Economics Letters 148 (2016) 59–62

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Efficiency comparison of random effects two stage least squares estimators

Chirok Han

Department of Economics, Korea University, 145 Anam-ro Seongbuk-gu, Seoul, Republic of Korea

HIGHLIGHTS

- Efficiency is compared between EC2SLS and G2SLS for panel data models.
- G2SLS is known to be as efficient as EC2SLS though the former uses fewer instruments.
- I establish asymptotic equivalence of G2SLS and EC2SLS under general conditions.
- I show that EC2SLS can be more efficient than G2SLS if the RF equations contain FE.
- Cornwell and Trumbull's (1994) model and data are examined,

ARTICLE INFO

Article history:
Received 9 June 2016
Received in revised form
21 August 2016
Accepted 7 September 2016
Available online 19 September 2016

JEL classification:
C23
C26

Keywords:
Panel data
Random effects
Two stage least squares
EC2SLS
G2SLS
Efficiency

ABSTRACT

For two stage least squares estimation (2SLS) with panel data, Baltagi and Li (1992) have shown that Baltagi's (1981) error-component 2SLS and Balestra and Varadharajan-Krishnakumar's (1987) generalized 2SLS are equivalent in terms of asymptotic variance under the random effects error-component assumption. In the present paper this asymptotic equivalence is extended to models with heteroskedastic and serial correlated errors. However, it is shown that the equivalence does not hold if fixed effects are present in the reduced-form equations. The theoretical claims are verified by simulations and are examined for Cornwell and Trumbull's (1994) model and data.

1. Introduction

For the structural error-component model $y_{it} = Z_{it}\beta + u_{it}$, $u_{it} = \mu_i + \nu_{it}$, with instruments X_{it} strictly exogenous to μ_i and ν_{it}, Baltagi's (1981) error-component two stage least squares (EC2SLS), and Balestra and Varadharajan-Krishnakumar's (1987, BV hereafter) generalized two stage least squares (G2SLS) are commonly used. Both methods begin with transforming the equation to $y_{it} - \hat{\theta}\bar{y}_i = (Z_{it} - \hat{\theta}\bar{Z}_i)\beta + (u_{it} - \hat{\theta}\bar{u}_i)$, where $\bar{y}_i$, $\bar{Z}_i$ and $\bar{u}_i$ are the individual-specific averages of y_{it}, Z_{it} and u_{it}, respectively, and $\hat{\theta}$ is chosen such that $u_{it} - (\text{plim}\,\hat{\theta})\bar{u}_i$ is serially uncorrelated under the random effects error component assumption (the "RE Assumption" hereafter) that $\text{E}(u_{it}u_{is}) = \sigma_\mu^2 + \sigma_\nu^2\{t = s\}$. After this transformation, EC2SLS estimates β by the pooled two stage least squares (2SLS) using instruments $(X_{it}, \bar{X}_i)$, while G2SLS uses a smaller subset $X_{it} - \hat{\theta}\bar{X}_i$ as instruments.

Because EC2SLS uses more instruments, it is naturally asymptotically efficient relative to G2SLS. Baltagi and Liu (2009) provides with a formal proof that the difference of the asymptotic variances of G2SLS and EC2SLS is positive-semidefinite. However, Baltagi and Li (1992) have shown that G2SLS does not lose asymptotic efficiency and is asymptotically as efficient as EC2SLS under the RE Assumption. In spite of this asymptotic equivalence, Baltagi and Liu (2009) find that the standard errors of EC2SLS are considerably smaller than those of G2SLS in a crime model using Cornwell and Trumbull's (1994) data: for the endogenous regressors, the standard errors for EC2SLS are approximately 60% smaller than those for G2SLS. Baltagi and Liu have attributed this substantial difference to small sample efficiency gain by EC2SLS, which will disap-

E-mail address: chirokhan@korea.ac.kr.

http://dx.doi.org/10.1016/j.econlet.2016.09.007

pear as the sample size increases to infinity. The present paper is aimed at providing more explanation to the evident efficiency gain by EC2SLS.

Two possibilities are explored in this regard. One is heteroskedasticity and serial correlation in the error term. But, as I formulate later, they do not make any difference. The equivalence in terms of asymptotic efficiency holds even in the presence of heteroskedasticity or serial correlation in the idiosyncratic error. This result is presented as Theorem 1 in Section 2. The second possibility that I explore is the regressor-instrument relationship, where I find an important source of asymptotic efficiency gain by EC2SLS. It turns out that G2SLS is not efficient if there are fixed effects in the reduced-form equations.

The organization of the paper is as follows: Section 2 presents the model, conventional assumptions, and formal definitions of EC2SLS and G2SLS. In this section, I will also present a theorem on the asymptotic equivalence of the two estimators when the RE Assumption does not hold. Section 3 presents a model in which both EC2SLS and G2SLS are consistent and EC2SLS is asymptotically more efficient than G2SLS. Simulation results are reported in each section. Cornwell and Trumbull's (1994) data and model are reexamined at the end of Section 3. Section 4 concludes the paper.

2. Asymptotic equivalence of EC2SLS and G2SLS

Consider the error-component model $y_{it} = Z_{it}\beta + u_{it}$, $u_{it} = \mu_i + \nu_{it}$, where μ_i are unobservable time-invariant random effects and ν_{it} are the time-varying idiosyncratic errors. Variables in Z_{it} are possibly correlated with the error term u_{it}, and we have a set of instrumental variables X_{it} exogenous to both error components. The panel data set is assumed to be balanced in order to avoid unnecessary complexity. Asymptotics will be considered as $n \to \infty$ with fixed T.

As explained in Section 1, after the transformation of the equation to $\ddot{y}_{it} = \ddot{Z}_{it}\beta + \ddot{u}_{it}$, where $\ddot{y}_{it} = y_{it} - \hat{\theta}\bar{y}_i$, etc., for a $\hat{\theta}$ which makes $u_{it} - (\text{plim}\,\hat{\theta})\bar{u}_i$ homoskedastic and serially uncorrelated under the RE Assumption, EC2SLS uses instruments $A_{it} = (\tilde{X}_{it}, \bar{X}_{it})$ and G2SLS uses $\ddot{X}_{it} = X_{it} - \hat{\theta}\bar{X}_i$. Clearly, $\ddot{X}_{it}$ is a linear combination of the elements of A_{it}, and thus EC2SLS is asymptotically efficient relative to G2SLS. But, Baltagi and Li (1992) have shown that G2SLS is asymptotically as efficient as EC2SLS under the RE Assumption. I will first examine whether this asymptotic equivalence also holds when the RE Assumption fails and both EC2SLS and G2SLS are inefficient.

For notational brevity, let X, $\tilde{X}$, $\bar{X}$ and $\ddot{X}$ denote the matrices of the nT rows of X_{it}, $\tilde{X}_{it}$, $\bar{X}_i$ and $\ddot{X}_{it}$ respectively. Similarly, let us define $\tilde{Z}$, $\bar{Z}$, $\ddot{Z}$, $\tilde{u}$, $\bar{u}$ and $\ddot{u}$ as matrices with nT rows. Throughout the paper, I will maintain the regularity that both $\tilde{X}_{it}$ and $\bar{X}_i$ have sufficient variability and explanatory power and that they are both exogenous, as follows:

Assumption A. (i) $(nT)^{-1}\tilde{X}'\tilde{X}$ and $(nT)^{-1}\bar{X}'\bar{X}$ are asymptotically finite and nonsingular; (ii) $\text{plim}(nT)^{-1}\tilde{X}'\tilde{Z}$ has full column rank; (iii) $(nT)^{-1/2}\tilde{X}'\tilde{u} = O_p(1)$ and $n^{-1/2}\sum_{i=1}^n \bar{X}_i'\bar{u}_i = O_p(1)$.

Note that EC2SLS and G2SLS are consistent under Assumption A. Condition (iii) does not require homoskedasticity or serial independence of u_{it}.

In what follows, I will present the result that the equivalence of the two estimators holds without the RE Assumption. In fact, there exists even a stronger form of equivalence that the difference between the two estimators, normalized by multiplying $\sqrt{nT}$, diminishes to zero as the sample size increases under acceptable regularity. The following theorem states it, where $\hat{\beta}_{EC}$ and $\hat{\beta}_G$ denote the EC2SLS and the G2SLS estimators of β respectively.

Theorem 1. *Let $\tilde{V}$ and $\bar{V}$ be some $(nT) \times k$ matrices of $\tilde{V}_{it}$ and $\bar{V}_i$ respectively, where $\bar{V}_i = T^{-1}\sum_{t=1}^T V_{it}$ and $\tilde{V}_{it} = V_{it} - \bar{V}_{it}$. If*

$$Z = X\Pi + V, \qquad (nT)^{-1}\tilde{X}'\tilde{V} = o_p(1) \quad \text{and} \\ (nT)^{-1}\bar{X}'\bar{V} = o_p(1), \tag{1}$$

then $(nT)^{1/2}(\hat{\beta}_{EC} - \hat{\beta}_G) = o_p(1)$ under Assumption A.

Condition in (1) is crucial for Theorem 1. BV (1987) consider a model that satisfies it; see their equation (1.9). Baltagi and Li (1992) also accept this reduced form in the proof of their Proposition 2. The second and third conditions in (1) hold as $n \to \infty$ if, for instance, $E(X_{is}'V_{it}) = 0$ for all s and t under cross-sectional independence, which is usually assumed in the random effects 2SLS context. The proof of Theorem 1 is rather straightforward and is given below.

Proof of Theorem 1. It is convenient to define

$$\breve{X} = \bar{X}(\bar{X}'\bar{X})^{-1} - (1-\hat{\theta})\tilde{X}(\tilde{X}'\tilde{X})^{-1}, \tag{2}$$

which is constructed to be orthogonal to $\ddot{X} = \tilde{X} + (1-\hat{\theta})\bar{X}$, i.e., $\breve{X}'\ddot{X} = 0$. Due to this orthogonality, we have $P_A = P_{[\tilde{X},\bar{X}]} = P_{[\ddot{X},\breve{X}]} = P_{\ddot{X}} + P_{\breve{X}}$, where A is the matrix of $A_{it} = [\tilde{X}_{it}, \bar{X}_i]$ and $P_H = H(H'H)^{-1}H'$ for any full column-rank matrix H. Let $m = nT$ in this proof. By (1), we have

$$m^{-1}\breve{X}'\ddot{Z} = m^{-1}\breve{X}'\ddot{X}\Pi + m^{-1}\breve{X}'\ddot{V} = m^{-1}\breve{X}'\ddot{V} = o_p(1). \tag{3}$$

We have $m^{-1}\ddot{Z}'P_{\breve{X}}\ddot{Z} = m^{-1}\ddot{Z}'\breve{X}(m^{-1}\breve{X}'\breve{X})^{-1}m^{-1}\breve{X}'\ddot{Z}$, where the middle term $m^{-1}\breve{X}'\breve{X}$ on the right-hand side is asymptotically nonsingular under Assumption A. Thus, due to (3), we have (a) $m^{-1}\ddot{Z}'P_A\ddot{Z} = m^{-1}\ddot{Z}'P_{\ddot{X}}\ddot{Z} + o_p(1)$. Similarly, we have (b) $m^{-1/2}\ddot{Z}'P_A\ddot{u} = m^{-1/2}\ddot{Z}'P_{\ddot{X}}\ddot{u} + o_p(1)$ because the remainder term equals

$$m^{-1/2}\ddot{Z}'P_{\breve{X}}\ddot{u} = m^{-1}\ddot{Z}'\breve{X}(m^{-1}\breve{X}'\breve{X})^{-1}m^{-1/2}\breve{X}'\ddot{u} = o_p(1)$$

due to (3) and Assumption A(iii). The result follows from (a), (b) and the identites $\sqrt{nT}(\hat{\beta}_G - \beta) = (m^{-1}\ddot{Z}'P_{\ddot{X}}\ddot{Z})^{-1}m^{-1/2}\ddot{Z}'P_{\ddot{X}}\ddot{u}$ and $\sqrt{nT}(\hat{\beta}_{EC} - \beta) = (m^{-1}\ddot{Z}'P_A\ddot{Z})^{-1}m^{-1/2}\ddot{Z}'P_A\ddot{u}$. □

Remark 1.1. Theorem 1 implies that $\sqrt{nT}(\hat{\beta}_{EC} - \beta)$ and $\sqrt{nT}(\hat{\beta}_G - \beta)$ have the same asymptotic distribution. Thus, the equivalence of the asymptotic variances of EC2SLS and G2SLS established by Baltagi and Li (1992, Proposition 2) and pointed out by Baltagi and Liu (2009, Remark 1) holds without the RE Assumption as long as (1) holds. □

Remark 1.2. The equivalence in Theorem 1 holds for any $\hat{\theta}$ value, not just the "correct" factor for the GLS-like transformation, as long as it is stochastically bounded. Under the conditions in Theorem 1, if one transforms the equation to $y_{it} - \hat{\theta}\bar{y}_i = (Z_{it} - \hat{\theta}\bar{Z}_i)\beta + (u_{it} - \hat{\theta}\bar{u}_i)$ for whatever $\hat{\theta}$, then the instruments can be transformed in the same way to $X_{it} - \hat{\theta}\bar{X}_i$ without sacrificing asymptotic efficiency by omitting $\bar{X}_i$. For instance, when $\hat{\theta} = 1$ is used, the corresponding G2SLS is the 2SLS on the within transformation using $\tilde{X}_{it}$ as instruments, and the corresponding EC2SLS is that using $(\tilde{X}_{it}, \bar{X}_i)$ as instruments. They are algebraically identical and Theorem 1 holds. As another example, if $\hat{\theta} = 0$, then the conclusion of Theorem 1 suggests that the pooled 2SLS (applied to the original equation) using X_{it} as instruments is asymptotically as efficient as that using $(X_{it}, \bar{X}_i)$ as instruments. □