An Exact Test for Multiple Inequality and Equality Constraints in the Linear Regression Model

FRANK A. WOLAK*

In this article we consider the linear regression model $y = X\beta + \varepsilon$, where ε is $N(0, \sigma^2 I)$. In this context we derive exact tests of the form H: $R\beta \geq r$ versus K: $\beta \in R^K$ for the case in which σ^2 is unknown. We extend these results to consider hypothesis tests of the form H: $R_1\beta \geq r_1$ and $R_2\beta = r_2$ versus K: $\beta \in R^K$. For each of these hypotheses tests we derive several equivalent forms of the test statistics using the duality theory of the quadratic programming. For both tests we derive their exact distribution as a weighted sum of Snedecor's F distributions normalized by the numerator degrees of freedom of each F distribution of the sum. A methodology for computing critical values as well as probability values for the tests is discussed.

The relationship between this testing framework and the multivariate one-sided hypothesis testing literature is also discussed. In this context we show that for any size of the hypothesis test H: $R\beta \geq r$ versus K: $\beta \in R^K$ the test statistic and critical value obtained are the same as those from the hypothesis test H: $\lambda = 0$ versus K: $\lambda \geq 0$, where λ is the expectation of the Lagrange multiplier arising from the estimation of β subject to the equality constraints $R\beta = r$. In this way we link the multivariate inequality constraints test to the much studied multivariate one-sided hypothesis test, H: $\mu = 0$ versus K: $\mu \geq 0$, where μ is the mean of a multivariate normal random vector. We also show that the test H: $R_1\beta \geq r_1$ and $R_2\beta = r_2$ versus K: $\beta \in R^K$ has the following equivalent test in terms of λ, H: $\lambda = 0$ versus K: $\lambda_1 \geq 0$, and $\lambda_2 \neq 0$, where λ_1 is the subvector of λ corresponding to $R_1\beta \geq r_1$ and λ_2 corresponds to $R_2\beta = r_2$. Extensions of recent work in one-sided hypothesis testing for the coefficients of the linear regression model are also derived. For the normal linear regression model we derive exact tests for the hypothesis testing problems H: $R\beta = r$ versus K: $R\beta \geq r$ and H: $R\beta = r$ versus K: $R_1\beta \geq r_1$ and $R_2\beta \neq r_2$.

KEY WORDS: Hypothesis testing; Multivariate one-sided tests; Order-restricted inference; Applications of duality theory.

1. INTRODUCTION

Tests for multivariate inequality constraints and combinations of multivariate inequality and equality constraints should have wide application in applied research. For example, in econometric modeling, economic theory often supplies the researcher with a priori information about some or all of the signs of the parameters of the regression, as well as information as to whether a coefficient or sum of coefficients is zero or not. In other cases economic theory provides information about only the signs of several linear combinations of the parameters. Conventional two-sided multivariate tests are not designed to test these null hypotheses implied by economic theory. The tests proposed here are explicitly designed for these purposes. Possible applications of a multivariate inequality constraints testing procedure are not confined to econometrics. In general, these testing procedures offer the researcher a way to statistically test a priori beliefs about the signs of regressions coefficients.

The comparison of a priori knowledge with the empirically estimated model is a specific example of a commonly performed ad hoc procedure. In this procedure the applied researcher has many variables that are believed to influence the dependent variable of the model and, for several of them, a strong belief about the sign of the parameter associated with that variable. In practice, the researcher runs the unconstrained model with all of the variables believed to influence the dependent variable included. Based on the signs of the estimated parameters of the model, one of the variables associated with an incorrectly signed parameter is deleted from the equation and the model is reestimated. If this new estimated equation has any incorrectly signed coefficients, then one of the corresponding variables is removed from the regression and the model is again reestimated. This procedure is repeated until all of the variables left in the equation about which the researcher has a priori beliefs have correctly signed estimated coefficients. A multivariate inequality constraints test allows the researcher to assess, in a hypothesis-testing framework, the validity of this ad hoc procedure for deleting variables from the unrestricted model, that is, whether or not the data is consistent with true values of the parameters satisfying the sign restrictions imposed on the estimated coefficients.

Yancey, Judge, and Bock (1981) discussed tests of the null hypothesis that a subset of the parameter vector lies in the positive orthant for the special case in which the design matrix in the linear regression model is orthogonal ($X'X = I$, the identity matrix), and the covariance matrix of the disturbance vector is scalar [$E(\varepsilon\varepsilon') = \sigma^2 I$]. Our results generalize their results to the case of the arbitrary design matrix and general equality and inequality constraints. The first generalization is essential to applying this testing procedure, because the case of an orthogonal design matrix rarely, if ever, occurs in empirical practice.

Robertson and Wegman (1978) tested order restrictions as a null hypothesis within the context of the exponential family of distributions. They considered hypothesis tests of the form H: $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_K$ versus an unrestricted alternative. Dykstra and Robertson (1983) extended this testing framework to cases in which a collection of independent normal means is, in their words, decreasing on the average. This allows reversals in the aforementioned inequalities over short ranges of the μ_j ($j = 1, \ldots, K$). The general methodology these researchers used to cal-

* Frank A. Wolak is Assistant Professor, Department of Economics, Stanford University, Stanford, CA 94305. This work was supported by National Science Foundation Grant SES-84-20262 at the Department of Economics, Harvard University. The very helpful comments of C. L. Cavanagh, A. K. Han, D. W. Jorgenson, and M. W. Watson on an earlier draft are appreciated. The author gratefully acknowledges the considerable assistance of the associate editor and two referees in providing many useful suggestions for improving the style, content, and clarity of this article. The associate editor deserves special mention for his very thoughtful and detailed editorial comments on several previous versions.

Journal of the American Statistical Association
September 1987, Vol. 82, No. 399, Theory and Methods

culate the null distribution of their likelihood ratio statistic for testing order restrictions can be extended to our problem of testing multivariate inequality constraints.

Gourieroux, Holly, and Monfort (1982), hereafter referred to as GHM, within the context of the linear regression model, dealt with a problem falling in the general class of multivariate one-sided hypothesis tests. They were interested in testing the multivariate equality constraints null hypothesis $R\beta = r$ against the restricted alternative $R\beta \geq r$. We are concerned with testing the null hypothesis of the inequality constraints $R\beta \geq r$ versus an unrestricted alternative.

The derivation of our test statistics relies heavily on the duality theory of quadratic programming. An understanding of the results of the multivariate one-sided hypothesis testing literature is very helpful to understanding our results. Perlman (1969) provided an excellent summary of this literature.

The outline of the article is as follows. In Section 2 we introduce the unconstrained, inequality-constrained, equality-constrained, and mixed inequality- and equality-constrained estimators of the linear regression model. Section 3 contains the derivation of the two likelihood ratio test statistics in their various forms. In Section 4 we take two different tacks to show that the distribution of the test statistics for the purposes of testing the null hypothesis is a weighted sum of Snedecor's F distributions, where entering into each distribution in these sums is a scale factor that is the numerator degrees of freedom of the F distribution. In Section 5 we discuss the computation of the critical value for each test statistic as well as a methodology for computing probability values for our testing procedure. In Section 6 we derive the conditions under which there exists a small sample distribution for the GHM test statistics and also extend their testing framework to consider two-sided equality constraints in conjunction with multivariate one-sided hypotheses. The Appendix gives the available closed-form solutions for the weights used in computing the null distribution.

2. THE THREE ESTIMATORS

Consider the linear regression model

$$y = X\beta + \varepsilon, \tag{2.1}$$

where y is a $(T \times 1)$ vector, X is a $(T \times K)$ matrix of rank K, and β is a $(K \times 1)$ vector. We assume that ε is a $(T \times 1)$ random vector that is $N(0, \sigma^2 I)$, where I is an identity matrix of rank T. We assume that σ^2 is unknown.

The matrix of constraints, R, is a $(P \times K)$ matrix of rank P, where $P \leq K$. The inequality constraints are expressed as $R\beta \geq r$, where r is a known $(P \times 1)$ vector. We should note here that throughout the remainder of the article, $\geq$, when applied to vectors, implies that $\geq$ applies for each element of the two vectors compared. We now define the inequality-constrained estimator for the general linear regression model.

The quadratic programming problem that yields the inequality-constrained least squares (ICLS) estimator is the following:

$$\min_b (y - Xb)'(y - Xb) \qquad \text{subject to } Rb \geq r.$$

There are several methods by which this problem can be solved. Gill, Murray, and Wright (1981) provided an excellent survey of these methods. We will write the solution to this optimization problem as $\tilde{b}$. The $(P \times 1)$ vector of Kuhn–Tucker multipliers for the constraints $Rb \geq r$ is represented by $\tilde{\lambda}$. The unconstrained estimator is the usual ordinary least squares (OLS) estimator, which is $\hat{b} = (X'X)^{-1}X'y$. For completeness, we associate with $\hat{b}$ a Kuhn–Tucker multiplier $\hat{\lambda}$, which is identically zero. We also find it useful to consider the equality-constrained estimator that is the solution to the following optimization problem:

$$\min_b (y - Xb)'(y - Xb) \qquad \text{subject to } Rb = r.$$

We will denote the solution to this optimization problem by $\bar{b}$ and its associated Lagrange multiplier by $\bar{\lambda}$.

Using a derivation by Gourieroux et al. (1982) and Liew (1976), it can be shown that all three of the estimators satisfy the following equation:

$$b_n = \hat{b} + (X'X)^{-1}R'\lambda_n/2, \tag{2.2}$$

where n indexes the unconstrained, inequality-constrained, or equality-constrained estimator. Using this equation for the inequality-constrained estimator yields the following equation:

$$\tilde{b} - \hat{b} = (X'X)^{-1}R'\tilde{\lambda}/2, \tag{2.3}$$

which will prove extremely useful in subsequent sections. Using the relationship (2.2) for the equality-constrained estimator yields the following:

$$R\bar{b} - R\hat{b} = r - R\hat{b} = R(X'X)^{-1}R'\bar{\lambda}/2. \tag{2.4}$$

This follows from the fact that $R\bar{b} = r$, by definition. This implies that

$$2(R(X'X)^{-1}R')^{-1}(r - R\hat{b}) = \bar{\lambda}. \tag{2.5}$$

We will find this relationship useful in later sections for relating $\bar{\lambda}$ to $\hat{b}$.

We can also modify our inequality-constrained estimator framework to consider the mixed inequality- and equality-constrained estimator, which is defined as follows:

$$\min_b (y - Xb)'(y - Xb)$$

$$\text{subject to } R_1 b \geq r_1 \quad \text{and} \quad R_2 b = r_2, \tag{2.6}$$

where R_1 is composed of the first L $(L \leq P)$ rows of R and R_2 is the remaining $P - L$ rows of R. Correspondingly, r_1 is the first L elements of r and r_2 is the remaining $P - L$ elements of r. Associated with the solution of this quadratic program (QP) is a vector of multipliers $\dot{\lambda}' = (\dot{\lambda}_1', \dot{\lambda}_2')$, where $\dot{\lambda}_1$ is associated with the inequality constraints and is hence restricted to be greater than or equal to zero. This is a Kuhn–Tucker multiplier. The subvector $\dot{\lambda}_2$ is associated with the equality constraints and is hence unrestricted. This is a Lagrange multiplier. By $\dot{b}$ we will denote the solution to QP (2.6). For our mixed constraints estimation problem (2.6), the equality-constrained form of this problem is, by construction, the equality-constrained problem considered previously. The unconstrained problem for this case is once again simply OLS estimation. It can be verified that the equality (2.2) continues to hold for the mixed constraints estimator and its associated vector of multipliers.

Note that all of the estimators and their associated Kuhn–Tucker or Lagrange multipliers will be denoted throughout the article by the symbols defined here.

3. DERIVATION OF TEST STATISTICS

For the regression model (2.1), we first derive the likelihood ratio test for the null hypothesis $R\beta \geq r$. The extension to including equality constraints follows once we have the framework for inequalities alone. Throughout this discussion $\hat{b}$ will denote the maximum likelihood estimate (MLE) that is equivalent to the OLS estimate of β and $\tilde{b}$ will denote the inequality-constrained MLE or ICLS estimate, as these two estimates are also equivalent. The Kuhn–Tucker multiplier will be denoted by $\tilde{\lambda}$. Consider the likelihood ratio (LR) test, which is defined in the usual fashion as

$$\text{LR} = -2\ln(\tilde{L}/\hat{L}) = 2(\ln \hat{L} - \ln \tilde{L}),$$

where $\tilde{L}$ and $\hat{L}$ are the maximum values of the likelihood function under the null hypothesis ($R\beta \geq r$) and maintained hypothesis ($\beta \in R^K$), respectively.

It follows that if σ^2 is known the LR statistic takes the following form:

$$\text{LR} = [(y - X\tilde{b})'(y - X\tilde{b}) - (y - X\hat{b})'(y - X\hat{b})]/\sigma^2.$$

In addition, the LR statistic is also the optimal value of the objective function from the following QP:

$$\min_b[(y - Xb)'(y - Xb) - (y - X\hat{b})'(y - X\hat{b})]/\sigma^2 \quad \text{subject to } Rb \geq r. \tag{3.1}$$

This QP can be rewritten as follows:

$$\min_b[y'X(X'X)^{-1}X'y - 2y'Xb + b'X'Xb]/\sigma^2 \quad \text{subject to } Rb \geq r. \tag{3.2}$$

This form will prove useful later, but for the present it puts (3.1) into the form of the standard QP:

$$\min_x a + c'x + \tfrac{1}{2}x'Qx \qquad \text{subject to } Ax \geq b. \tag{3.3}$$

The dual of this standard QP (3.3) can be written in the following form:

$$\max_\lambda \lambda'(b + AQ^{-1}c) - \tfrac{1}{2}\lambda'AQ^{-1}A'\lambda - \tfrac{1}{2}c'Q^{-1}c + a \quad \text{subject to } \lambda \geq 0. \tag{3.4}$$

See Luenberger (1969, chap. 8) or Avriel (1976, chap. 7) for a discussion of the duality theory relevant to this context. If we define problem (3.2) and its equivalent form (3.1) as the primal, then the dual of the optimization problem (3.1), using (3.3) and (3.4), is

$$\max_\lambda[\lambda'(r - R\hat{b}) - \tfrac{1}{4}\lambda'R(X'X)^{-1}R'\lambda]/\sigma^2 \quad \text{subject to } \lambda \geq 0. \tag{3.5}$$

We define the Kuhn–Tucker test statistic (KT) as the optimal value of the dual problem, QP (3.5), which is

$$\text{KT} = \tilde{\lambda}'R(X'X)^{-1}R'\tilde{\lambda}/4\sigma^2.$$

From the theory of quadratic programming we know that the optimal value of the objective function of the primal equals that same value for the dual problem, subject to certain regularity conditions (Gill et al. 1981, p. 76). Necessary conditions are that $(X'X)$ is nonsingular and $R(X'X)^{-1}R'$ is positive definite. As both of these conditions are true by assumption, we have KT = LR. The following two statistics are also equivalent to the KT and LR statistics:

$$W = (R\tilde{b} - R\hat{b})'(R(X'X)^{-1}R')^{-1}(R\tilde{b} - R\hat{b})/\sigma^2 \tag{3.6}$$

and

$$\overline{W} = (\tilde{b} - \hat{b})'(X'X)(\tilde{b} - \hat{b})/\sigma^2. \tag{3.7}$$

Equation (2.3) multiplied on both sides by R allows us to show the equivalence of (3.6) to the KT statistic. Equation (2.3) shows that (3.7) is equivalent to the KT statistic. Both statistics (3.6) and (3.7) are a form of what we define as a Wald statistic for testing multivariate inequality constraints. These statistics are so named for their resemblance to the Wald (1943) test for multivariate equality constraints, because they are defined in a similar fashion, as the magnitude of the difference between the unrestricted estimate and the restricted estimate evaluated in the norm of the covariance matrix of the unrestricted estimate. The $\overline{W}$ statistic is also the optimal value of the objective function from the following QP:

$$\min_b(b - \hat{b})'(X'X)(b - \hat{b})/\sigma^2 \qquad \text{subject to } Rb \geq r. \tag{3.8}$$

This implies that $\tilde{b}$, the ICLS estimator, is also the value of b that solves (3.8). To see this, expand the objective function of (3.8) using the fact that $\hat{b} = (X'X)^{-1}X'y$ and note that the objective function of this problem is the same as that from QP (3.2) and, therefore, QP (3.1). Hence, in the case in which σ^2 is known, the LR, KT, W, and $\overline{W}$ forms of the likelihood ratio statistic for testing multivariate inequality constraints are all equivalent. Therefore, it is clear that all of these forms of the LR statistic possess the same distribution. They will continue to possess the same distribution if we use the same estimate of σ^2 in their computation when σ^2 is unknown.

For our estimate of σ^2 we will use the standard unbiased estimate, s^2, written in the usual fashion as

$$s^2 = (y - X\hat{b})'(y - X\hat{b})/(T - K), \tag{3.9}$$

where $\hat{b}$ is the OLS estimate of β. We know that $s^2(T -$

$K)/\sigma^2$ is distributed as a χ^2_{T-K} random variable. In addition, s^2 is distributed independently of $\hat{b}_i$, for $i = 1, \ldots, K$. Wilks (1962, chap. 10) provided a detailed discussion of the sampling distribution results for the normal linear regression model. The likelihood ratio statistic for the case in which σ^2 is unknown simply replaces σ^2 with s^2 in its LR, KT, W, and $\overline{W}$ forms calculated previously. Theil (1971, pp. 141–145) showed that the classical F test for multivariate equality constraints when σ^2 is unknown may be regarded as a likelihood ratio test for the equality constraints null hypothesis. Using similar logic, the LR, KT, W, and $\overline{W}$ statistics can be shown to be equivalent forms of the likelihood ratio statistic for multivariate inequality constraints.

For the mixed hypothesis test, for the case in which σ^2 is known, the LR form of the likelihood ratio statistic is the optimal value of the objective function from the following QP:

$$\min_b[(y - Xb)'(y - Xb) - (y - X\hat{b})'(y - X\hat{b})]/\sigma^2$$

$$\text{subject to } R_1 b \geq r_1 \quad \text{and} \quad R_2 b = r_2. \tag{3.10}$$

Using the duality theory discussed previously, it can be rewritten in terms of the vector of Kuhn–Tucker and Lagrange multipliers as the optimal value of the objective function from the following QP:

$$\max_\lambda[\lambda'(r - R\hat{b}) - \tfrac{1}{4}\lambda'R(X'X)^{-1}R\lambda]/\sigma^2$$

$$\text{subject to } \lambda_1 \geq 0, \tag{3.11}$$

where λ_1 is composed of the first L elements of λ. The elements of λ_2, the remaining $P - L$ elements of λ, are unrestricted. The W and $\overline{W}$ forms of these test statistics are exactly the same as (3.6) and (3.7) except that $\tilde{b}$ is replaced by $\dot{b}$, where $\dot{b}$ is the solution to QP (3.10) as well as QP (2.6). It follows that the KT form of the statistic is the optimal value of the objective function of (3.11). By the same logic used for the case of the inequality constraints alone, the optimal objective function value of (3.10) equals that from (3.11). Using (2.2) for $\dot{b}$ and $\dot{\lambda}$, we can also show that the W, $\overline{W}$, KT, and LR forms of the likelihood ratio statistic are all equivalent for this case as well. To construct the various forms of our likelihood ratio statistic for the case that σ^2 is unknown we once again use the estimate s^2 given in (3.9) in place of σ^2.

4. DISTRIBUTION OF TEST STATISTICS UNDER NULL HYPOTHESIS

We first consider the case of only inequality constraints. The derivation of the distribution of our test statistic under the null hypothesis is complicated by the fact that our null hypothesis does not specify a unique value for β. It only requires that β satisfy a system of linear inequalities. However, a monotonicity property of the power function of the test and the results of the multivariate one-sided hypothesis testing literature allow us to derive the null distribution of our test statistic for the least favorable configuration (and, therefore, any size test) of our null hypothesis.

Before proceeding with the derivation of the null distribution of our test statistic we first summarize the results of the multivariate one-sided hypothesis testing literature and modify them slightly to fit our framework. This literature deals with the following hypothesis testing problem:

$$H\colon \xi = 0 \text{ versus } K\colon \xi \geq 0, \qquad \xi \in R^P,$$

$$\overline{\xi} = \xi + \eta, \qquad \eta \text{ is a } (P \times 1) \text{ vector that is } N(0, \sigma^2\Delta). \tag{4.1}$$

We assume that Δ is of full rank P, positive definite, and known. For the moment we will assume that σ^2 is also known. This hypothesis testing problem has a long history in the mathematical statistics literature. Bartholomew (1959a,b, 1961) considered a related problem of testing order restrictions between independent normal means. Kudo (1963) extended Bartholomew's results to the specific case considered in (4.1). At around this same time Nuesch (1966) also treated this problem. Perlman (1969) dealt with (4.1) as well as several other problems within the general class of one-sided multivariate hypothesis tests. Perlman considered both null and alternative hypotheses where the mean vector of a multivariate normal random vector lies in a positive homogenous set. We will use the approach of Perlman (1969), as it is the most general of all approaches presented.

Perlman formulated the likelihood ratio test for (4.1) as the maximum value of the objective function from the following QP:

$$\max_\xi[\overline{\xi}'\Delta^{-1}\overline{\xi} - (\xi - \overline{\xi})'\Delta^{-1}(\xi - \overline{\xi})]/\sigma^2$$

$$\text{subject to } \xi \geq 0. \tag{4.1a}$$

Intuitively, this statistic is the difference between two distances: the distance between the unconstrained estimate of ξ, which we have defined previously as $\overline{\xi}$, and its null value of $\hat{\xi}$ (which by hypothesis is zero) and the distance between the unconstrained estimate and the positive orthant. All distances are defined in the norm of the covariance matrix, $\sigma^2\Delta$. Let $\tilde{\xi}$ be the value of ξ that satisfies (4.1a). Define the following statistic:

$$U = [\overline{\xi}'\Delta^{-1}\overline{\xi} - (\tilde{\xi} - \overline{\xi})'\Delta^{-1}(\tilde{\xi} - \overline{\xi})]/\sigma^2$$
$$= \tilde{\xi}'\Delta^{-1}\tilde{\xi}/\sigma^2.$$

From Perlman (1969) we have the following:

Theorem 4.1. Under the null hypothesis $H\colon \xi = 0$, the distribution of the LR statistic U for any $c > 0$ is

$$\Pr_{0,\sigma^2\Delta}[U \geq c] = \sum_{k=1}^{P} \Pr[\chi^2_k \geq c]w(P, k, \Delta).$$

$$\Pr_{0,\sigma^2\Delta}[U = 0] = w(P, 0, \Delta).$$

Proof. See Kudo (1963), Nuesch (1966), or Perlman (1969).

The notation $\Pr_{0,\sigma^2\Delta}[U \geq c]$ denotes the probability of the event $[U \geq c]$ assuming $\bar{\xi}$ is normally distributed with mean 0 and covariance matrix $\sigma^2\Delta$. This notation will be used throughout the remainder of the article. The distribution of U is weighted sum of chi-squared distributions ranging from zero to P df. Note that a χ_k^2 for $k = 0$ is simply a point mass at the origin. Hence for any $c > 0$, $\Pr(\chi_0^2 \geq c) = 0$. Thus if we want the null distribution for all $c \geq 0$ the summation can begin at $k = 0$. The role of the χ_0^2 random variable will become clear in the calculation of the critical value for our testing problem. The sum of the weights, $w(P, k, \Delta)$, from $k = 0$ to P is 1. These weights, as noted previously, depend explicitly on Δ and are the probability that the P-dimensional vector $\bar{\xi}$ has exactly k positive elements.

Closed-form solutions for the weights are available for the cases in which $P \leq 4$ (Kudo 1963). Shapiro (1985) provided alternative closed-form expressions for these weights for the case in which $P = 4$. In the Appendix these expressions are reproduced using our notation. Various numerical methods are available for the cases in which $P \geq 5$. Bohrer and Chow (1978) gave an algorithm that is designed to calculate these weights up to the case in which $P = 10$. Siskind (1976) computed a Taylor expansion of the null distribution of the test statistic for Bartholomew's (1959a,b) hypothesis test for cases in which $P > 4$ and, therefore, avoided the numerical methods necessary to compute the weights and critical values for this testing procedure. Unfortunately, his technique is not straightforward to apply to general problems and it only applies to the cases in which $P \leq 7$.

For some special cases of the covariance matrix of $\bar{\xi}$, Robertson and Wright (1983) gave approximations for the weights used in the computation of the null distribution for Bartholomew's ordered means hypothesis-testing problem for various configurations of the relative magnitudes of the weights. This discussion of the weights shows that their computation is a major stumbling block to the widespread application, to higher-dimensional problems ($P \geq 8$), of this testing framework.

A final methodology for computing these weights for the cases in which $P \geq 8$ is to use Monte Carlo techniques. Here the researcher takes, say 1,000 draws from a multivariate normal distribution with mean zero and covariance matrix Δ. For each draw he computes $\bar{\xi}$ and counts the number of elements of the vector greater than zero. In this case $w(P, k, \Delta)$ is computed as the proportion of the 1,000 draws in which $\bar{\xi}$ has exactly k elements greater than zero. This technique has the following advantages. No expensive numerical integration techniques are required. There are no limits on the values of P for which it is applicable. Because it is a Monte Carlo technique, however, the resulting weights are not exact. Preliminary comparisons of this technique with exact techniques are very encouraging in terms of the degree of agreement with the exact procedure.

Before considering the null distribution of our test statistics from Section 3, we extend Perlman's results for conditions closely related to our testing problem. Suppose that all of the assumptions of the hypothesis-testing problem (4.1) continue to hold except that σ^2 is unknown. In addition, suppose that their exists an unbiased estimate of σ^2, s^2, which is independent of $\bar{\xi}$, such that s^2v/σ^2 is distributed as a chi-squared random variable with v df. Using this estimate of σ^2, our test statistic becomes

$$U^* = \bar{\xi}'\Delta^{-1}\bar{\xi}/s^2. \tag{4.2}$$

As shown previously, the numerator of U^* is a mixture of chi-squared random variables, which are all independent of s^2. Note that s^2/σ^2 is a chi-squared random variable divided by its degrees of freedom. By definition, the ratio of two independent chi-squared random variables divided by their respective degrees of freedom is distributed with Snedecor's F distribution, with its two parameters the degrees of freedom of the numerator and denominator chi-squared random variables. Using this definition we state the null distribution of U^* in the form of a corollary whose proof follows from that of Theorem 4.1.

Corollary 4.1. Consider the case in which σ^2 is unknown but there exists an unbiased estimate of it distributed independently of $\bar{\xi}$ such that s^2v/σ^2 is distributed as χ_v^2. Under the null hypothesis H: $\xi = 0$, the distribution of the modified LR statistic U^* is, for all $c^* > 0$,

$$\Pr_{0,\sigma^2\Delta}[U^* \geq c^*] = \sum_{k=1}^{P} \Pr[F_{k,v} \geq c^*/k]w(P, k, \Delta),$$

$$\Pr_{0,\sigma^2\Delta}[U^* = 0] = w(P, 0, \Delta).$$

Note that the null distribution of U^* depends only on the elements of Δ, which are assumed to be known parameters. We are now able to proceed with the derivation of the null distribution of our LR statistic for the case that σ^2 is unknown.

To do this we first consider the following testing problem:

$$H: \mu \geq 0 \quad \text{versus} \quad K: \mu \in R^P, \qquad \hat{\mu} = \mu + \nu,$$
$$\nu \text{ is a } (P \times 1) \text{ vector that is } N(0, \sigma^2\Omega). \tag{4.3}$$

We assume that Ω is of full rank P and known. For the moment assume that σ^2 is also known. For this problem our sample space in the Neyman–Pearson framework is $\Theta = R^P$. The positive orthant and its boundary in P-dimensional space is the subset of Θ in which μ lies under the null hypothesis. We denote this by Θ_H, and its relative complement under Θ is denoted by Θ_K. Following Lehmann (1959), let s be the test statistic for our hypothesis test and S the rejection region. If

$$\sup_{\mu \in \Theta_H} \Pr_{\mu,\sigma^2\Omega}(s \in S) = \alpha,$$

then S is the rejection region for a size α test of our null hypothesis. We will now show how to construct a rejection region for a level α test of our testing problem (4.3).

In Perlman (1969) a related form of this hypothesis-testing problem was considered. Using Perlman's logic, the LR test for this problem is the minimum value of the objective function from the following QP:

$$\min_{\mu}(\hat{\mu} - \mu)'\Omega^{-1}(\hat{\mu} - \mu)/\sigma^2 \qquad \text{subject to } \mu \geq 0.$$

Denote by $\tilde{\mu}$ the solution to this QP. Define

$$Z = (\hat{\mu} - \tilde{\mu})'\Omega^{-1}(\hat{\mu} - \tilde{\mu})/\sigma^2. \tag{4.4}$$

A special case of lemma 8.2 in Perlman (1969) is given below.

Lemma 4.1. For any $\mu \geq 0$ and $c \in R_+$ the following is true:

$$\Pr_{\mu,\sigma^2\Omega}[Z \geq c] \leq \Pr_{0,\sigma^2\Omega}[Z \geq c].$$

Proof. See Perlman (1969, pp. 562–563). Note that this lemma continues to hold for the case in which σ^2 is unknown and is replaced by an estimate of it that is independent of the numerator of (4.4) in the computation of the statistic Z.

As an immediate corollary we have

$$\sup_{\mu\in\Theta_H} \Pr_{\mu,\sigma^2\Omega}[Z \geq c] = \Pr_{0,\sigma^2\Omega}[Z \geq c].$$

We now have identified the unique least favorable value of μ to specify in order to compute the null distribution for a size α test of our composite null hypothesis. This monotonicity property gives results similar to those for the univariate inequality-constraint test. In that case the least favorable value of the mean of a normal random variable is zero for a test of the null hypothesis that this mean is greater than or equal to zero. We should note here that Lemma 4.1 and its corollary continue to hold for any Θ_H that takes the form of a convex cone.

We state the distribution under the null hypothesis of our test statistic (4.4) in the form of a theorem whose proof is given in Wolak (1987).

Theorem 4.2. Under the null hypothesis $\mu \geq 0$ the distribution of the likelihood ratio statistic Z has the following property for all $c > 0$:

$$\sup_{\mu\in\Theta_H} \Pr_{\mu,\sigma^2\Omega}[Z \geq c] = \Pr_{0,\sigma^2\Omega}[Z \geq c] = \sum_{k=1}^{P} \Pr[\chi^2_k \geq c]w(P, P - k, \Omega),$$

$$\sup_{\mu\in\Theta_H} \Pr_{\mu,\sigma^2\Omega}[Z = 0] = \Pr_{0,\sigma^2\Omega}[Z = 0] = w(P, P, \Omega).$$

The weights, $w(P, m, \Omega)$, are of the same functional form as those calculated for the multivariate one-sided testing problem described previously.

If we are faced with a case in which σ^2 is unknown but we have an unbiased estimate of it that is distributed independently of $\hat{\mu}$, we can then apply the same logic used to get Corollary 4.1 from Theorem 4.1 to Theorem 4.2. Let

$$Z^* = (\hat{\mu} - \tilde{\mu})'\Omega^{-1}(\hat{\mu} - \tilde{\mu})/s^2$$

be the modified LR statistic for our inequality-constraints test.

Corollary 4.2. Consider the case in which σ^2 is unknown but there exists an unbiased estimate of it distributed independently of $\hat{\mu}$ such that s^2v/σ^2 (where s^2 is the unbiased estimate of σ^2) is distributed as χ^2_v. The distribution of the modified LR statistic Z^* has the following property for all $c^* > 0$:

$$\sup_{\mu\in\Theta_H} \Pr_{\mu,\sigma^2\Omega}[Z^* \geq c^*] = \Pr_{0,\sigma^2\Omega}[Z^* \geq c^*] = \sum_{k=1}^{P} \Pr[F_{k,v} \geq c^*/k]w(P, P - k, \Omega),$$

$$\sup_{\mu\in\Theta_H} \Pr_{\mu,\sigma^2\Omega}[Z^* = 0] = \Pr_{0,\sigma^2\Omega}[Z^* = 0] = w(P, P, \Omega).$$

Consider the following special case of our problem (4.3) where we define

$$\mu = R\beta - r, \qquad \hat{\mu} = R\hat{b} - r, \qquad \text{and} \quad \Omega = R(X'X)^{-1}R'. \tag{4.5}$$

As stated in Section 3, we use $s^2 = (y - X\hat{b})'(y - X\hat{b})/(T - K)$ as our estimate of σ^2 used in the computation of the LR statistic. In this case our test statistic is

$$\min_{\mu}(\hat{\mu} - \mu)'[R(X'X)^{-1}R']^{-1}(\hat{\mu} - \mu)/s^2 \qquad \text{subject to } \mu \geq 0. \tag{4.6}$$

Using (3.4) note that the dual of (4.6) is

$$\max_{\lambda}[-\lambda'\hat{\mu} - \tfrac{1}{4}\lambda'R(X'X)^{-1}R'\lambda]/s^2 \qquad \text{subject to } \lambda \geq 0. \tag{4.7}$$

By definition $-\hat{\mu} = r - R\hat{b}$; therefore, we can rewrite (4.7) as

$$\max_{\lambda}[\lambda'(r - R\hat{b}) - \tfrac{1}{4}\lambda'R(X'X)^{-1}R'\lambda]/s^2 \qquad \text{subject to } \lambda \geq 0. \tag{4.8}$$

Recall the $\overline{W}$ form of the likelihood ratio statistic that was defined as the optimal value of (3.8). Using (3.4) we can show that the dual of (3.8) is also

$$\max_{\lambda}[\lambda'(r - R\hat{b}) - \tfrac{1}{4}\lambda'R(X'X)^{-1}R'\lambda]/s^2 \qquad \text{subject to } \lambda \geq 0. \tag{4.9}$$

Because the duals of (3.8) and (4.6) are exactly the same optimization problem the optimal objective function values of (3.8) and (4.6) are equivalent. In addition, because the Kuhn–Tucker multipliers from (3.8) and (4.6) are equivalent we know that the solution to (4.6) is $\tilde{\mu} = R\tilde{b} - r$, where $\tilde{b}$ is the solution to (3.8) as well as the ICLS estimator. Thus our test statistic Z^* and the four forms of the likelihood ratio statistic considered in Section 3 are equivalent. Hence their distribution under the null hypothesis $R\beta \geq r$ is equivalent to the null distribution of Z^*. This distribution is a weighted sum of F distributions as given in Corollary 4.2.

Theorem 4.3. For the case in which σ^2 is unknown but replaced by s^2, under the null hypothesis $R\beta \geq r$ the distribution of LR, the likelihood ratio statistic, has the following property for all $c > 0$:

$$\sup_{R\beta \geq r} \Pr_{\beta,\sigma^2(X'X)^{-1}}[\text{LR} \geq c]$$
$$= \Pr_{\beta^*,\sigma^2(X'X)^{-1}}[\text{LR} \geq c]$$
$$= \sum_{k=1}^{P} \Pr[F_{k,T-K} \geq c/k] w(P, P-k, A),$$
$$\sup_{R\beta \geq r} \Pr_{\beta,\sigma^2(X'X)^{-1}}[\text{LR} = 0] = \Pr_{\beta^*,\sigma^2(X'X)^{-1}}[\text{LR} = 0]$$
$$= w(P, P, A),$$

where β^* is any value of β such that $R\beta^* = r$ and $A = [R(X'X)^{-1}R']$.

Proof. Given the distributional properties of s^2 and $\hat{b}$, this result follows directly from corollary 4.3 of Wolak (1987).

To relate our testing problem to the multivariate one-sided hypothesis test, we now consider our testing problem in terms of the vector of dual variables. Recall that, as shown in Theorem 4.3, we choose a least favorable value of β such that $R\beta = r$ for any size test of our null hypothesis. This implies that for this value of β the vector of multipliers arising from the equality-constrained estimation of β, $\bar{\lambda}$ is $N(0, 4\sigma^2[R(X'X)^{-1}R']^{-1})$.

Before proceeding we define some notation. Let $\hat{v} = r - R\hat{b}$. Recall the dual of QP (3.1), QP (3.5), which was used to calculate the KT form of the LR statistic. Written using our new notation it becomes

$$\max_{\lambda} [\lambda'\hat{v} - \tfrac{1}{4}\lambda' R(X'X)^{-1}R'\lambda]/s^2 \qquad \text{subject to } \lambda \geq 0. \tag{4.10}$$

We can complete the square of this objective function by adding and subtracting $\hat{v}'(R(X'X)^{-1}R')^{-1}\hat{v}/s^2$ into it. We can rewrite (4.10) as follows:

$$\max_{\lambda} - [(\lambda - 2(R(X'X)^{-1}R')^{-1}\hat{v})' R(X'X)^{-1}R',$$
$$(\lambda - 2(R(X'X)^{-1}R')^{-1}\hat{v})]/4s^2 + \hat{v}'(R(X'X)^{-1}R')^{-1}\hat{v}/s^2$$
$$\text{subject to } \lambda \geq 0. \tag{4.11}$$

Recall Equation (2.5). From that we note that

$$\bar{\lambda} = 2(R(X'X)^{-1}R')^{-1}\hat{v}.$$

This allows us to show the following equivalence:

$$\bar{\lambda}' R(X'X)^{-1}R'\bar{\lambda}/4 = \hat{v}'(R(X'X)^{-1}R')^{-1}\hat{v}. \tag{4.12}$$

Using (2.5) and (4.12) we can rewrite (4.10) as

$$\max_{\lambda} \bar{\lambda}'(R(X'X)^{-1}R')\bar{\lambda}/4s^2$$
$$- (\lambda - \bar{\lambda})'(R(X'X)^{-1}R')(\lambda - \bar{\lambda})/4s^2$$
$$\text{subject to } \lambda \geq 0. \tag{4.13}$$

If we replace ξ by λ and Δ by $4(R(X'X)^{-1}R')^{-1}$, then the problem (4.13) is exactly the same as problem (4.1a) for the case in which σ^2 is unknown but estimated by s^2. Hence our statistic U^* from the hypothesis-testing problem (4.1) is equal to the four equivalent forms of the LR statistic derived in Section 3. The solution to QP (4.13) is $\tilde{\lambda}$, the KT multiplier vector defined in Section 2. Using the dual approach to the multivariate inequality-constraints test we can derive the null distribution by an application of the results of the multivariate one-sided hypothesis test. Because $R\beta = r$ for the purposes of computing the null distribution for any size test of our inequality-constraints null hypothesis, this test procedure results in the same critical value as the test in terms of the dual variables of the null hypothesis that the true value of the Lagrange multiplier is zero versus the restricted alternative that it is greater than or equal to zero. The true value of the Lagrange multiplier, λ, is the expectation of $\bar{\lambda}$. Taking the expectation of both sides of (2.5) yields $\lambda = 2(R(X'X)^{-1}R')^{-1}(r - R\beta)$, so λ is in fact defined in terms of β as well as being the expected value of $\bar{\lambda}$. Recall s^2, our unbiased estimate of σ^2. From the results of normal linear regression theory, s^2 is independent of $\hat{b}$ and also $\bar{\lambda}$, as $\bar{\lambda}$ is $N(0, 4\sigma^2[R(X'X)^{-1}R']^{-1})$ and by (2.5) equal to a known matrix times $\hat{b}$. Thus our estimate of σ^2 satisfies the requirement of Corollary 4.1. We summarize our result about the null distribution of our LR statistics with s^2 as our estimate of σ^2 in the following theorem.

Theorem 4.4. For the hypothesis-testing problem H: $\lambda = 0$ versus K: $\lambda \geq 0$ (which by Lemma 4.1 is equivalent to, for the purpose of computing critical values for any size test, the testing problem H: $R\beta \geq r$ versus K: $\beta \in R^K$), the null distribution of the LR statistic with σ^2 replaced by $s^2 = (y - X\hat{b})'(y - X\hat{b})/(T - K)$ is

$$\Pr_{0,4\sigma^2\Lambda}[\text{LR} \geq c^*] = \sum_{k=1}^{P} \Pr[F_{k,T-K} \geq c^*/k] w(P, k, 4\Lambda),$$
$$\Pr_{0,4\sigma^2\Lambda}[\text{LR} = 0] = w(P, 0, 4\Lambda),$$

where $\Lambda = (R(X'X)^{-1}R')^{-1}$.

Note that the weights in Theorem 4.3 depend on $P - k$ and $[R(X'X)^{-1}R']$ and those in Theorem 4.4 depend on k and $4[R(X'X)^{-1}R']^{-1}$. Wolak (1987) showed that $w(P, k, \Gamma) = w(P, P - k, b\Gamma^{-1})$, $b > 0$ for $k = 0$ to P. Hence the two null distributions in Theorem 4.3 and Theorem 4.4 are the same weighted sum of F distributions. Thus by either the primal or dual methodology we can show the same distribution for our test statistic for any size test of our null hypothesis.

We will now consider the null distribution of our mixed equality- and inequality-constraints test statistics. We can construct the null distribution of the mixed constraints test statistics in terms of the vector of multipliers using the same logic used to construct the null distribution of the inequality-constraints LR statistics. Recall our mixed constraints hypothesis-testing problem:

$$H\colon R_1\beta \geq r_1 \quad \text{and} \quad R_2\beta = r_2 \quad \text{versus} \quad K\colon \beta \in R^K. \tag{4.14}$$

By applying Lemma 4.1 in the same fashion as it is used in Theorem 4.3 to the inequality constraints $R_1\beta \geq r_1$, we know that for the purposes of testing our null hypothesis we choose a β such that $R_1\beta = r_1$ as well as $R_2\beta = r_2$. This implies that in terms of λ, the expected value of $\bar{\lambda}$, following the logic of Theorem 4.4, the equivalent dual testing problem takes the following form:

$$H\colon \lambda = 0 \quad \text{versus} \quad K\colon \lambda_i \geq 0, \quad i = 1, \ldots, L \leq P;$$
$$\lambda_i \neq 0, \quad i = L + 1, \ldots, P, \tag{4.15}$$

where $\bar{\lambda} = \lambda + \eta$ and η is distributed as $N(0, 4\sigma^2 (R(X'X)^{-1}R')^{-1})$. For the moment, we assume that σ^2 is known. For the inequality constraints, λ_i is constrained to be greater than or equal to zero with strict inequality for at least one element under the alternative. This is a multivariate one-sided hypothesis test in terms of the λ_i ($i = 1, \ldots, L$). For the equality constraints, λ_i is unconstrained under the alternative. In other words, for the equality constraints we are considering a two-sided or unrestricted alternative for the λ_i ($i = L + 1, \ldots, P$).

Recall optimization problem (2.6), which defines the mixed equality- and inequality-constrained estimator. By following the logic used to derive Theorem 4.4 we can show that the likelihood ratio statistic for the mixed constraint case is also the optimal value of the objective function from the following QP:

$$\max_{\lambda} \bar{\lambda}'(R(X'X)^{-1}R')\bar{\lambda}/4\sigma^2 - (\lambda - \bar{\lambda})'(R(X'X)^{-1}R')(\lambda - \bar{\lambda})/4\sigma^2$$
$$\text{subject to } \lambda_i \geq 0, \quad i = 1, \ldots, L. \tag{4.16}$$

The remaining elements of λ are unrestricted. Denote this optimal objective function value by Y. This is the same form of the LR test statistic Kudo (1963) derived for the hypothesis-testing problem (4.15). Recall that $\bar{\lambda}$, the Lagrange multiplier arising from the equality constrained estimation procedure, is distributed as $N(0, 4\sigma^2 (R(X'X)^{-1}R')^{-1})$. Kudo also derived the null distribution of the LR test statistic for the hypothesis-testing problem (4.15). We state his result, which holds for our test statistics, in the notation of our mixed equality- and inequality-constraints testing framework in the following lemma.

Lemma 4.2. For the hypothesis-testing problem (4.15), under the null hypothesis $H\colon \lambda = 0$, the LR test statistic Y has the following distribution:

$$\Pr(Y \geq c) = \sum_{k=0}^{L} \Pr(\chi^2_{P-L+k} \geq c) w(L, k, \Psi),$$

where Ψ is the submatrix of $(R(X'X)^{-1}R')^{-1}$ corresponding to λ_i ($i = 1, \ldots, L$).

Proof. See Kudo (1963).

To extend this result to the case in which σ^2 is unknown we note that $\bar{\lambda}$ is independent of our estimate of σ^2, s^2. Thus if we substitute s^2 for σ^2 in our test statistic (4.16) and call this statistic Y^* we have the following lemma.

Lemma 4.3. Consider the hypothesis-testing problem $H\colon \lambda = 0$ versus $\lambda_1 \geq 0$ and $\lambda_2 \neq 0$ (which by Lemma 4.1 is equivalent to, for any size test, the hypothesis test $H\colon R_1\beta \geq r_1$ and $R_2\beta = r_2$ versus $K\colon \beta \in R^K$). For the case in which σ^2 is unknown, by replacing it with s^2, we have the following distribution, for the purpose of testing our null hypothesis, of the statistic Y^*:

$$\Pr(Y^* \geq d) = \sum_{k=0}^{L} \Pr[F_{P-L+k,T-K} \geq d/(P - L + k)] w(L, k, \Psi),$$

where Ψ is the submatrix of $(R(X'X)^{-1}R')^{-1}$ corresponding to λ_i ($i = 1, \ldots, L$).

We can also consider the null distribution of the mixed equality- and inequality-constraints test statistics in terms of the primal problem by extending Theorem 4.3 and recalling that the appropriate monotonicity property of the power function in β obtained from Lemma 4.1 continues to hold in this case. Using this logic we have the following lemma.

Lemma 4.4. For the hypothesis-testing problem $H\colon R_1\beta \geq r_1$ and $R_2\beta = r_2$ versus $K\colon \beta \in R^K$, in the case in which σ^2 is known but replaced by s^2, the distribution of the LR statistic Y^* satisfies the following property:

$$\sup_{\beta \in B} \Pr_{\beta,\sigma^2(X'X)^{-1}}(Y^* \geq d) = \Pr_{\beta^*,\sigma^2(X'X)^{-1}}(Y^* \geq d)$$
$$= \sum_{k=0}^{L} \Pr[F_{P-L+k,T-K} \geq d/(P - L + k)] w(L, L - k, \Pi),$$

where Π is the covariance matrix of $R_1\ddot{b}$ divided by σ^2 and β^* is any β such that $R\beta = r$.

We define B as follows: $B = \{\beta \mid R_1\beta \geq r_1 \text{ and } R_2\beta = r_2, \beta \in R^K\}$. We define $\ddot{b}$ as the estimate of β calculated assuming that $R_2\beta = r_2$. We use the covariance matrix of $R_1\ddot{b}$ in the computation of the weights because under our null hypothesis the unrestricted estimate of β assumes that $R_2\beta = r_2$. From Silvey (1970) the covariance matrix of $\ddot{b}$ divided by σ^2 is

$$\text{var}(\ddot{b})/\sigma^2 = (X'X)^{-1} - (X'X)^{-1}R_2'(R_2(X'X)^{-1}R_2')^{-1}R_2(X'X)^{-1}.$$

This implies that

$$\text{var}(R_1\ddot{b})/\sigma^2 = R_1(X'X)^{-1}R_1' - R_1(X'X)^{-1}R_2'(R_2(X'X)^{-1}R_2')^{-1}R_2(X'X)^{-1}R_1'.$$

We have shown that $\bar{\lambda}$, the equality-constrained estimate of λ; that value of λ arising from the computation of an estimate of β given the constraints $R_1\beta = r_1$ and $R_2\beta = r_2$ (which means that $R\beta = r$) has the following covariance matrix:

$$\operatorname{var}(\bar{\lambda}) = 4\sigma^2(R(X'X)^{-1}R')^{-1}.$$

Recalling that

$$R = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix},$$

we can rewrite $R(X'X)^{-1}R'$ as follows:

$$\begin{bmatrix} R_1(X'X)^{-1}R_1' & R_1(X'X)^{-1}R_2' \\ R_2(X'X)^{-1}R_1' & R_2(X'X)^{-1}R_2' \end{bmatrix}.$$

By the partitioned matrix inversion lemma (see Theil 1971, p. 18) the element of $(R(X'X)^{-1}R')^{-1}$ corresponding to $R_1(X'X)^{-1}R_1'$ is

$$Q = (R_1(X'X)^{-1}R_1' - R_1(X'X)^{-1}R_2'(R_2(X'X)^{-1}R_2')^{-1}R_2(X'X)^{-1}R_1')^{-1}.$$

Note that $k[\operatorname{var}(R_1\ddot{b})]^{-1} = Q$, $k > 0$. Note in addition that the weights of Lemma 4.4 depend on $\operatorname{var}(R_1\ddot{b})/\sigma^2$ and those in Lemma 4.3 depend on Q. As shown in Wolak (1987) and claimed here previously, $w(J, m, A) = w(J, J - m, aA^{-1})$ ($a > 0$). Hence the two weighted sums of F distributions given in Lemmas 4.3 and 4.4 are exactly the same. Our primal–dual relationship used to compute the null distribution for the case of multiple inequality constraints continues to hold for the case of combinations of multiple inequality and equality constraints.

We now have the exact null distribution for both the inequality-constraints test statistic and the mixed equality- and inequality-constraints test statistic. We should note here that these results hold for more general forms of the covariance matrix of the errors than those of the form $\sigma^2 I$. If the covariance matrix of ε is $\sigma^2\Lambda$ where Λ is known and positive definite, then all of our small sample results continue to hold for both sets of test statistics modified appropriately. Note that for the case in which $L = 0$ our mixed inequality- and equality-constraints test reduces to the standard multivariate equality-constraints test, the well-known F test.

It is natural at this point to discuss the power of these tests. For the case in which the matrix $(X'X) = I$, the identity matrix, the null hypothesis is $\beta \geq 0$, and $\beta \in R^2$; power calculations are reported in Yancey et al. (1982). The power of this test, for testing our one-sided null hypothesis, is at least as great in all cases as the two-sided test, H: $R\beta = r$ versus K: $R\beta \neq r$, because it takes into account the fact in our case that $\lambda > 0$ under the alternative. One would also expect the mixed inequality and equality test statistic to have superior power properties for mixed null hypotheses when compared with standard multivariate equality-constraints tests, for the same reason. For a discussion of power for multivariate one-sided hypothesis tests see Bartholomew (1961) and Barlow, Bartholomew, Bremner, and Brunk (1972). Wolak (1987) discussed the computation of the power function for the multivariate inequality-constraints test. Goldberger (1986) provided an excellent study of the comparative power properties of multivariate one-sided and inequality-constraints tests.

At this point we should also note some of the properties possessed by our test. As stated in Perlman (1969) for all testing problems of this class, the power of the test approaches 1 uniformly in σ^2 and β as the distance, in the norm of the covariance matrix of $\hat{b}$, between β and where it lies under the null hypothesis tends to infinity. The tests are not unbiased. The least favorable distribution is obtained at a β such that $R\beta = r$. The power is smaller for values of β elsewhere in the region defined by the null hypothesis. By continuity of the power function of the test statistic, there are values of β not in the region defined by the null hypothesis where the power is smaller than when β is such that $R\beta = r$. Nevertheless, our tests are consistent. For values of β such that $R\beta \ngeq r$ the power of our test approaches 1 as T tends to infinity.

5. APPLYING TEST STATISTICS

Because both test statistics have a null distribution that is a weighted sum of F distributions, the calculation of the critical value for a hypothesis-testing problem no longer is as simple as looking up the relevant number in the tables of quantiles of the F distribution. The widespread availability, however, of FORTRAN subroutine libraries, such as the IMSL library (International Mathematical and Statistical Libraries, Inc. 1982), make the task substantially easier.

For the level α test in the pure inequality-constraints case the critical value is the solution in x of the following equation:

$$\alpha = \sum_{k=1}^{P} \Pr[F_{k,T-K} \geq x/k]w(P, P - k, \Xi), \quad (5.1)$$

where $\Xi = (R(X'X)^{-1}R')$ in our notation. This problem can be solved by any method for finding the zeros of a univariate function. The IMSL library has several such codes available for this purpose.

For the level α test in the mixed equality- and inequality-constraints case the critical value is the solution in x of

$$\alpha = \sum_{k=0}^{L} \Pr[F_{P-L+k,T-k} \geq x/(P - L + k)] \times w(L, L - k, \Pi), \quad (5.2)$$

where Π is as defined in Lemma 4.4. This problem also involves finding zeros of a univariate function.

There is another methodology that can be used in cases in which the IMSL library or other such subroutine libraries are unavailable. In this case we calculate the probability of getting a value greater than or equal to the likelihood ratio statistic from a random variable with the null distribution of our test statistic. If $G(x)$ is the distribution of our test statistic under the null hypothesis and LR is our test statistic, we calculate $1 - G(\mathrm{LR})$ as follows for the inequality-constrained case:

$$1 - G(\mathrm{LR}) = \sum_{k=1}^{P} \Pr[F_{k,T-K} \geq \mathrm{LR}/k] w(P, P - k, \Xi).$$

For the mixed equality- and inequality-constrained case $1 - G(\mathrm{LR})$ is

$$1 - G(\mathrm{LR}) = \sum_{k=0}^{L} \Pr[F_{P-L+k,T-K} \geq \mathrm{LR}/(P - L + k)] \times w(L, L - k, \Pi).$$

In calculating these probabilities for the F distributions we can either use numerical integration codes or interpolate the relevant probabilities from available tables of the F distribution. There are also various series expansions methodologies for calculating these probabilities (see Lackritz 1984). In this instance an investigator rejects the null hypothesis if $1 - G(\mathrm{LR}) < \alpha$, where α is the size of the hypothesis test.

6. CONNECTIONS TO GHM AND MULTIVARIATE ONE-SIDED TESTS

In this section we extend the framework of Gourieroux et al. (1982) to consider hypotheses tests combining one-sided and two-sided hypotheses. We also detail the conditions on the covariance matrix of errors such that there is a small sample, exact distribution for the various forms of their test statistic.

As mentioned earlier, Gourieroux et al. were concerned with the testing problem

$$H\colon R\beta = r \qquad \text{versus} \qquad K\colon R\beta \geq r,$$

with the inequality strict in at least one element, for the linear regression model (2.1) with ε in our notation $N(0, \Sigma)$. Gourieroux et al. assumed that Σ is positive definite and unknown. They derived three asymptotically equivalent tests for their problem. If we assume that Σ takes the form $\sigma^2\Lambda$ where Λ is known and positive definite but σ^2 is unknown, we can derive the small sample null distribution for their test statistics. Replace σ^2 in their statistics with $s^2 = (y - X\hat{b})'\Lambda^{-1}(y - X\hat{b})/(T - K)$, which is derived from the unconstrained generalized least squares regression. Different from the results derived earlier for testing multivariate inequality constraints, the modified statistics presented in this section are not in the strict sense LR statistics. As shown in Hillier (1986), for these statistics to be considered LR statistics the estimate of σ^2 used in their computation would have to be based on the estimate of β derived under the alternative hypothesis as opposed to the unrestricted estimate of β. It is unclear, however, whether or not the use of s^2 as an estimate of σ^2 in computing the test statistic will result in a test with power properties inferior to those of the likelihood ratio test. Proceeding under this caveat, the LR form of the test statistic for their problem in this case is the maximum value of the objective function from the following QP:

$$\max_b \; -(y - Xb)'\Lambda^{-1}(y - Xb)/s^2 + (y - X\bar{b})'\Lambda^{-1}(y - X\bar{b})/s^2 \quad \text{subject to } Rb \geq r, \tag{6.1}$$

where $\bar{b}$ is the equality-constrained estimate of β. The solution to this QP is $\tilde{b}$, the ICLS estimator. The dual of (6.1), the KT form of the GHM statistic, is in our notation:

$$\min_\lambda (\lambda - \bar{\lambda})'R(X'\Lambda^{-1}X)^{-1}R'(\lambda - \bar{\lambda})/4s^2 \quad \text{subject to } \lambda \leq 0. \tag{6.2}$$

The LR form of their statistic can be rewritten as the optimal value of the objective function from the following QP:

$$\max_b \; -(b - \hat{b})'(X'\Lambda^{-1}X)(b - \hat{b})/s^2 + (\bar{b} - \hat{b})'(X'\Lambda^{-1}X)(\bar{b} - \hat{b})/s^2 \quad \text{subject to } Rb \geq r.$$

The solution to this QP is $\tilde{b}$. The Wald forms of the statistic for this case of the GHM hypothesis test replaces $\hat{b}$ by $\tilde{b}$ and $\tilde{b}$ by $\bar{b}$ in Equations (3.6) and (3.7).

By a straightforward application of Corollary 4.1, the modified GHM statistic, which we denote by LR, has the following exact null distribution:

$$\Pr(\mathrm{LR} \geq c) = \sum_{k=1}^{P} \Pr(F_{k,T-K} \geq c/k) \times w(P, k, R(X'\Lambda^{-1}X)^{-1}R'),$$

$$\Pr(\mathrm{LR} = 0) = w(P, 0, R(X'\Lambda^{-1}X)^{-1}R').$$

We can also extend the GHM framework to consider two-sided hypotheses in conjunction with their one-sided hypothesis for the same covariance matrix structure for the errors, as was assumed previously. We are now interested in the following hypothesis-testing problem:

$$H\colon R\beta = r \qquad \text{versus} \qquad K\colon R_1\beta \geq r_1 \quad \text{and} \quad R_2\beta \neq r_2,$$

where R_1 and R_2 are as defined in QP (2.6) and r is partitioned in the same manner. The LR form of the test statistic for this hypothesis test is the optimal objective function value from the following QP:

$$\max_b \; -(y - Xb)'\Lambda^{-1}(y - Xb)/s^2 + (y - X\bar{b})'\Lambda^{-1}(y - X\bar{b})/s^2 \quad \text{subject to } R_1 b \geq r_1. \tag{6.3}$$

The KT form of this statistic is the optimal value of the objective function of the dual optimization problem,

$$\min_\lambda (\lambda - \bar{\lambda})'R(X'\Lambda^{-1}X)^{-1}R'(\lambda - \bar{\lambda})/4s^2 \quad \text{subject to } \lambda_1 \leq 0 \quad \text{and} \quad \lambda_2 = 0, \tag{6.4}$$

where λ_1 corresponds to the one-sided equality constraints and λ_2 corresponds to the two-sided equality constraints. The Wald form of the statistic for this mixed hypothesis test replaces $\hat{b}$ by the optimal value of b from QP (6.3) and $\tilde{b}$ by $\bar{b}$ in Equations (3.6) and (3.7). The four forms of the combination one-sided and two-sided statistic are all equivalent by the same logic as given previously. By LR* we denote the optimal objective function value of (6.3), which by duality theory is equal to the optimal objective function value of (6.4). The null distribution of LR* is

$$\Pr(\mathrm{LR}^* \geq c) = \sum_{k=0}^{L} \Pr[F_{P-L+k,T-K} \geq c/(P - L + k)]w(L, k, \Gamma),$$

where $\Gamma = R_1(X'\Lambda^{-1}X)^{-1}R_1'$.

To calculate the critical values or probability values for either the GHM statistic for small samples or the extended GHM statistic with mixtures of one-sided and two-sided tests we use the same procedure as described in Section 5.

7. CONCLUSIONS AND EXTENSIONS

In this article we devised an exact, small sample methodology for testing general linear inequality restrictions within the context of the linear regression model. We extended these results to consider equality and inequality restrictions. In addition we illustrated the relationship between our testing framework and the multivariate one-sided hypothesis-testing literature. In the process we derived the conditions under which the existing multivariate one-sided hypothesis-testing framework for the linear regression model derived by Gourieroux et al. has exact distribution results, as opposed to asymptotic results. We also extended this framework to consider both one-sided and two-sided hypotheses jointly.

In Wolak (1986) a framework for testing multivariate nonlinear inequality constraints in nonlinear models was derived. This framework is also extended to enable testing combinations of nonlinear inequality constraints. Asymptotic results for local [as defined in Wolak (1986)] inequality-constraints tests similar to the exact distribution results of this article obtain, although several complications arise in the derivation of the results because of the nonlinearity of the parameters in the constraints and model.

APPENDIX: EXPRESSIONS FOR WEIGHTS

In this Appendix we give closed-form expressions for the weights, $w(P, k, \Sigma)$, used in the computation of the null distributions of our test statistics for dimensions of the multivariate inequality constraints test ranging from 2 to 4. Wolak (1987) gave an illustrative application of this testing technique to show how it is actually implemented.

Using expressions derived by Kudo (1963) for $P = 2$ and 3 and Shapiro (1985) for $P = 4$, the weights, $w(P, k, \Sigma)$, are given here. For $P = 2$ we have

$$w(2, 0, \Sigma) = \tfrac{1}{2}\pi^{-1} \arccos(\rho_{12}), \qquad w(2, 1, \Sigma) = \tfrac{1}{2},$$
$$w(2, 2, \Sigma) = \tfrac{1}{2} - \tfrac{1}{2}\pi^{-1} \arccos(\rho_{12}),$$

where ρ_{12} is the correlation coefficient associated with the (2×2) covariance matrix Σ. For $P = 3$ we have

$$w(3, 0, \Sigma) = \tfrac{1}{2} - w(3, 2, \Sigma), \qquad w(3, 1, \Sigma) = \tfrac{1}{2} - w(3, 3, \Sigma),$$
$$w(3, 2, \Sigma) = \tfrac{1}{4}\pi^{-1}(3\pi - \arccos(\rho_{12.3}) - \arccos(\rho_{13.2}) - \arccos(\rho_{23.1})),$$
$$w(3, 3, \Sigma) = \tfrac{1}{4}\pi^{-1}(2\pi - \arccos(\rho_{12}) - \arccos(\rho_{13}) - \arccos(\rho_{23})),$$

where ρ_{ij} is the ijth element of the correlation matrix associated with the (3×3) covariance matrix Σ. If $X \in R^P$ is $N(\mu, \Sigma)$, then $\rho_{ij.k}$ is the partial correlation between X_i and X_j holding X_k fixed. Finally for $P = 4$ we have

$$w(4, 0, \Sigma) = \tfrac{1}{2} - w(4, 4, \Sigma) - w(4, 2, \Sigma),$$
$$w(4, 1, \Sigma) = \tfrac{1}{8}\pi^{-1}\left(-4\pi + \sum_{i>j:i,j\neq k} \arccos(\rho_{ij.k})\right),$$
$$w(4, 2, \Sigma) = \tfrac{1}{4}\pi^{-2} \sum_{i>j,k>l:k,l\neq i,j} (\arccos(\rho_{ij}))(\pi - \arccos(\rho_{kl.ij})),$$
$$w(4, 3, \Sigma) = \tfrac{1}{8}\pi^{-1}\left(8\pi - \sum_{i>j:i,j\neq k} \arccos(\rho_{ij.k})\right).$$

The weight $w(4, 4, \Sigma)$ is the probability that $X \in R^4$, as defined previously, has all positive elements. This probability can be obtained by numerically integrating a multivariate normal distribution function. The notation $\rho_{kl.ij}$ is the partial correlation coefficient between X_k and X_l holding X_i and X_j fixed. The remainder of the notation is as defined previously for the case in which $P = 3$. Anderson (1984, pp. 35–43) provided a detailed discussion of the computation of the partial correlation coefficients for an arbitrary covariance matrix Σ.

We should note that for the case in which $\Sigma = \sigma^2 I$ the weights exist in closed form for all P. They take the following form:

$$w(P, k, \sigma^2 I) = \frac{1}{2^P}\begin{bmatrix} P \\ k \end{bmatrix} = \frac{1}{2^P}\begin{bmatrix} P \\ P - k \end{bmatrix} = w(P, P - k, \sigma^2 I).$$

[*Received April 1985. Revised January 1987.*]

REFERENCES

Anderson, T. W. (1984), *An Introduction to Multivariate Statistical Analysis* (2nd ed.), New York: John Wiley.

Avriel, M. (1976), *Nonlinear Programming: Analysis and Methods,* Englewood Cliffs, NJ: Prentice-Hall.

Barlow, R. E., Bartholomew, D. J., Bremner, J. N., and Brunk, H. D. (1972), *Statistical Inference Under Order Restrictions,* New York: John Wiley.

Bartholomew, D. J. (1959a), "A Test of Homogeneity for Ordered Alternatives, I," *Biometrika,* 46, 36–48.

——— (1959b), "A Test of Homogeneity for Ordered Alternatives, II," *Biometrika,* 46, 328–335.

——— (1961), "A Test of Homogeneity of Means Under Restricted Alternatives," *Journal of the Royal Statistical Society,* Ser. B, 23, 239–281.

Bohrer, R., and Chow, W. (1978), "Weights for One-Sided Multivariate Inference," *Applied Statistics,* 27, 100–104.

Dykstra, R. L., and Robertson, T. (1983), "On Testing Monotone Tendencies," *Journal of the American Statistical Association,* 78, 342–350.

Gill, P. E., Murray, W., and Wright, M. H. (1981), *Practical Optimization,* New York: Academic Press.

Goldberger, A. S. (1986), "One-Sided and Inequality Tests for a Pair of Means," Workshop Series Paper 8629, Social Science Research Institute, University of Wisconsin-Madison.

Gourieroux, C., Holly, A., and Monfort, A. (1982), "Likelihood Ratio, Wald Test, and Kuhn–Tucker Test in Linear Models With Inequality Constraints on the Regression Parameters," *Econometrica,* 50, 63–80.

Hillier, G. (1986), "Joint Tests of Zero Restrictions on Nonnegative Regression Coefficients," *Biometrika,* 73, 657–669.

International Mathematical and Statistical Libraries, Inc. (1982), *IMSL Library 1* (9th ed.), Houston, TX: Author.
Kudo, A. (1963), "A Multivariate Analogue of the One-Sided Test," *Biometrika,* 50, 403–418.
Lackritz, J. R. (1984), "Exact p Values for F and t Tests," *The American Statistician,* 38, 312–314.
Lehmann, E. L. (1959), *Testing Statistical Hypothesis,* New York: John Wiley.
Liew, C. K. (1976), "Inequality Constrained Least-Squares Estimation," *Journal of the American Statistical Association,* 71, 746–751.
Luenberger, D. G. (1969), *Optimization by Vector Space Methods,* New York: John Wiley.
Nuesch, P. E. (1966), "On the Problem of Testing Location in Multivariate Problems for Restricted Alternatives," *Annals of Mathematical Statistics,* 37, 113–119.
Perlman, M. D. (1969), "One-Sided Problems in Multivariate Analysis," *Annals of Mathematical Statistics,* 40, 549–567.
Robertson, T., and Wegmen, E. J. (1978), "Likelihood Ratio Tests for Order Restrictions in Exponential Families," *Annals of Statistics,* 6, 485–505.
Robertson, T., and Wright, F. T. (1983), "On Approximation of the Level Probabilities and Associated Distributions in Order Restricted Inference," *Biometrika,* 70, 597–606.
Shapiro, A. (1985), "Asymptotic Distribution of Test Statistics in the Analysis of Moment Structures Under Inequality Constraints," *Biometrika,* 72, 133–144.
Silvey, S. D. (1970), *Statistical Inference,* London: Chapman & Hall.
Siskind, V. (1976), "Approximate Probability Integrals and Critical Values for Bartholomew's Test for Ordered Means," *Biometrika,* 63, 647–654.
Theil, H. (1971), *Principles of Econometrics,* New York: John Wiley.
Wald, A. (1943), "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations Is Large," *Transactions of the American Mathematical Society,* 54, 426–482.
Wilks, S. S. (1962), *Mathematical Statistics,* New York: John Wiley.
Wolak, F. A. (1986), "Testing Nonlinear Inequality Constraints in the Maximum Likelihood Model," Technical Report 21, Stanford University, Econometric Workshop, Dept. of Economics.
——— (1987), "Testing Inequality Constraints in Linear Econometric Models," Technical Report 25, Stanford University, Econometric Workshop, Dept. of Economics.
Yancey, T. A., Bohrer, R., and Judge, G. G. (1982), "Power Function Comparisons in Inequality Hypothesis Testing," *Economics Letters,* 9, 161–167.
Yancey, T. A., Judge, G. G., and Bock, M. E. (1981), "Testing Multiple Equality and Inequality Hypothesis in Economics," *Economics Letters,* 7, 249–255.