MPRA
Munich Personal RePEc Archive

# Implicit Nulls and Alternatives for Hypothesis Tests

Dutta, Jayasri and Zaman, Asad

Barnard College, Columbia University, Katholieke Universiteit Leuven, Center for Operations Research and Econometrics

April 1989

Online at https://mpra.ub.uni-muenchen.de/113344/
MPRA Paper No. 113344, posted 21 Jun 2022 06:46 UTC

# Implicit Nulls and Alternatives for Hypothesis Tests*

Jayasri Dutta† Asad Zaman‡

April 1989, revised December 1992

**Abstract**

A test can be said to 'detect' an alternative hypothesis only if its power against this alternative exceeds its size. We use this principle to define the implicit null and alternative for a test. We analyze the performance of several tests for location and scale parameters.

*We would like to thank Chris Cavanaugh, Phoebus Dhrymes, Christian Gourieroux, Michael Mcaleer, and particularly Alain Trognon for comments; and seminar participants at Columbia, Bonn, and Tilburg Universities, and at INSEE for discussions.

†Barnard College, Columbia University , Katholieke Universiteit Leuven and Center for Operations Research and Econometrics, Louvain

‡Department of Economics, Columbia University

# 1 Definition of the Implicit Null and Alternative

## 1.1 Introduction

Given any statistic $T$ with a known distribution under the null hypothesis, we obtain a hypothesis test by rejecting the null for improbable values of $T$. For example, if $f_0(x)$ is the density of the observation $X$ under the null, the classical tail test rejects the null for small values of $T_0(X) \equiv f_0(X)$. Since any statistic with known distribution can be used, it is easy to generate numerous different ways of testing the same null hypothesis. Typically different tests have superior power in different regions of the parameter space. This method of generating a test pays no attention to the alternative hypotheses, and it is frequently unclear against which alternatives a given test procedure will perform best. Our object in this paper is to associate with each test statistic $T$ two sets of distributions, *the implicit null* denoted $\Theta_0^*$ and the *implicit alternative* denoted $\Theta_1^*$ such it is natural to view $T$ as a test statistic for the null hypothesis $H_0 : \theta \in \Theta_0^*$ versus the alternative $H_1 : \theta \in \Theta_1^*$. Our analysis reveals, for example, that the tail test $T_0$ described above is appropriate when the family of alternatives is the set of translates of $f_0$ but not so when alternatives include rescaled versions of the original density.

With few exceptions, theoretical work on hypothesis tests in econometrics is confined to asymptotic analysis and based on local alternatives. For example Davidson and Mackinnon(1987) define the terms implicit null and alternative to be directions in the tangent space of the manifold defined by the null hypothesis. Thus their definition does not apply to any fixed point in the parameter space at all. Our definitions use finite sample properties and are global — that is, they take into consideration the structure of the full set of alternatives (rather than just those close to the null). Except in cases where a uniformly most powerful (UMP) or a UMP Invariant test exists, it is very difficult to analyze finite sample global properties of hypothesis tests. Most investigations of this kind use numerical methods to compute and compare powers of various tests. Our methods are an improvement in the sense that we get analytical results. However, our results merely indicate regions on which tests will have power greater than the level of the test, without giving more precise information about the power. Thus numerical methods will provide more information. Our results will provide guidance on the appropriate regions of the parameter space to explore via numerical methods.

We apply our definitions to several tests in the regression model to determine the implicit nulls and alternatives. While most tests perform as advertised (i.e. the implicit null and alternatives are as expected), some, for example the Goldfeld-Quandt test, the Breusch-Pagan test, and others do not have the expected properties.

## 1.2 The Implicit Null and Alternative

Suppose we observe $X \sim F^X(x, \theta)$, where $\theta \in \Theta$ the parameter space. We propose to use a statistic $T(X)$ to test the null hypothesis $H_0 : \theta \in \Theta_0$, rejecting for large values of $T$. For every constant $c$, we can evaluate $P(T(x) > c|\theta)$ for $\theta \in \Theta_0$. Choose $c_\alpha$ to be such that

$$\sup_{\theta \in \Theta_0} \mathbf{P}(T(X) > c_\alpha|\theta) = \alpha.$$

A test of level $\alpha$ rejects $H_0$ whenever $T(x) > c_\alpha$. We will consider tests defined by the test statistic $T(x)$; obviously, this defines rejection regions for each $\alpha \in (0, 1)$.

For small $\alpha$, the strategy of rejecting for $T(X) > c_\alpha$ ensures that when the null is true, the probability of rejection is low. However, for any given $\theta_1 \notin \Theta_0$, the probability of rejection may be even lower than $\alpha$; in this case, the test is biased. Define the power of a level-$\alpha$ test based on $T(x)$, at $\theta$, as

$$\beta(\theta, \alpha) = P(T(x) > c_\alpha|\theta).$$

We define the *implicit $\alpha$-alternative* to be the set $\Theta_a^\alpha = \{\theta \in \Theta : \beta(\theta, \alpha) > \alpha\}$; the *implicit $\alpha$-null* is the complement of this set $\Theta_0^\alpha = \{\theta : \theta \notin \Theta_a^\alpha\}$. Note that $\Theta_0 \subseteq \Theta_0^\alpha$. By definition the test which rejects for $T(X) > c_\alpha$ is *unbiased* for $H_0 : \theta \in \Theta_0^\alpha$ versus the alternative $H_1 : \theta \in \Theta_a^\alpha$.

It is possible that for some $\theta \in \Theta$, a test based on $T(x)$ is unbiased at some level $\alpha$, and biased at another, say $\alpha'$. Since the level the test, $\alpha$, cannot be chosen on theoretical grounds, it seems reasonable to require that the alternative hypothesis of a test, which is a qualitative feature, should be invariant to the choice of $\alpha$ .

**Definition 1** *A parameter $\theta \in \Theta$ is an implicit alternative of the test $T(x)$ if and only if for each $\alpha \in (0, 1)$,*

$$\beta(\theta, \alpha) = \mathbf{P}(T(X) > c_\alpha|\theta) > \alpha,$$

*where*

$$\sup_{\theta_0 \in \Theta_0} \mathbf{P}(T(x) > c_\alpha) = \alpha.$$

*It is in the implicit null if and only if, for each $\alpha \in (0, 1)$,*

$$\beta(\theta, \alpha) \leq \mathbf{P}(T(x) > c_\alpha|\theta) \leq \alpha.$$

Define $\Theta_1^*$ as the set of implicit alternatives; and $\Theta_0^*$ the set of implicit nulls of a test. Clearly,

$$\Theta_0^* = \cap_{0<\alpha<1} \Theta_0^\alpha;$$
$$\Theta_a^* = \cap_{0<\alpha<1} \Theta_a^\alpha.$$

With this definition, it is obvious that hypothesis tests based on the statistic $T(X)$ are unbiased at *all* levels for the null $H_0 : \theta \in \Theta_0^*$ versus the alternative $H_1 : \theta \in \Theta_a^*$.

## 1.3 Relation to Stochastic Dominance and MLR

A useful characterization of the implicit alternative can be given in terms of *stochastic dominance*. Recall that random variable $X$ is said to be stochastically larger than $Y$ if for all $t \in \mathbf{R}$, $\mathbf{P}(X > t) \geq \mathbf{P}(Y > t)$. Since the following lemma is just a restatement of the definition of the implicit null and alternatives, we omit the proof.

**Lemma 1** *The distribution of the test statistic $T(x)$ must be stochastically larger under the implicit alternative than under the implicit null; that is, if $\theta_1 \in \Theta_a^*$ and $\theta_0 \in \Theta_0^*$,*

$$\mathbf{P}(T(x) > c|\theta_1) \geq \mathbf{P}(T(x) > c|\theta_0),$$

*for each $c \in \mathbf{R}$.*

Stochastic dominance induces a partial ordering of probability distributions; hence, $\Theta_a^*$ and $\Theta_0^*$ are disjoint. The ordering is not complete, so that some $\theta$ may be neither in the null nor the alternative. The implicit null is always non-empty, since it contains at least $\Theta_0$.

Our definitions are well behaved in the situation where $T(x)$ is the uniformly most powerful test for $H_0 : \theta = \theta_0$. The Monotone Likelihood Ratio or MLR, property implies that a uniformly most powerful test exists. MLR is stronger than stochastic dominance. Suppose $\theta$ is a real parameter, and the densities $f^X(x, \theta)$ have MLR. By Lemma 2 of Chapter 3 in Lehmann(1986), the distribution of the likelihood ratio $T(X) = f^X(x, \theta_1)/f^X(x, \theta_0)$ is stochastically increasing in $\theta$. It follows that for the uniformly most powerful test based on $T(X)$, the implicit null is $\Theta_0^* = \{\theta \leq \theta_0\}$ and the implicit alternative is $\Theta_a^* = \{\theta > \theta_0\}$).

Also note that if a test statistic $T(X)$ provides a UMP Invariant test for $H_0 : \theta \in \Theta_0$ versus the alternative $H_1 : \theta \in \Theta_1$, then it is easily seen that the implicit null is $\Theta_0$ and the implicit alternative is $\Theta_1$. See for example exercise 6 of Chapter 1 of Lehmann(1986) which establishes that the test based on the maximal invariant is unbiased at all levels. It follows immediately that the implicit null and alternative are as stated. From this general result, we can obtain implicit nulls and alternatives for many common tests. For example, in the linear regression model $y = X\beta + \epsilon$ with $\epsilon \sim N(0, \sigma^2\mathbf{I})$, the usual $F$ test for the general linear restriction $H_0 : R\beta = r$ versus $H_1 : R\beta \neq r$ is a UMP invariant test. This is established in Chapter 7 of Lehmann(1986); for an elementary exposition see Zaman(1989). It follows immediately that the implicit null and alternative are exactly as stated.

All examples cited above are related to the fact that Monotone Likelihood Ratio implies stochastic dominance. In the next section we discuss an example where stochastic dominance holds but MLR fails and there is no UMP or UMP invariant test.

# 2 Location Parameters

## 2.1 Pitman's Test

Suppose $X \in \mathbf{R}^k$ has distribution $F^X(x-\theta)$, and density $f^X(x-\theta)$ for some $\theta \in \mathbf{R}^k$. Consider testing the null hypothesis $H_0 : \theta = 0$ versus the alternative $H_1 : \theta \neq 0$.

**Lemma 2 (Pitman(1938))** *The test which rejects for large values of* $T(X) = 1/f^X(X)$ *has the implicit null* $\Theta_0^* = \{0\}$*, and the implicit alternative* $\Theta_a^* = \{\theta \neq 0\}$*.*

*Proof*: Let $A_\alpha = \{x \in \mathbf{R}^k : T(X) \leq c_\alpha\}$ be the *acceptance* region of the $\alpha$-level test based on $T$. For any $\theta_1 \neq 0$, The probability of acceptance is smaller than the same probability under the null:

$$\begin{aligned}\mathbf{P}(X \in A_\alpha|\theta_1) &= \mathbf{P}(X - \theta_1 \in A_k - \theta_1|\theta_1) \\ &= \mathbf{P}(X \in A_k - \theta_1|\theta_0 = 0) \\ &\leq \mathbf{P}(X \in A_k|\theta_0 = 0)\end{aligned}$$

The last inequality is due to the fact that under the null hypothesis, all translates $A_k - \theta_1$ have smaller probability than $A_k$. □

This establishes that $T(X)$ is stochastically larger under the alternative $\theta_1 \neq 0$ than under the null. Clearly, this holds whether or not the family $F$ possesses the monotone likelihood ratio property. It is easily established that the test based on $T$ is a Bayes test when we place the natural invariant prior, Lebesgue measure, on the space of alternatives.

Note that for a null hypothesis $H_0 : X \sim f_0(x)$, the traditional tail test rejects for $f_0(X) \leq c$ of equivalently $T_0(X) = (1/f_0(X)) \geq c'$. Even though alternatives are not explicitly considered, it is clear that this test works well against alternatives which are translates of $f_0$. However if alternative specifications for the density of $X$ include scale changes $H_1 : X \sim (1/\sigma)f_0(x/\sigma)$, the tail test will not do well against all such alternatives. In fact, for symmetric unimodal $f_0$, values of $\sigma \leq 1$ will be part of the implicit null of the test, while $\sigma > 1$ will be the implicit alternative.

## 2.2 Extension of Pitman's test by Projection

Now suppose $X$ has density $f^X(x-\theta)$ for $\theta = (\theta_1, \ldots, \theta_k) \in \mathbf{R}^k$. To test the null hypothesis $H_0 : \theta_1 = \theta_2 = \cdots = \theta_r = 0$ where $r < n$, it seems natural to eliminate the nuisance parameters $\theta_{r+1}, \ldots, \theta_n$ by integration. The marginal density of $X_1, \ldots, X_r$ is simply:

$$f^{(X_1,\ldots,X_r)}(x_1, \ldots, x_r) = \int_{\mathbf{R}^{(n-r)}} \cdots \int f^X(x_1 - \theta_1, \ldots, x_n - \theta_r)\, dx_{r+1} \cdots dx_n$$

By making the change of variables $z_j = x_j - \theta_j$ for $j = r+1, \ldots, n$, it is easily verified that the marginal density does not depend on the unknown nuisance paramters $\theta_{r+1}, \ldots, \theta_n$ and that $\theta_1, \ldots, \theta_r$ occur as translation parameters. Thus after reducing to the marginal densities, we can apply Pitman's original procedure to get a test which is unbiased at all levels. This test, which rejects $H_0$ for $f^{(X_1,\ldots,X_r)} < c$ has implicit null exactly the same as $H_0$ and the implicit alternative every parameter value not in $H_0$.

A further extension of this technique will prove valuable. Let $V$ be an $r$-dimensional vector subspace of $\mathbf{R}^n$ and consider the null hypothesis that the projection of $\theta$ in $V$ (denoted $\Pi_V(\theta)$) is zero. Let $v_1, \ldots, v_r$ be an orthonormal basis for $V$ and let $v_{r+1}, \ldots, v_n$ be an orthonormal basis for the orthogonal complement of $V$. The following result is worth stating as a theorem:

**Theorem 1** *The test based on rejecting for small values of the statistic $T(y)$ given below is unbiased at all levels:*

$$T(y) = \int f^Y(y - [\psi_{r+1}v_1 + \cdots + \psi_n v_n])\, d\psi_{r+1} \cdots d\psi_n$$

**Proof**: Let $V$ stand for the orthogonal $n \times n$ matrix with $j$-th column $v_j$ and consider the transformation $x = V'y$ $\psi = V'\theta$. Note that $y = Vx$ and $\theta = V\psi$. In terms of the parameters $\psi$, the null hypothesis has the simple form $H_0 : \psi_1 = \cdots = \psi_r = 0$. The marginal density of $X_1, \ldots, X_r$ can be written as

$$f^{(X_1,\ldots,X_r)}(x_1, \ldots, x_r, \psi) = \int_{\mathbf{R}^{n-r}} \cdots \int f^Y(V(x - \psi)) dx_{r+1} \cdots dx_n$$

We reject $H_0$ for small values of this density evaluated at $\psi_1 = \cdots = \psi_r = 0$. By changing variables from $x_j$ to $x_j - \psi_j$ we can write the marginal density as

$$f(x_1, \ldots, x_r, \psi) = \int f^Y(Vx - [\psi_1 v_1 + \cdots + \psi_n v_n]))d\psi_{r+1} \cdots d\psi_n$$

Evaluating at $H_0$, we get the statistic $T(y)$ of the theorem

## 2.3 Testing for Equality of Location Parameters

As an application of the test of the previous section, consider testing for equality of the location parameters. Given $y$ with density $f^Y(y - \theta)$, we wish to test $H_0 : \theta_1 = \cdots = \theta_n$. Let $\mathbf{e}$ be the vector of 1's and let $V$ be the space orthogonal to $\mathbf{e}$. The null hypothesis is equivalent to $H_0 : \Pi_V(\theta) = 0$. By the theorem of the previous section, and unbiased test is obtained by rejecting for small values of the statistic

$$T(y) = \int f^Y(y - \psi_n e)\, d\psi_n$$

If the density of $y-\theta$ is standard normal ($y \sim N(\theta, \mathbf{I}_n)$), $T(y)$ is easily calculated:

$$\begin{aligned}
T(y) &= \int \frac{1}{\left(\sqrt{2\pi}\right)^n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(y_i-\psi)^2\right) d\psi \\
&= \int \frac{1}{\left(\sqrt{2\pi}\right)^n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(y_i-\overline{y}+\overline{y}-\psi)^2\right) d\psi \\
&= \frac{\sqrt{2\pi}}{\sqrt{n}\left(\sqrt{2\pi}\right)^n} \exp\left(-\frac{1}{2}\sum(y_i-\overline{y})^2\right) \int \frac{\sqrt{n}}{\sqrt{2\pi}} \exp\left(-\frac{n}{2}(\overline{y}-\psi)^2\right) d\psi \\
&= \frac{1}{\sqrt{n}\left(\sqrt{2\pi}\right)^{n-1}} \exp\left(-\frac{1}{2}\sum(y_i-\overline{y})^2\right)
\end{aligned}$$

Rejecting for low values of $T(y)$ is equivalent to rejecting $H_0$ for large values of $\sum(y_i-\overline{y})^2$.

Another interesting case is that of the lognormal distribution. Suppose $y_i = \log z_i^2$ where $z_i \sim N(0, \sigma_i^2)$ are independent. Then $\theta_i = \log \sigma_i^2$ becomes a location parameter for the $y_i$ which has density:

$$f^{Y_i} = \frac{e^{(y_i-\theta_i)}}{\sqrt{2\pi}} \exp\left(-(1/2)e^{(y_i-\theta_i)}\right)$$

The hypothesis of homoskedasticity (i.e. $H_0 : \sigma_1^2 = \cdots = \sigma_n^2$) is the hypothesis of equality for the translation parameter $\theta$. The statistic $T(y)$ for this case is (with all sums being from 1 to $n$ over index $i$):

$$\begin{aligned}
T(y) &= (2\pi)^{-n/2} \int_{-\infty}^{\infty} e^{\sum(y_i-\theta)} \exp\left(-\frac{1}{2}\sum e^{(y_i-\theta)}\right) d\theta \\
&= (2\pi)^{-n/2} \exp\left(\sum y_i\right) \int_{-\infty}^{\infty} e^{-n\theta} \exp\left(-\frac{1}{2}e^{-\theta}\sum e^{y_i}\right) d\theta
\end{aligned}$$

Making the change of variables $x = e^{-\theta}\sum e^{y_i}$ in the integral, and substituting $\log z_i^2$ for $y_i$, and combining constants into $C$, we get

$$\begin{aligned}
T(y) &= Ce^{\left(\sum y_i\right)} \int \left(\sum e^{y_i}\right)^{-n/2} x^{-(n-2)/2} e^{-x}\, dx \\
&= C' \frac{e^{\sum y_i}}{\left(\sum e^{2y_i}\right)^{-n/2}} \\
&= C' \frac{\left(\prod z_i^2\right)^{1/2}}{\left(\sum \log z_i^2\right)^{-n/2}}
\end{aligned}$$

From this the test statistic based on the ratio of the geometric mean (GM) to the arithmetic mean (AM) of the quantities $z_i^2$ provides a test for homoskedasticity which is unbiased against *all* heteroskedastic alternatives. Note that the test depends on the explicit assumption of normality for the $z_i$.

## 2.4 The Case of Unknown Variance: Impossibility

Now suppose $y \sim N(\theta, \sigma^2 \mathbf{I}_T)$ and consider the null hypothesis $H_0 : \theta = 0$. If $\sigma^2$ is known then a the test statistic $T(y) = \|y\|^2/\sigma^2$ has the correct implicit null and alternative. It worth noting that this statistic will remain the same for *any* unimodal spherically symmetric distribution — let us term this *robustness of the shape of the test* or *shape robustness.* This is not robustness in the usual sense, which refers to robustness of the size; that is, the choice of the constant $c$ for which $T(y) > c$ gives an $\alpha$ level test will depend on the particular distribution. If $\sigma^2$ is unknown then $H_0$ *cannot* be tested. That is, there exists no test which has implicit null $H_0$ and implicit alternative all parameters not in the null.

**Lemma 3 (Impossibility)** *Let $R \subset \mathbf{R}^T$ be any rejection region of size $\alpha \in (0,1)$:*

$$\mathbf{P}(y \in R | y \sim N(0, \sigma^2 \mathbf{I})) \leq \alpha.$$

*Then there exist $\theta \neq 0$ and $\sigma^2$ such that $\mathbf{P}(y \in R | y \sim N(\theta, \sigma^2 \mathbf{I}) < \alpha$.*

*Proof:* The lemma asserts the existence of nonzero $\theta$ in the implicit null. If the rejection region $R$ is a closed set, pick a point $x \notin R$ and a small radius $r > 0$ such the sphere of radius $r$ around $x$, $N(x,r) = \{y : \|y - x\| < r\}$ does not intersect $R$. Setting $\theta = x$ and driving $\sigma$ to zero will make the probability of $N(x,r)$ go to 1 and hence the probability of $R$ go to zero, yielding the desired contradiction. For general measurable rejection regions a more complex argument based on the same idea yields the same result.

## 2.5 Shape Robustness of the F-test

Next consider the general location-scale problem with arbitrary density $f_0$. Suppose we observe $y \sim (1/\sigma) f((y - \theta)/\sigma)$ and wish to test the null hypothesis $H_0 : \theta = 0$. As demonstrated, for $f_0$ normal, if $\sigma$ is unknown, we cannot find a test with implicit alternative $H_1 : \theta \neq 0$. This impossibility result is valid for arbitrary densities $f_0$. We can develop a test when the alternative is restricted, as we now show.

Let $X$ be a vector subspace of $\mathbf{R}^T$. While it is not possible to devise tests with implicit alternative $H_1 : \theta \neq 0$ with unknown $\sigma^2$, it is possible when the alternative is restricted to the set $H_1 : \Pi_X(\theta) \neq 0$. In this case, the condition that $\Pi_{X^\perp}(\theta) = 0$ becomes a maintained hypothesis (valid both under the null and the alternative). Thus it is possible to obtain an estimate of the variance $\sigma^2$ by using $\|\Pi_{X^\perp}(y)\|^2$ for example. The usual $F$ statistic for testing $\Pi_X(\theta) = 0$

versus the alternative $H_1 : \Pi_X(\theta) \neq 0$ is

$$F = c\frac{\|\Pi_X(y)\|^2}{\|\Pi_{X^\perp}(y)\|^2}$$

This statistic has implicit null $\theta = 0$ and implicit alternative $\theta \in X$ for a large class of densities $f_0$ and hence is shape robust. This is the content of our next theorem.

**Theorem 2** *Suppose the random variable $y$ taking values in $\mathbf{R}^T$ has density $(1/\sigma)f_0((y-\theta)/\sigma)$, for some $\theta \in \mathbf{R}^T$ and $\sigma > 0$. IF $f_0$ is symetric and unimodal, the hypothesis test based on the $F$ statistic defined above has implicit null $\Theta_0 : \theta = 0$ and implicit alternative $\theta \in X$.*

**Remarks**: Note that it is well known that $F$ is not size robust; that is, for any constant $c$, the probability of rejection, $\mathbf{P}(F > c|f_0)$ depends on the density $f_0$. Our result shows that implicit alternative and the null do not depend on the density $f_0$. This leads naturally to the idea of whether it is possible to 'estimate' the constant $c$ in some way so as to adapt to the density. One possibility is to use the Bootstrap. This way we can exploit the shape robustness of the $F$ test while avoiding the difficulty of non-robustness of level. We hope to explore these ideas in further research.

*Proof*:

# 3 Scale Parameters

$$y_t = x_t'\beta + u_t; t = 1, \cdots, N;$$

where we assume the errors are symmetric around zero, and there exists a sequence of constants $\sigma_t$ and a distribution $F$ such that $u_t/\sigma_t \overset{i.i.d.}{\sim} F$. In this model, the null hypothesis of homoskedasticity can be stated as:

$$H_0 : \sigma_1^2 = \cdots = \sigma_t^2 = \cdots = \sigma_N^2;$$

this is a hypothesis involving $N - 1$ restrictions. A more geometric description is useful for the intuition. Let $\sigma_N$ stand for the vector $(\sigma_1^2, \ldots, \sigma_N^2)$ in $\mathbf{R}^N$, and let $\mathbf{e}_N$ represent the vector of ones. Let $V(\mathbf{e})$ be the set of all scalar multiples of $\mathbf{e}$, or equivalently, the vector subspace spanned by $\mathbf{e}$. Then the null hypothesis is $H_0 : \sigma_N \in V(\mathbf{e})$. The alternative hypothesis $H_1 : \sigma_n \notin V(\mathbf{e})$ is called non-specific heteroskedasticity. If the set of alternatives is constrained in some way (usually to lie in some low dimensional subspace), we will call this specific heteroskedasticity.

### 3.0.1 Exact Tests

Consider first the family of exact tests. The construction of the test statistic is a straightforward application of the scale- invariant testing procedure suggested by Lehmann and Stein(1948); Kadiyala(1970) discusses properties of this test, and applications to the regression model.

This class of tests is defined as follows. Let $\Delta_N$ be a $N \times N$ diagonal matrix, with distinct non-negative elements $\delta_{Nt}$. The test statistic is

$$S(u_N; \Delta_N) = \frac{u'_N \Delta_N u_N}{u'_N u_N} = \frac{\sum_{t=1}^{N} \delta_{Nt} u_t^2}{\sum_{t=1}^{N} u_t^2} \tag{1}$$

The test rejects for large values of $S$; the critical region of size $\alpha$ can be chosen, in finite samples, by the numerical evaluation of the probability distribution of $S(u_N, \Delta_N)$; the distribution cannot be derived explicitly, even if $u_N$ are normally distributed, except for particular choices of $\delta_{Nt}$.

For each choice of $\Delta_N$, we have a different test. Several authors have suggested different choices of $\Delta_N$, with particular attention to numerical properties, both for the calculation of critical values, as well as power considerations in finite sample simulations.

This test was originally derived for the normal distribution by Lehman and Stein(op. cit.). Suppose that under the null hypothesis, $H_0 : u_N \sim N(0, \sigma^2 I_N), \sigma^2 > 0$. Consider the alternative hypothesis $H_1 : u_N \sim N(0, \sigma^2 \Sigma_N)$. Restrict attention to scale-invariant tests, whose size and power are invariant to $\sigma^2$. Then, the test which rejects for small values of $\frac{u'_N \Sigma_N^{-1} u_N}{u'_N u_N}$ is the most powerful invariant test for testing $H_0$ against $H_1$. Obviously, this test is equivalent to $S(u_N; \Delta_N)$, where

$$\Delta_N = I - \Sigma_N^{-1}.$$

This indicates the explicit alternative of this test in the normal model. As it happens, the test is robust, so that its implicit alternatives and nulls are the same in all distributions $F$. In addition, the implicit alternative is much larger than the class $\Sigma(\Delta_N) = (I - \Delta_N)^{-1}$, the explicit alternative. Nevertheless, we can examine different choices of $\Delta_N$ made by different authors, by considering the properties of $\Sigma(\Delta_N)$.

The main application of the Lehmann-Stein lemma is in testing for autocorrelation; the most celebrated application being the Durbin- Watson statistic (Durbin and Watson (1971));in addition exact tests for unit roots were also constructed by this method (Berenblutt and Webb (1973); Sargan and Bhargava (1983)). One of the earliest applications to heteroskedasticity testing was proposed by Goldfeld and Quandt(1965). They suggested a choice of $\Delta_{Nt} \in \{0, 1\}$. Consider a sample ordered according to some a priori criterion, and construct the test based on $S(u_N, \Delta_N)$ where

$$\begin{aligned} \delta_{Nt} &= 0 \quad \text{if} \quad t < N^* < N \\ &= 1 \quad \text{if} \quad t \geq N^* \end{aligned}$$

Clearly, $\Sigma(\Delta_N)$ has variances constant within, but not across, the two subsamples $(1, \cdots, N^*-1)$ and $(N^*, \cdots, N)$, so that the Goldfeld-Quandt test explicitly tests for a regime change in variances.

Szroeter(1978) examined the properties of this class of tests for heteroskedasticity in the regression model, with $\delta_{Nt}$ set as a monotone increasing function of $\frac{t}{N}$. Among his suggestion were

$$\delta_{Nt} = \frac{t}{N}$$

as well as

$$\delta_{Nt} = 1 - \cos(\pi \frac{t}{N});$$

the latter being suggested for reasons of computation of the finite sample distribution. Evans and King(1988) examined these tests, and other similar ones, for their power against alternative increasing variance sequences. The form of $\Sigma(\Delta_N)$ immediately suggests that the particular choice of $\delta(\frac{t}{N})$ has to be guided by the rate of growth of variances suspected in the alternative.

Information on an auxiliary variable, $x_t$ can be incorporated by choosing $\delta_{Nt} = d(x_t)$. Evans and King(1985) examine the properties of point-optimal tests, where the class of tests has $\delta_{Nt}(\theta) = (1 + \theta x_t)$; for each choice of $\theta \neq 0$, the test $S(u_N, \Delta_N(\theta))$ is the most powerful test against a point alternative $\Sigma_N(\theta)$.

# 4 An Evaluation of Heteroskedasticity Tests

In this Section, we evaluate the two classes of heteroskeasticity tests by characterizing their implicit nulls and alternatives. We also examine the issue of robustness. This shows that robust tests exist for *specific* alternatives, i.e when the alternative hypothesis is restricted to a parameter space, restricted to of lower dimension than the sample size. For example, the class of exact quadratic form tests $S(u_N, \Delta_N)$ are robust, even though they were derived in the context of a given data distribution ( the normal). These tests restrict the class of alternatives to a monotone order .

## 4.1 Exact Tests

Consider the test based on $S(u_N, \Delta_N)$ for fixed $\Delta$. To examine the properties of this test in finite samples, we will drop the subscript $N$, since $\Delta_N$ is fixed for each $N$

We will say that a sequence of variances $\sigma_t^2$ is *monotone increasing (decreasing)* in $\Delta$ if for all $t, t'$ such that $\delta_t < \delta_{t'}$, we have $\sigma_t^2 \leq \sigma_{t'}^2$ ( $\sigma_t^2 \geq \sigma_{t'}^2$ ). As before, $u_t^2 = \sigma_t^2 w_t$ where $w_t \overset{i.i.d.}{\sim} F$. The null hypothesis of homoskedasticity is that $\sigma_t^2 = \sigma^2$ for all $t$. Our first result characterizes the class of implicit alternatives for the test based on $S$ as being precisely the set of heteroskedastic variance sequences which are monotonic in $\delta$. The fact that the test based on $S(u;\Delta)$ , with ordered $\delta_t$ is unbiased for any monotone ordered variance sequence was first noted by Szroeter (op. cit.) in the normal maintained hypothesis. Clearly, this is a robust property which follows from the structure of the test statistic itself. By requiring that the alternative hypothesis be discriminated against by any size $\alpha$-test, we are able to get this monotonicity property to be equivalent to the implicit alternative.

**Theorem 3** *The implicit alternative of the test based on $S(u;\Delta)$ contains all heteroskedastic variance sequences which are monotone increasing in $\Delta$. This condition is necessary and sufficient for a variance sequence to be an implicit alternative for all $F$. A variance sequence is in the implicit null for all $F$ if and only if it is monotone non-increasing in $\Delta$.*

*Proof*: Let $H_1$ be the set of all heteroskedastic variance sequences which are monotone in $\Delta$. We will show that $S(u,\Delta) < S(\Sigma^{1/2}u,\Delta)$ whenever $\sigma_t^2$ are monotone increasing in $\Delta$. Note that for any $c$, the inequality $S(u,\Delta) > c$ is equivalent to

$$\sum_{t=1}^{N}(\delta_t - c)\sigma^2 w_t > 0, \tag{2}$$

and $S(\Sigma^{1/2}u,\Delta) > c$ is equivalent to

$$\sum_{t=1}^{N}(\delta_t - c)\sigma_t^2 w_t > 0 \tag{3}$$

To prove the first assertion of the theorem, we will show that (2) implies (3) under monotone varinces.

Let $T^-$ be the set of all indices such that $\delta_t < c$ and let $T^+$ be the set of all indices such that $\delta_t > c$. Because $\sigma_t^2$ is monotonic increasing in $\Delta$ it must be the case that

$$\max_{t\in T^-}\sigma_t^2 = s^- \leq s^+ = \min_{t\in T^+}\sigma_t^2$$

Multiply both sides of (2) by a constant $k$ such that $k\sigma^2 \in [s^-, s^+]$ and subtract the RHS of (2) from (3) to get:

$$\sum_{t=1}^{N}(\delta_t - c)(\sigma_t^2 - k\sigma^2)w_t > 0 \tag{4}$$

By our choice of $k$, the terms $\delta_t - c$ and $\sigma_t^2 - k\sigma^2$ have matching signs, so the inequality asserted in (4) holds, implying (3).

A similar argument proves the second assertion. If $\sigma_t$ are monotone decreasing in $\Delta$, we can choose $k$ so that $\delta_t - c$ and $\sigma_t^2 - k\sigma^2$ have opposite signs and reverse the inequality in (4). Now if $S(u, \Delta) < c$ under the null, then it must be the case that $S(u, \Delta) < c$ under the monotone decreasing alternative.

We now show that these conditions are necessary and sufficient. To show that a sequence of variances does not belong to the implicit alternative it suffices to show that for some set of values of $w_t$ of nonzero probability measure we have

$$S_1 \equiv \frac{\sum_{t=1}^N \sigma_t^2 \delta_t w_t}{\sum_{t=1}^N \sigma_t^2 w_t} < \frac{\sum_{t=1}^N \delta_t w_t}{\sum_{t=1}^N w_t} \equiv S_0$$

From this it will follow that the statistic $S(u, \Delta)$ cannot be stochastically larger under the alternative (where it equals $S_1$ ) then under the null (where it equals $S_0$ ). Define weights $g_i = \frac{w_i}{\sum_t w_t}$ and $g_i' = \frac{\sigma_i^2 w_i}{\sum_t \sigma_t^2 w_t}$; note that $S_1 = \sum_t g_t' \delta_t$ while $S_0 = \sum_t g_t \delta_t$ . Assume for some $t, t'$, $\delta_t < \delta_{t'}$ while $\sigma_t > \sigma_t'$ so that the variances are not monotonic increasing in $\Delta$. Then $g_t' > g_t$ and $g_{t'}' < g_t'$. Thus under the alternative hypothesis, greater weight (relative to the null) is attached to the smaller number $\delta_t$ and less to the larger one, $\delta_t'$. For outcomes of $w_t$ such that the remaining terms in both sums are more or less in balance, this will force the previous inequality to hold. If $w$'s have full support, the set of such outcomes will have probability greater than zero. A parallel argument for necessity of monotonic decrease in $\Delta$ for a sequence to belong to the implicit null completes the proof. □

This class of tests, which were suggested in the context of the normal distribution, are robust. If the ordering of the variance sequence is known, a priori, the rejection region can be derived for any underlying distribution.

The fact that the implicit alternative contains all variance sequences monotone in $\delta_t$ implies that two tests, $S(u, \Delta_1)$ and $S(u, \Delta_2)$ have the same nulls and alternatives if they are monotone increasing functions of each other. For example, all members of the Szroeter class with $\delta_{i,t} = f_i(t)$ have the same nulls and alternatives as long as $\partial f_i / \partial t > 0$. The actual power achieved against a particular alternative will differ; this is similarly true of point-optimal tests within this class. The Goldfeld-Quandt test has $\delta_t \in \{0, 1\}$. Since this has $\delta_t$ weakly monotone in $t$, the class of implicit alternatives is larger than that of the Szroeter class. Similarly, tests based on auxiliary variables such as $\delta_t = (1 + \theta z_t)$ have implicit nulls and alternatives which depend on the sign of $\theta$, but not its value. In fact, any Szroeter type test, after ordering the sample according to $z_t$, will have the same property.

Heuristically, it is easy to describe the finite sample implicit alternative for the Breusch-Pagan test. Consider an arbitrary vector of variances $\sigma_N = (\cdots, \sigma_t^2, \cdots)'$. This can always be written as

$$\sigma_N = X_N \gamma_N + q_N, \tag{5}$$

where

$$\gamma_N = (X_N' X_N)^{-1} X_N' \sigma_N;$$

$$q_N = (I_N - X_N(X_N' X_N)^{-1} X_N')\sigma_N.$$

Obviously, $q_N$ is the vector of omitted heteroskedasticity, which is the part of $\sigma_N$ orthogonal to $X_N$. Define , for any positive vector $w_N$, the quantities

$$v_N^1 = P_N(\cdots, x_t'\gamma_N w_t, \cdots)';$$

and

$$v_N^2 = P_N(\cdots, q_t w_t, \cdots)'.$$

The test statistic writes as

$$BP_N = \frac{||v_N^1||}{||v_N^1 + v_N^2||}.$$

Note that the test statistic is monotonically declining in $h_N = ||v_N^2||/||v_N^1||$, the size of omitted heteroskedasticity relative to the explained component. This has a surprising implication: a large amount of heteroskedasticity correlated with $X$ will not be detected by the Breusch-Pagan statistic if it is small relative to the total heteroskedasticity.

Presumably, this class of tests is of interest in a situation where the suspected hetroskedasticity is correlated with $X$. However, the test detects such an alternative only if this collinear component is a large *proportion* of actual heteroskedasticity. This is disturbing, since the statistic can fail to be significant even though there is a large amount of heteroskedasticity correlated with $X$. This is seen from the fact that , for each non-zero $\gamma_N$, and a fixed, positive $w_N$, the quantity $h_N$ can be increased arbitrarily by choosing $q_N$ large enough.

There is a second problem, which is that this particular test is not robust, in the following sense. Suppose that $\sigma_t^2$ is a variance sequence satisfying the condition (**??**). Asymptotically, the test detects this form of heteroskedasticity with probability 1, irrespective of the form of the finite sample distribution. The lack of robustness implies that for some data distributions $F$, the induced test will be biased against this alternative in all finite samples. Consider the model exactly as in (5), with $q_N = 0$. To prove that such data distributions exist, it suffices to demonstrate that for some open set $W \in R^{+N}$ , $w \in W$ implies that

$$\frac{\|w'\Sigma_N X_N\|}{\|w_N' X_N\|} < \frac{\|w_N'\Sigma_N P_N\|}{\|w' P_N\|}.$$

Notice that $\Sigma_N X_N$ and $X_N$ are not linearly related, so that we can choose $w$ such that $w'\Sigma_N X_N$ is arbitrarily small, holding $w_N' X_N$ to any preassigned value.

This lack of robustness arises because the test statistic is designed to detect linear correlation between $v$ and $x$; however, with heteroskedasticity, the effect of $x$ on $v$ is multiplicative. As we show next, this can be remedied by an appropriate choice of the function $v(u)$. As we show , the choice $v(u) = \ln u^2$ (instead of $v(u) = u^2$) in Bickel's statistic $B_2(v; X)$ is a robust test. With this modification, the finite sample implicit alternative for the test consists of *all* variances satisfying $\ln \sigma_t^2 = \sigma_0 + x_t'\gamma$ with $\gamma \neq 0$.

**Lemma 4** *Define* $v_N = (\cdots \ln u^2, \cdots)'$. *For* any *distribution* $F$, *consider the test which rejects the null hypothesis of homoskedasticity for large values of the statistic*

$$DZ(u) = \frac{v_N' P_N X_N (X_N' X_N)^{-1} P_N v_N}{v_N' P_N v_N}.$$

*The finite sample implicit alternative for this test includes all variance sequences satisfying* $\ln \sigma_t^2 = \sigma_0 + x_t'\gamma$ *with* $\gamma \neq 0$

*Proof*: We must show that the distribution of $DZ(u)$ is stochastically larger under heteroskedasticity than under the null hypothesis of homoskedasticity. We will drop the subscript $N$ since the result is for a sample of fixed size.Let $w_t^* = \ln(\sigma_0 u_t^2/\sigma_t) = \ln u_t^2 - \theta_t$, (with $\theta_t = x_t'\gamma$ ) so that $w_t^* \overset{i.i.d.}{\sim} F^*$ under both the null and the alternative hypothesis. Since $\theta = 0$ under the null, it suffices to show that

$$\frac{(w^* + \theta)' PX(X'X)^{-1}X'P(w^* + \theta)}{(w^* + \theta)'P(w^* + \theta)} > \frac{w^{*'} PX(X'X)^{-1}X'Pw^*}{w^{*'}Pw^*}$$

This inequality will establish that $DZ(u)$ is *numerically* larger, and not just stochastically larger, under the alternative hypothesis. To prove it, note that by assumption the variables $X$ have been centered at zero, so that $PX = X$. Since $\theta = X\gamma$, we also have $X(X'X)^{-1}X'\theta = \theta$. Simplifying, and multiplying both sides by the denominator of the RHS and dividing by the numerator of the LHS, the inequality becomes

$$\frac{w^{*'} PX(X'X)^{-1}X'Pw^* + 2\theta' w^* + \theta'\theta}{w^{*'}PX(X'X)^{-1}X'Pw^*} > \frac{w^{*'}Pw + 2\theta' w^* + \theta'\theta}{w^{*'}Pw^*}$$

This follows immediately from the fact that $w^{*'} PX(X'X)^{-1}X'Pw^* \leq w{*}'Pw^*$, since projections can only reduce the length of $w^*$. Because the inequality is numerical, stochastic inequality holds simultaneously for all distributions $F$, proving the lemma. □

This statistic avoids one of the difficulties with the Breusch-Pagan test; if the (multiplicative) heteroskedasticity is generated by the $x$'s, the test *will* have power greater than size, regardless of the underlying distribution $F$. The fact that in this version , the test detects exponential rather than linear relations between $\sigma_t^2$ and $x_t$ matters in finite samples, but not asymptotically.

The problem that a large enough $q_N$ will bias the test , persists in this form as well. Define $\theta_N = X_N\gamma_N + q_N$ as before; note that $DZ_N$ is monotonically declining in the quantity $||q_N|/||X_N\gamma_N||$. This a quantity which does not involve $w_N$. Since $||\theta_N|| = ||X_N\gamma_N|| + ||q_N||$, we can construct sequences of $\sigma_N$ such that $\gamma_N$ is constant and $||q_N||$ increases. These are sequences of *increasing* heteroskedasticity on which the power of the test is monotonically declining, no matter what the distribution $F$.

## 4.2 Interpreting Nonspecific Tests

The construction of the Pitman test can be derived in another way, which also clarifies the exact role of distributional assumptions in non-specific testing. [1]

By Assumption 1, the maintained hypothesis has $w_t = u_t^2/\sigma_t^2$ distributed i.i.d ; this must also be true of $v_t = \ln w_t = \ln u_t^2 - \theta_t$ where $\theta_t = \ln \sigma_t^2$. Write $F^v$ as the common distribution of $v_t$. Any distributional assumption for $w$ automatically specifies $F^v$ completely. Write

$$v_t = \mu_F + \theta_t + \epsilon_t; t = 1, 2, \cdots \tag{6}$$

where $\mu_F = \mathbf{E}_F v_t$. Clearly, $\epsilon_t$ are i.i.d with zero mean. The transformation rewrites the problem in familiar regression form.

Consider testing the null hypothesis

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_N = \theta > 0.$$

If $\mu_F$ is known, this hypothesis is testable. In particular, we can estimate $\theta$ under the null hypothesis either as

$$\hat{\theta} = \ln(\frac{\sum_{t=1}^N u_t^2}{N});$$

or as

$$\tilde{\theta} = \frac{\sum_{t=1}^N v_t}{N} - \mu_F.$$

The Pitman test rejects for large values of $\hat{\theta} - \tilde{\theta}$, which is, of course, $\ln(AM : GM)$. If $F$, and hence $\mu_F$, is unknown, the hypothesis is not identified.

Tests against specific, or parametric families such as $\theta_t = h(x_t, \gamma_t)$ can be implemented without knowledge of $\mu_F$. For example, the linear family of alternatives is $\theta_t = x_t'\gamma$. The hypothesis tested is $H_0 : \gamma = 0$; and the chi-squared test yields the test statistic $DZ_N$. Obviously, other models of variable dependent heteroskedasticity could be tested as parametric restrictions.

Next, consider testing subgroup constancy. Divide the sample into $k$ subgroups. The hypothesis to be tested is

$$\theta_i \equiv \frac{\sum_{t_i=1}^{N_i} \theta_{t_i}}{N_i} = \theta_0; i = 1, \cdots, k.$$

[1] We are grateful to A.Trognon for this insight.

This implies that $\sum Ev_i/N_i$ is constant over subgroups i. The hypothesis can be tested without knowledge of $\mu_F$. Bartlett's test (Bartlett(1937)) is exactly a test of this hypothesis.

## 5 Conclusions

In many econometric applications, it is important to test for the presence of heteroskedasticity. The large array of tests for heteroskedasticity currently available bear evidence to this. While the nature of heteroskedasticity – and hence, the most likely direction of the alternative hypothesis – may be evident in some applications , it is not obvious in many others : typically, in cross sectional applications, where the data does not follow a natural order. One would like to have a test which detects all departures from the null hypothesis of homoskedasticity , even if this means giving up precision in any particular direction. A related problem is that the exact nature of the data distribution is never known, so that we would like such a test to perform "reasonably" in all distributions. As our result shows, one cannot have both, at least in finite samples. A test is either robust, or non-specific.

Which of the two must be sacrificed surely depends on the particular problem. We have demonstrated that the standard tests are all "specific": there are heteroskedastic sequences they necessarily fail to detect. Even though some of them (e.g. the LM tests) are not robust, modifications can achieve robustness for a large class of alternatives. We demonstrated this correction for the Breusch and Pagan (1979) test. The basic principle is likely to carry over to alternative applications of LM testing. If we are explicitly ready to give up robustness, the Pitman (1939) principle can be used to derive a truly non-specific test. We used this to derive a test for the normal distribution. This has the attractive property of being able to detect discrepancies in the arithmetic and geometric means of the variance sequence. The trade-off between specificity and robustness is a finite sample phenomenon. Asymptotically, the Pitman test acquires robustness; while the Bickel tests do not become non-specific.

Part of our interest here was to develop methods of comparison for tests not based only on asymptotic properties. Asymptotics, even local asymptotics, glosses over many of the important differences between test procedures. While global comparisons are usually difficult to interpret in high dimensional parameter spaces, it is possible to compare tests at a cruder level. In developing these comparisons, we have given up precision; for example, in comparing finite sample implicit alternatives and nulls, all members of the Szroeter (1978) appear equivalent. They are clearly not, since different tests achieve optimal power against different directions.

## References


[1] Andrews, D.W.K (1989) "Power in Econometric Applications" Econometrica , 57, 5 (1059-1090)

[2] Bahadur, R.R. (1971) " Some Limit Theorems in Statistics" Regional Conference Series in Applied Mathematics, No. 4, National Science Foundation.

[3] Bartlett, M.S (1937) :"Properties of Sufficiency of Statistical Tests" Proceedings of the Royal Statistical Society (Series A), 160, 268- 282

[4] I.I.Berenblutt and G.Webb (1973) : "A New Test for Autocorrelated Errors in the Linear Regression Model" Journal of the Royal Statistical Society, Series B,34,1 33-50

[5] Bickel, Peter J. (1978) "Using Residuals Robustly I:Tests for Heteroscedasticity, Nonlinearity" ,The Annals of Statistics, 6,2, 266-91

[6] Breusch , T and A. Pagan (1979) "A Simple Test of Heteroskedasticity and Random Coefficient Variation" Econometrica, 47, 1287-1294

[7] Chesher, A. (1984)"Testing for Neglected Heterogeneity" Econometrica 52, 865-872.

[8] Davidson,R.L and James G. MacKinnon (1987) Implicit Alternatives and the Local Power of Test Statistics" Econometrica ,55,6,1305-1329.

[9] J.Durbin and G.S.Watson (1971) "Testing for serial Correlation in Least Squares regressions.III" Biometrika ,58,1,1-18

[10] Engle , R.F (1984) : "Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics" in Handbook of Econometrics, Volume II, (ed.) Z.Grilliches and M.Intrilligator, North-Holland, New York

[11] Evans, M.A. and M.L.King (1985) "A Point Optimal Test for Heteroscedastic Disturbances" Journal of Econometrics, 27, 163- 178.

[12] Evans, M.A. and M.L.King (1988) " A Further Class of Tests for Heteroscedasticity", Journal of Econometrics, 37,..

[13] Holly , Alberto (1987) :"Specification Tests: An Overview" in *Advances in Econometrics Fifth World Congress, vol.I.* edited by T. Bewley; Cambridge University press.

[14] Huber , Peter J. (1981) : *"Robust Statistics"* , J.Wiley and sons, New York

[15] Kadiyala ,K.R. (1970) "Testing for The Independence of Regression Disturbances" Econometrica,38,1, 97-117.

[16] Kendall, M.G.M and Stuart ,A. (1973) The Advanced Theory of Statistics Vol 2. 3rd Edition , Griffin, London.

[17] King, M.L (1988) "Towards a Theory of Point Optimal Testing" Econometric Reviews

[18] Koenker, R (1981) A Note on Studentizing a Test for Heteroskedasticity , Journal of Econometrics , 17, 107-112

[19] Lehmann, E and C.Stein (1948): "Most Powerful Tests of Composite Hypotheses I: Normal Distributions" Annals of Mathematical Statistics 19, 495-516

[20] Lehmann, E.L (1986) Testing Statistical Hypotheses ; 2nd Edition John Wiley , New York

[21] Pitman, E.J.G (1938) Tests of Hypotheses Concerning Location and Scale Parameters ,Biometrika ,31, pp 200 231

[22] Ramsey, J.B. (1969) Tests for Specification Errors in Classical Linear Least Squares Regression Analysis ", Journal of the Royal Statistical Society , B , 31, 350-371

[23] Sargan, J.D. and A.Bhargava (1983) "Testing Residuals from the Least Squares Regression for Being Generated by the Gaussian Random walk", Econometrica, 51,1, 153-174.

[24] Szroeter, Jerzy (1978) "A Class of Parametric Tests for Heteroscedasticity in Linear Econometric Models" Econometrica, 46,6 1311-1327.

[25] White , H (1980) A Heteroskedasticity Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity', Econometrica 48: 817-838

[26] White , H (1982) Maximum Likelihood Estimation of Misspecified Models , Econometrica, 50, 1-26