Economics Bulletin

Volume 35, Issue 1

Partial efficient estimation of SUR models

Hailong Qian
Saint Louis University

Heather L. Bednarek
Saint Louis University

Abstract

In this paper, we consider efficient estimation of coefficients of interest in seemingly unrelated regressions (SUR) models. Using the GMM interpretation of the usual OLS and GLS/FGLS estimation of regression coefficients in SUR models, we derive the necessary and sufficient condition for the equal asymptotic efficiency of the OLS and FGLS estimators of a subset of regression coefficients. As a result, our paper extends the current SUR literature on the numerical equality of the OLS and GLS/FGLS estimators of the whole coefficient vector (see for example, Dwivedi and Srivastava, 1978) to the asymptotic equivalence of the OLS and GLS/FGLS estimators of a subset of the coefficient vector.

Acknowledgement: The authors are very grateful to the Editor and two anonymous referees for their very constructive suggestions that significantly improved the quality of the paper.
Citation: Hailong Qian and Heather L. Bednarek, (2015) "Partial efficient estimation of SUR models", *Economics Bulletin*, Volume 35, Issue 1, pages 338-348
Contact: Hailong Qian - qianh@slu.edu, Heather L. Bednarek - bednarhl@slu.edu.
Submitted: June 25, 2014. **Published:** March 11, 2015.

1. Introduction

Since Zellner's (1962) seminal paper introducing the seemingly unrelated regressions (SUR) model, SUR models have found a wide range of applications. For examples, SUR models have been applied to consumer demands (e.g. Bewley (1986)), regional economics (e.g. LeGallo and Chasco (2008)), public economics (e.g. Gebremariam, *et al.* (2012)), urban economics (e.g. Baltagi and Bresson (2011)), finance (e.g. Hodgson, *et al.* (2002)) and international trade (e.g. Moon and Perron (2004)). The basic SUR model has been extended to cases with heteroscedasticity and serial correlation in the error terms (e.g. Creel and Farell (1996)). It has also been applied to panel data models with individual effects (e.g. Baltagi (1980)). SUR models have been further applied to nonstationary VAR models for testing panel unit roots and cointegration with cross-sectional correlations; see for examples, Moon (1999), Chang (2004), Moon and Perron (2004) and Mark, *et al.* (2005). Lastly, the newest research development in SUR models is concerned with applications to cross-sectional or panel datasets that have both temporal and spatial correlations; see for examples, Anselin (1988), Lundberg (2006), Zhou and Kockelman (2009), Baltagi and Pirotte (2011) and Gebremariam, *et al.* (2012). For more complete references of the recent SUR literature, interested readers are further referred to two excellent surveys by Fiebig (2001) and Moon and Perron (2006).

It is also well-known in the SUR literature that the generalized least squares (GLS) or feasible GLS (FGLS) estimation applied to the whole system of SUR is asymptotically more efficient than the equation-by-equation ordinary least squares (OLS) estimation; see for examples Zellner (1962) and Wooldridge (2010, Chapter 7). However, Zellner (1962, p. 351) shows that when the covariance matrix is diagonal or each equation in the system of SUR includes the same explanatory variables, GLS/FGLS and OLS are numerically identical. Motivated by Zellner's sufficient conditions and using the optimality condition for the OLS estimator in general linear models (see for examples, Zyskind (1967), Kruskal (1968), Rao (1973) and Gourieroux and Monfort (1980)), several subsequent papers derive different (but equivalent) forms of necessary and sufficient conditions for the numerical equality of OLS and GLS/FGLS estimators of SUR models; see for examples, Dwivedi and Srivastava (1978), Gourieroux and Monfort (1980), Baltagi (1988), Baksalary and Trenkler (1989), and an excellent monograph on SUR models by Srivastava and Giles (1987). Revanker (1974) and Schmidt (1978) show that for the estimation of the regression coefficients of the overidentified equations, it is numerically identical whether GLS/FGLS estimation is applied to the whole system of SUR or just to the sub-system of overidentified equations.[1] Gourieroux and Monfort (1980) also establishes the necessary and sufficient condition for the numerical equivalence of OLS and GLS estimators of a subset of regression coefficients of SUR models.

Using the GMM interpretation of OLS and GLS/FGLS estimators of SUR models, Qian (2008) derives the necessary and sufficient condition for the equal asymptotic efficiency of OLS and GLS/FGLS estimation of the whole system of SUR. Surprisingly, it appears that no paper has so far considered when the equation-by-equation OLS estimation of a subset of regression coefficients in a general SUR model is asymptotically as efficient as the GLS/FGLS estimation applied to the whole system of SUR. In this paper, we seek to fill this gap. More precisely, using the GMM interpretation of OLS and GLS/FGLS estimation of SUR models and the partial redundancy of moment conditions of Breusch, *et al.* (1999), we will establish the necessary and sufficient condition for the OLS estimator of a subset of regression coefficients of a general SUR model to be asymptotically efficient. To be more specific, suppose that we have a system of household demand equations for durable goods and nondurable goods and that our main interest is in the estimation of the parameters in the demand equations for nondurable goods. Then the current paper seeks to find the necessary and sufficient condition for the equation-by-equation OLS estimation of the regression coefficients in the demand equations for nondurable goods to be asymptotically as efficient as the FGLS estimation applied to the whole system of demand equations. Thus, the current paper extends the full asymptotic efficiency result of Qian (2008) to the partial asymptotic efficiency of OLS estimation of SUR models. Here we want to point out the main difference between this paper and the current literature on efficient estimation of SUR models. Our paper is concerned with the comparison of asymptotic efficiency of OLS and FGLS estimators of a subset of regression coefficients of SUR models, while almost all of the published papers on efficient estimation of SUR models, with the exception of Ravankar (1974), Schmidt (1978) and Gourieroux and Monfort (1980),[2] focused on the numerical equality of OLS and GLS estimators of the whole vector of regression coefficients. As such, our necessary and sufficient condition for the partial asymptotic efficiency of the OLS estimator (Theorem 2 in the next section) generalizes various sufficient (and necessary in some cases) conditions for the numeral equality of OLS and GLS/FGLS estimators of SUR models.

The rest of the paper is organized as follows. Section 2 presents the main results, while Section 3 contains a small Monte Carlo simulation. Section 4 briefly concludes.

2. Partial Efficiency of OLS Estimation of SUR Models

Consider a system of G seemingly unrelated regressions:

$$y_{gt} = x_{gt}'\beta_g + \varepsilon_{gt},\ \ g = 1, 2, \ldots, G;\ \ t = 1, 2, \ldots, T, \tag{1}$$

where the subscripts g and t index equations and observations, respectively, y_{gt} is the dependent variable of the g-th equation, x_{gt} is a $k_g \times 1$ vector of explanatory variables, β_g is a $k_g \times 1$ vector of unknown regression coefficients, ε_{gt} is the disturbance term, and T is the sample size.

Stacking over equations for a given observation t, we can rewrite (1) as:

$$y_t = X_t'\beta + \varepsilon_t,\ \ t = 1, 2, \ldots, T, \tag{2}$$

where $y_t = (y_{1t},\ldots,y_{Gt})'$, $X_t = \mathrm{diag}(x_{1t},\cdots,x_{Gt})$, $\beta = (\beta_1',\ldots,\beta_G')'$ and $\varepsilon_t = (\varepsilon_{1t},\ldots,\varepsilon_{Gt})'$. We assume that $E(X_tX_t')$ is nonsingular so that the system of SUR in (1) is identified.

Let x_t be an $M \times 1$ vector of distinct explanatory variables appearing in system (1). Now, to be consistent with the SUR literature (see for example, chapter 7 of Wooldridge (2010)), we make the following two standard assumptions for the system of SUR in (1):

Assumptions: (SUR.1) $E(\varepsilon_t \otimes x_t) = 0$, where $\otimes$ denotes the Kronecker product.
(SUR.2) $E[(I_G \otimes x_t)\varepsilon_t\varepsilon_t'(I_G \otimes x_t')] = E[(I_G \otimes x_t)\Sigma(I_G \otimes x_t')]$, with $\Sigma \equiv E(\varepsilon_t\varepsilon_t')$.

In the rest of the paper, we will assume that the covariance matrix Σ is positive definite (*p.d.*). Because the covariance matrix is rarely known in applications, we only consider FGLS estimation of (1) in this paper. Of course, when Σ is known or diagonal, the result obtained in this paper also applies to GLS. Also, for simplicity of derivation, we assume that $(y_t', x_t')'$ is independent and identically distributed over observations.

To define the moment conditions implied by Assumption (SUR.1), we further define x_{gt}^* as a $k_g^* \times 1$ (with $k_g^* = M - k_g$) vector of distinct regressors appearing in system (1) other than the g-th equation. For example, if G=2, $x_{1t} = (w_{1t}', w_{2t}')'$, $x_{2t} = (w_{2t}', w_{3t}')'$ and there are no common elements in w_{1t}, w_{2t} and w_{3t}, then $x_t = (w_{1t}', w_{2t}', w_{3t}')'$, $x_{1t}^* = w_{3t}$ and $x_{2t}^* = w_{1t}$. Also, when $x_{gt} = x_t$, we define $x_{gt}^* = 0$.

Then, as explained in Qian (2008), Assumption (SUR.1) implies the following two sets of moment conditions:

$$E[g_{1t}(\beta)] \equiv E[X_t(y_t - X_t'\beta)] = E(\begin{bmatrix} x_{1t}(y_{1t} - x_{1t}'\beta_1) \\ \vdots \\ x_{Gt}(y_{Gt} - x_{Gt}'\beta_G) \end{bmatrix}) = 0 \qquad \text{(MC-W)}$$

$$E[g_{2t}(\beta)] \equiv E[X_t^*(y_t - X_t'\beta)] = E(\begin{bmatrix} x_{1t}^*(y_{1t} - x_{1t}'\beta_1) \\ \vdots \\ x_{Gt}^*(y_{Gt} - x_{Gt}'\beta_G) \end{bmatrix}) = 0, \qquad \text{(MC-C)}$$

where X_t^* is the block-diagonal matrix with x_{gt}^* on the g-th block. Hereafter, we will refer to (MC-W) and (MC-C) as the within- and cross-equation moment (or orthogonality) conditions, respectively.

Let $\hat{\beta}$ and $\tilde{\beta}$ be the equation-by-equation OLS and the FGLS estimators of β in (1), respectively. Now, notice that the equation-by-equation OLS estimator of β is just the (optimal) GMM estimator based on (MC-W), and the FGLS estimator of β is asymptotically equivalent to the optimal GMM estimator based on the joint moment conditions (MC-W) and (MC-C); see, for example, Theorem 8.4 of Wooldridge (2010, p. 221). Thus, to show that the OLS and FGLS estimators of β in (1) have the same asymptotic efficiency is equivalent to showing that the GMM estimator of β using the moment conditions in (MC-W) has the same asymptotic efficiency as the optimal GMM estimator of β using the moment conditions in both (MC-W) and (MC-C). This is equivalent to showing that the cross-equation orthogonality conditions (MC-C) is redundant (in the sense of Breusch, *et al.* (1999)) given the within-equation orthogonality conditions (MC-W) for the efficient estimation of β. In fact, Qian (2008) derives the necessary and sufficient condition for the asymptotic optimality of the equation-by-equation OLS estimator of the whole parameter vector β in the system of SUR model (1). To facilitate comparison with the new result of this paper (Theorem 2 below), we summarize the main result of Qian (2008) in Theorem 1.

Theorem 1. The equation-by-equation OLS estimator of β in the SUR model (1) is asymptotically as efficient as the FGLS estimator of β, if and only if:

$$\sigma_{ij} E(x_{it}^* r_{jit}') = 0 \text{, for i, j = 1, 2, ..., G, and } i \neq j, \qquad (3)$$

where σ_{ij} is the (i, j)-element of the covariance matrix (Σ) and $r_{jit} \equiv x_{jt} - E(x_{jt}x_{it}')[E(x_{it}x_{it}')]^{-1}x_{it}$ is the population linear projection error of x_{jt} on x_{it}.

Proof: See the theorem of Qian (2008, p. 1458). ■

When the covariance matrix is diagonal or each regression equation in the system of SUR includes the same set of explanatory variables, condition (3) is obviously satisfied. When the regressors in the SUR are treated as non-stochastic, we can easily verify that various necessary and sufficient conditions for the numerical equality of OLS and GLS estimators of regression coefficients in SUR models are sufficient but not necessary for (3); see for examples, Gourieroux and Monfort's (1980, p. 1088), Baltagi (1988, 1989), Baksalary and Trenkler (1989) and Bartels and Fiebig (1991).

We now turn to deriving the necessary and sufficient condition for the equal asymptotic efficiency of OLS and FGLS estimation for a subset of regression coefficients in (1). For this purpose and without loss of generality, suppose that we are now only interested in estimating the regression coefficients of the first g ($1 \leq g < G$) equations in (1). We thus further define:

$$\theta_1 \equiv (\beta_1', \cdots, \beta_g')',$$

$$X_{1t} \equiv \begin{bmatrix} x_{1t} & & \\ & \ddots & \\ & & x_{gt} \end{bmatrix}, \qquad X_{2t} \equiv \begin{bmatrix} x_{g+1,t} & & \\ & \ddots & \\ & & x_{Gt} \end{bmatrix},$$

$$X_{1t}^* \equiv \begin{bmatrix} x_{1t}^* & & \\ & \ddots & \\ & & x_{gt}^* \end{bmatrix}, \qquad X_{2t}^* \equiv \begin{bmatrix} x_{g+1,t}^* & & \\ & \ddots & \\ & & x_{Gt}^* \end{bmatrix}.$$

We also partition the covariance matrix Σ as

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix},$$

with Σ_{11} $g \times g$ and Σ_{22} $(G-g) \times (G-g)$. Given these notations and the partial redundancy condition of Breusch *et al.* (1999, Theorem 7) or Qian (2002, Theorem 2), we obtain Theorem 2.

Theorem 2. Let $\hat{\beta}$ and $\tilde{\beta}$ be, respectively, the OLS and FGLS estimators of β in the system of SUR model (1). Then, under assumptions (SUR.1)-(SUR.2), the OLS estimator $\hat{\theta}_1 = (\hat{\beta}_1', \cdots, \hat{\beta}_g')'$ of $\theta_1 = (\beta_1', \cdots, \beta_g')'$ is asymptotically as efficient as the FGLS estimator $\tilde{\theta}_1 = (\tilde{\beta}_1', \cdots, \tilde{\beta}_g')'$, if and only if the following two conditions hold:

$$E(X_{1t}^* X_{1t}') = E(X_{1t}^* \Sigma_{11} X_{1t}')[E(X_{1t} \Sigma_{11} X_{1t}')]^{-1} E(X_{1t} X_{1t}'), \tag{4}$$

$$E(X_{2t}^* \Sigma_{21} X_{1t}') = E(X_{2t}^* X_{2t}')[E(X_{2t} X_{2t}')]^{-1} E(X_{2t} \Sigma_{21} X_{1t}'). \tag{5}$$

These two conditions are also equivalent to

$$\sigma_{ij} E(x_{it}^* r_{jit}') = 0, \text{ for } i = 1, 2, ..., G,\ j = 1, 2, ..., g, \text{ and } i \neq j, \tag{6}$$

where σ_{ij} is the covariance between the disturbances of equations i and j, and $r_{jit} \equiv x_{jt} - E(x_{jt} x_{it}')[E(x_{it} x_{it}')]^{-1} x_{it}$ is the population linear projection error of x_{jt} on x_{it}.

Proof: In order to save space, the proof is omitted but is available from the authors on request. ■

Now, comparing Theorems 1 and 2, we can easily see that condition (3) is sufficient but not necessary for condition (6). Thus, Theorem 2 extends the full efficiency of OLS estimation of SUR models to the partial efficiency of OLS estimation for a subset of regression coefficients. Here we also notice that condition (4) only involves the first g equations of the system of SUR, while condition (5) partially depends on the correlation between the first g equations and the remaining (G-g) equations. In fact, using Theorem 1 of Breusch, *et al.* (1999), we can easily verify that condition (4) is just the full redundancy condition of $m_{2t}(\theta_1)$ given $m_{1t}(\theta_1)$ for the estimation of θ_1, where the moment conditions are defined as,

$$E[m_{1t}(\theta_1)] \equiv E(\begin{bmatrix} x_{1t}(y_{1t} - x_{1t}'\beta_1) \\ \vdots \\ x_{gt}(y_{gt} - x_{gt}'\beta_g) \end{bmatrix}) = 0, \; E[m_{2t}(\theta_1)] \equiv E(\begin{bmatrix} x_{1t}^*(y_{1t} - x_{1t}'\beta_1) \\ \vdots \\ x_{gt}^*(y_{gt} - x_{gt}'\beta_g) \end{bmatrix}) = 0.$$

When the regressors in (1) are treated as non-stochastic, we can verify that Gourieroux and Monfort's (1980, p. 1089) necessary and sufficient condition for the numerical equality of OLS and GLS estimators of a subset of regression coefficients in SUR models is sufficient but not necessary for (6). Condition (6) is also related to Schmidt (1978). More specifically, Schmidt (1978) shows that, for the estimation of the regression coefficients in overidentified equations, GLS/FGLS estimation applied to the group of overidentified equations results in the same estimator as GLS/FGLS estimation applied to the whole system. To compare his finding with (6), let's consider a three-equation system and assume that the third equation is just-identified and that we are only interested in estimating the coefficients in the first two equations (that is, g=2). For this case, condition (6) becomes $\sigma_{ij}E(x_{it}^* r_{jit}') = 0$ for i = 1, 2, 3 and j = 1, 2. Now, note that this condition is satisfied when i=3, because when the third equation is just-identified, $x_{3t}^* = 0$ and $r_{j3t} = x_{jt} - L(x_{jt} | x_{3t}) = 0$ for j = 1, 2. Then, by Theorem 2, the OLS estimation of the regression coefficients in the first two equations is asymptotically as efficient as FGLS applied to the whole system of three equations, if and only if $\sigma_{ij}E(x_{it}^* r_{jit}') = 0$ for i, j = 1, 2. This according to Theorem 1 is just the necessary and sufficient condition for OLS of the first two equations to be as efficient as the FGLS applied to the first two equations. Thus, the combination of Theorems 1 and 2 includes Schmidt's (1978) finding as a special case.

Given Theorem 2, we can now provide several useful sufficient conditions for the partial asymptotic optimality of the OLS estimation of SUR models.

Corollary. The equation-by-equation OLS estimator of $(\beta_1', \cdots, \beta_g')'$ in the SUR model (1) is asymptotically as efficient as the FGLS estimator of $(\beta_1', \cdots, \beta_g')'$, if one of the following conditions is true:

(A) The first g regression equations include the same explanatory variables and each of the remaining (G-g) regression equations is just-identified;

(B) The disturbances of the first g-equations are uncorrelated with each other and are also uncorrelated with any of the disturbances of the remaining (G-g) equations;

(C) The first g regression equations include the same explanatory variables and the disturbances of the first g-equations are uncorrelated with any of the disturbances of the remaining (G-g) equations;

(D) The disturbances of the first g-equations are uncorrelated with each other and each of the remaining (G-g) regression equations is just-identified.

Proof: It is easy to verify that any one of these conditions is sufficient for condition (6). ■

Condition (A) is related to Schmidt (1978). More precisely, Schmidt (1978) shows that FGLS applied to the sub-system of overidentified equations is asymptotically as efficient as the FGLS applied to the entire system of SUR. Then, under Condition (A), FGLS applied to the first g equations is asymptotically equivalent to FGLS applied to the entire system (for the estimation of the regression coefficients in the first g equations). Now, if each of the first g equations includes the same explanatory variables, it is well known that the equation-by-equation OLS estimation is algebraically identical to the FGLS applied to the first g equations. Condition (B) says that when the first g equations are uncorrelated with the rest of the equations in the system, then FGLS applied to the first g equations results in asymptotically efficient estimation of the regression coefficients in the first g equations. Now, if the disturbances of the first g equations are further uncorrelated with each other, OLS applied to each of the first g equations becomes asymptotically efficient. Thus, Condition (B) extends one of Zellner's (1962, p. 351) two well-known sufficient conditions for the numerical equality of OLS and GLS estimators of the whole coefficient vector to a subset of coefficients. Conditions (C) and (D) are various mixings of Conditions (A) and (B). Thus, our corollary extends two well-known sufficient conditions for the numerical equality of OLS and GLS estimators of the whole parameter vector to the equal asymptotic efficiency of OLS and FGLS estimators for a subset of parameters.

3. Monte Carlo Simulations

The equivalence of the OLS and FGLS estimators of parameters of interest established in Theorem 2 is an asymptotic result. In this section, we conduct a small Monte Carlo simulation to compare their finite sample performances when the condition of Theorem 2 is satisfied. For this purpose, we consider the following three-equation system of SUR:[3]

$$y_{1t} = 10 + 2x_{11t} + \varepsilon_{1t},\ y_{2t} = -10 + 6x_{21t} + \varepsilon_{2t},\ y_{3t} = 20 + 5x_{31t} + 3x_{32t} + \varepsilon_{3t} \tag{7}$$

Following Schmidt (1977), we set the variances of the disturbances equal to 1 and consider four alternative values of the correlations (ρ_{21}, ρ_{31} and ρ_{32}), 0, 0.3, 0.6 and 0.9, where ρ_{ij} is the correlation between ε_{it} and ε_{jt} for i, j = 1, 2, 3. We also assume that we are only interested in estimating the regression coefficients of the first two equations, $\theta_1 \equiv (\beta_{10},\beta_{11},\beta_{20},\beta_{22})'$, where $(\beta_{10},\beta_{11})'$ and $(\beta_{20},\beta_{21})'$ denote the regression coefficients of the first and second equations, respectively. Similarly, we use $(\beta_{30},\beta_{31},\beta_{32})'$ to denote the regression coefficients of the third equation. For our various data generating processes (DGPs) in this simulation, we generate $\varepsilon_t \equiv (\varepsilon_{1t},\varepsilon_{2t},\varepsilon_{3t})'$ by *i.i.d.* N(0, Σ), where Σ is the covariance matrix. (Σ is also the correlation matrix under our normalization of unit variances). For our simulations, we consider four cases.

Case 1: ρ_{21}=0, ρ_{31}=0 and ρ_{32}=0. x_{11t} is generated by an *i.i.d.* uniform distribution over [0, 20] (that is, x_{11t} ~ *i.i.d.* U(0, 20)), x_{21t} ~ *i.i.d.* U(0, 50), x_{31t} ~ *i.i.d.* U(0, 100), and x_{32t} ~ *i.i.d.* U(0, 200).

Case 2: ρ_{31}=0, ρ_{32}=0 and $x_{11t} = x_{21t}$. $\rho_{21} \in \{0.3, 0.6, 0.9\}$. $x_{11t} = x_{21t}$ ~ *i.i.d.* U(0, 20), x_{31t} ~ *i.i.d.* U(0, 100), and x_{32t} ~ *i.i.d.* U(0, 200).

Case 3: ρ_{21}=0, $x_{11t} = x_{31t}$ and $x_{21t} = x_{32t}$. $(\rho_{31},\rho_{32}) \in \{0.3, 0.6, 0.9\} \times \{0.3, 0.6, 0.9\}$. $x_{11t} = x_{31t}$ ~ *i.i.d.* U(0, 20), and $x_{21t} = x_{32t}$ ~ *i.i.d.* U(0, 50).

Case 4: $x_{11t} = x_{21t} = x_{31t}$. $(\rho_{21},\rho_{31},\rho_{32}) \in \{0.3, 0.6, 0.9\} \times \{0.3, 0.6, 0.9\} \times \{0.3, 0.6, 0.9\}$. $x_{11t} = x_{21t} = x_{31t}$ ~ *i.i.d.* U(0, 20), and x_{32t} ~ *i.i.d.* U(0, 200).

Table 1. Ratios of the MSE of OLS Estimator to the MSE of FGLS Estimator

ρ_{21}	ρ_{31}	ρ_{32}	T	β_{11}	β_{21}	β_{31}	β_{32}
Case 1							
0	0	0	20	0.9253	0.9226	0.9284	0.9241
			50	0.9642	0.9644	0.9661	0.9692
			100	0.9772	0.9796	0.9817	0.9778
Case 2: $x_{11t} = x_{21t}$							
0.3	0	0	20	0.9590	0.9612	0.9265	0.9266
			50	0.9788	0.9800	0.9597	0.9655
			100	0.9882	0.9897	0.9801	0.9761
0.6	0	0	20	0.9626	0.9649	0.9172	0.9283
			50	0.9842	0.9849	0.9691	0.9589
			100	0.9899	0.9908	0.9814	0.9802
0.9	0	0	20	0.9663	0.9664	0.9321	0.9283
			50	0.9810	0.9806	0.9713	0.9665
			100	0.9903	0.9902	0.9811	0.9753
Case 3: $x_{11t} = x_{31t}$, $x_{21t} = x_{32t}$							
0	0.3	0.3	20	0.9577	0.9573	1.0576	1.0487
			50	0.9808	0.9831	1.0781	1.0869
			100	0.9932	0.9896	1.0879	1.0968
0	0.3	0.9	20	0.9538	0.9656	4.4687	1.0556
			50	0.9835	0.9808	5.0269	1.0896
			100	0.9892	0.9901	5.0026	1.0885
0	0.6	0.6	20	0.9554	0.9615	1.4402	1.4587
			50	0.9811	0.9831	1.5505	1.5669
			100	0.9924	0.9890	1.5250	1.5528
0	0.9	0.3	20	0.9614	0.9490	1.0490	4.3854
			50	0.9825	0.9763	1.0914	5.1045
			100	0.9917	0.9910	1.0910	5.1732
Case 4: $x_{11t} = x_{21t} = x_{31t}$							
0.3	0.3	0.3	20	1.0000	1.0000	0.9999	1.0649
			50	1.0000	1.0000	1.0001	1.1263
			100	1.0000	1.0000	1.0032	1.1318
0.6	0.6	0.6	20	1.0000	1.0000	1.0086	1.5742
			50	1.0000	1.0000	1.0010	1.7507
			100	1.0000	1.0000	1.0020	1.7644
0.9	0.9	0.9	20	1.0000	1.0000	1.0261	4.7436
			50	1.0000	1.0000	1.0076	6.2470
			100	1.0000	1.0000	1.0001	6.5542

It is easy to verify that, for the first two equations of (7), each of the four cases above is sufficient for condition (6) of Theorem 2. Thus, for each case, the equation-by-equation OLS estimation of the coefficients in the first two equations of (7) is asymptotically as efficient as the FGLS estimation applied to the whole system of (7). The main purpose of this simulation is to gauge how well this asymptotic equivalence result performs in finite samples. Our simulation is carried out in GAUSS 14. For each case we consider three sample sizes, T = 20, 50, 100, and for each sample size we use 10,000 replications. Following Schmidt (1977) and Baltagi, *et al.* (1989) and to save space, we report in Table 1 the ratios of the mean square errors (MSEs) of OLS estimators of β_{11}, β_{21} β_{31} and β_{32} to the corresponding MSEs of the FGLS estimators. Due to space limitations, we only report selected combinations of (ρ_{21}, ρ_{31}, ρ_{32}) in Table 1, as other combinations exhibit similar patterns of the MSE ratios.

From Table 1, we observe that when the sample size is larger than 20, the ratios of the MSEs of OLS estimators of $\hat{\beta}_{11}$ and $\hat{\beta}_{21}$ to the corresponding MSEs of FGLS estimators are very close to 1 for Cases 1-3 and exactly equal to 1 for Case 4 (because in this case the OLS and FGLS estimators of β_{11} and β_{21} are numerically identical). We thus conclude that, under the DGP designs of this simulation, the asymptotic equivalence established in Theorem 2 appears to hold well in finite samples. However, it is worth pointing out that, as expected, Table 1 also indicates that the FGLS estimation applied to the whole system of SUR in (7) results in more efficient estimation of β_{31} (in Case 3) and β_{32} (in Cases 3 and 4) than the OLS estimation applied to the third equation alone.

4. Conclusions

In this paper, using the partial redundancy condition of Breusch, *et al.* (1999) and the moment conditions implicitly exploited by OLS and GLS/FGLS estimators of standard SUR models, we derived the necessary and sufficient condition for the asymptotic efficiency of the equation-by-equation OLS estimation of parameters of interest in systems of SUR. The main result of this paper advances the current SUR literature that usually focuses on the efficient estimation of the whole system to the efficient estimation of a sub-system. The four sufficient conditions provided in the Section 2 also generalize various sufficient conditions for the algebraic equivalence of OLS and GLS/FGLS estimators of the whole coefficient vector in a system of SUR to a subset of coefficients. The results of this paper can also be applied to panel data models with small time dimension and a large number of cross-sections. One possible extension of the current paper is to simultaneous equations models. More specifically, the current literature on the relationship between 2SLS and 3SLS estimators has mainly focused on the estimation of all regression coefficients appearing in the system. It would be of interest to see whether we could find a general necessary and sufficient condition for the 2SLS estimator of a subset of regression coefficients in a system of simultaneous equations to be asymptotically as efficient as the corresponding 3SLS estimator applied to the whole system. This is left for future research.

Notes

1. An equation in a SUR model is said to be overidentified if it does not include all of the explanatory variables appearing in the system. Otherwise, the equation is said to be just-identified. See Schmidt (1978).

2. Ravankar (1974) and Schmidt (1978) consider SUR models in which one group of equations is just-identified and the remaining equations are overidentified. They show that the GLS/FGLS estimation applied to the group of overidentified equations is as efficient as the GLS/FGLS applied to the whole system of equations (for the estimation of the coefficients in the overidentified equations). Gourieroux and Monfort (1980) considers, among other things, the numerical equality of the equation-by-equation OLS estimator and the GLS estimator of SUR models. They also consider the numerical equality of OLS and GLS estimators for a subset of regression coefficients in SUR models.

3. The parameters values are chosen partially based on the DGPs in Kmenta and Gilbert (1968) and Schmidt (1977). Also, as shown in Breusch (1980), the Monte Carlo results are invariant to the specific values of the regression coefficients.

Acknowledgement: The authors are very grateful to the Associate Editor and two anonymous referees for their very constructive suggestions that significantly improved the quality of the paper.

References

Anselin, L. (1988) "A Test for Spatial Autocorrelation in Seemingly Unrelated Regressions" *Economics Letters* 28, 335-341.

Baksalary, J. and G. Trenkler (1989) "The Efficiency of OLS in a Seemingly Unrelated Regressions Model" *Econometric Theory* 5, 463-465.

Baltagi, B. (1980) "On Seemingly Unrelated Regressions with Error Components" *Econometrica* 48, 1547-1551.

Baltagi, B. (1988) "The Efficiency of OLS in a Seemingly Unrelated Regressions Model" *Econometric Theory* 4, 536-537.

Baltagi, B. (1989) "On the Efficiency of Two-Stage and Three-Stage Least Squares Estimators" *Econometric Reviews* 7, 165-169.

Baltagi, B. and G. Bresson (2011) "Maximum Likelihood Estimation and Lagrange Multiplier Tests for Panel Seemingly Unrelated Regressions with Spatial Lag and Spatial Errors: An Application to Hedonic Housing Prices in Paris" *Journal of Urban Economics* 69, 24-42.

Baltagi, B., S. Garvin and S. Kerman (1989) "Further Monte Carlo Evidence on Seemingly Unrelated Regressions with Unequal Number of Observations" *Annales D'Economie et de Statistique* 14, 103-115.

Baltagi B. and A. Pirotte (2011) "Seemingly Unrelated Regressions with Spatial Error Components" *Empirical Economics* 40, 5-49.

Bartels, R. and D. Fiebig (1991) "A Simple Characterization of Seemingly Unrelated Regressions Models in Which OLS is BLUE" *The American Statistician* 45, 137-140.

Bewley, R. (1986) *Allocation Models: Specification, Estimation and Applications*. Cambridge, MA: Ballinger Publishing Company.

Breusch, T. (1980) “Useful Invariance Results for Generalized Regression Models” *Journal of Econometrics* 13, 327-340.

Breusch, T., H. Qian, P. Schmidt and D. Wyhowski (1999) “Redundancy of Moment Conditions” *Journal of Econometrics* 91, 89-111.

Chang, Y. (2004) “Bootstrap Unit Root Tests in Panels with Cross-Sectional Dependence” Journal of Econometrics 120, 263-293.

Creel, M. and M. Farell (1996) “SUR Estimation of Multiple Time-Series Models with Heteroscedascity and Serial Correlation of Unknown Form” *Economics Letters* 53, 239-245.

Dwivedi, T. and V. Srivastava (1978) “Optimality of Least Squares in the Seemingly Unrelated Regression Equation Model” *Journal of Econometrics* 7, 391-395.

Fiebig, D. (2001) “Seemingly Unrelated Regressions” in *A Companion to Theoretical Econometrics* by B. Baltagi, Ed., Blackwell Publishers, 101-121.

Gebremariam, G., T. Gebremedhin, P. Schaeffer (2012) “County-level Determinants of Local Public Services in Appalachia: A Multivariate Spatial Autoregressive Model Approach” *Annals of Regional Science* 49, 175-190.

Gourieroux, C. and A. Monfort (1980) “Sufficient Linear Structures: Econometric Applications” *Econometrica* 48, 1083-1097.

Hodgson, D., O. Linton and K. Vorkink (2002) “Testing the Capital Asset Pricing Model Efficiently under Elliptical Symmetry: A Semiparametric Approach” *Journal of Applied Econometrics* 17, 619-639.

Kmenta, J. and R. Gilbert (1968) “Small Sample Properties of Alternative Estimators for Seemingly Unrelated Regressions” *Journal of the American Statistical Association* 63, 1180-1200.

Kruskal, W. (1968) “When Are Gauss-Markov and Least Squares Estimators Identical? A Coordinate-free Approach” *Annals of Mathematical Statistics* 39, 70-75.

LeGallo, J. and C. Chasco (2008) “Spatial Analysis of Urban Growth in Spain, 1900-2001” Empirical Economics 34, 59-80.

Lundberg, J. (2006) “Spatial Interaction Model of Spillovers from Locally Provided Public Services” *Regional Studies* 40, 631-644.

Mark, N., M. Ogaki and D. Sul (2005) "Dynamic Seemingly Unrelated Cointegrating Regressions" *Review of Economic Studies* 72, 797-820.

Moon, H. (1999) "A Note on Fully-Modified Estimation of Seemingly Unrelated Regressions Models with Integrated Regressors" *Economics Letters* 65, 25-31.

Moon, H. and B. Perron (2004) "Efficient Estimation of the SUR Cointegration Regression Model and Testing for Purchasing Power Parity" *Econometric Reviews* 23, 293-323.

Moon, H. and B. Perron (2006) "Seemingly Unrelated Regressions," in *The New Palgrave Dictionary of Economics*, 2nd ed., by L. Blume and S. Durlauf, Eds., Palgrave and Macmillan.

Qian, H. (2002) "Partial Redundancy of Moment Conditions" *Econometric Theory* 18, 531-539.

Qian, H. (2008) "Redundancy of Moment Conditions and the Efficiency of OLS Estimation in Seemingly Unrelated Regressions Models" *Econometric Theory* 24, 1456-1460.

Rao, C. R. (1973) *Linear Statistical Inference and Its Applications*, Wiley: New York.

Revankar, N. (1974) "Some Finite Sample Results in the Context of Two Seemingly Unrelated Regression Equations" *Journal of the American Statistical Association* 69, 187-190.

Schmidt, P. (1977) "Estimation of Seemingly Unrelated Regressions with Unequal Numbers of Observations" *Journal of Econometrics* 5, 365-377.

Schmidt, P. (1978) "A Note on the Estimation of Seemingly Unrelated Regression Systems" *Journal of Econometrics* 7, 259-261.

Srivastava, V. and D. Giles (1987) *Seemingly Unrelated Regression Equations Models*, Marcel Dekker: New York.

Wooldridge, J. (2010) *Econometric Analysis of Cross Section and Panel Data*, The MIT Press.

Zellner, A. (1962) "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias" *Journal of the American Statistical Association* 57, 348-368.

Zhou, B. and K. Kockelman (2009) "Predicting the distribution of households and employment: a seemingly unrelated regression model with two spatial processes" *Journal of Transport Geography* 17, 369-376.

Zyskind, G. (1967) "On Canonical Forms, Nonnegative Covariance Matrices and Best and Simple Least Squares Linear Estimators in Linear Models" *Annals of Mathematical Statistics* 38, 1092-1109.