ON CONSTRAINED LEAST-SQUARES ESTIMATION WITH DISTORTED DATA

Autor(es): Frank, Luis.

Fuente: Cuadernos del CIMBAGE, Nº 20 (Diciembre 2018), pp 27-38

Publicado por: Facultad de Ciencias Económicas de la Universidad de Buenos Aires.

Vínculo: http://ojs.econ.uba.ar/ojs/index.php/CIMBAGE/issue/view/179

Cuadernos del CIMBAGE N°20, 2°Edición (2018) 27-38

ON CONSTRAINED LEAST-SQUARES ESTIMATION WITH DISTORTED DATA

Luis Frank

Instituto Nacional de Estadísticas y Censos. Dirección de Cuentas Nacionales.
Av. Pte. Julio A. Roca 609- CABA- C1067ABB- Argentina

lfrank@indec.gob.ar

Received June 26th 2018, accepted August 21st 2018

Abstract

The paper presents a constrained least squares estimator for “distorted” data and analyses its asymptotic properties. The proposed estimator is unbiased (although inefficient) and asymptotically normal, provided that the true data were distorted by replacing them by a code that meets the Grenander conditions. Nevertheless, the properties of the estimator in the context of small samples remain unknown.

Keywords: constrained least squares, missing data, regression, linear models.

JEL Codes: C130, C180, C360.

ACERCA DE LA ESTIMACIÓN MÍNIMO CUADRÁTICA RESTRINGIDA CON DATOS FALSEADOS

Luis Frank

Instituto Nacional de Estadísticas y Censos. Dirección de Cuentas Nacionales.
Av. Pte. Julio A. Roca 609- CABA- C1067ABB- Argentina

lfrank@indec.gob.ar

Recibido el 26 de junio de 2018, aceptado el 21 de agosto de 2018

Resumen

El artículo presenta un estimador de mínimo cuadrático restringido para datos "falseados" y analiza sus propiedades asintóticas. El estimador propuesto es insesgado (aunque ineficiente) y asintóticamente normal, siempre que los datos verdaderos hayan sido falseados reemplazándolos por un código que satisfaga las condiciones de Grenander. Sin embargo, las propiedades del estimador en el contexto de muestras pequeñas permanecen aún desconocidas.

Palabras clave: mínimos cuadrados restringidos, datos faltantes, regresión, modelos lineales.

Códigos JEL: C130, C180, C360.

1. INTRODUCTION

Consider the "true" model

$$\mathbf{y} = \mathbf{X}_0\boldsymbol{\beta}_0 + \boldsymbol{\varepsilon}_0,\ \boldsymbol{\varepsilon}_0 \sim (\mathbf{0},\ \sigma_0^2\boldsymbol{\Omega}_0),$$

where $\mathbf{y}$ is a $n \times 1$ vector of observations, $\mathbf{X}_0$ is a matrix of constants of size $n \times k$ and full column rank, $\boldsymbol{\beta}_0$ is a vector of $k \times 1$ unknown fixed parameters, $\boldsymbol{\varepsilon}_0$ is an unobservable vector of $n \times 1$ independent but not identically distributed random variables, and $\sigma_0^2\boldsymbol{\Omega}_0$ is a known symmetric positive definite matrix. The parameter vector $\boldsymbol{\beta}_0$ is subject to the system of linear constraints $\mathbf{R}\boldsymbol{\beta}_0 = \mathbf{r}$ where $\mathbf{R}$ and $\mathbf{r}$ are, respectively, a full row rank matrix of dimension $q \times k$ and a vector of dimension $q \times 1$. If $\mathbf{X}_0$ is a sample matrix, then it satisfies the Grenander conditions (Greene, 2006, p. 65).

Now suppose that for some reason not all elements of $\mathbf{X}_0$ have been properly recorded. Instead, a matrix $\mathbf{X} = \mathbf{X}_0 + \mathbf{U}$ is provided, with each element $u_{ij} = (1 - \delta_{x=x0})(c_{ij} - x_{0ij})$, where $\delta_{x=x0}$ is a Kronecker delta that equals 1 if $x_{ij} = x_{0ij}$ or 0 otherwise. The constant c_{ij} is a substitute code for the true x_{0ij}. That is, c_{ij} is simply a code indicating that the true value is unknown, either because it was not recorded or because its record is faulty or "distorted". This means that both $\mathbf{X}_0$ and $\mathbf{U}$ are not fully observable, unlike $\mathbf{X}$ which is a matrix of known constants. For reasons that will become apparent below, we assume that $\mathbf{X}$ is a full column rank matrix and that the distortion of data occurs completely at random. We estimate $\boldsymbol{\beta}_0$ under this specification.

2. UNCONSTRAINED SOLUTIONS

2.1 Deductions

Let

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},\ \text{where } \boldsymbol{\varepsilon} \sim (\boldsymbol{\mu},\ \sigma_\varepsilon^2\boldsymbol{\Omega})$$

be the distorted model. As in the true model $\sigma_\varepsilon^2\boldsymbol{\Omega}$ is a known matrix as well as $cov(\boldsymbol{\varepsilon},\boldsymbol{\varepsilon}_0) = \sigma_0^2\sigma^2\boldsymbol{\Psi}$. If we minimize (Bera, 1994) the sum of squares $L = (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'\boldsymbol{\Omega}^{-1}(\mathbf{y}-\mathbf{X}_0\boldsymbol{\beta}_0)$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_0$, we get the system of normal equations

$$\partial L/\partial \mathbf{b}_0 = -\mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{y} + \mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{X}\mathbf{b} = \mathbf{0}$$

$$\partial L/\partial \mathbf{b} = -\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{y} + \mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0\mathbf{b}_0 = \mathbf{0}$$

where the solutions for $\mathbf{b}$ and $\mathbf{b}_0$ are

$$\mathbf{b}_{GLS} = (\mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{y}$$
$$\mathbf{b}_{0,GLS} = (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{y} \quad (1)$$

The existence of $(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}$ is guaranteed because $\mathbf{X}$ and $\mathbf{X}_0$ are full column rank matrices. Simple inspection of the solution $\mathbf{b}_0$ supports the finding that (1) is an unbiased, though not efficient, estimator of $\boldsymbol{\beta}_0$:

$$E(\mathbf{b}_0) = \boldsymbol{\beta}_0 + (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}E(\boldsymbol{\varepsilon}_0\,|\,\mathbf{X},\mathbf{X}_0)$$
$$= \boldsymbol{\beta}_0$$
$$E[(\mathbf{b}_0 - \boldsymbol{\beta}_0)(\mathbf{b}_0 - \boldsymbol{\beta}_0)'] = (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}E(\boldsymbol{\varepsilon}_0\boldsymbol{\varepsilon}_0'\,|\,\mathbf{X},\mathbf{X}_0)\boldsymbol{\Omega}^{-1}\mathbf{X}(\mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}$$
$$= \sigma_0^2(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Omega}_0\boldsymbol{\Omega}^{-1}\mathbf{X}(\mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}$$
$$E[(\mathbf{b} - \boldsymbol{\beta})(\mathbf{b}_0 - \boldsymbol{\beta}_0)'] = (\mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}_0'\boldsymbol{\Omega}^{-1}E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}_0'\,|\,\mathbf{X},\mathbf{X}_0)\boldsymbol{\Omega}^{-1}\mathbf{X}(\mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}$$
$$= \sigma_0^2\sigma^2\,(\mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}_0'\boldsymbol{\Omega}^{-1}\boldsymbol{\Psi}\boldsymbol{\Omega}^{-1}\mathbf{X}(\mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}. \quad (2)$$

The solutions in (1) are called Generalized Least Squares (GLS). Alternatively, we get a Least Squares (LS) solution by solving the normal equations which emerge after minimizing $L = (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}_0\boldsymbol{\beta}_0)$,

$$\mathbf{b}_{LS} = (\mathbf{X}_0'\mathbf{X})^{-1}\mathbf{X}_0'\mathbf{y}$$
$$\mathbf{b}_{0,LS} = (\mathbf{X}'\mathbf{X}_0)^{-1}\mathbf{X}'\mathbf{y}. \quad (3)$$

In this case $\mathbf{b}_0$ also appears unbiased but inefficient:

$$E(\mathbf{b}_0) = \boldsymbol{\beta}_0 + (\mathbf{X}'\mathbf{X}_0)^{-1}\mathbf{X}'E(\boldsymbol{\varepsilon}_0\,|\,\mathbf{X},\mathbf{X}_0)$$
$$= \boldsymbol{\beta}_0$$
$$E[(\mathbf{b}_0 - \boldsymbol{\beta}_0)(\mathbf{b}_0 - \boldsymbol{\beta}_0)'] = (\mathbf{X}'\mathbf{X}_0)^{-1}\mathbf{X}'E(\boldsymbol{\varepsilon}_0\boldsymbol{\varepsilon}_0'\,|\,\mathbf{X},\mathbf{X}_0)\mathbf{X}(\mathbf{X}_0'\mathbf{X})^{-1}$$
$$= \sigma_0^2(\mathbf{X}'\mathbf{X}_0)^{-1}(\mathbf{X}'\boldsymbol{\Omega}_0^{-1}\mathbf{X})(\mathbf{X}_0'\mathbf{X})^{-1}$$
$$E[(\mathbf{b} - \boldsymbol{\beta})(\mathbf{b}_0 - \boldsymbol{\beta}_0)'] = (\mathbf{X}_0'\mathbf{X})^{-1}\mathbf{X}_0'E(\boldsymbol{\varepsilon},\boldsymbol{\varepsilon}_0'\,|\,\mathbf{X},\mathbf{X}_0)\mathbf{X}(\mathbf{X}_0'\mathbf{X})^{-1}$$
$$= \sigma_0^2\sigma^2(\mathbf{X}_0'\mathbf{X})^{-1}\mathbf{X}_0'\boldsymbol{\Psi}\mathbf{X}(\mathbf{X}_0'\mathbf{X})^{-1}. \quad (4)$$

Note that pre-multiplying the normal equations for $\mathbf{b}_{GLS}$ by $(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{X}'(\mathbf{X}_0\mathbf{X}_0')^{-1}\mathbf{X}_0$ and the normal equations for $\mathbf{b}_{LS}$ by $(\mathbf{X}'\mathbf{X}_0)^{-1}\mathbf{X}'(\mathbf{X}_0\mathbf{X}_0')^{-1}\mathbf{X}_0$ in (1) and (3), respectively, allows the identities

$$\mathbf{b}_{0,GLS} = (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}\mathbf{b}_{GLS}$$
$$\mathbf{b}_{0,LS} = (\mathbf{X}'\mathbf{X}_0)^{-1}\mathbf{X}'\mathbf{X}\mathbf{b}_{LS}, \quad (5)$$

provided $(\mathbf{X}_0\mathbf{X}_0')^{-1}$ exists. These identities, however, are not required for the forthcoming analysis.

2.2 Another IV estimator?

At a first sight, (1) and (3) may resemble the well-known instrumental variables (IV) estimators. However, this is only apparent because the latter require knowledge of a closely-related-to-$\mathbf{X}_0$ matrix $\mathbf{Z}$ that satisfies $\mathbf{Z}'\varepsilon_0 = \mathbf{0}$ although the IV approach does not provide a way of finding such a matrix. Instead our approach requires only the replacement of the distorted values in $\mathbf{X}_0$ by any code that satisfies the Grenander conditions (see section 4), provided the distorted data appear completely at random. The relevance of this issue can be seen using an example. Consider the matrix of instruments $\mathbf{Z} = \mathbf{\Omega}^{-1}\mathbf{X}$ and the transformed model

$$\mathbf{Z'y} = \mathbf{Z'X}_0\boldsymbol{\beta}_0 + \boldsymbol{\nu} \text{ where } \boldsymbol{\nu}\sim(\mathbf{0},\ \sigma_0^2\mathbf{Z}'\mathbf{\Omega}_0\mathbf{Z}).$$

Then $\mathbf{b}_{0,\mathrm{IV}} = (\mathbf{Z'X}_0)^{-1}\mathbf{Z'y}$ is an IV estimator equivalent to $\mathbf{b}_{0,\mathrm{GLS}}$ in (1). However, the role of $\mathbf{Z}$ as a matrix of instruments is obscure because there is no way to check if $\mathbf{Z}$ is close enough to $\mathbf{X}_0$ and the fulfillment of the orthogonality condition $\mathbf{z}_j'\varepsilon_0 = 0$ without knowing $\mathbf{X}$ and $\mathbf{\Omega}$, which (from an IV perspective) are completely unknown. In our approach $\mathbf{X}$ is a matrix of true and assigned values completely defined and $\mathbf{\Omega}^{-1}$ is just an arbitrary weighting matrix that may be estimated by standard procedures. Accordingly, while through our approach we ensure that $\mathbf{b}_0$ will converge asymptotically to $\boldsymbol{\beta}_0$, through the IV approach we do not know that for certain.

3. CONSTRAINED SOLUTIONS

We propose here the Lagrangean

$$L = (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'\mathbf{\Omega}^{-1}(\mathbf{y}-\mathbf{X}_0\boldsymbol{\beta}_0) + \boldsymbol{\lambda}'(\mathbf{T}\boldsymbol{\beta}-\mathbf{t}) + \boldsymbol{\lambda}_0'(\mathbf{R}\boldsymbol{\beta}_0-\mathbf{r}),$$

where $\mathbf{T}$ is a linear transformation of $\mathbf{R}$ which allows us to express the system $\mathbf{R}\boldsymbol{\beta}_0 = \mathbf{r}$ as $\mathbf{T}\boldsymbol{\beta} = \mathbf{t}$, and vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}_0$ are two Lagrange multipliers. Then, the first order conditions are,

$$\partial L/\partial \mathbf{b}_0^* = -\mathbf{X}_0'\mathbf{\Omega}^{-1}\mathbf{y} + \mathbf{X}_0'\mathbf{\Omega}^{-1}\mathbf{X}\mathbf{b}^* + \mathbf{R}'\boldsymbol{\lambda}_0^* = \mathbf{0}$$

$$\partial L/\partial \mathbf{b}^* = -\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{y} + \mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X}_0\mathbf{b}_0^* + \mathbf{T}'\boldsymbol{\lambda}^* = \mathbf{0}$$

$$\partial L/\partial \boldsymbol{\lambda}^* = \mathbf{T}\mathbf{b}^* - \mathbf{t} = \mathbf{0}$$

$$\partial L/\partial \boldsymbol{\lambda}_0^* = \mathbf{R}\mathbf{b}_0^* - \mathbf{r} = \mathbf{0}$$

with solutions

$$\mathbf{b}_{GLS}^{*} = \mathbf{b} - (\mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{R}'[\mathbf{T}(\mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{Tb} - \mathbf{t})$$

$$\mathbf{b}_{0,GLS}^{*} = \mathbf{b}_0 - (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{T}'[\mathbf{R}(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{T}']^{-1}(\mathbf{Rb}_0 - \mathbf{r}). \quad (6)$$

These solutions are called Restricted Generalized Least Squares (RGLS). Similarly, a Restricted Least Squares (RLS) solution may be obtained by minimizing the constrained LS Lagrangean

$$L = (\mathbf{y}-\mathbf{X}\beta)'(\mathbf{y}-\mathbf{X}_0\beta_0) + \lambda'(\mathbf{T}\beta-\mathbf{t}) + \lambda_0'(\mathbf{R}\beta_0-\mathbf{r}),$$

and solving the first order conditions, resulting in

$$\mathbf{b}_{LS}^{*} = \mathbf{b} - (\mathbf{X}_0'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{T}(\mathbf{X}_0'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{Tb} - \mathbf{t})$$

$$\mathbf{b}_{0,LS}^{*} = \mathbf{b}_0 - (\mathbf{X}'\mathbf{X}_0)^{-1}\mathbf{T}'[\mathbf{R}(\mathbf{X}'\mathbf{X}_0)^{-1}\mathbf{T}']^{-1}(\mathbf{Rb}_0 - \mathbf{r}). \quad (7)$$

Frank (2007, 2008a and 2009) showed that

$$\mathbf{b}_0^{*} = \mathbf{b}_0 - \boldsymbol{\Sigma}^2\mathbf{R}'(\mathbf{R}\boldsymbol{\Sigma}^2\mathbf{R}')^{-1}(\mathbf{Rb}_0-\mathbf{r})$$

is a general form for RGLS and RLS, where $\mathbf{b}_0$ is the matching unconstrained estimator and $\boldsymbol{\Sigma}^2$ is $var(\mathbf{b}_0)$ as in (2) or (4). Then, by equating

$$\boldsymbol{\Sigma}^2_{GLS}\mathbf{R}' = (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{T}' \text{ and } \boldsymbol{\Sigma}^2_{LS}\mathbf{R}' = (\mathbf{X}'\mathbf{X}_0)^{-1}\mathbf{T}'$$

we get, respectively, the restriction matrices

$$\mathbf{T} = \mathbf{R}(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{\Omega}_0\boldsymbol{\Omega}^{-1}\mathbf{X}\sigma_0^2$$

$$\mathbf{T} = \mathbf{R}(\mathbf{X}_0'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Omega}_0^{-1}\mathbf{X}\sigma_0^2.$$

The scalars σ_0^2 may be omitted since they cancel out in the expression of $\mathbf{b}_0^*$. The expectation, variance and covariance of $\mathbf{b}_0^*$ are

$$E(\mathbf{b}_0^*) = E(\mathbf{b}_0\,|\,\mathbf{X},\mathbf{X}_0,\mathbf{y}) - \boldsymbol{\Sigma}^2\mathbf{R}'(\mathbf{R}\boldsymbol{\Sigma}^2\mathbf{R}')^{-1}[\mathbf{R}E(\mathbf{b}_0\,|\,\mathbf{X},\mathbf{X}_0,\mathbf{y})-\mathbf{r}]$$

$$= \beta_0$$

$$E[(\mathbf{b}_0^* - \beta_0)(\mathbf{b}_0^* - \beta_0)'] = var\{[\mathbf{I} - \boldsymbol{\Sigma}^2\mathbf{R}'(\mathbf{R}\boldsymbol{\Sigma}^2\mathbf{R}')^{-1}\mathbf{R}]\mathbf{b}_0\}$$

$$= \boldsymbol{\Sigma}^2 - \boldsymbol{\Sigma}^2\mathbf{R}'(\mathbf{R}\,\boldsymbol{\Sigma}^2\mathbf{R}')^{-1}\mathbf{R}\boldsymbol{\Sigma}^2$$

$$E[(\mathbf{b}^* - \beta)(\mathbf{b}_0^* - \beta_0)'] = (\mathbf{I} - \mathbf{A})cov(\mathbf{b},\mathbf{b}_0)(\mathbf{I} - \mathbf{A}_0)$$

$$= \sigma_0^2\sigma^2\,(\mathbf{I} - \mathbf{A})\mathbf{B}\boldsymbol{\Psi}\mathbf{B}_0'(\mathbf{I} - \mathbf{A}_0)', \quad (8)$$

where $\mathbf{A}_0$ is $\boldsymbol{\Sigma}^2\mathbf{R}'(\mathbf{R}\boldsymbol{\Sigma}^2\mathbf{R}')^{-1}\mathbf{R}$ and $\mathbf{B}_0$ is either $(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}$ or $(\mathbf{X}'\mathbf{X}_0)^{-1}\mathbf{X}'$ depending if we are in a GLS or LS context, respectively. Matrices **A** and **B** are the counterpart for **b**.

4. ASYMPTOTICS

In this section, we derive the sampling distribution of $\mathbf{b}_0$ and $\mathbf{b}_0^*$. So, if $\mathbf{X}$ comes from a sample, $\mathbf{b}_0$ will converge to β_0 in probability (which we write $\mathbf{b}_0 \rightarrow_p \beta_0$) if

$$\begin{aligned} \text{plim}_{n\to\infty}\, \mathbf{b}_0 &= \beta_0 + \text{plim}_{n\to\infty}\, (\mathbf{S}/n)^{-1}\mathbf{X}'\mathbf{\Omega}^{-1}\varepsilon_0/n \qquad \text{(by Slutsky's th.)} \\ &= \beta_0 + \text{plim}_{n\to\infty}\, (\mathbf{S}/n)^{-1}\, \text{plim}_{n\to\infty}\, \mathbf{X}'\mathbf{\Omega}^{-1}\varepsilon_0/n \\ &= \beta_0 \end{aligned}$$

where $\mathbf{S} = \mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X}_0$. This condition is satisfied if

$$\text{(i) plim}_{n\to\infty}\mathbf{S}/n = \text{plim}_{n\to\infty}\, \mathbf{X}_0'\mathbf{\Omega}^{-1}\mathbf{X}_0/n + \text{plim}_{n\to\infty}\, \mathbf{U}'\mathbf{\Omega}^{-1}\mathbf{X}_0/n = \mathbf{Q}_0 + \mathbf{Q}_{\mathrm{U}}, \quad (9)$$

where $\mathbf{Q}_0$ and $\mathbf{Q}_{\mathrm{U}}$ are finite and nonsingular matrices, and

$$\text{(ii) plim}_{n\to\infty}\, \mathbf{X}_0'\mathbf{\Omega}^{-1}\varepsilon_0/n + \text{plim}_{n\to\infty}\, \mathbf{U}'\mathbf{\Omega}^{-1}\varepsilon_0/n = \mathbf{0}. \quad (10)$$

In (9) the convergence of $\mathbf{X}_0'\mathbf{\Omega}^{-1}\mathbf{X}_0/n$ to $\mathbf{Q}_0$ is guaranteed by imposing the Grenander conditions on $\mathbf{X}_0$. In (10) the convergence of $\mathbf{X}_0'\mathbf{\Omega}^{-1}\varepsilon_0/n$ to $\mathbf{0}$ is also guaranteed by imposing the Grenander conditions and by the definition of ε_0. The convergences of $\mathbf{U}'\mathbf{\Omega}^{-1}\mathbf{X}_0/n$ to $\mathbf{Q}_{\mathrm{U}}$ and $\mathbf{U}'\mathbf{\Omega}^{-1}\varepsilon_0/n$ to $\mathbf{0}$, however, require additional conditions.

The expression $\mathbf{U}'\mathbf{\Omega}^{-1}\mathbf{X}_0/n$ converges to $\mathbf{Q}_{\mathrm{U}}$ if and only if for any two columns of $\mathbf{U}$ and $\mathbf{X}_0$ (which we shall call $\mathbf{u}$ and $\mathbf{x}_0$, respectively), $\text{plim}_{n\to\infty}$ $\mathbf{u}'\mathbf{\Omega}^{-1}\mathbf{x}_0\ /n = q_{\mathrm{U}ij}$. Nevertheless, each element $u_{ij} = (1 - \delta_{x=x0})\ (c_{ij} - x_{0ij})$, whence

$$\begin{aligned} \text{plim}_{n\to\infty}\, \mathbf{u}_j'\mathbf{\Omega}^{-1}\mathbf{x}_{0j'}\ /n &= \text{plim}_{n,\, m\to\infty}\, [(\mathbf{I} - \Delta_j)(\mathbf{c}_j - \mathbf{x}_{0j})]'\mathbf{\Omega}^{-1}\mathbf{x}_{0j'}/n \\ &= \text{plim}_{n,\, m\to\infty}\, \mathbf{c}_j'(\mathbf{I} - \Delta_j)\mathbf{\Omega}^{-1}\mathbf{x}_{0j'}/n - const. \\ &\Rightarrow \text{plim}_{m\to\infty}\, \mathbf{c}_j'\mathbf{\Omega}^{-1}\mathbf{x}_{0j'}/m = const. \qquad (11) \end{aligned}$$

$$\begin{aligned} \text{plim}_{n\to\infty}\, \mathbf{u}_j'\mathbf{\Omega}^{-1}\varepsilon_0\ /n &= \text{plim}_{n,\, m\to\infty}\, [(\mathbf{I} - \Delta_j)(\mathbf{c}_j - \mathbf{x}_{0j})]'\mathbf{\Omega}^{-1}\varepsilon_0/n \\ &\Rightarrow \text{plim}_{m\to\infty}\, \mathbf{c}_j'\mathbf{\Omega}^{-1}\varepsilon_0\ /m = 0 \qquad (12) \end{aligned}$$

where Δ_j is a diagonal matrix of $\delta_{x=x0}$ and m is the number of distorted values in the j-th column of $\mathbf{X}_0$.[1] Clearly, (11) and (12) are satisfied by imposing the Grenander conditions on the code c_{ij}. With regard to $\mathbf{b}_0^*$ we write

[1] Note that in (11) we assumed that m grows at the same rate as n.

$$\begin{aligned}\operatorname{plim}_{n\to\infty}\mathbf{b}_0^* &= \boldsymbol{\beta}_0 - \operatorname{plim}_{n\to\infty}(\mathbf{S}/n)^{-1}\mathbf{T}'[\mathbf{R}(\mathbf{S}/n)^{-1}\mathbf{T}']^{-1}(\mathbf{R}\mathbf{b}_0 - \mathbf{r})\\ &= \boldsymbol{\beta}_0 - \mathbf{Q}^{-1}\mathbf{T}'[\mathbf{R}\mathbf{Q}^{-1}\mathbf{T}']^{-1}\,\mathbf{R}\mathbf{Q}^{-1}\operatorname{plim}_{n\to\infty}\mathbf{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon}_0/n\\ &= \boldsymbol{\beta}_0,\end{aligned}$$

where $\mathbf{b}_0^*$ is either the RGLS solution in (6) or the RLS solution in (7) and $\mathbf{Q} = \mathbf{Q}_0 + \mathbf{Q}_U$.

5. CONCLUSION

The unconstrained estimators presented in (1) and (3), as well as their constrained counterparts in (6) and (7), seem to have been overlooked in the econometric literature despite their simplicity (see e.g. Judge *et al.*, 1985; Greene, 2006; Rao *et al.* 2008) as alternative estimators for distorted datasets. According to the literature, the common practice to deal with distorted datasets is to replace $\mathbf{X}_0$ in e.g. the GLS solution by an estimator $\mathbf{X}_0^+$ obtained by some imputation criterion (Little, 1992). This approach, however, leads to an inconsistent estimator of $\boldsymbol{\beta}_0$ whenever the convergence in probability of $\mathbf{G}_0^+/n\to\mathbf{Q}$ is not guaranteed. Moreover, if an estimated $\boldsymbol{\Omega}_0$ is to be plugged into $\mathbf{G}_0^+$ the consistency of $\boldsymbol{\Omega}_0^+$ cannot be ensured as $\boldsymbol{\Omega}_0^+$ depends on the properties of the residuals of the feasible LS estimator which in turn depends on $\mathbf{X}_0^{+\prime}\mathbf{X}_0^+$. Instead, if we introduce $\mathbf{X}_0^+$ in our solution it still holds that $\operatorname{plim}_{n\to\infty}\mathbf{S}^+/n = \mathbf{Q}$, provided $\mathbf{X}$ remains non-random. The same argument is extendable to the constrained counterparts (6) and (7).

Although we introduced the concept of distorted data as more general than just missing data, our development appears to treat distorted data as missing. That is because we did not try to exploit the information contained in the true values by constructing a code c_{ij} after them. For example, we could have defined the perfectly valid code $c_{ij} = \mathbf{1}'\mathbf{x}_{0j}/(n-m)$ for the even i-s and $c_{ij} = -\mathbf{1}'\mathbf{x}_{0j}/(n-m)$ for the odd i-s, if $\mathbf{X}_0$ were not a sample matrix. If $\mathbf{X}_0$ were a sample matrix this code may presumably improve the convergence rate of $\mathbf{b}_0$ to $\boldsymbol{\beta}_0$ as $n\to\infty$, but progress in this is direction is beyond the scope of the paper.

Frank (2008b) arrived to (1) minimizing the sums of squares of the distorted model $L = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}$, subject to $\mathbf{X} = \mathbf{X}_0 + \mathbf{U}$ as shown in the appendix. In this paper we show that it is possible to obtain the same estimator minimizing the covariance between the true and the distorted models. In fact, the latter procedure is preferable to the first because it leads directly to $\mathbf{b}_0$ rather than to a feasible $\mathbf{b}_0$ after discarding an additive linear transformation of the error term $\boldsymbol{\varepsilon}_0$. The findings (6) and (7) are new, but it should be pointed out that they rely heavily on the existence of a general form for constrained linear models (see Frank, 2008a) although a proof of its uniqueness is (to the best knowledge of the author) still not available.

As a final remark, note that combining previous asymptotic results and the Central Limit Theorem (e.g. in Lindberg-Feller's version) we get easily

$$n^{1/2}(\mathbf{b}_0 - \boldsymbol{\beta}_0) \rightarrow_d N(\mathbf{O},\ \sigma_0^2\ \boldsymbol{\Sigma}^2)$$

$$n^{1/2}(\mathbf{b}_0^* - \boldsymbol{\beta}_0) \rightarrow_d N(\mathbf{O},\ \sigma_0^2\ \boldsymbol{\Sigma}^2 - \sigma_0^2\boldsymbol{\Sigma}^2\mathbf{R}'(\mathbf{R}\ \boldsymbol{\Sigma}^2\mathbf{R}')^{-1}\mathbf{R}\boldsymbol{\Sigma}^2),$$

where $\boldsymbol{\Sigma}^2 = \mathbf{A}\boldsymbol{\Omega}_0\mathbf{A}'$ and $\mathbf{A}_{\mathrm{GLS}} = \mathbf{S}^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}$ or $\mathbf{A}_{\mathrm{LS}} = (\mathbf{X}'\mathbf{X}_0)^{-1}\mathbf{X}'$. Note that the matrix

$$\boldsymbol{\Delta} = \sigma_0^2\mathbf{A}(\boldsymbol{\Omega}_0 - \mathbf{X}_0\mathbf{G}_0^{-1}\mathbf{X}_0')\mathbf{A}' = \sigma_0^2(\mathbf{AP})\mathbf{I}_n(\mathbf{AP})'$$

is symmetric positive semi-definite (Lütkepohl, 1996, pp. 156-157 and 151-152) as there exists a matrix $\mathbf{P}$ such that $\boldsymbol{\Omega}_0 - \mathbf{X}_0\mathbf{G}_0^{-1}\mathbf{X}_0'$ may be written as $\mathbf{PP}'$. Then, $\mathbf{b}_0$ is an inefficient estimator of $\boldsymbol{\beta}_0$ as would be expected due to the loss of information caused by the distortion of $\mathbf{X}_0$. However, the properties of $\mathbf{b}_0$ and $\mathbf{b}_0^*$ in the context of small samples remain unknown.

REFERENCES

Bera, A. K. (1994). Review: Econometric Analysis by William H. Greene. *Journal of the American Statistical Association* 89(428), 1567-1569.

Frank, L. (2007). A Comparison of Restricted Heteroskedastic Estimators for Small Samples ($n \leq 50$). *2007 Proceedings of the American Statistical Association, Business and Economic Statistics Section* [CD-ROM], Alexandria, VA: American Statistical Association: 872-875.

Frank, L. (2008a). Restricted Heteroskedastic IV Estimators: A Simulation with Small Samples. *2008 Proceedings of the American Statistical Association* [CD-ROM], Alexandria, VA: American Statistical Association.

Frank, L. (2008b). Missing Values in **X**: A Small Sample Simulation. *Trabajos de CLATSE 2008*. Montevideo, Uruguay.

Frank, L. (2009). Restricted Linear Models: Which Estimator Performs Better? *2009 Proceedings of the American Statistical Association, Business and Economic Statistics Section*, Alexandria, VA: American

Statistical Association: 2049-2054. Available at https://www.amstat.org.

Henderson, H. and Searle, S. (1981). On Deriving the Inverse of a Sum of Matrices. *SIAM Review* 23(1), 53-60.

Greene, W. (2006). Econometric Analysis. 6th. Edition. Prentice Hall.

Judge, G. Griffiths, W. Carter Hill, R. Lütkepohl, H. Tsoung-Chao Lee (1985). The Theory and Practice of Econometrics. Wiley.

Little, R. (1992). Regression with Missing **X**'s: A Review. *Journal of the American Statistical Association* 87(420), 1227-1237.

Lütkepohl, H. (1996). Handbook of Matrices. Wiley. New York.

Rao, C. R., Toutenburg, H. Shalabh and Heumann, C. (2008). Linear Models and Generalizations. Least Squares and Alternatives. 3rd Edition. Springer Verlag.

APPENDIX

We minimize the Lagrangean

$$L = f(\beta\,|\,\mathbf{X},\mathbf{y}) = (\mathbf{y} - \mathbf{X}\beta)'\boldsymbol{\Omega}^{-1}(\mathbf{y} - \mathbf{X}\beta) + 2\mathbf{g}'(\mathbf{X}-\mathbf{X}_0-\mathbf{U})$$

and solve the first order conditions $\partial L/\partial\mathbf{b} = \mathbf{0}$ and $\partial L/\partial\mathbf{g} = \mathbf{0}$, in order to get the well-known GLS solution

$$\begin{aligned}\mathbf{b} &= [(\mathbf{X}_0 + \mathbf{U})'\boldsymbol{\Omega}^{-1}(\mathbf{X}_0 + \mathbf{U})]^{-1}(\mathbf{X}_0 + \mathbf{U})'\boldsymbol{\Omega}^{-1}\mathbf{y}\\ &= (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{y} \qquad (A.1)\end{aligned}$$

From this solution we shall deduce an expression for $\mathbf{b}_0$. Therefore, we define the following notation:

(a) $\mathbf{G} = \mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}$, $\mathbf{G}_0 = \mathbf{X}_0'\boldsymbol{\Omega}_0^{-1}\mathbf{X}_0$ and $\hat{\mathbf{G}}_0 = \mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{X}_0$

(b) $\mathbf{H} = \mathbf{G}^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}$, $\mathbf{H}_0 = \mathbf{G}_0^{-1}\mathbf{X}_0'\boldsymbol{\Omega}_0^{-1}$ and $\hat{\mathbf{H}}_0 = \hat{\mathbf{G}}_0^{-1}\mathbf{X}_0'\boldsymbol{\Omega}^{-1}$

where the existence of $\mathbf{H}$, $\mathbf{H}_0$ and $\hat{\mathbf{H}}_0$ is guaranteed because $\mathbf{G}$, $\mathbf{G}_0$ and $\hat{\mathbf{G}}_0$ are symmetric matrices (Lütkepohl, 1996, pp. 156-157) positive semidefinite (Lütkepohl, 1996, pp. 151-152). Then, adding and subtracting $\hat{\mathbf{H}}_0$ in (A.1) we write

$$\mathbf{b} = \hat{\mathbf{G}}_0^{-1}\mathbf{X}_0'\boldsymbol{\Omega}^{-1}\mathbf{y} + (\mathbf{G}^{-1}\mathbf{X}' - \hat{\mathbf{G}}_0^{-1}\mathbf{X}_0')\boldsymbol{\Omega}^{-1}\mathbf{y} \qquad (A.2)$$

However, $\boldsymbol{\Omega} = \boldsymbol{\Omega}_0 + (\boldsymbol{\Omega} - \boldsymbol{\Omega}_0)$ which implies that $\boldsymbol{\Omega}^{-1} = [\boldsymbol{\Omega}_0 + (\boldsymbol{\Omega} - \boldsymbol{\Omega}_0)]^{-1}$. Applying the result of (Henderson and Searle, 1981) on the inverse of a sum of matrices we get

$$\begin{aligned}\boldsymbol{\Omega}^{-1} &= \boldsymbol{\Omega}_0^{-1} - [\boldsymbol{\Omega}_0(\boldsymbol{\Omega} - \boldsymbol{\Omega}_0)^{-1}\boldsymbol{\Omega}_0 + \boldsymbol{\Omega}_0]^{-1}\\ &= \boldsymbol{\Omega}_0^{-1} - \mathbf{B}^{-1},\end{aligned}$$

in turn meaning that

$$\begin{aligned}\hat{\mathbf{G}}_0^{-1} &= [\mathbf{X}_0'(\boldsymbol{\Omega}_0^{-1} - \mathbf{B}^{-1})\mathbf{X}_0]^{-1}\\ &= (\mathbf{X}_0'\boldsymbol{\Omega}_0^{-1}\mathbf{X}_0)^{-1} + \mathbf{C}\\ &= \mathbf{G}_0^{-1} + \mathbf{C}.\end{aligned}$$

Now replacing these equalities in (A.2) yields

$$\begin{aligned}\mathbf{b} &= (\mathbf{G}_0^{-1} + \mathbf{C})\mathbf{X}_0'(\boldsymbol{\Omega}_0^{-1} - \mathbf{B}^{-1})\mathbf{y} + (\mathbf{H} - \hat{\mathbf{H}}_0)\,\mathbf{y}\\ &= \mathbf{b}_0 + [-\mathbf{G}_0^{-1}\mathbf{X}_0'\mathbf{B}^{-1} + \mathbf{C}\mathbf{X}_0'\boldsymbol{\Omega}^{-1} + \mathbf{G}^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1} - \hat{\mathbf{G}}_0^{-1}\mathbf{X}_0'\boldsymbol{\Omega}^{-1}]\,\mathbf{y},\end{aligned}$$

but recalling that $\mathbf{C} = \hat{\mathbf{G}}_0^{-1}-\mathbf{G}_0^{-1}$ and $\mathbf{B}^{-1}= \boldsymbol{\Omega}_0^{-1} - \boldsymbol{\Omega}^{-1}$

$$\begin{aligned}\mathbf{b} &= \mathbf{b}_0 + [-\mathbf{G}_0^{-1}\mathbf{X}_0'\mathbf{B}^{-1} + (\mathbf{C}- \hat{\mathbf{G}}_0)\mathbf{X}_0'\boldsymbol{\Omega}^{-1} + \mathbf{G}^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}]\,\mathbf{y}\\ &= \mathbf{b}_0 + [-\mathbf{G}_0^{-1}\mathbf{X}_0'(\mathbf{B}^{-1}+ \boldsymbol{\Omega}^{-1}) + \mathbf{G}^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}]\,\mathbf{y}\end{aligned}$$

$$= \mathbf{b}_0 + [\mathbf{G}^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1} - \mathbf{G}_0{}^{-1}\mathbf{X}_0'\boldsymbol{\Omega}_0{}^{-1}]\,\mathbf{y} \qquad \text{(A.3)}$$

where $\mathbf{b}_0 = \mathbf{G}_0{}^{-1}\mathbf{X}_0'\boldsymbol{\Omega}_0{}^{-1}\mathbf{y}$ is the GLS solution for the true model. Note that in addition to the initial assumptions, we require that $\boldsymbol{\Omega} - \boldsymbol{\Omega}_0$ is a nonsingular matrix. Next we replace $\mathbf{y}$ in (A.3) by the expression of the true model, and add and subtract the term $(\mathbf{HX}_0 - \mathbf{I})\mathbf{b}_0$ on the right hand side of equality. So we obtain

$$\mathbf{b} = \mathbf{b}_0 + (\mathbf{HX}_0 - \mathbf{I})\beta_0 + (\mathbf{H} - \mathbf{H}_0)\varepsilon_0$$

$$= \mathbf{HX}_0\mathbf{b}_0 - (\mathbf{HX}_0 - \mathbf{I})(\mathbf{b}_0 - \beta_0) + (\mathbf{H} - \mathbf{H}_0)\varepsilon_0.$$

Pre-multiplying both sides of the expression by $(\mathbf{HX}_0)^{-1}$, canceling identities and rearranging terms yields

$$\mathbf{b}_0 = (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{y} + \boldsymbol{\delta} \qquad \text{(A.4)}$$

where we assume that $(\mathbf{HX}_0)^{-1} = (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}$ exists. Expanding the term $\boldsymbol{\delta}$ and recalling that $\mathbf{b}_0 - \beta_0 = \mathbf{H}_0\varepsilon_0$ we see that

$$\boldsymbol{\delta} = [\mathbf{I} - (\mathbf{HX}_0)^{-1}]\,(\mathbf{b}_0 - \beta_0) - (\mathbf{HX}_0)^{-1}(\mathbf{H} - \mathbf{H}_0)\varepsilon_0$$

$$= [\mathbf{H}_0 - (\mathbf{HX}_0)^{-1}\mathbf{H}]\,\varepsilon_0$$

or, more explicitly,

$$\boldsymbol{\delta} = [\mathbf{G}_0{}^{-1}\mathbf{X}_0'\boldsymbol{\Omega}_0{}^{-1} - (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}]\,\varepsilon_0$$

$$= -\,(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}\varepsilon_0$$

if $(\boldsymbol{\Omega}_0{}^{-1/2}\mathbf{X}_0)'(\boldsymbol{\Omega}_0{}^{-1/2}\varepsilon_0) = \mathbf{0}$. However, $\boldsymbol{\delta}$ is not observable, so we propose the feasible solution

$$\mathbf{b}_0{}^* = (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}_0)^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{y}.$$