

A Comparison of Alternative Estimators of Linearly Aggregated Macro Models

Fikri Akdeniz
Department of Statistics
Faculty of Arts and Sciences
Çukurova University, Adana, Turkey
fikriakdeniz@gmail.com

Hans Joachim Werner
Department of Statistics
Faculty of Law & Economics
University of Bonn, Adenauerallee 24-42
D-53113 Bonn, Germany
hjlw.de@uni-bonn.de

Abstract

This paper deals with the linear aggregation problem. For the true underlying micro relations, which explain the micro behavior of the individuals, no restrictive rank conditions are assumed. Thus the analysis is presented in a framework utilizing generalized inverses of singular matrices. We investigate several estimators for certain linear transformations of the systematic part of the corresponding macro relations. Homogeneity of micro parameters is discussed. Best linear unbiased estimation for micro parameters is described.

Keywords: Aggregation bias, Best linear unbiased estimator, Linear aggregation, Macro parameter, Micro parameter, Specification error.

AMS classifications: 62J05; JEL: C13

1. Introduction

The linear aggregation problem considered consists of obtaining consistent and economically meaningful micro economic (from individuals) and macro economic (from groups of individuals) relationships.

The problem of aggregation over micro units has had a long tradition in the econometric literature, stretching back to the pioneering work of (Theil 1954). (Kloek 1961) used matrix notation to explain aggregation over one set of individuals. (Misra 1969, Gupta 1970, Moriguchi 1970, Lutjohann 1972, Wu 1973, Akdeniz and Milliken 1975), Akdeniz and Werner (1999) also utilized matrix notation to study micro and macro economic relations. Sasaki (1978) has analyzed empirically the aggregation problems. Lee *et al.* (1990) considered the problem of 'aggregation bias' defined by the deviation of the macro parameters from the average of the corresponding micro parameters. In addition to an aggregation of a real units, one may also distinguish an aggregation of models or

equations pertaining to a real units. In this regard, one may analyze the impact of a particular aggregation level of individual units upon the explanatory power of a model or relationship (see Blommestein and Nijkamp,1986).

In this paper, we considered a more general linear aggregation model by allowing for collinearity between the explanatory variables. Thus, the analysis is presented in a framework using generalized inverses of singular matrices. In this case, we obtained a consistent relationship between the micro and macro parameters.

We are concerned with a comparison of two estimators of the macro parameters (actually, due to the collinearity, of the structural portion $X\gamma$). One obvious estimate is obtained directly from the application of ordinary least squares to the macro equation. It is well known (indeed, because of the specification error above) that the expectation of this estimate generally differs from the aggregate of the expectation of ordinary least squares estimates obtained from the micro equations. Therefore, comparing these two estimators will necessarily involve considerations of bias as well as efficiency. Nevertheless, this is the comparison which has traditionally been made. In the present study the direct ordinary least squares estimate (OLSE) of the macro equation is retained, but a rather different aggregate of the micro estimates is employed. Instead of simply aggregating the micro estimates, the projection into the space spanned by the macro explanatory variables is taken.

This paper is organized as follows. The next section describes the general linear aggregation problem. A comparison of two estimators of the macro parameters is given in Section 3. The OLSE and weighted least squares estimator (WLSE) of the macro parameters is compared in Section 4. Homogeneous micro parameters and aggregation bias are discussed in Section 5. Section 6 describes the best linear unbiased estimation of micro parameters.

2. The general linear aggregation problem

Let the economic relationship for the i -th economic unit (general disaggregated model) be given by

$$y_i = X_i\beta_i + u_i, \quad i = 1, 2, \dots, N \quad (2.1)$$

or

$$L_i := \{y_i, X_i\beta_i; V_i\}, \quad i = 1, 2, \dots, N \quad (2.2)$$

where y_i is a vector of observations on the micro dependent variable; X_i is a $T \times K$ matrix of explanatory micro variables; β_i is a vector of K micro parameters and u_i is a disturbance vector, N is the number of economic units and $\text{cov}(y_i, y_j) = 0$ for $i \neq j$; X_i and $V_i \geq 0$ are fixed and known matrices.

By letting

$$y^* = [y'_1, y'_2, \dots, y'_N]', \quad X^* = \text{diag}(X_1, X_2, \dots, X_N)$$

$$\beta^* = (\beta'_1, \beta'_2, \dots, \beta'_N)', \quad u^* = (u'_1, u'_2, \dots, u'_N)', \quad V^* = \text{diag}(V_1, V_2, \dots, V_N)$$

we can obviously combine our N individual micro relations in the *single linear micro model*

$$y^* = X^* \beta^* + u^* \tag{2.3}$$

or

$$L^* := \{y^*, X^* \beta^*, V^*\} \tag{2.4}$$

The following assumptions are used throughout this paper.

Assumption 1. The matrix X_i has rank $r_i \leq K$ and consists of nonstochastic elements.

Assumption 2. u_i 's have zero means, are independent of (X_1, X_2, \dots, X_N) i.e.

$$E(u_i / X_i) = 0, E(u^* / X^*) = 0, E(u_i u_j' / X_i, X_j) = \begin{cases} V_i & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases} \text{ and } E(u^* u^{*'} / X^*) = V^*$$

where V^* is block-diagonal matrix, consisting of matrices V_i . Define the macro variables

$$y := \sum_{i=1}^N y_i \text{ and } X := \sum_{i=1}^N X_i = J_T X^* J_K' \tag{2.5}$$

as unweighted aggregates of the corresponding micro variables over the N individuals. For convenience, put

$$V := \sum_{i=1}^N V_i, J_T = j_N' \otimes I_T, J_K = j_N' \otimes I_K \tag{2.6}$$

where I_T is a $T \times T$ identity matrix j_N is a $N \times 1$ column vector of unit elements and $A \otimes B$ denotes the Kronecker product of the matrices A and B . Then $y = J_T y^*$ and it is clear that the *disaggregate model* or *true linear macro model (derived aggregate model)*

$$y = J_T X^* \beta^* + J_T u^* \tag{2.7}$$

or

$$L := \{y, J_T X^* \beta^*, V\} \tag{2.8}$$

is implied by the micro model

$$L^* := \{y^*, X^* \beta^*, V^*\},$$

where $Var(J_T u^*) = J_T Var(u^*) J_T' = J_T V^* J_T' = \sum_{i=1}^N V_i = V$. Using the aggregate

observations the *artificial macro model (aggregate model)* can be written as

$$y = X\gamma + u, \tag{2.9}$$

or

$$L_a = \{y, X\gamma, V\} \tag{2.10}$$

The macro relations in (2.9) generally differ from the true macro relations, implying there is possible *specification error* in (2.9). $E(y) \neq X\gamma$ can occur. We are trying essentially to describe the vector $J_T X^* \beta^*$ in the best possible way by $X\gamma$. The best description is the projection of $J_T X^* \beta^*$ onto the vector space spanned by the columns of X . The projection of onto the vector space is

$$X\gamma = XX^+ \sum_{i=1}^N X_i \beta_i = XX^+ J_T X^* \beta^*. \tag{2.11}$$

where X^+ denotes the Moore-Penrose inverse of X (Styan, 1983)). Akdeniz and Milliken (1975) has shown that the system of equations relating to γ and β^* , $X\gamma = XX^+J_T X^* \beta^*$ is consistent. We don't need the Moore-Penrose inverse of X to establish the relationship between the quantities in the micro and macro equations. We can use any generalized inverse, but we need the Moore-Penrose inverse in Section 3.

Thus the specification error \tilde{u} is in the vector space spanned by the columns of X , i.e.,

$$\tilde{u} = (I - XX^+)J_T X^* \beta^*, \tag{2.12}$$

or

$$\tilde{u} = J_T X^* \beta^* - X\gamma. \tag{2.13}$$

The specification error is the vector between $J_T X^* \beta^*$ and its projection onto $C(X)$. It is also called the "Aggregation bias" of u (Green, 1964). Where $C(X)$ denotes the column space of X . Direct estimation of aggregate model parameters generates in general aggregation bias.

The macro disturbance vector will be equal to $\sum_{i=1}^N u_i = J_T u^*$ only if the "perfect aggregation" condition $H_\xi : J_T X^* \beta^* - X\gamma = 0$ is satisfied (Pesaran *et al.* 1989). Hence the macro disturbance vector u is expressed as follows:

$$u = \tilde{u} + J_T u^* \tag{2.14}$$

or

$$u = S\beta^* + J_T u^* \tag{2.15}$$

Where $S := J_T X^* - XP$, $P := X^+ J_T X^*$ (Wu,1973; Terzi, 2009). S is a $T \times KN$ matrix of least squares residuals and P is a $K \times KN$ matrix of least squares "auxiliary regression" coefficients. The estimated auxiliary equation for the i -th micro unit is written in the form (Kloek,1961):

$$X_i = XB_i + W_i \tag{2.16}$$

where $B_i := X^+ X_i$ and $W_i := (I - XX^+)X_i$. Thus we have $X = \sum_{i=1}^N X_i = X \sum_{i=1}^N B_i + \sum_{i=1}^N W_i$

and $\sum_{i=1}^N B_i = X^+ X$, $\sum_{i=1}^N W_i = 0$.

Therefore, (2.14) is expressed as follows:

$$u = \sum_{i=1}^N W_i \beta_i + \sum_{i=1}^N u_i. \tag{2.17}$$

The aggregation procedure complicates the structure of macro errors. In equation (2.15) there are two terms; the first term is the linear combination of all micro coefficients, weighted by components of the S matrix, and the second term is the sum of corresponding micro errors (Wu,1973; Maddala, 1977). When the macro relation is specified as the regression model (2.9), the macro parameter γ is implicitly defined as a

vector of constants independent of X (Lutjohann (1972)). For full rank model γ is defined as a mathematical expectation of its least squares estimator (OLSE) (see Kloek, 1961). For less than full rank model $X\gamma$ is defined as the expectation of the least squares estimates when y is regressed on X i.e.,

$$X\gamma = E(XX^+y / X^*) = E(X\hat{\gamma} / X^*), \tag{2.18}$$

where $\hat{\gamma} = X^+y$ is the OLSE of γ from (2.9). It can be shown that (2.11) and (2.18) are equivalent.

3. Comparison of two estimators of the macro parameters

By using the relationship $X\gamma = XX^+J_T X^* \beta^*$ between γ and β^* one can obtain two estimators of $X\gamma$. Therefore, we consider $X\gamma$ here rather than γ , since we don't assume that X necessarily has full column rank. Many widespread uses of the linear model $y = X\gamma + u$ are such that γ can not be estimated uniquely. But linear combinations of elements of $X\gamma$ can be estimated uniquely. Confining attention to estimating $X\gamma$ therefore avoid that lack of uniqueness.

The least squares estimator, to be denoted $OLSE(X\gamma)$ is

$$\hat{\eta}_1 = X\hat{\gamma} = OLSE(X\gamma) = XX^+y. \tag{3.1}$$

But another estimator exists by using the relationship $X\gamma = XX^+J_T X^* \beta^*$. Let X^*b^* be the vector of least squares estimate of the micro parameters obtained from (2.1). Then the other unbiased estimate of $X\gamma$ is

$$\hat{\eta}_2 = Xb = XX^+J_T X^* b^*. \tag{3.2}$$

The expectation of Xb is

$$E(Xb / X^*) = XX^+J_T X^* \beta^* = XP\beta^* = X\gamma, \tag{3.3}$$

$$E(Xb / X^*) = E[XX^+J_T X^* b^* / X^*]. \tag{3.4}$$

Thus we have two unbiased estimators of $X\gamma$, $\hat{\eta}_1$ and $\hat{\eta}_2$. These two estimators will, in general, not identical. The problem of equality $\hat{\eta}_1 = \hat{\eta}_2$ is discussed in the following theorem.

THEOREM 3.1. The two estimators of $X\gamma$, $\hat{\eta}_1$ and $\hat{\eta}_2$ are equal if and only if

$$C(X) \subseteq C(X_i) \tag{3.5}$$

for $i = 1, 2, \dots, N$, where $C(X_i)$ denotes the column space of X_i .

Proof of necessity: Let $\hat{\eta}_1 = \hat{\eta}_2$. Observe that the two estimators under study can be expressed in terms of micro dependent variable as $\hat{\eta}_1 = XX^+J_T y^*$ and $\hat{\eta}_2 = XX^+J_T X^* X^{*+} y^*$. Letting $H := XX^+$ and $H^* := X^* X^{*+}$, we therefore get

$$\hat{\eta}_1 - \hat{\eta}_2 = HJ_T(I_T - H^*)y^* \tag{3.6}$$

Consequently, $\hat{\eta}_1 = \hat{\eta}_2$ if and only if $HJ_T(I_T - H^*)y^* = 0$ for all y^* or equivalently,

$$HJ_T(I_T - H^*) = 0$$

The latter happens if and only if we have $H = HX_iX_i^+, i = 1, 2, \dots, N$. Premultiplying $H = HX_iX_i^+$ by X' gives $X' = X'X_iX_i^+$ or $X = X_iX_i^+X$ and therefore $C(X) \subseteq C(X_i)$ for each i completing the proof on necessity.

Proof of sufficiency: To that end, let (3.5) hold true for each $i = 1, 2, \dots, N$. Then there exist matrices $A_i (i = 1, 2, \dots, N)$ such that

$$X = \sum_{j=1}^N X_j = X_i A_i. \tag{3.7}$$

The matrix product $HJ_T(I_T - H^*)$ can then be expressed as

$$\begin{aligned} & X'^+ \{ X'(I - X_1 X_1^+), \dots, X'(I - X_i X_i^+), \dots, X'(I - X_N X_N^+) \} = \\ & X'^+ \{ A'_1 X'_1 (I - X_1 X_1^+), \dots, A'_i X'_i (I - X_i X_i^+), \dots, A'_N X'_N (I - X_N X_N^+) \} = 0. \end{aligned}$$

Thus $HJ_T(I - H^*) = 0$ implies that $HJ_T(I - H^*)y^* = 0$ which implies that

$$\hat{\eta}_1 = \hat{\eta}_2.$$

The experimental situations where these two estimators are identical will be the exception rather than the rule, thus we wish to determine which of the two estimators is better when they are not equal. To do that we examine the covariance structure of each estimator to determine which estimator has the smaller variance. The following Theorem describes this relationship.

THEOREM 3.2. The variances of the components of $\hat{\eta}_2$ are smaller than or equal to the corresponding variances of the components of $\hat{\eta}_1$, thus $\hat{\eta}_2$ is the better of the two estimators.

Proof. The proof consists of deriving the dispersion matrices of $\hat{\eta}_1$ and $\hat{\eta}_2$ and then examining the diagonal elements of the difference of the two dispersion matrices. The dispersion matrix of $\hat{\eta}_1$ is denoted by Σ_1 and is defined as

$$\begin{aligned} \Sigma_1 &= E \left[(X\hat{\gamma} - XX^+ \sum_{i=1}^N X_i \beta_i)(X\hat{\gamma} - XX^+ \sum_{i=1}^N X_i \beta_i)' \right], \\ &= E(X\hat{\gamma} - X\gamma)(X\hat{\gamma} - X\gamma)' \\ &= E(XX^+ y - X\gamma)(XX^+ y - X\gamma)' \\ &= XX^+ E(uu')XX^+. \tag{3.8} \\ &= HE(uu')H' \end{aligned}$$

We use the expression for u given in (2.15) to evaluate $E(uu')$ as

$$E(uu') = J_T V^* J_T' + MRR'M' \tag{3.9}$$

where $M := I - XX^+$ and $R := J_T X^* \beta^*$. Thus replacing $E(uu')$ by (3.9), Σ_1 becomes

$$\begin{aligned} \Sigma_1 &= H(J_T V^* J_T' + MRR'M')H' \\ &= HJ_T V^* J_T' H' \\ &= H\left(\sum_{i=1}^N V_i\right)H' = HVH' \end{aligned} \tag{3.10}$$

The dispersion matrix of $\hat{\eta}_2 = Xb$ is

$$\begin{aligned} \Sigma_2 &= E\left[(Xb - XX^+ \sum_{i=1}^N X_i \beta_i)(Xb - XX^+ \sum_{i=1}^N X_i \beta_i)'\right] \\ &= [(Xb - X\gamma)(Xb - X\gamma)'] \\ &= E[(HJ_T X^* b^* - X\gamma)(HJ_T X^* b^* - X\gamma)'] \\ &= HJ_T E[(X^* b^* - X^* \beta^*)(X^* b^* - X^* \beta^*)'](HJ_T)' \\ &= HJ_T H^* V^* H^* J_T' H'. \end{aligned} \tag{3.11}$$

To compare the two estimators, we examine $\Sigma_1 - \Sigma_2$

$$\begin{aligned} \Sigma_1 - \Sigma_2 &= HJ_T (V^* - H^* V^* H^*) (HJ_T)' \\ &= HJ_T Q (HJ_T)' \end{aligned} \tag{3.12}$$

where $Q = (V^* - H^* V^* H^*) = V^* - X^* X^{*+} V^* X^* X^{*+}$ is block-diagonal matrix, consisting of matrices $V_i - X_i X_i^+ V_i X_i X_i^+$.

We assumed that each V_i is positive definite, thus V^* is also positive definite. To substantiate the theorem, it is sufficient to show that the matrix Q is nonnegative. As V^* is positive definite there exists a non singular matrix Δ such that $V^* = \Delta' \Delta$ or that $\Delta'^{-1} V^* \Delta^{-1} = I$. The matrix Q can be expressed as

$$Q = \Delta'(I - (\Delta'^{-1} H \Delta')(\Delta H \Delta^{-1}))\Delta. \tag{3.13}$$

Thus it is sufficient to show $I - N'N$ is nonnegative, where $N = \Delta H \Delta^{-1}$. But $H^2 = H$ and $N^2 = N$, hence the characteristic roots of N are either +1 or 0 as are the characteristic roots of $N'N$. The matrix $N'N$ is nonnegative, thus there exists an orthogonal matrix Λ such that $\Lambda' N' N \Lambda = D_\lambda$ where D_λ has the characteristic roots of $N'N$ on the diagonal, i.e.,

$$D_\lambda = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

Hence

$$\Lambda'(I - N'N)\Lambda = I - \Lambda'N'N\Lambda = I - D_\lambda = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}$$

Thus $I - N'N$ has characteristic roots of +1 and 0, which implies $I - N'N$ is idempotent and hence a nonnegative matrix. This means that Q is a nonnegative matrix and thus

$$\Sigma_1 - \Sigma_2 = HJ_T Q (HJ_T)' \tag{3.14}$$

is also a nonnegative matrix. The matrix $\Sigma_1 - \Sigma_2$ being nonnegative implies that its diagonal elements are greater than or equal to zero, thus the variance of each component of $\hat{\eta}_1 = X\hat{\gamma}$ is greater than or equal to the variance of the corresponding component of $\hat{\eta}_2 = Xb$. That completes the proof.

4. Comparison of OLSE and WLSE of the macro parameters

In model (2.9), $X\gamma$ is defined as a mathematical expectation of its weighted least squares estimator (WLSE)

$$X\gamma = E[X(X'V^{-1}X)^+ X'V^{-1}y / X^*] = E(X\tilde{\gamma} / X^*), \tag{4.1}$$

where $X\tilde{\gamma} = X(X'V^{-1}X)^+ X'V^{-1}y$ and $V = E[(u - E(u))(u - E(u))']$. The aggregation problems which arise from the use of macro equation (2.9) instead of the micro equations (2.1). The macro relations in (2.9) are generally different from the true macro relations in (2.7). The "aggregate" method presumably involves a specification error in (2.9). Essentially we are trying to describe the vector $J_T X^* \beta^*$ in the best possible way by $X\gamma$ weighted by the respective covariance matrices. This best description is the projection of $J_T X^* \beta^*$ onto the vector space spanned by the columns of X . The projection of $J_T X^* \beta^*$ onto the vector space weighted by V is

$$\begin{aligned} X\gamma &= X(X'V^{-1}X)^+ X'V^{-1}J_T X^* \beta^* \\ &= P^* J_T X^* \beta^*, \end{aligned} \tag{4.2}$$

where $P^* = X(X'V^{-1}X)^+ X'V^{-1}$ is the orthogonal projector onto $C(X)$, which is unique for any choice of the generalized inverse involved in (4.1) (Rao,1973)). It is shown that equations (4.1) and (4.2) are equivalent.

Thus the specification error u^0 is in the vector space orthogonal to the vector space spanned by the columns of a matrix X , i.e.,

$$u^0 = (I - P^*)J_T X^* \beta^* \tag{4.3}$$

Considering (2.7), (2.9) and (4.3) then the macro disturbance vector u can be written as

$$\begin{aligned} u &= J_T u^* + u^0, \\ &= J_T u^* + (I - P^*)J_T X^* \beta^*. \end{aligned} \tag{4.4}$$

By utilizing the relationship between γ and β^* two estimators of $X\gamma$ can be obtained. The WLSE of $X\gamma$ from (2.9) is

$$WLSE(X\gamma) = X\tilde{\gamma} = X(X'V^{-1}X)^+ X'V^{-1}y, \tag{4.5}$$

so that $X\gamma$ is uniquely determined. But another estimator exists by utilizing the relationship

$$X\gamma = P^* J_T X^* \beta^* . \tag{4.6}$$

Let $\tilde{\beta}^*$ be the vector of WLSE of the micro parameters obtained from (2.4), then an estimate of $X\gamma$ is

$$Xb^* = P^* J_T X^* \tilde{\beta}^* \tag{4.7}$$

where $X^* \tilde{\beta}^* = X^* (X^{*'} V^{*-1} X^*)^+ X^{*'} V^{*-1} y^*$. The expectation of Xb^* is

$$E(Xb^* / X^*) = X\gamma . . \tag{4.8}$$

THEOREM 4.1. The variances of the components of $X\tilde{\gamma}$ are greater than or equal to the corresponding variances of the components of Xb^* , thus Xb^* is the better of the two estimators.

Proof. The proof consists of deriving the dispersion matrices of $X\tilde{\gamma}$ and Xb^* and then examining the diagonal elements of the difference of the two dispersion matrices.

The dispersion matrix of $X\tilde{\gamma}$ is denoted by W_1 and is defined to be

$$W_1 = E[(X\tilde{\gamma} - E(X\tilde{\gamma}))(X\tilde{\gamma} - E(X\tilde{\gamma}))'] \tag{4.9}$$

The equation (4.4) is then used to evaluate (4.9). The expectation of $X\tilde{\gamma}$ is

$$\begin{aligned} E(X\tilde{\gamma} / X^*) &= E[X(XV^{-1}X)^+ XV^{-1}y / X^*] \\ &= X\gamma + P^* E(u) \\ &= X\gamma + P^* (I - P^*) J_T X^* \beta^* + P J_T E(u^* / X^*) \\ &= X\gamma , \end{aligned} \tag{4.10}$$

where $P^* (I - P^*) = 0$ and $E(u^* / X^*) = 0$. Then

$$W_1 = P^* E(uu')(P^*)' \tag{4.11}$$

where $E(uu') = (I - P^*)RR'(I - P^*)' + J_T V^* J_T'$ and $R := J_T X^* \beta^*$ as before. Thus W_1 becomes

$$\begin{aligned} W_1 &= P^* [(I - P^*)RR'(I - P^*)' + J_T V^* J_T'] (P^*)' \\ &= P^* [J_T V^* J_T'] (P^*)' = P^* V (P^*)' . \end{aligned}$$

The dispersion matrix of Xb^* is

$$\begin{aligned} W_2 &= E[(Xb^* - X\gamma)(Xb^* - X\gamma)'] \\ &= E(P^* J_T X^* \tilde{\beta}^* - P^* J_T X^* \beta^*)(P^* J_T X^* \tilde{\beta}^* - P^* J_T X^* \beta^*)' \\ &= P^* J_T X^* (X^{*'} V^{*-1} X^*)^+ (P^* J_T X^*)' . \end{aligned} \tag{4.12}$$

To compare the two estimators, we examine $W_1 - W_2$ or

$$W_1 - W_2 = P^* J_T [V^* - X^* (X^{*'} V^{*-1} X^*)^+ X^{*'}] (P^* J_T)' . \tag{4.13}$$

In order to substantiate the theorem, it is sufficient to show that $W_1 - W_2$ is a n.n.d. matrix. Since V^* is p.d. it can be written as $V^* := FF'$, where F is a nonsingular matrix. Thus (4.13) becomes

$$W_1 - W_2 = (P^* J_T F) M (P^* J_T F)', \quad (4.14)$$

Where $M := I - B(B'B)^+ B'$ and $B := F^{-1} X^*$. It is well known that M is n.n.d., then M can be written as $M := CC'$. Thus

$$W_1 - W_2 = (P^* J_T F C) (P^* J_T F C)' \quad (4.15)$$

is n.n.d. matrix. This completes the proof. Thus above theorem implies that the better estimator of the macro relations is via the consistent equation expressing the relationship between γ and β^* .

5. Homogeneous micro parameters and aggregation bias

The problem of 'aggregation bias', as originally discussed by Theil (1954) is defined in terms of the deviations of macro parameters from the averages of the corresponding micro parameters. Theil (1971) discusses the nature of aggregation bias in estimating a linear model with aggregate data to make macro inferences when the micro models are linear. Lee, Pesaran and Pierse (1990) develop direct tests of aggregation bias in contrast to the indirect test proposed by Zellner (1962) which tests the hypothesis that all the disaggregated coefficients are equal. Offutt (1988) performed two tests to investigate the hypothesis of aggregation bias for the random coefficient model. Pieraccini (2005) shows that direct estimation of the aggregate model leads in general to aggregation bias and that even in presence of perfect aggregation the disaggregate model is to be preferred. Monteforte (2007) assess the existence and characteristics of the aggregation error in estimated macroeconomic models. It should be recalled that for any macro parameter the aggregation bias due to any micro parameter is absent when all micro parameters are equal among micro units. Suppose the null hypothesis that all micro parameters are equal among micro units is tested.

It should be noted that the homogeneity of micro regression coefficients is not a necessary condition but a sufficient condition for the absence of aggregation bias. Hence, the rejection the hypothesis of homogeneity does not necessarily mean the presence of aggregation bias (Sasaki, 1978). Zellner (1962) developed the two alternative methods for test of homogeneity of micro regression coefficients.

5.1. Homogeneity of Micro Parameters

It is possible to establish test of a variety of hypothesis about coefficient vector in (2.2). Now, if we impose the restriction

$$\beta_1 = \beta_2 = \dots = \beta_N = \beta_* \quad (5.1)$$

or

$$L\beta = 0$$

where

$$L = \begin{bmatrix} 1 & -1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & -1 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & 1 & -1 & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & 1 & -1 \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \beta_{N-1} \\ \beta_N \end{bmatrix}$$

then the system in (2) becomes

$$y^* = Z^* \beta_* + v^* \tag{5.2}$$

where $Z^* = [X' \quad X'_2 \quad \dots \quad X'_{N-1} \quad X'_N]$. Different micro-units are characterized by the same regression coefficient. Pre multiplying of (5.2) by $J_T = j'_N \otimes I_T$ we have the macro relations are:

$$\tilde{y} = Z \beta_* + \tilde{v} \tag{5.3}$$

where $\tilde{y} := J_T y^*$, $Z := J_T Z^*$ and $\tilde{v} := J_T v^*$ i.e., there is no aggregation bias. It is shown that $Z \hat{\beta}_*$ is unbiased least squares estimator of $Z \beta_*$. According to the micro relations the expected value of $Z \hat{\beta}_*$ is

$$\begin{aligned} E(Z \hat{\beta}_* / Z^*) &= E(Z Z^+ y / Z^*) = E[Z Z^+ J_T Z^* \beta_* + J_T v^* / Z^*] \\ &= Z Z^+ J_T Z^* \beta_* = Z Z^+ Z \beta_* = Z \beta_*. \end{aligned} \tag{5.4}$$

In this case, the dispersion matrix of $Z \hat{\beta}_*$ is

$$\Sigma_1^* = Z Z^+ J_T \Sigma^* (Z Z^+ J_T)' \tag{5.5}$$

where $\Sigma^* = E(v^* v^{*'})$. The least squares estimator of β_* from (5.2) is any solution of the normal equations $Z^* Z^* \tilde{b} = Z^* y^*$. It is shown that $Z \tilde{b}$ is unbiased least squares estimator of $Z \beta_*$. The dispersion matrix of $Z \tilde{b}$ is

$$\Sigma_2^* = (Z Z^+ J_T Z^* Z^{*+}) \Sigma^* (Z Z^+ J_T Z^* Z^{*+})' \tag{5.6}$$

Finally, using the Theorem 3.2 one may shows that

$$\begin{aligned} \Sigma_1^* - \Sigma_2^* &= Z Z^+ J_T \Sigma^* (Z Z^+ J_T)' - (Z Z^+ J_T Z^* Z^{*+}) \Sigma^* (Z Z^+ J_T Z^* Z^{*+})' \\ &= Z Z^+ J_T [\Sigma^* - Z^* Z^{*+} \Sigma^* Z^* Z^{*+}] (Z Z^+ J_T)' \end{aligned} \tag{5.7}$$

is nonnegative definite.

6. Best linear unbiased estimation for micro parameters

Fundamental problem is to find the "best" estimate of the parameter vector β^* in Eq. (2.3) based on the y_1, y_2, \dots, y_N . There are many equivalent forms of the best linear

unbiased estimator of $X^* \beta^*$ in Eq. (2.3), to be denoted $BLUE(X^* \beta^*)$. When the dispersion matrix V^* is positive definite, then it is well known that

$$\begin{aligned} BLUE(X^* \beta^*) &= X^* (X^{*'} V^{*-1} X^*)^{-1} X^{*'} V^{*-1} y^* = X^* \tilde{\beta}^* = G' y^* \\ &= OLSE(X^* \beta^*) + H^* \Sigma M^* (M^* \Sigma M^*)^+ M^* y^* \end{aligned} \tag{6.1}$$

where $G' := X^* (X^{*'} V^{*-1} X^*)^{-1} X^{*'} V^{*-1}$ and $(X^{*'} V^{*-1} X^*)^{-1}$ is any generalized inverse of $(X^{*'} V^{*-1} X^*)$ and $H^* := X^* X^{*+}$, $M^* := I - H^*$ (Puntanen and Styan, 1989; Searle, 1994)

The least squares estimate of β^* is given by $b^* = X^{*+} y^*$ as before. $X^* b^*$ is also a linear unbiased estimate of $X^* \beta^*$. On the other hand, using the Eq. (2.6) we have the following expression $X \gamma^* = XX^+ J_T X^* \tilde{\beta}^*$, where $X \gamma^*$ is an estimate of $X \gamma$ in model $L_a = \{y, X \gamma, V\}$. Thus we have the following dispersion matrix of $X \gamma^*$

$$\begin{aligned} D_1 &= E[(X \gamma^* - E(X \gamma^*))(X \gamma^* - E(X \gamma^*))'] \\ &= E(XX^+ J_T P_0^* y^* - XX^+ J_T X^* \beta^*)(XX^+ J_T P_0^* y^* - XX^+ J_T X^* \beta^*)' \\ &= E(C^* u^* u^{*'} C^{*'}) = C^* E(u^* u^{*'}) C^{*'} = C^* V^* C^{*'} \end{aligned} \tag{6.2}$$

where $P_0^* := X^* (X^{*'} V^{*-1} X^*)^+ X^{*'} V^{*-1}$, $C^* := XX^+ J_T P_0^*$ and $P_0^* X^* = X^*$. The dispersion matrix of $XX^+ J_T X^* b^*$ is given by

$$\begin{aligned} D_2 &= E[XX^+ J_T (X^* b^* - X^* \beta^*)(X^* b^* - X^* \beta^*)(XX^+ J_T)'] \\ &= (XX^+ J_T X^* X^{*+}) V^* (XX^+ J_T X^* X^{*+})' \end{aligned} \tag{6.3}$$

Since $X^* \tilde{\beta}^*$ is BLUE of $X^* \beta^*$ we conclude that

$$D_2 - D_1 = (XX^+ J_T X^* X^{*+}) [V^* - X^* (X^{*'} V^{*-1} X^*)^+ X^{*'}] (XX^+ J_T X^* X^{*+})' \tag{6.4}$$

is nonnegative definite matrix by Theorem 4.1. Following Puntanen *et al.* (2000) we will state the following Theorem 6.1.

THEOREM 6.1 Let V be $p \times p$ a nonnegative definite and symmetric matrix. Moreover, let A be a $p \times p$ symmetric idempotent matrix of rank $r \leq p$. Then $\Sigma = V - AVA' = V - AVA$ is nonnegative definite if and only if

$$AV(I-A) = 0. \tag{6.5}$$

Proof. Let $[T: U]$ be a block-partitioned orthogonal matrix such that the columns of T and U constitute a basis for the range of A and $I-A$, respectively. Then we have $V - AVA' \geq 0$ if and only if

$$\begin{bmatrix} T' \\ U' \end{bmatrix} (V - AVA') [T: U] \geq 0 \tag{6.6}$$

Since $T'(V - AVA')T = 0$, $T'(V - AVA')U = T'VU$, $U'(V - AVA')T = U'VT$, and $U'(V - AVA')U = U'VU$, (6.6) reduces to

$$\begin{pmatrix} 0 & T'VU \\ U'VT & U'VU \end{pmatrix} \geq 0. \tag{6.7}$$

Evidently, the matrix (6.7) and so $V^{-1}AVA'$ is nonnegative definite if and only if $T'VU = 0$. The latter is equivalent to (6.5), and so the proof is complete.

THEOREM 6.2. Under the Gauss-Markov model (2.4), $L^* := \{y^*, X^* \beta^*, V^*\}$, $G'y^* = BLUE(X^* \beta^*)$ for all $y \in C(X^* : V^*)$ if and only if the following conditions

- a) $G'y^*$ is unbiased., i.e., $G'X^* = X^*$.
 - b) $C(G'V^*) \subseteq C(X^*)$., $G'V^* = X^*K$ for some K
- hold along with any one of the following two conditions:
- c₁) $G'V^* = V^*G$
 - c₂) Using the dispersion matrices we have the following Lowner partial ordering:
 $D(G'y^*) \leq D(y^*)$ i.e., $V^* - G'V^*G$ is n.n.d. (see Baksalary and Puntanen, 1990).

Proof: Using the Theorem 6.1 we get the following results for model $L^* := \{y^*, X^* \beta^*, V^*\}$.

- a) $G'y^* = X^*(X^{*'}V^{*-1}X^*)^{-1}X^{*'}V^{*-1}y^*$ is unbiased estimator of $X^* \beta^*$ i.e., $G'X^* = X^*$.

We know that $X^*(X^{*'}V^{*-1}X^*)^{-1}X^{*'}V^{*-1}X^* = X^*$ (see Rao, 1973), thus we get $G'X^* = X^*$.

- b) $G'V^* = X^*(X^{*'}V^{*-1}X^*)^{-1}X^{*'}V^{*-1}V^* = X^*K$; where $K = (X^{*'}V^{*-1}X^*)^{-1}X^{*'}$. Thus we have $C(G'V^*) \subseteq C(X^*)$.

- c₁) $G'V^* = X^*(X^{*'}V^{*-1}X^*)^{-1}X^{*'}V^{*-1}V^* = V^*V^{*-1}X^*(X^{*'}V^{*-1}X^*)^{-1}X^{*'}$, we have $G'V^* = V^*G$.

- c₂) $D(G'y^*) \leq D(y^*)$ i.e., $V^* - G'V^*G$ is n.n.d. It is seen that

$$\begin{aligned} V^* - G'V^*G &= V^* - X^*(X^{*'}V^{*-1}X^*)^{-1}X^{*'}V^{*-1}V^*V^{*-1}X^*(X^{*'}V^{*-1}X^*)^{-1}X^{*' \\ &= V^* - X^*(X^{*'}V^{*-1}X^*)^{-1}X^{*' \end{aligned}$$

is n.n.d. by Theorem 4.1, where $(X^{*'}V^{*-1}X^*)^{-1}$ reflexive generalized inverse of $X^{*'}V^{*-1}X^*$. On the other hand, since G' is idempotent matrix, then using the Theorem 6.1 $V^* - G'V^*G$ is n.n.d. if and only if $G'V^*(I - G) = 0$. It is seen that $G'V^* = X^*(X^{*'}V^{*-1}X^*)^{-1}X^{*' = G'V^*G$.

7. Conclusions

In this paper our primary concern has been with the problem of choice between macro and micro regression equations for the purpose of predicting macro variables. We are concerned with a comparison of two estimators of the macro parameters (actually, due to the collinearity, of the structural portion $X\gamma$). We consider $X\gamma$ rather than γ , since we do not assume that X necessarily has full column rank. The analysis is presented in a framework utilizing generalized inverses. The OLSE and WLSE estimates of the macro parameters is compared. Homogeneity of micro parameters is discussed. Best linear unbiased estimation for micro parameters is described.

References

1. Akdeniz, F. and Milliken, G. A.(1975).The relationship between macro and micro parameters. *International Economic Review* 16: 511-515.
2. Akdeniz, F. and Werner, H. J. (1999). More on the linear aggregation problem (Abstract) *Eight International Workshop on Matrices and Statistics*, 6-7 August, Tampere, Finland.
3. Baksalary, J. K. and Puntanen, S.(1990). Characterizations of the best linear unbiased estimator in the general Gauss-Markov model with the use of matrix partial ordering. *Linear Algebra and Its Appl.*127: 363-370.
4. Blommestein, H. and Nijkamp, P. (1986). Testing the spatial scale and dynamic structure in regional models (A contribution to spatial econometric specification analysis) *Journal of Regional Science* 26(1), 1-17.
5. Green, H. A. J. (1964). *Aggregation in Economic Analysis: Introductory Survey*, Princeton University Press.
6. Gupta, K. L. (1971). Aggregation bias in linear economic models. *International Economic Review*, 12(2): 293-305.
7. Kloek, T. (1961). Note on convenient matrix notations in multivariate statistical analysis and in the theory of linear aggregation. *International Economic Review*2: 351-360.
8. Lee, K. C. Pesaran, M. H. and Pierse, R. G. (1990). Testing for aggregation bias in linear models *The Economic Journal*, 100: 137-150.
9. Lütjohann, H. (1972). Another note on linear aggregation of economic relations. *International Economic Review* 13: 426-428.
10. Maddala, G. S. (1977). *Econometrics*, McGraw Hill, International Editions.
11. Misra, P.N. (1969). A note on linear aggregation of economic relations. *International Economic Review* 10: 247-249.
12. Monteforte, L. (2007) Aggregation bias in macro models: Does it matter for the euro area? *Economic Modelling* 24, 236-361.
13. Moriguchi, C. (1970). Aggregation over time in macroeconomic relations. *International Economic Review*, 11(3): 427-440.
14. Offutt, S. (1988) Aggregating crop production data: A random coefficient approach, *The Journal of Agricultural Economics Research*, 40(2): 1-18.
15. Pesaran, M. H., Pierse, R.G. and Kumar, M. S. (1989). Econometric analysis of aggregation in the context of linear prediction models, *Econometrica*57: 861-888.
16. Pieraccini, L.(2005). Is aggregation ever necessary?*Quaderni di Statistica*, Vol. 7: 1.

17. Puntanen, S. and Styan, G. P. H. (1989). The equality of the ordinary least squares estimator and the best linear unbiased estimator. *The American Statistician* 43:151-164.
18. Puntanen, S. Styan, G. P. H. and Werner, H. J. (2000). Two matrix based proofs that the linear estimator Gy is the best linear unbiased estimator. *Journal of Statistical Planning and Inference* 88(2): 173-179.
19. Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, Wiley, New York.
20. Sasaki, K. (1978). An empirical analysis of linear aggregation problems. *Journal of Econometrics* 7: 311-331.
21. Searle, S. R. (1994). Extending some results and proofs for the singular linear model. *Linear Algebra and Its Appl.* 210:139-151.
22. Styan, G. P. H. (1983). Generalized inverses, Encyclopedia of Statistical Sciences (S. Kotz, N.L. Johnson, C.B. Read, Eds.), Vol. 3, pp. 334-337 Wiley, New York.
23. Terzi, S. (2009) Perfect aggregation in linear models: a geometrical insight. *Quaderni di Statistica*, 11:1-14.
24. Theil, H. (1954). *Linear aggregation of economic relations*, North-Holland Publishing Company, Amsterdam.
25. Theil, H. (1971). *Principles of Econometrics*, Wiley, New York.
26. Wu, Y. R. (1973). On some aspects of linearly aggregated macro models. *International Economic Review* 14: 785-788.
27. Zellner, A. (1962). An efficiency method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*, 57: 348-368.