E C O N O M I C S   B U L L E T I N

# Fully restricted linear regression: A pedagogical note

Harry Haupt
*University of Regensburg*

Walter Oberhofer
*University of Regensburg*

## *Abstract*

This paper presents a comprehensive approach to estimation and hypothesis testing under a set of full restrictions, some of these arising from adding−up conditions on the endogenous variable. In contrast to the existing statistical literature, this paper uses an argumentation style familiar from classical econometric textbooks, to provide an insightful, straightforward, and nevertheless rigorous exposition of this topic.

We are very indebted to Adrian Pagan and an anonymous referee for many unusually helpful comments and suggestions.
**Citation:** Haupt, Harry and Walter Oberhofer, (2002) "Fully restricted linear regression: A pedagogical note." *Economics Bulletin*, Vol. 3, No. 1 pp. 1−7
**Submitted:** August 20, 2001. **Accepted:** January 15, 2002.
**URL:** http://www.economicsbulletin.com/2002/volume3/EB−01C10005A.pdf

## 1. Introduction

This note takes up the problem of estimation in the classical linear model subject to linear restrictions when some of these restrictions arise from adding-up conditions, which imply the singularity of a covariance matrix. Singularity in a covariance matrix might be thought of as arising in one of two ways. First, when a particular variable is allocated among a number of equations. Dhrymes and Schwarz (1987) give a brief survey of the economic origins of this mechanism and of the corresponding literature; for an extensive survey and applications see also the monograph of Bewley (1986). Typically, allocation models contain several behavioral equations and an additional identity. Imposing this identity or adding-up condition on the regression equations causes *implicit restrictions* on the parameters, resulting in what might be called a *model derived singularity*. Due to this singularity, the usual generalized least squares (GLS) estimator does not exist. Thus, a modified GLS procedure has been derived, using the Moore-Penrose inverse of the covariance matrix.

A second approach to singularity is to recognize that it can occur for unknown reasons and needs to be accounted for with appropriate statistical methods. Hence it is common practice to construct the class of estimable functions and estimators with certain desirable properties for different constellations of the linear regression model. Thus this case, where there is a singular disturbance covariance matrix implying restrictions on the parameters, may well be described as *singularity by assumption*. The properties of the corresponding estimators were first systematically investigated by Rao (1965, 1973), Rao and Mitra (1973), Theil (1971), and Kreijger and Neudecker (1977). The last two are often grouped together in the literature and termed the Theil-Kreijger-Neudecker (TKN) estimator. Magnus and Neudecker (1988) summarized and completed the results on estimation in this model class.

The aim of this paper is to simplify and to synthesize existing theory by bringing together some well known findings. After introducing the set-up of the model (section 2), the paper continues to impose the complete set of implicit and *explicit* (i.e. exogenous) restrictions on the regression equations implied by the singularity. This is done by using the well known transformation of a model with nonspherical disturbances — subject to implicit and explicit constraints — to an unconstrained regression model with spherical disturbances (section 3). Then, for the transformed model, the usual results from the basic textbook literature apply, and this is done in a straightforward way in section four.

## 2. Problem Setting

Consider the classical linear model given by

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{u}, \tag{1}$$

where $\mathbf{y}$ is the $T \times 1$ vector of observations on the endogenous variable, $\mathbf{b}$ is the $K \times 1$ vector of unknown parameters, $\mathbf{X}$ is the $T \times K$ regressor matrix and $\mathbf{u}$ is the $T \times 1$ vector of disturbances. We assume that all regressors are nonstochastic, $\mathrm{rk}(\mathbf{X}) = m \leq K$ and $\mathrm{E}(\mathbf{u}) = \mathbf{0}$. The covariance matrix of $\mathbf{u}$ is given by $\mathcal{V}(\mathbf{u}) = \sigma^2\boldsymbol{\Omega}$, where we assume that $\boldsymbol{\Omega}$ is known and $\sigma^2 \neq 0$.

Two constraints might apply to the regression equation (1).

(i) The matrix $\mathbf{\Omega}$ may be positive semidefinite and singular, due to the imposition of adding-up conditions on the endogenous variable.

(ii) The parameters $\mathbf{b}$ might be subject to linear restrictions given by

$$\mathbf{Hb} = \mathbf{h}, \tag{2}$$

with some of these restrictions arising from the adding-up conditions.

Sections three and four of the paper will incorporate both of these constaints using an appropriate set-up in line with Theil (1971) which will now be described. Initially however, we focus on the issues responsible for problem (ii), which will naturally lead us to problem (i).

The parameters in (1) are taken to obey the exogenous restrictions

$$\mathbf{Rb} = \mathbf{r}, \tag{3}$$

where both the $q \times K$ matrix $\mathbf{R}$ and the $q \times 1$ vector $\mathbf{r}$ are known and $\text{rk}(\mathbf{R}) = \text{rk}\begin{pmatrix} \mathbf{R} & \mathbf{r} \end{pmatrix}$.

In addition to the *explicit* restrictions on $\mathbf{b}$ in (3) we consider *implicit* restrictions, represented by adding-up conditions on the endogenous variables $\mathbf{A}'\mathbf{y} = \mathbf{s}$, where both the $T \times p$ matrix $\mathbf{A}$ and the $p \times 1$ vector $\mathbf{s}$ are fixed and known. This implies

$$\text{var}(\mathbf{A}'\mathbf{y}) = \text{var}(\mathbf{A}'\mathbf{u}) = \mathbf{0}, \tag{4}$$

which leads to the parameter restrictions

$$\mathbf{s} = \mathbf{A}'\mathbf{y} = \mathbf{A}'\mathbf{Xb} \qquad a.s. \tag{5}$$

It immediately follows from equation (4) that there is a singularity in the covariance matrix $\mathbf{\Omega}$ since $\mathbf{\Omega A} = \mathbf{0}$, where we assume that the $p < T$ linearly independent columns of $\mathbf{A}$ constitute a basis of the null space of $\mathbf{\Omega}$. Without loss of generality we also assume that $\mathbf{A}'\mathbf{A} = \mathbf{I}_p$. Consequently $\mathbf{\Omega}$ can be diagonalized as

$$\mathbf{\Omega} = \begin{pmatrix} \mathbf{A} & \mathbf{F} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda} \end{pmatrix} \begin{pmatrix} \mathbf{A}' \\ \mathbf{F}' \end{pmatrix} = \mathbf{F\Lambda F}', \tag{6}$$

where $\begin{pmatrix} \mathbf{A} & \mathbf{F} \end{pmatrix}$ is an orthogonal $T \times T$ matrix and $\mathbf{\Lambda}$ is the diagonal matrix of the positive eigenvalues of $\mathbf{\Omega}$ of the order $(T - p)$. The Moore-Penrose inverse representation of $\mathbf{\Omega}$ with $\mathbf{\Omega}^+ = \mathbf{F\Lambda}^{-1}\mathbf{F}'$ follows from (6).

Because there is a potential for an identification problem, we have to assume that the matrix

$$\begin{pmatrix} \mathbf{R} \\ \mathbf{X} \end{pmatrix}$$

has full column rank. If this is not true then identification is problematic. This is easily seen by assuming that the above matrix is rank deficient, implying the existence of a nonzero vector $\mathbf{g}$, with $\mathbf{Xg} = \mathbf{0}'$and $\mathbf{Rg} = \mathbf{0}$. Hence $\mathbf{X}(\mathbf{b} + \mathbf{g}) = \mathbf{Xb}$, $\mathbf{R}(\mathbf{b} + \mathbf{g}) = \mathbf{r}$ and $\mathbf{A}'\mathbf{X}(\mathbf{b} + \mathbf{g}) = \mathbf{A}'\mathbf{Xb}$. The latter three equations imply that $\mathbf{b}$ and $\mathbf{b} + \mathbf{g}$ are observationally equivalent, and consequently $\mathbf{b}$ cannot be identified.

Finally, by setting

$$\mathbf{H} \equiv \begin{pmatrix} \mathbf{R} \\ \mathbf{A}'\mathbf{X} \end{pmatrix} \qquad \text{and} \qquad \mathbf{h} \equiv \begin{pmatrix} \mathbf{r} \\ \mathbf{s} \end{pmatrix},$$

the restrictions (3) and (5) combine to the set of full restrictions (2).

In order to guarantee the consistency of (2) we further assume that $\text{rk}(\mathbf{H}) = \text{rk}\begin{pmatrix} \mathbf{H} & \mathbf{h} \end{pmatrix}$.

## 3. GLS Estimation

It is well known from the literature that the GLS estimator corresponding to the regression problem described in section 2 is not defined and that there are several ways of circumventing this problem. One method to deal with problem (i) is to premultiply both sides of (1) by the nonsingular matrix

$$\begin{pmatrix} \mathbf{A}' \\ \mathbf{\Lambda}^{-1/2}\mathbf{F}' \end{pmatrix}.$$

This results in an equation system which is statistically equivalent to (1) and is represented by constraints (5) and the regression equation

$$\mathbf{\Lambda}^{-1/2}\mathbf{F}'\mathbf{y} = \mathbf{\Lambda}^{-1/2}\mathbf{F}'\mathbf{X}\mathbf{b} + \mathbf{\Lambda}^{-1/2}\mathbf{F}'\mathbf{u},$$

or

$$\mathbf{y}^* = \mathbf{X}^*\mathbf{b} + \mathbf{u}^*, \tag{7}$$

where the covariance matrix of $\mathbf{u}^*$ is, due to (6), given by $\sigma^2\mathbf{I}_{T-p}$.

Having neutralized problem (i) we are left with the transformed equation (7) and the set of explicit and implicit restrictions (2). The unconstrained GLS-type system of normal equations corresponding to (7) could then be written as $\mathbf{X}'\mathbf{\Omega}^+\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{\Omega}^+\mathbf{y}$. Imposing the full set of restrictions (2), the restricted GLS estimation of (7) leads to the normal equations

$$\begin{pmatrix} \mathbf{X}'\mathbf{\Omega}^+\mathbf{X} & \mathbf{H}' \\ \mathbf{H} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{b}} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{\Omega}^+\mathbf{y} \\ \mathbf{h} \end{pmatrix}, \tag{8}$$

where $\boldsymbol{\lambda}$ contains the Lagrange multipliers.

Now the problem is to find an intuitive, simple, and straightforward approach to the derivation of $\hat{\mathbf{b}}$ and to establish the properties of this estimator.

To start with we need a basis for the null space of $\mathbf{H}$. The vectors of this basis will be combined as columns of the $K \times m$ matrix $\mathbf{N}$, where $m = K - \text{rk}(\mathbf{H})$. Without loss of generality we assume $\mathbf{N}'\mathbf{N} = \mathbf{I}_m$. Thus the matrix

$$\begin{pmatrix} \mathbf{H} \\ \mathbf{N}' \end{pmatrix}$$

has full column rank and

$$\mathbf{H}\mathbf{N} = \mathbf{0}.$$

The complete set of restrictions (2) can be inverted to find the parameter values satisfying them, namely

$$\mathbf{b} = \mathbf{H}^+\mathbf{h} + \mathbf{N}\mathbf{c}, \tag{9}$$

where $\mathbf{H}^+$ is the Moore-Penrose inverse of $\mathbf{H}$ and $\mathbf{c}$ is an $m \times 1$ vector containing the free parameters.

Note that any pseudo-inverse of $\mathbf{H}$ could be used in (9). This parameter reduction is due to Rao (1965, 4a.9) and represents an important result in the regularization of singularity problems. Replacing $\mathbf{b}$ in (7) by the right-hand side of (9) yields

$$\mathbf{y}^{**} = \mathbf{X}^{**}\mathbf{c} + \mathbf{u}^{*}, \tag{10}$$

with

$$\mathbf{y}^{**} \equiv \mathbf{y}^{*} - \mathbf{X}^{*}\mathbf{H}^{+}\mathbf{h} \qquad \text{and} \qquad \mathbf{X}^{**} \equiv \mathbf{X}^{*}\mathbf{N}.$$

Obviously, model (10) is a classical linear regression model with $(T-p)$ observations and $m$ parameters that are free of restrictions. The OLS normal equations of (10) are given by

$$\mathbf{N}'\mathbf{X}'\boldsymbol{\Omega}^{+}\mathbf{X}\mathbf{N}\hat{\mathbf{c}} = \mathbf{N}'\mathbf{X}'\boldsymbol{\Omega}^{+}\mathbf{y} - \mathbf{N}'\mathbf{X}'\boldsymbol{\Omega}^{+}\mathbf{X}\mathbf{H}^{+}\mathbf{h}. \tag{11}$$

Now define

$$\mathbf{S} \equiv (\boldsymbol{\Lambda}^{-1/2}\mathbf{F}'\mathbf{X}\mathbf{N})'(\boldsymbol{\Lambda}^{-1/2}\mathbf{F}'\mathbf{X}\mathbf{N}) = \mathbf{N}'\mathbf{X}'\boldsymbol{\Omega}^{+}\mathbf{X}\mathbf{N} = \mathbf{X}'^{**}\mathbf{X}^{**}.$$

Due to the orthogonality of $\mathbf{H}$ and $\mathbf{N}$

$$\begin{pmatrix} \mathbf{H} \\ \boldsymbol{\Lambda}^{-1/2}\mathbf{F}'\mathbf{X} \end{pmatrix} \mathbf{N} = \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\Lambda}^{-1/2}\mathbf{F}'\mathbf{X}\mathbf{N} \end{pmatrix}. \tag{12}$$

Then, obviously

$$\text{rk}\left[\begin{pmatrix} \mathbf{R} \\ \mathbf{A}'\mathbf{X} \\ \boldsymbol{\Lambda}^{-1/2}\mathbf{F}'\mathbf{X} \end{pmatrix} \mathbf{N}\right] = \text{rk}\left[\begin{pmatrix} \mathbf{R} \\ \mathbf{A}'\mathbf{X} \\ \mathbf{F}'\mathbf{X} \end{pmatrix} \mathbf{N}\right] = \text{rk}\left[\begin{pmatrix} \mathbf{R} \\ \mathbf{X} \end{pmatrix} \mathbf{N}\right] = \text{rk}(\mathbf{N}) = m. \tag{13}$$

Consequently, due to (12) and (13), we get $\text{rk}(\boldsymbol{\Lambda}^{-1/2}\mathbf{F}'\mathbf{X}\mathbf{N}) = \text{rk}(\mathbf{X}^{**}) = m$ and this proves

LEMMA 1. *The matrix* $\mathbf{S}$ *is invertible.*

Thus, a unique solution to the normal equations (11) is given by

$$\hat{\mathbf{c}} = \mathbf{S}^{-1}(\mathbf{N}'\mathbf{X}'\boldsymbol{\Omega}^{+}\mathbf{y} - \mathbf{N}'\mathbf{X}'\boldsymbol{\Omega}^{+}\mathbf{X}\mathbf{H}^{+}\mathbf{h}), \tag{14}$$

with $\text{var}(\hat{\mathbf{c}}) = \sigma^2\mathbf{S}^{-1}$. Of course, we are not interested in these parameters, but the ones in $\mathbf{b}$. Hence, an explicit solution for $\hat{\mathbf{b}}$ is given by

$$\hat{\mathbf{b}} = \mathbf{N}\mathbf{S}^{-1}\mathbf{N}'\mathbf{X}'\boldsymbol{\Omega}^{+}\mathbf{y} + (\mathbf{I}_K - \mathbf{N}\mathbf{S}^{-1}\mathbf{N}'\mathbf{X}'\boldsymbol{\Omega}^{+}\mathbf{X})\mathbf{H}^{+}\mathbf{h}, \tag{15}$$

and the parameter covariance matrix has the simple form $\text{var}(\hat{\mathbf{b}}) = \sigma^2\mathbf{N}\mathbf{S}^{-1}\mathbf{N}'$, where $\hat{\mathbf{b}}$ satisfies the complete set of restrictions.

An equivalent expression to (15) would be

$$\hat{\mathbf{b}} = \bar{\mathbf{b}} + \mathbf{B}(\mathbf{y} - \mathbf{X}\bar{\mathbf{b}}),$$

where

$$\bar{\mathbf{b}} = \mathbf{H}^{+}\mathbf{h}$$

is a special solution to the complete set of restrictions, and

$$\mathbf{B} = \mathbf{N}\mathbf{S}^{-1}\mathbf{N}'\mathbf{X}'\boldsymbol{\Omega}^{+}$$

is a weighting matrix. Note that these results can be easily transferred to the estimation of linear combinations $\mathbf{W}\mathbf{b}$.

## 4. BLU Properties of $\hat{\mathbf{b}}$, Unbiased Estimation of $\sigma^2$, and Wald Tests

After the derivation of the restricted GLS estimator we are of course interested in its properties. In an intriguingly simple manner the transformations done in the previous section lead us to the proof of the BLU properties of (15). Due to

$$\begin{pmatrix} \mathbf{c} \\ \mathbf{h} \end{pmatrix} = \begin{pmatrix} \mathbf{N}' \\ \mathbf{H} \end{pmatrix} \mathbf{b} \qquad \text{and} \qquad \mathbf{b} = \mathbf{H}^{+}\mathbf{h} + \mathbf{N}\mathbf{c},$$

there is a one-to-one-correspondence between either the true parameters $\mathbf{c}$ and $\mathbf{b}$, and the corresponding estimators $\hat{\mathbf{c}}$ and $\hat{\mathbf{b}}$. The fact that $\hat{\mathbf{b}}$ inherits the properties of $\hat{\mathbf{c}}$ becomes evident by considering $\hat{\mathbf{b}}$ as a linear function of $\mathbf{y}^{**}$. Thus, since there is also a one-to-one-correspondence between $\mathbf{y}^{**}$ and $\mathbf{y}$, namely

$$\begin{pmatrix} \mathbf{y}^{**} \\ \mathbf{s} \end{pmatrix} = \begin{pmatrix} \mathbf{\Lambda}^{-1/2}\mathbf{F}' \\ \mathbf{A}' \end{pmatrix} \mathbf{y} - \begin{pmatrix} \mathbf{\Lambda}^{-1/2}\mathbf{F}'\mathbf{X}\mathbf{H}^{+}\mathbf{h} \\ \mathbf{0} \end{pmatrix},$$

implying

$$\mathbf{y} = \mathbf{F}\mathbf{\Lambda}^{1/2}\mathbf{y}^{**} + \mathbf{X}\mathbf{H}^{+}\mathbf{h},$$

quite obviously $\hat{\mathbf{b}}$ is BLU as a function of $\mathbf{y}$.

It is important to note, however, that despite its general uniqueness, estimator (15) has no unique representation.[1] This can be easily seen by considering an arbitrary (affine) linear unbiased estimator $\tilde{\mathbf{b}}$ in $\mathbf{y}$, with

$$\tilde{\mathbf{b}} = \mathbf{L}\mathbf{y} + \mathbf{d} = (\mathbf{L} + \mathbf{D}\mathbf{A}')\mathbf{y} + \mathbf{d} - \mathbf{D}\mathbf{s} \qquad a.s.,$$

where $\mathbf{D}$ is an arbitrary $K \times p$ matrix. Now, since $\mathbf{L}$ and $\mathbf{d}$ have to be calculated when proving the minimal variance property of $\tilde{\mathbf{b}}$ directly, there are no unique solutions, as has been pointed out by Rao (1973, Corollary 3.6). Due to $\mathbf{A}'\mathbf{y} = \mathbf{s}$, the unique representation (14) of $\hat{\mathbf{c}}$ in $\mathbf{y}^{**}$ is lost in the transformation to represent $\hat{\mathbf{b}}$ as a function of $\mathbf{y}$.[2]

---

[1] Rao (1965) considers the case of a singular regressor matrix and explicit restrictions, but does not allow for a singular covariance matrix. Rao and Mitra (1971) and Rao (1973) discuss the case of a singular covariance matrix and the resulting restrictions without considering further constraints. This allows linear homogeneity in $\mathbf{y}$ of the resulting estimator $\hat{\mathbf{b}} = \mathbf{L}\mathbf{y}$, whereas in the case of further constraints, as pointed out by Magnus and Neudecker (1988, p.255), $\hat{\mathbf{b}}$ must be linear nonhomogeneous (affine) in $\mathbf{y}$, i.e. $\hat{\mathbf{b}} = \mathbf{L}\mathbf{y} + \mathbf{d}$.

What the work of the cited authors has in common, is that, from an economic point of view, their approach is rather unnatural. That means, instead of discussing adding-up conditions, they assume the singularity of the covariance matrix and their aim is to *construct* BLU estimators instead of deriving GLS estimators. Moreover, their use of conditions on linear manifolds is quite useless in applied work.

Rao (1973, p.278, Lemma 2.1) shows that from $\mathrm{E}(\mathbf{y}) = \mathbf{X}\mathbf{b}$ and $\mathcal{V}(\mathbf{y}) = \sigma^2\mathbf{\Omega}$ follows $\mathbf{y} \in \mathcal{M}(\mathbf{X} : \mathbf{\Omega})$. He further states that *this is the only statement that can be made when* $\mathbf{y}$ *is not observed*. If, however, $\mathbf{y}$ is observable, he proves (p.279, Lemma 2.4) that *the knowledge of an observation on* $\mathbf{y}$ *enables us to specify the particular subspace of* $\mathcal{M}(\mathbf{X} : \mathbf{\Omega})$ *to which the random variable belongs.*

Magnus and Neudecker (1988, p.272, Proposition 4) postulate $\mathbf{y} \in \mathcal{M}(\mathbf{X} : \mathbf{\Omega})$, *a.s.*, or, in the case of exogenous restrictions $\mathbf{y} \in \mathcal{M}\begin{pmatrix} \mathbf{X} & \mathbf{\Omega} \\ \mathbf{R} & \mathbf{0} \end{pmatrix}$ *a.s.*, to guarantee model consistency. However, this condition can neither be verified nor is it necessary, because one of the model's preliminaries (p.254, equation 1) is $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{u}$ and, consequently $\mathrm{E}(\mathbf{y}) = \mathbf{X}\mathbf{b}$ and $\mathcal{V}(\mathbf{u}) = \sigma^2\mathbf{\Omega}$. Then, by following Rao, we automatically get $\mathbf{y} \in \mathcal{M}(\mathbf{X} : \mathbf{\Omega})$.

[2] Thus, the statement of Magnus and Neudecker (1988, p.272), that the model $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{u}$ with $\mathbf{A}'\mathbf{y} = \mathbf{s}$ and $\mathcal{V}(\mathbf{u}) = \sigma^2\mathbf{\Omega}$ is equivalent to the model $\mathbf{F}'\mathbf{y} = \mathbf{F}'\mathbf{X}\mathbf{b} + \mathbf{F}'\mathbf{u}$ with $\mathcal{V}(\mathbf{F}'\mathbf{u}) = \sigma^2\mathbf{\Lambda}$, is not true with respect to a unique representation of $\hat{\mathbf{b}}$ as a function of $\mathbf{y}$, and a function of $\mathbf{F}'\mathbf{y}$, respectively.

Obviously a best quadratic estimate of $\sigma^2$ as a function of $\mathbf{y}^{**}$ is given by

$$\hat{\sigma}^2 = \hat{\mathbf{u}}^{*\prime}\hat{\mathbf{u}}^*/(T-p-m), \tag{16}$$

where $\hat{\mathbf{u}}^* = \mathbf{y}^{**} - \mathbf{X}^{**}\hat{\mathbf{c}}$. Consequently, due to the one-to-one correspondence between $\mathbf{y}$ and $\mathbf{y}^{**}$, $\hat{\sigma}^2$ is a best quadratic estimate of $\sigma^2$ as a function of $\mathbf{y}$. Under the assumption of normality $(T-p-m)\hat{\sigma}^2/\sigma^2$ is distributed as $\chi^2$ with $(T-p-m)$ degrees of freedom.

Although the discussion has been couched in terms of a single index model with adding-up restrictions, it has been pointed out by several authors that the results so derived are perfectly applicable in the context of systems of regression equations. Moreover, with (15) to hand, extra linear restrictions on $\mathbf{b}$ such as $\mathbf{Gb} = \mathbf{g}$ can be tested straightforwardly. This has has been recently done by Ravikumar, Ray and Savin (2000) in the SUR context. Without loss of generality we assume that $\mathbf{G}$ has full row rank, say $n$, and that $\mathbf{Gb} = \mathbf{g}$ is neither contradictory nor includes any already implemented restrictions. Under the assumption of normality the test statistic is

$$(\mathbf{G}\tilde{\mathbf{b}} - \mathbf{g})'(\sigma^2\mathbf{GNS^{-1}N'G'})^{-1}(\mathbf{G}\tilde{\mathbf{b}} - \mathbf{g}), \tag{17}$$

which is distributed as $\chi^2$ with $n$ degrees of freedom under $H_0 : \mathbf{Gb} = \mathbf{g}$. It can be easily shown that the matrix $\mathbf{GNS^{-1}N'G'}$ is nonsingular. Replacing $\sigma^2$ by $\hat{\sigma}^2$, (17) will be distributed as the $F$ distribution with $n$ and $(T-p-m)$ degrees of freedom.

For details concerning the estimation of $\boldsymbol{\Omega}$, when it is unknown, by iteratively applying feasible GLS, as well as further elaborations see Haupt and Oberhofer (2000).

## References


Bewley, R. (1986) *Allocation Models: Specifications, Estimation, and Application.* Ballinger Publishing Company.

Dhrymes, P.J., and Schwarz, S. (1987) "On the Invariance of Estimators for Singular Systems of Equations." *Greek Economic Review* **9**(1): 231-250.

Haupt, H., and Oberhofer W. (2000) "Estimation of Constrained Singular Seemingly Unrelated Regression Models." *Econometric Society World Congress 2000 Contributed Papers, No. 398; (http://fmwww.bc.edu/RePEc/es2000/0398.pdf).*

Kreijger, R.G., and Neudecker, H. (1977) "Exact Linear Restrictions on Parameters in the General Linear Model with a Singular Covariance Matrix." *Journal of the American Statistical Association* **72**: 430-432.

Magnus, J.R., and Neudecker, H. (1988) *Matrix Differential Calculus with Applications in Statistics and Econometrics.* J. Wiley & Sons.

Rao, C.R. (1965) *Linear Statistical Inference and Its Applications.* J. Wiley & Sons.

Rao, C.R. (1973) "Representations of Best Linear Unbiased Estimators in the Gauss-Markoff Model with a Singular Dispersion Matrix". *Journal of Multivariate Analysis* **3**: 276-292.

Rao, C.R., and Mitra, S.K. (1971) *Generalized Inverse of Matrices and its Applications.* J.Wiley & Sons.

Ravikumar, B., Ray, S., and Savin, N.E. (2000) "Robust Wald Tests in SUR Systems with Adding-Up Restrictions." *Econometrica* **68**(3): 715-719.

Theil, H. (1971) *Principles of Econometrics.* J. Wiley & Sons.