

Redundancy of Moment Conditions for Linear Transformation of Parameters

Heather Bednarek, Hailong Qian*

Department of Economics, Saint Louis University, Saint Louis, USA

Email: bednarhl@slu.edu, gianh@slu.edu

Received April 2014

Abstract

In this paper, we consider the redundancy of an extra set of moment conditions, given an initial set of moment conditions, for the efficient estimation of an arbitrary linear transformation of an original parameter vector. The redundancy condition derived in the current paper unifies the full and partial redundancy of moment conditions of Breusch, Qian, Schmidt and Wyhowski (*Journal of Econometrics*, 1999).

Keywords

GMM Estimation, Moment Conditions, Redundancy, Partial Redundancy

1. Introduction

In their seminal paper, [1] derived, respectively, the necessary and sufficient condition for the full and partial redundancy of an extra set of moment conditions, given an original set of moment conditions, for the efficient generalized method of moments (GMM) estimation of parameters of interest. However, they treated full and partial redundancy of moment conditions as independent conditions and derived them separately.

In this paper, we will extend their redundancy concept to estimation of an arbitrary linear transformation of an original parameter vector and derive the necessary and sufficient condition for an extra set of moment conditions to be redundant, given an initial set of moment conditions, for the efficient estimation of the linear transformation of parameters. More specifically, the current paper makes three new contributions. Firstly, we extend the moment redundancy concept of [1] to estimation of an arbitrary linear transformation of original parameters. Secondly, the redundancy condition derived in the current paper unifies the full and partial redundancy conditions of [1]. As a result, our redundancy condition given in the theorem of Section 2 includes as two special cases the full and partial redundancy conditions of [1]. Lastly, we use a much simpler approach based on matrix ranks to deriving our main results instead of using “brute-force” matrix algebra, as adopted by [1] and [2].

The rest of the paper is organized as follows. Section 2 presents the main results, while Section 3 briefly concludes.

*Corresponding author.

2.Redundancy of Moment Conditions for Linear Transformation of Parameters

Let $\hat{\theta}$ be the optimal GMM estimator of θ_0 , based on moment conditions:

$$E[g_1(w_t, \theta_0)] = 0, \quad t=1, 2, \dots, T, \quad (1)$$

where w_t is a vector of observable variables, θ_0 is a $p \times 1$ vector of unknown parameters to be estimated, $g_1(w_t, \theta)$ is an $m \times 1$ vector of moment functions, and T is the sample size. For identification purpose, we assume $m \geq p$. We also assume that the vector of moment functions $g_1(w_t, \theta)$ satisfies the usual regularity conditions; for a list of usual regularity conditions, see for example [3]. As adopted by [1] and for simplicity of the derivation below, we make three usual assumptions, as follows.

Assumptions: (A.1) The sample $\{w_t : t = 1, 2, \dots, T\}$ is independent and identically distributed.

(A.2) $\Omega_{11} \equiv \text{var}[g_1(w_t, \theta_0)]$ is positive-definite.

(A.3) $D_1 \equiv E[\partial g_1(w_t, \theta_0) / \partial \theta']$ has full column rank.

Under these assumptions, θ_0 is identified and the optimal GMM estimator $\hat{\theta}$ of θ_0 is consistent and asymptotically normal, with its asymptotic variance matrix equal to $AV(\hat{\theta}) = (D_1' \Omega_{11}^{-1} D_1)^{-1}$; see for example [4].

Now, suppose that we have available an extra set of moment conditions:

$$E[g_2(w_t, \theta_0)] = 0, \quad t = 1, 2, \dots, T, \quad (2)$$

where $g_2(w_t, \theta_0)$ is an $n \times 1$ vector of moment functions satisfying usual regularity conditions. Then, it is well-known in the GMM literature that the optimal GMM estimator of θ_0 using both sets of moment conditions (1) and (2) is usually asymptotically more efficient than the optimal GMM estimator using moment conditions (1) alone. However, there are circumstances when adding an extra set of moment conditions (2) to the existing set of moment conditions (1) does not improve the asymptotic efficiency of the resultant GMM estimator of θ_0 . This is the so-called (full) redundancy of moment conditions (2), given moment conditions (1), for the efficient estimation of θ_0 ; [1] derived the necessary and sufficient condition for the set of moment conditions (2) to be (fully) redundant given the set of moment conditions (1). There are also circumstances when adding the second set of moment conditions (2) onto the first set of moment conditions (1) does not improve the asymptotic efficiency of the resultant GMM estimator of a sub-vector of θ_0 . This is the so-called partial redundancy of moment conditions (2), given moment conditions (1), for the efficient estimation of a sub-vector of θ_0 ; [1] used “brute-force” matrix algebra to derive the necessary and sufficient condition for the partial redundancy of an extra set of moment conditions, while [5] used moment projection approach to deriving the same partial-redundancy condition.

One main weakness of [1] is that it treats the full and partial redundancy of moment conditions as independent of each other and derives them separately. In what follows, we will extend the moment redundancy concept of [5] to the estimation of an arbitrary linear transformation of an original parameter vector and derive the necessary and sufficient condition for an extra set of moment conditions to be redundant, given an initial set of moment conditions, for the efficient estimation of the linear transformation of original parameters. The necessary and sufficient condition given in the theorem below unifies the full and partial redundancy of moment conditions.

More specifically, suppose that we are interested in the estimation of a linear transformation of the original parameter vector θ_0 :

$$\delta_0 = A\theta_0, \quad (3)$$

where A is a $q \times p$ matrix of known constants, with full row rank. Then, since $\hat{\theta}$ is the optimal GMM estimator of θ_0 based on moment conditions (1), it is easy to verify that $\hat{\delta} = A\hat{\theta}$ is a consistent and efficient estimator of δ_0 , with its asymptotic variance equal to:

$$AV(\hat{\delta}) = A(D_1' \Omega_{11}^{-1} D_1)^{-1} A'. \quad (4)$$

Now, let $\tilde{\theta}$ be the optimal GMM estimator of θ_0 , based on the joint moment conditions (1) and (2). [1] show that the optimal GMM estimator $\tilde{\theta}$ based on moment conditions (1) and (2) is equivalent to the optimal

GMM estimator based on the following set of orthogonalized moment conditions:

$$\mathbb{E}\begin{pmatrix} \mathbf{g}_1(\mathbf{w}_t, \theta_0) \\ \mathbf{r}_2(\mathbf{w}_t, \theta_0) \end{pmatrix} = \mathbf{0}, \quad (5)$$

where $\mathbf{r}_2(\mathbf{w}_t, \theta) \equiv \mathbf{g}_2(\mathbf{w}_t, \theta) - \Omega_{21}\Omega_{11}^{-1}\mathbf{g}_1(\mathbf{w}_t, \theta)$, with $\Omega_{21} \equiv \mathbb{E}[\mathbf{g}_2(\mathbf{w}_t, \theta_0)\mathbf{g}_1(\mathbf{w}_t, \theta_0)']$. As a result, we can express the asymptotic variance of $\tilde{\theta}$ as:

$$\text{AV}(\tilde{\theta}) = (\mathbf{G}'\Sigma^{-1}\mathbf{G})^{-1}, \quad (6)$$

where

$$\mathbf{G} = \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{G}_2 \end{bmatrix} \equiv \mathbb{E}\begin{pmatrix} \partial\mathbf{g}_1(\mathbf{w}_t, \theta_0)/\partial\theta' \\ \partial\mathbf{r}_2(\mathbf{w}_t, \theta_0)/\partial\theta' \end{pmatrix} = \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 - \Omega_{21}\Omega_{11}^{-1}\mathbf{D}_1 \end{bmatrix}, \quad (7A)$$

$$\Sigma \equiv \text{var}\begin{pmatrix} \mathbf{g}_1(\mathbf{w}_t, \theta_0) \\ \mathbf{r}_2(\mathbf{w}_t, \theta_0) \end{pmatrix} = \begin{bmatrix} \Omega_{11} & \mathbf{0} \\ \mathbf{0} & \Sigma_{22} \end{bmatrix}, \quad (7B)$$

and where $\mathbf{D}_2 \equiv \mathbb{E}[\partial\mathbf{g}_2(\mathbf{w}_t, \theta_0)/\partial\theta']$ and $\Sigma_{22} \equiv \Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12}$.

Now, given $\tilde{\theta}$, we can also estimate δ_0 by $\tilde{\delta} = \mathbf{A}\tilde{\theta}$. Then, it is easy to verify that $\tilde{\delta}$ is a consistent estimator of δ_0 and its asymptotic variance is equal to:

$$\text{AV}(\tilde{\delta}) = \mathbf{A}(\mathbf{G}'\Sigma^{-1}\mathbf{G})^{-1}\mathbf{A}'. \quad (8)$$

Since $\tilde{\theta}$ is asymptotically at least as efficient as $\hat{\theta}$, it is easy to verify that $\tilde{\delta}$ is also asymptotically no less efficient than $\hat{\delta}$. Then, an interesting question is when $\hat{\delta}$ is asymptotically as efficient as $\tilde{\delta}$, or equivalently, under what circumstances will adding the extra set of moment conditions (2) to moment conditions (1) not improve the asymptotic efficiency of the GMM estimator of the transformed parameter vector, $\delta_0 = \mathbf{A}\theta_0$? When $\text{AV}(\hat{\delta}) = \text{AV}(\tilde{\delta})$, we will say that the extra set of moment conditions (2) is redundant, given the original set of moment conditions (1), for the (efficient) estimation of the transformed parameter vector δ_0 .

We now proceed to find the necessary and sufficient condition for $\text{AV}(\hat{\delta}) = \text{AV}(\tilde{\delta})$. To this end, we first state a well-known rank formula for partitioned-matrices; see for example Masarglia and [6]).

Lemma. Let \mathbf{A} be a nonsingular matrix of order r_1 , \mathbf{B} , \mathbf{C} and \mathbf{D} be $r_1 \times c_2$, $r_2 \times r_1$, and $r_2 \times c_2$ matrices, respectively. Then,

$$\text{rk}\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \text{rk}(\mathbf{A}) + \text{rk}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}), \quad (9A)$$

or equivalently,

$$\text{rk}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}) = \text{rk}\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} - \text{rk}(\mathbf{A}). \quad (9B)$$

Then, using (9B), we have:

$$\begin{aligned} & \text{rk}[\text{AV}(\hat{\delta}) - \text{AV}(\tilde{\delta})] \\ &= \text{rk}[\mathbf{A}(\mathbf{D}_1'\Omega_{11}^{-1}\mathbf{D}_1)^{-1}\mathbf{A}' - \mathbf{A}(\mathbf{G}'\Sigma^{-1}\mathbf{G})^{-1}\mathbf{A}'] \\ &= \text{rk}\begin{pmatrix} \mathbf{G}'\Sigma^{-1}\mathbf{G} & \mathbf{A}' \\ \mathbf{A} & \mathbf{A}(\mathbf{D}_1'\Omega_{11}^{-1}\mathbf{D}_1)^{-1}\mathbf{A}' \end{pmatrix} - \text{rk}(\mathbf{G}'\Sigma^{-1}\mathbf{G}) \end{aligned} \quad (10A)$$

$$= \text{rk} \left(\begin{bmatrix} \mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1 + \mathbf{G}_2' \boldsymbol{\Sigma}_{22}^{-1} \mathbf{G}_2 & \mathbf{A}' \\ \mathbf{A} & \mathbf{A}(\mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1)^{-1} \mathbf{A}' \end{bmatrix} \right) - p, \quad (10B)$$

Using $\mathbf{G}' \boldsymbol{\Sigma}^{-1} \mathbf{G} = \mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1 + \mathbf{G}_2' \boldsymbol{\Sigma}_{22}^{-1} \mathbf{G}_2$ from (7A)-(7B) and $\text{rk}(\mathbf{G}' \boldsymbol{\Sigma}^{-1} \mathbf{G}) = p$. Note that,

$$\begin{aligned} & \begin{bmatrix} \mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1 + \mathbf{G}_2' \boldsymbol{\Sigma}_{22}^{-1} \mathbf{G}_2 & \mathbf{A}' \\ \mathbf{A} & \mathbf{A}(\mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1)^{-1} \mathbf{A}' \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{G}_2' & \mathbf{I}_p \\ 0 & \mathbf{A}(\mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1)^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{22}^{-1} & 0 \\ 0 & \mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1 \end{bmatrix} \begin{bmatrix} \mathbf{G}_2 & 0 \\ \mathbf{I}_p & (\mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1)^{-1} \mathbf{A}' \end{bmatrix}. \end{aligned}$$

Substituting it into (10B), we have:

$$\begin{aligned} & \text{rk}[\mathbf{A}\mathbf{V}(\hat{\delta}) - \mathbf{A}\mathbf{V}(\tilde{\delta})] \quad (11A) \\ &= \text{rk} \left(\begin{bmatrix} \mathbf{G}_2' & \mathbf{I}_p \\ 0 & \mathbf{A}(\mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1)^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{22}^{-1} & 0 \\ 0 & \mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1 \end{bmatrix} \begin{bmatrix} \mathbf{G}_2 & 0 \\ \mathbf{I}_p & (\mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1)^{-1} \mathbf{A}' \end{bmatrix} \right) - p \\ &= \text{rk} \left(\begin{bmatrix} \mathbf{G}_2 & 0 \\ \mathbf{I}_p & (\mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1)^{-1} \mathbf{A}' \end{bmatrix} \right) - p \end{aligned}$$

(using $\text{rk}(\mathbf{C}'\mathbf{B}\mathbf{C}) = \text{rk}(\mathbf{C})$, with \mathbf{B} positive definite)

$$\begin{aligned} &= \text{rk} \left(\begin{bmatrix} \mathbf{I}_p & (\mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1)^{-1} \mathbf{A}' \\ \mathbf{G}_2 & 0 \end{bmatrix} \right) - p \quad (\text{exchanging rows}) \\ &= \text{rk}(\mathbf{I}_p) + \text{rk}(0 - \mathbf{G}_2 \mathbf{I}_p^{-1} (\mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1)^{-1} \mathbf{A}') - p \quad (\text{using (9A)}) \\ &= \text{rk}(\mathbf{G}_2 (\mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1)^{-1} \mathbf{A}'). \quad (11B) \end{aligned}$$

Now, given the equality of expressions (11A) and (11B), we are ready to establish the main result of the paper.

THEOREM. The extra set of moment conditions (2) is redundant, given the set of moment conditions (1), for the efficient estimation of the linear transformation of $\delta_0 = \mathbf{A}\theta_0$, if and only if $\mathbf{G}_2 (\mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1)^{-1} \mathbf{A}' = 0$.

Proof: By the definition, moment conditions (2) is redundant, given moment conditions (1), for the efficient estimation of $\delta_0 = \mathbf{A}\theta_0$, if and only if $\mathbf{A}\mathbf{V}(\hat{\theta}) = \mathbf{A}\mathbf{V}(\tilde{\theta})$, or equivalently, if and only if $\text{rk}[\mathbf{A}\mathbf{V}(\hat{\theta}) - \mathbf{A}\mathbf{V}(\tilde{\theta})] = 0$ (using the fact that the rank of a matrix is zero if and only of the matrix itself is a zero matrix). Then, by using the equality of expressions (11A) and (11B), the result of this theorem follows immediately. ■

Given this theorem, we can now easily show that the full and partial redundancy conditions obtained by Breusch *et al.* (1999) are just two special cases of it.

COROLLARY 1. When the transformation matrix \mathbf{A} is nonsingular, the extra set of moment conditions (2) is redundant, given the set of moment conditions (1), for the efficient estimation of $\delta_0 = \mathbf{A}\theta_0$, if and only if $\mathbf{G}_2 = 0$, or equivalently, $\mathbf{D}_2 = \boldsymbol{\Omega}_{21} \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1$.

Proof: When \mathbf{A} is nonsingular, the redundancy condition $\mathbf{G}_2 (\mathbf{D}_1' \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1)^{-1} \mathbf{A}' = 0$ of the theorem above is equivalent to $\mathbf{G}_2 = 0$. Also, by the definition of $\mathbf{G}_2 = \mathbf{D}_2 - \boldsymbol{\Omega}_{21} \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1$ in (7A), $\mathbf{G}_2 = 0$ is the same as $\mathbf{D}_2 = \boldsymbol{\Omega}_{21} \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1$. ■

A special case of Corollary 1 is when the transformation matrix \mathbf{A} is an identity matrix; that is, $\delta_0 = \theta_0$. Then, the condition of $\mathbf{D}_2 = \boldsymbol{\Omega}_{21} \boldsymbol{\Omega}_{11}^{-1} \mathbf{D}_1$ is just the full redundancy condition of Theorem 1 of Breusch *et al.* (1999, p. 94). Thus, the full redundancy condition of [1] is just a special case of the necessary and sufficient con-

dition of the theorem above.

We now turn to using the theorem above to derive the partial redundancy of moment conditions (2), given moment conditions (1), for the estimation of a sub-vector of θ_0 . For this purpose, we partition the parameter vector θ_0 into $\theta_0 = (\theta_{10}', \theta_{20}')$, where θ_{10} is $p_1 \times 1$ and θ_{20} is $p_2 \times 1$, with $p_1 + p_2 = p$. We also partition the expected derivative matrices accordingly:

$$D_1 = (D_{11}, D_{12}), \text{ with } D_{1j} = E[\partial g_1(w_t, \theta_0) / \partial \theta_j'] \text{ for } j = 1, 2 \quad (12A)$$

$$G_2 = (G_{21}, G_{22}), \text{ with } G_{2j} = E[\partial r_2(w_t, \theta_0) / \partial \theta_j'], \text{ for } j = 1, 2. \quad (12B)$$

Without loss of generality, suppose that we are now mainly interested in estimating the first subset of parameters, θ_{10} . That is, we wish to efficiently estimate $\delta_0 = A\theta_0 = \theta_{10}$, with the transformation matrix A defined as $A = [I_{p_1}, 0_{p_1 \times p_2}]$. Now, substituting (12A)-(12B) and $A = [I_{p_1}, 0_{p_1 \times p_2}]$ into the redundancy condition of the theorem above, we have:

$$\begin{aligned} G_2(D_1' \Omega_{11}^{-1} D_1)^{-1} A' &= (G_{21}, G_{22}) \begin{bmatrix} D_{11}' \Omega_{11}^{-1} D_{11} & D_{11}' \Omega_{11}^{-1} D_{12} \\ D_{12}' \Omega_{11}^{-1} D_{11} & D_{12}' \Omega_{11}^{-1} D_{12} \end{bmatrix}^{-1} \begin{bmatrix} I_{p_1} \\ 0 \end{bmatrix} \\ &= (G_{21}, G_{22}) \begin{bmatrix} E^{-1} \\ - (D_{12}' \Omega_{11}^{-1} D_{12})^{-1} D_{12}' \Omega_{11}^{-1} D_{11} E^{-1} \end{bmatrix} \\ &= [G_{21} - G_{22} (D_{12}' \Omega_{11}^{-1} D_{12})^{-1} D_{12}' \Omega_{11}^{-1} D_{11}] E^{-1}, \end{aligned}$$

where we used the formula for partitioned-matrix inverse in the second equation, with

$E \equiv D_{11}' \Omega_{11}^{-1} D_{11} - D_{11}' \Omega_{11}^{-1} D_{12} (D_{12}' \Omega_{11}^{-1} D_{12})^{-1} D_{12}' \Omega_{11}^{-1} D_{11}$. That is, we establish:

$$G_2(D_1' \Omega_{11}^{-1} D_1)^{-1} A' = \quad (13)$$

Now, combining this expression with the theorem above, we obtain Corollary 2.

COROLLARY 2. The extra set of moment conditions (2) is partially redundant, given the set of moment conditions (1), for the efficient estimation of θ_{10} , if and only if:

(A) $G_2(D_1' \Omega_{11}^{-1} D_1)^{-1} A' = 0$, with $A = [I_{p_1}, 0_{p_1 \times p_2}]$, or equivalently,

(B) $G_{21} = G_{22} (D_{12}' \Omega_{11}^{-1} D_{12})^{-1} D_{12}' \Omega_{11}^{-1} D_{11}$.

Proof: Condition (A) is just the redundancy condition of the theorem above, with the special transformation-matrix $A = [I_{p_1}, 0_{p_1 \times p_2}]$. The equivalence of conditions (A) and (B) follows directly from the equality in (13). ■

Here we note that condition (B) of Corollary 2 is identical to the partial redundancy condition of [1] (Theorem 7) and [2], Theorem 2). Both of these papers used very tedious and brute-force matrix algebra to derive it, while we straightforwardly derived it as a special case of the general redundancy condition for linear transformation of parameters.

Before we conclude this paper, we want to briefly compare the redundancy conditions for transformed parameters of the current paper with the redundancy conditions in restricted GMM estimation of [5]. More specifically, [5] considers efficient GMM estimation of θ_0 based on moment conditions (1), subject to a possibly nonlinear set of q restrictions, $r(\theta_0) = 0$. [5] Theorem 1) shows that the extra set of moment conditions (2) is redundant, given moment conditions (1), for the efficient estimation of θ_0 subjected to $r(\theta_0) = 0$, if and only if

$$G_2 M_{[F]} = 0, \quad (14)$$

where $F \equiv \partial r(\theta_0) / \partial \theta'$. Comparing (14) with $G_2(D_1' \Omega_{11}^{-1} D_1)^{-1} A' = 0$ (the redundancy condition for the linear transformation of parameters of the current paper), we can easily verify that they are the same as $G_2 = 0$ (the full redundancy condition of [1], when A is nonsingular and F is defined as the $q \times p$ zero matrix for no restrictions. However, we want to emphasize that the redundancy condition for restricted GMM esti-

mation, given in Theorem 1 of [5], is substantially different from the redundancy condition for transformed parameters, given in the theorem of the current paper. To see the difference between them, let's consider a special case, as follows. Let $\theta_0 = (\theta_{10}', \theta_{20}')$, as defined before. Now, suppose that we are only interested in estimating θ_{10} (that is, the transformation matrix is $A = [I_{p_1}, 0_{p_1 \times p_2}]$, as we considered in Corollary 2 above) and that the set of restrictions in the restricted GMM estimation is given by $r(\theta_0) = \theta_{20} = 0$ (that is, θ_{20} is known). For this special case, we can easily verify that the redundancy condition (14) for the restricted GMM estimation of θ_{10} becomes $G_{21} = 0$, which is actually the full redundancy condition of moment conditions (2), given moment conditions (1), for the efficient estimation of θ_{10} , because θ_{20} is known. On the other hand, the redundancy condition for the transformed parameter vector $\delta_0 = A\theta_0 = \theta_{10}$ becomes the partial redundancy condition, $G_{21} = G_{22}(D_{12}'\Omega_{11}^{-1}D_{12})^{-1}D_{12}'\Omega_{11}^{-1}D_{11}$, as shown in Corollary 2. Thus, we can interpret the redundancy condition of the current paper, $G_2(D_1'\Omega_{11}^{-1}D_1)^{-1}A' = 0$, as the partial redundancy of moment conditions (2), given moment conditions (1), for the efficient estimation of parameters of interest, $\delta_0 = A\theta_0$, while $G_2M_{[F]} = 0$ (the condition for the redundancy of moment conditions (2), given moment conditions (1), in the restricted GMM estimation considered by [5]), can be interpreted as the full redundancy condition of moment conditions (2), given moment conditions (1), for the efficient estimation of the free parameters (after substituting out the restrictions).

3. Conclusions

In this paper, using the idea of redundancy of moment conditions for estimating a linear transformation of parameters, we successfully unify full and partial redundancy of moment conditions. As a consequence, the necessary and sufficient condition for an extra set of moment conditions to be redundant for the estimation of a linear transformation of original parameters encompasses the full and partial redundancy conditions of [1].

Two possible applications of the results of the current paper are the efficient estimation of parameters of interest in panel data models and systems of equations. [7] (Section 8.4.2), for example, compares the relative efficiency of GMM, generalized instrumental variables (GIV) and the traditional 3SLS estimators of the whole coefficient vector of a system of linear regressions. Using the results of the current paper and appropriately defined sets of moment conditions, we could find useful conditions for GIV and the traditional 3SLS estimators of a sub-vector of regression coefficients (e.g. the coefficient vector of the first equation in a system) to be as efficient as the optimal GMM estimator applied to the entire system. This is a topic for future research.

References

- [1] Breusch, T., Schmidt, P., Qian, H. and Wyhowski, D. (1999) Redundancy of Moment Conditions. *Journal of Econometrics*, **91**, 89-111. [http://dx.doi.org/10.1016/S0304-4076\(98\)00050-5](http://dx.doi.org/10.1016/S0304-4076(98)00050-5)
- [2] Qian, H. (2002) Partial Redundancy of Moment Conditions. *Econometric Theory*, **18**, 531-539. <http://dx.doi.org/10.1017/S0266466602182132>
- [3] Newey, W. and McFadden, D. (1994) Large Sample Estimation and Hypothesis Testing. In: Engle, R. and McFadden, D., Eds., *Handbook of Econometrics*, Elsevier, Berlin, 2111-2245.
- [4] Hansen, L.P. (1982) Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, **50**, 1029-1054. <http://dx.doi.org/10.2307/1912775>
- [5] Qian, H. (2013) Redundancy of Moment Conditions in Restricted GMM Estimation. Invited for Revision by *Econometric Theory*; Being Revised.
- [6] Marsaglia, G. and Styan, G. (1974) Equalities and Inequalities for Ranks of Matrices. *Linear and Multilinear Algebra*, **2**, 269-292.
- [7] Wooldridge, J. (2010) *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge.