackknife	.057	.034	.016
EGLS	.066	.039	.024
PRE/ols or egls	.066	.041	.024
PRE/var=ols or jackknife	.054	.031	.014

Table 10
Empirical Size Using Park's method of EGLS

$\sigma_t^2 = 5 + 10x_t^2$; CV = .075

Nominal Size	.050	.025	.010
T=30			
OLS	.048	.023	.007
GLS	.050	.026	.005
OLS/functional form Ω	.052	.027	.010
OLS/jackknife	.060	.032	.017
EGLS	.066	.039	.016
PRE/ols or egls	.060	.033	.013
PRE/var=ols or jackknife	.045	.022	.007
T=60			
OLS	.054	.027	.009
GLS	.050	.027	.010
OLS/functional form Ω	.055	.032	.014
OLS/jackknife	.059	.027	.011
EGLS	.066	.036	.015
PRE/ols or egls	.057	.029	.010
PRE/var=ols or jackknife	.052	.026	.009

Table 11

Empirical Size Using Park's method of EGLS

$\sigma_t^2 = 1 + 10x_t$; CV = .20

Nominal Size	.050	.025	.010
T=30			
OLS	.044	.021	.008
GLS	.051	.023	.007
OLS/functional form Ω	.044	.024	.009
OLS/jackknife	.061	.035	.016
EGLS	.060	.029	.014
PRE/ols or egls	.054	.028	.013
PRE/var=ols or jackknife	.046	.021	.009
T=60			
OLS	.044	.021	.007
GLS	.048	.019	.008
OLS/functional form Ω	.048	.032	.010
OLS/jackknife	.048	.026	.009
EGLS	.057	.027	.011
PRE/ols or egls	.056	.024	.010
PRE/var=ols or jackknife	.043	.022	.007

Table 12
Empirical Size Using Park's method of EGLS
σ_t^2 = 1 or 4.84 (structural change model); CV = .432

Nominal Size	.050	.025	.010
T=30			
OLS	.058	.024	.011
GLS	.046	.029	.011
OLS/functional form Ω	.058	.039	.019
OLS/jackknife	.070	.037	.019
EGLS	.080	.049	.027
PRE/ols or egls	.075	.046	.026
PRE/var=ols or jackknife	.059	.026	.011
T=60			
OLS	.051	.021	.010
GLS	.051	.022	.010
OLS/functional form Ω	.066	.036	.016
OLS/jackknife	.063	.032	.014
EGLS	.085	.050	.026
PRE/ols or egls	.084	.046	.026
PRE/var=ols or jackknife	.054	.025	.011

Table 13
Empirical Power Using Park's method of EGLS

$\sigma_t^2 = 1 + 10x_t^2$; CV = .432

Empirical Size	.050	.025	.010
T=30			
OLS	.181	.126	.047
GLS	.236	.128	.096
OLS/functional form Ω	.200	.121	.051
OLS/jackknife	.171	.122	.080
EGLS	.232	.131	.077
PRE/ols or egls	.216	.128	.079
PRE/var=ols or jackknife	.178	.120	.062
T=60			
OLS	.220	.135	.076
GLS	.288	.162	.088
OLS/functional form Ω	.255	.142	.057
OLS/jackknife	.239	.129	.077
EGLS	.286	.157	.077
PRE/ols or egls	.275	.155	.076
PRE/var=ols or jackknife	.232	.132	.083

Table 14
Empirical Power Using Park's method of EGLS

$\sigma_t^2 = 5 + 10x_t^2$; CV = .075

Empirical size	.050	.025	.010
T=30			
OLS	.133	.077	.035
GLS	.149	.099	.051
OLS/functional form Ω	.129	.072	.027
OLS/jackknife	.135	.072	.025
EGLS	.125	.083	.026
PRE/ols or egls	.144	.089	.033
PRE/var=ols or jackknife	.138	.074	.033
T=60			
OLS	.170	.114	.063
GLS	.188	.107	.056
OLS/functional form Ω	.175	.109	.049
OLS/jackknife	.150	.115	.057
EGLS	.177	.110	.059
PRE/ols or egls	.164	.108	.065
PRE/var=ols or jackknife	.168	.109	.060

Table 15

Empirical Power Using Park's method of EGLS

$\sigma_t^2 = 1 + 10x_t$; CV = .20

Empirical Size	.050	.025	.010
T=30			
OLS	.157	.086	.048
GLS	.160	.102	.054
OLS/functional form Ω	.158	.105	.047
OLS/jackknife	.146	.081	.044
EGLS	.158	.113	.053
PRE/ols or egls	.163	.096	.045
PRE/var=ols or jackknife	.159	.083	.049
T=60			
OLS	.228	.151	.075
GLS	.259	.186	.119
OLS/functional form Ω	.229	.134	.076
OLS/jackknife	.245	.151	.082
EGLS	.238	.171	.091
PRE/ols or egls	.234	.163	.086
PRE/var=ols or jackknife	.234	.148	.079

Table 16

Empirical Power Using Park's method of EGLS

σ_t^2 = 1 or 4.84 (structural change model); CV = .432

Empirical Size	.050	.025	.010
T=30			
OLS	.215	.139	.070
GLS	.268	.128	.079
OLS/functional form Ω	.221	.139	.047
OLS/jackknife	.220	.146	.072
EGLS	.233	.125	.057
PRE/ols or egls	.247	.122	.061
PRE/var=ols or jackknife	.211	.137	.067
T=60			
OLS	.367	.257	.147
GLS	.385	.308	.198
OLS/functional form Ω	.372	.244	.142
OLS/jackknife	.370	.264	.117
EGLS	.381	.250	.168
PRE/ols or egls	.374	.235	.167
PRE/var=ols or jackknife	.377	.258	.149

Conclusions

The results of these Monte Carlo studies suggest that inferences using EGLS estimates are superior in terms of power, and equivalent in terms of size, to several other procedures over a variety of heteroskedasticity generating processes. However, the superior power of EGLS inferences depends upon the functional form used to model the heteroskedasticity. When the functional form of the heteroskedasticity is known a priori, EGLS is unequivocably the best technique. When the functional form is unknown, some caution in applying EGLS is required, but these results have shown that EGLS inferences are somewhat robust to errors in the selection of the functional form. In particular, the method of Park yielded inferences of superior or equivalent power even in the presence of mixed heteroskedasticity with relatively large constant terms. In practice, a researcher examining the residuals would very likely have detected the mixed heteroskedasticity by comparing the fit of functional forms, or by the use of a rule of thumb such as Kennedy's (1985), and would have produced superior EGLS inferences using the modified Glejser method.

In the case of structural change heteroskedasticity, the method of Park yielded inferences of power very similar to the OLS/jackknife inferences, but the correct EGLS procedure was superior in terms of power. Even a visual inspection of residuals arising from a case of structural change heteroskedasticity would likely have indicated to the practitioner the inappropriateness of the method of Park. It is also common in econometric studies to have some other information which could identify the exact point (or date) at which the structural change switching takes place (eg. a policy change, tax changes, price controls, war, etc.).

These results indicate that the best procedure is to make inferences using EGLS estimates after some preliminary diagnostic work on the OLS residuals whenever there is some empirical evidence or theoretical reason to relate the variance of the errors to an exogenous variable via a simple functional form. This procedure should be followed even if the functional form only approximates the true heteroskedasticity generating process.

In defense of the OLS/jackknife inferences, it should be noted that the technique yields inferences which are superior in terms of both size and power to simple OLS inferences, and also that each of the cases of heteroskedasticity tested here involved a known exogenous variable (either x_t or a 0-1 dummy for structural change) related to the variance of the errors in a rather simple way. The practitioner may well be faced with cases of heteroskedasticity which are unrelated to the independent variables in any simple way, and would therefore be well advised to use the jackknife estimator and ignore the functional form. MacKinnon and White suggest that the jackknife may be applied even when errors are homoskedastic, however they refer only to size and not power results. These results show a small reduction in power as a consequence of using jackknife inferences with homoskedastic data.

The use of critical values from a t distribution when the small sample statistic is only approximately distributed as a t distribution presents no problems for any of the techniques examined. In agreement with the findings of MacKinnon and White for the jackknife, the differences between empirical and nominal sizes are not large, even for sample sizes as small as 30, and disappear as the sample size increases beyond 100. Moreover, the empirical size of EGLS inferences is apparently not affected by choosing the incorrect functional form.

The results show no reason to use the technique of making inferences using OLS coefficient estimates and the functional form estimate for Ω. For inferences using OLS coefficients, the jackknife estimate of the variance yielded higher power than did this technique, and in any case, where good estimates of Ω are possible, EGLS is by far the recommended procedure for estimation and inference. This study therefore reject the hypothesis suggested by earlier studies (Nozari, 1984) that inferences are more accurate in terms of size when made using OLS coefficients.

The use of the Goldfeld-Quandt pre-test to choose between OLS and EGLS inferences in the presence of heteroskedasticity resulted in inferences almost identical to EGLS , but resulted in a significant improvement in power over simple EGLS procedures for the case of homoskedastic errors. The conclusion from this is to recommend the use of the Goldfeld-Quandt pre-test whenever there is any possibility of heteroskedasticity, since the costs of ignoring heteroskedasticity, in terms of reduced power and bias, are high, while the cost of the type I error in the pre-test are low.

Bibliography

Deaton, M.L., Reynolds, M.R., and Myers, R.H. (1983). Estimation and Hypothesis Testing in Regression in the Presence of Nonhomogeneous Error Variances. Communications in Statistics, Simulation and Computation, 12 No. 1, 45-66.

Efron, B. (1982). The Jackknife, the Bootstrap and Other Resampling Plans. Philadelphia, Society for Industrial and Applied Mathematics.

Goldfeld, S.M. and Quandt, R.E. (1972). Non-linear Methods in Econometrics. Amsterdam, North-Holland.

Harvey, A.C. (1976). Estimating Regression Models with Multiplicative Heteroscedasticity. Econometrica, 44, No. 3, 461-465.

Judge, G. et al., (1985). The Theory and practice of Econometrics. New York, Wiley.

Kennedy, P.E., (1985). A Suggestion for Measuring Heteroskedasticity. Communications in Statistics, Simulation and Computation, 14, No. 4, 845-851.

Kennedy, P.E., (1985). A Rule of Thumb for Mixed Heteroskedasticity. Economics Letters, 18, 157-159.

Kleijnen, J.P.C., Cremers, P., and van Belle, F., (1985). The Power of Weighted and Ordinary Least Squares with Estimated Unequal Variances in Experimental Design. Communications in Statistics, Simulation and Computation, 14, No. 1, 85-102.

MacKinnon, J.G., and White, H., (1985). Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties. Journal of Econometrics, 29, 305-325.

Nozari, A., (1984). Generalized and Ordinary Least Squares with Estimated and Unequal Variances. Communications in Statistics, Simulation and Computation, 13, No. 4, 521-537.

Tukey, J., (1958). Bias and Confidence in Not Quite Large Samples, abstract. Annals of Mathematical Statistics, 29 614.

White, H., (1980). A heteroskedasticity-consistent Covariance Matrix and a direct test for heteroskedasticity: Econometrica, 48, 817-838.