Iranian Journal of Economic Research
Vol. 17, No. 50, Spring 2012, pp. 83-98

Relative Performance of Components Variance Estimators in Random Effects Models

Teimour Mohammadi*

Received: 18 September 2011
Accepted: 15 February 2012

This paper presents an assessment of the small-sample performance of the three well-known estimators of components variance in random effects model for panel data. The estimators considered are Swamy-Arora, Wansbeek-Kaptayn and Wallace-Hussain. To this end, by simulating a one-way error component model in the form of random effects, small sample performance of three variance estimators is studied. The implications of these results for indentifying the model and its estimation are specified. In these simulations, conditions under which Swamy-Arora estimator is inferior to alternatives are expressed. It is shown that in small samples the estimator thus obtained can give highly wrong guidance. In one-way error component model this small sample size refers to the number of cross-sections.

Keywords: Panel Data, Random Effects, Component Variance Estimators, Simulation.
JEL Clasiification: C1, C33, C5, C51.

1. Introduction

Three well-known components variance estimators in random effects models for panel data, are Swamy-Arora, Wansbeek-Kaptayn and Wallace-Hussain. The aim of this paper is to give guidelines as to where the researcher is advised to use which one. Traditionally, the default option of well-known softwares for estimating variance components of random effects models is Swamy-Arora. In this paper, after presenting the theoretical back-ground in Section 1, three famous alternative estimators including Swamy-Arora are re-examined. In contrast to the well-known arguments in favor of it, which have large-

* Assistant Professor, Faculty of Economics, Allameh Tabatabaie University, Email: atmahmadi@gmail.com

sample justifications and are presented in Section 2, we have mentioned in Section 3 that, by simulating a random effects model and Monte-Carlo experiment based on it, there are cases in which, Swamy-Arora estimator (and the corresponding FGLS estimator of the mean equation) does not have a high mark, hence the alternatives may outperform it. In other words, the paper aims to turn the attention to small-sample merits of two other alternatives. Section 5 concludes the paper.

2. Theoretical Foundations

The basic formulation of one-way random effects model is

$$Y_i = \alpha + X_i \beta + \eta_i$$

where

$$\eta_i = \varepsilon_{it} + u_i$$

cross-section index $i = 1, \ldots, n$ (= number of cross-section)
period index $t = 1, \ldots, T$ (= number of periods)
$X_i =$ matrix of observations of nonstochastic independent variables for cross section i
$Y_i =$ vector of observations of dependent variable for cross-section i
ε_{it} is the time-variant (idiosyncratic) random error term and u_i refers to the cross-section random component. This standard error component model, satisfies the following assumptions,

$$u_i \sim IID(0, \sigma_u^2)$$
$$\varepsilon_{it} \sim IID(0, \sigma_\varepsilon^2)$$

For estimation, it is suitable to write the stacked form of the model, i.e.,

$$Y = Z\delta + \eta$$

where $Z = (i_{nT} \vdots X)$, $\delta = (\alpha \vdots \beta)$, $X = [X_1' \vdots \ldots \vdots X_n']'$, $Y = [Y_1' \vdots \ldots \vdots Y_n']'$, $\eta = [\eta_1' \vdots \ldots \vdots \eta_n']'$. It is well-known that OLS estimates, although unbiased, are inefficient; in contrast, GLS estimates are BLUE, which in turn require the availability the var-cov matrix of η denoted by Ω,

$$Var - Cov(\eta) = \Omega.$$

The matrix Ω is computed as follows,

$$\Omega = E(\eta\eta') = \sigma_u^2 (I_n \otimes J_T) + \sigma_\varepsilon^2 (I_n \otimes I_T)$$
$$= T\sigma_u^2 (I_n \otimes \bar{J}_T) + \sigma_\varepsilon^2 (I_n \otimes E_T) + \sigma_\varepsilon^2 (I_n \otimes \bar{J}_T),$$

where $J_T = I_T - \frac{1}{T} i_T i'_T$, $\bar{J}_T = \frac{1}{T} J_T$, $E_T = I_T - \bar{J}_T$.

By defining $\sigma_1^2 = T\sigma_u^2 + \sigma_\varepsilon^2$, we have

$$\Omega = \sigma_1^2 H + \sigma_\varepsilon^2 R,$$

where $H = I_n \otimes \bar{J}_T$ and $R = I_n \otimes E_T$.

GLS estimation requires Ω^{-1} which by spectral decomposition of Ω, based on characteristic roots and vectors of Ω, we have

$$\Omega^{-1} = \frac{1}{\sigma_1^2} H + \frac{1}{\sigma_\varepsilon^2} R.$$

Premultiplying the model $Y = Z\delta + \eta$ by $\Omega^{-\frac{1}{2}}$ gives

$$\Omega^{-\frac{1}{2}} = \sigma_1^{-\frac{1}{2}} H + \sigma_\varepsilon^{-\frac{1}{2}} R.$$

Applying OLS on this transformed model gives GLS estimates. Transformed vector of the dependent variables observations is $y_{it} - \theta\bar{y}_{i0}$, $\bar{y}_{i0} = \frac{\sum_{t=1}^{T} y_{it}}{T}$, $\theta = 1 - \frac{\sigma_\varepsilon}{\sigma_1}$. Similar result holds for the matrix of observations for independent variables. In the next step, we must estimate σ_1^2 and σ_ε^2, which if the terms η_{it} *were known*, then

$$E(\eta' H\eta) = E(tr\ \eta' H\eta) = E(tr\ \eta\eta' H) = tr\ E(\eta\eta')H$$
$$= tr(\Omega H) = tr(\sigma_1^2 H + \sigma_\varepsilon^2 R)H$$
$$= tr(\sigma_1^2 H) + tr(RH)\ \sigma_\varepsilon^2 = \sigma_1^2 tr(H) = n\sigma_1^2 .$$

we have

(1) $$\hat{\sigma}_1^2 = \frac{\eta' H\eta}{tr(H)}.$$

Since, HR=0, H and R are idempotent matrices, hence

$$E(\hat{\sigma}_1^2) = E\left(\frac{\eta' H\eta}{tr(H)}\right) = \frac{n\sigma_1^2}{n} = \sigma_1^2 .$$

Similarly, since

$$E(\eta' R\eta) = E(tr\ \eta' R\eta) = E(tr\ \eta\eta' R) = tr\ E(\eta\eta')R = tr(\Omega R)$$
$$= tr(\sigma_1^2 H + \sigma_\varepsilon^2 R)R = \sigma_1^2 tr(HR) + \sigma_\varepsilon^2 tr(RR) = \sigma_\varepsilon^2 tr(R) = n(T-1)\sigma_\varepsilon^2 .$$

The Best Quadratic Unbiased estimator (BQU) for σ_ε^2 is

(2) $$\hat{\sigma}_\varepsilon^2 = \frac{\eta' R\eta}{tr(R)}$$

since $E(\hat{\sigma}_\varepsilon^2) = E\left(\frac{\eta' R\eta}{tr(R)}\right) = \frac{n(T-1)\sigma_\varepsilon^2}{n(T-1)} = \sigma_\varepsilon^2 .$

To sum up, the BQU estimators of σ_1^2 and σ_ε^2 based on 1 and 2 are

$$\hat{\sigma}_1^2 = \frac{\eta' H\eta}{tr(H)} \qquad , \qquad \hat{\sigma}_\varepsilon^2 = \frac{\eta' R\eta}{tr(R)}.$$

To give a more operational formulae for (2) and (1) we note that

$$\eta' H\eta = (H\eta)'(H\eta) = \eta' HH'\eta = \eta' H'H\eta$$

$$= \left[\overline{\eta}_{10}\overline{\eta}_{10} \ldots \overline{\eta}_{10}\overline{\eta}_{20} \ldots \overline{\eta}_{20} \ldots \overline{\eta}_{n0} \ldots \overline{\eta}_{n0}\right]\begin{bmatrix} \overline{\eta}_{10} \\ \vdots \\ \overline{\eta}_{10} \\ \overline{\eta}_{20} \\ \vdots \\ \overline{\eta}_{20} \\ \vdots \\ \overline{\eta}_{n0} \\ \vdots \\ \overline{\eta}_{n0} \end{bmatrix},$$

$$= \sum_{t=1}^{T} \overline{\eta}_{1}^{2} + \cdots + \sum_{t=1}^{T} \overline{\eta}_{n0}^{2} = T\overline{\eta}_{1}^{2} + \cdots + T\overline{\eta}_{n0}^{2} = T\sum_{i=1}^{n} \overline{\eta}_{i0}^{2}$$

where $\overline{\eta}_{i0} = \dfrac{\sum_{t=1}^{T} \eta_{it}}{T}$,

and also,

$$\eta' R\eta = (R\eta)'(R\eta) = \sum_{t=1}^{T} (\eta_{1t} - \overline{\eta}_{10})^2 + \cdots + \sum_{t=1}^{T} (\eta_{nt} - \overline{\eta}_{n0})^2$$

$$= \sum_{t=1}^{T} \sum_{i=1}^{n} (\eta_{it} - \overline{\eta}_{i0})^2$$

since $R\eta = \begin{bmatrix} \eta_{11} - \overline{\eta}_{10} \\ \vdots \\ \eta_{1T} - \overline{\eta}_{10} \\ \eta_{21} - \overline{\eta}_{20} \\ \vdots \\ \eta_{2T} - \overline{\eta}_{20} \\ \vdots \end{bmatrix}$.

Thus if the compound error terms η_{it} *were known*, the operational formulae for the variance estimators would be

(1)′
$$\hat{\sigma}_1^2 = \frac{\eta' H \eta}{tr(H)} = T \sum_{i=1}^{n} \frac{\overline{\eta}_{i0}^2}{n},$$

(2)′
$$\hat{\sigma}_\varepsilon^2 = \frac{\eta' R \eta}{tr(R)} = \frac{\sum_{i=1}^{n} \sum_{t=1}^{T} (\eta_{it} - \overline{\eta}_{i0})^2}{n(T-1)}.$$

Unfortunately, the population error terms, u_i and ε_{it}, and hence η_{it}, *are not known*, hence (1)′ and (2)′ are not operationally feasible. This is the critical point on which the paper is focused. The point is that there are several ways to estimate (1)′ and (2)′ in practice. Wallace and Hussain (1969) use OLS residuals $\hat{\eta}_{OLS}$ instead of the true η. Amemiya (1971) suggests using fixed-effects or within estimates residual instead of OLS residual. Following the work of Wansbeek and Kapteyn (1978) a number of softwares refer to the estimates of $\hat{\sigma}_1^2$ and $\hat{\sigma}_\varepsilon^2$ as Wansbeek and Kapteyn estimators of variance components. Swamy and Arora (1972) propose running two regressions (within and between) to estimate $\hat{\sigma}_\varepsilon^2$ and $\hat{\sigma}_1^2$ from respective mean square errors. In the followings, we refer to these estimators by the abbreviations WH, WK and SA, respectively.

3. Empirical Results

As to these estimators of variance components of random effects model, there are a vast literature, which compare their large-sample properties. Some of the well-known references are Wallace-Hussain (1969), Amemiya (1971), Swamy-Arora (1972), Fuller-Battese (1974), Rao-Kleffe (1980) and Baltagi (1981). All the writers emphasize that these variance components estimators are consistent but may be biased in finite samples.

It should also be noted that if the lagged values of the dependent variable are used as explanatory variables, these estimators of the variance components may be inconsistent. Matyas and Sevestre (1992) point out that these estimators are MINQUE and they are asymptotically (nT$\rightarrow\infty$ or n$\rightarrow\infty$) equivalent to estimators (1)′ and (2)′

if within residuals are used. If in (1)′ and (2)′ the OLS residuals are used then these estimators are less efficient than the MINQUEs ones. With these observations, the usefulness of SA estimators of variances are heavily emphasized in the literature so that the default option of the well-known softwares is SA.

As to the FGLS (or EGLS) estimator of β vector in the basic model, Maddala-Mount (1973) and Baltagi (1981) show that the estimation method used to obtain the estimated variance components has little effect on the behavior of the FGLS (or any other two-step estimation method). The basic requirement is that the method must be consistent. Recently, some authors have warned against the negative variance estimates in panel data models. Magazzini and Calzolari (2010) re-examine this neglected point in their research work.

4. Relative Performance of Components Variance Estimators in Small Samples

Our aim is a reappraisal of the relative advantage and disadvantage of WH, WK and SA estimators especially in small samples. In contrast to the large-sample merits of SA outlined in the last section, in small samples, several cases exist which lower its ranking relative to WH and WK. These cases are as follows:

i) In section (1) the mathematical formulae for computing $\hat{\sigma}_1^2$ and $\hat{\sigma}_\varepsilon^2$ have been stated. Some of the disadvantages of SA estimator relative to WH and WK originate from those formulae. The point is that since SA uses between and within estimates for computing $\hat{\sigma}_1^2$ and $\hat{\sigma}_\varepsilon^2$ and the between regression reduces the sample size, the SA estimates may go wrong in small samples (with small n). Specifically with one parameter in β and only two cross-sections, whether or not, the variance of cross-section random term u_i is big, although the basic model is truely random effects, SA estimator give a $\hat{\sigma}_u^2$ close to zero and the situation is worse if the variance of the cross section term (u_i) is large. The following simulation will make the point clear. In this simulation the basic model is

$$y_{it} = \alpha + \beta x_{it} + \varepsilon_{it} + u_i .$$

To break down the mathematical formula of SA, assume there are only 2 cross-sections. Also, assume u_i has only two realizations, with 10 period in the panel, the computer output for WH, WK and SA are as follows,

Table 1.

Dependent Variable: Y?
Method: Pooled EGLS (Cross-section random effects)
Date: 01/27/12 Time: 14:34
Sample: 1 10
Included observations: 10
Cross-sections included: 2
Total pool (balanced) observations: 20
Swamy and Arora estimator of component variances

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.29E+08	0.034907	3.69E+09	0.0000
X?	-7518791.	0.002914	-2.58E+09	0.0000
Random Effects (Cross)				
1--C	0.000000			
2--C	0.000000			
	Effects Specification			
			S.D.	Rho
Cross-section random			0.000000	0.0000
Idiosyncratic random			0.075145	1.0000
	Weighted Statistics			
R-squared	0.751879	Mean dependent var		50000057
Adjusted R-squared	0.738095	S.D. dependent var		51298887
S.E. of regression	26253055	Sum squared resid		1.24E+16
F-statistic	54.54538	Durbin-Watson stat		0.086340
Prob(F-statistic)	0.000001			
	Unweighted Statistics			
R-squared	0.751879	Mean dependent var		50000057
Sum squared resid	1.24E+16	Durbin-Watson stat		0.086340

Table 2.

Dependent Variable: Y?
Method: Pooled EGLS (Cross-section random effects)
Date: 01/27/12 Time: 14:34
Sample: 1 10
Included observations: 10
Cross-sections included: 2
Total pool (balanced) observations: 20
Wallace and Hussain estimator of component variances

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	50000005	49999995	1.000000	0.3306
X?	5.007942	0.196561	25.47775	0.0000
Random Effects (Cross)				
1--C	49999995			
2--C	-49999995			
	Effects Specification			
			S.D.	Rho
Cross-section random			70710671	1.0000
Idiosyncratic random			2.524876	0.0000
	Weighted Statistics			
R-squared	0.998439	Mean dependent var		0.564580
Adjusted R-squared	0.998352	S.D. dependent var		14.76943
S.E. of regression	0.599583	Sum squared resid		6.470996
F-statistic	11510.76	Durbin-Watson stat		0.036333
Prob(F-statistic)	0.000000			
	Unweighted Statistics			
R-squared	-0.000001	Mean dependent var		50000057
Sum squared resid	5.00E+16	Durbin-Watson stat		4.70E-18

Table 3.

Dependent Variable: Y?
Method: Pooled EGLS (Cross-section random effects)
Date: 01/27/12 Time: 14:35
Sample: 1 10
Included observations: 10
Cross-sections included: 2
Total pool (balanced) observations: 20
Wansbeek and Kapteyn estimator of component variances

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	49999995	50000001	1.000000	0.3306
X?	5.007942	0.005850	856.0498	0.0000
Random Effects (Cross)				
1--C	50000005			
2--C	-49999985			
	Effects Specification			
			S.D.	Rho
Cross-section random			70710671	1.0000
Idiosyncratic random			0.075145	0.0000
	Weighted Statistics			
R-squared	0.999975	Mean dependent var		0.016803
Adjusted R-squared	0.999974	S.D. dependent var		14.75808
S.E. of regression	0.075145	Sum squared resid		0.101643
F-statistic	732821.3	Durbin-Watson stat		2.313117

Prob(F-statistic)	0.000000		
	Unweighted Statistics		
R-squared	-0.000001	Mean dependent var	50000057
Sum squared resid	5.00E+16	Durbin-Watson stat	4.70E-18

Although the true model is random effects, the SA estimator wrongly rejects it. This fact stems from a zero $\hat{\sigma}_u^2$. On the contrary, nonzero $\hat{\sigma}_u^2$ from WH and WK truly accepts the random effects model. The result is that the desirable large sample performance of SA estimators, which has made it the default option in softwares, does not linearly generalize to small samples. Unfortunately, small sample performance may be so divergent that make WK or WH preferable.

ii) In contrast to the indifference quoted in the last section as to the variance estimator used in FGLS estimates of β, the bias in small samples with large variance of u_i may be considerable. The above tables show the FGLS estimates of β from the simulation. Note that the true value of β is 5. It is evident that the bias of the FGLS estimator based on SA is drastically large. Of course, only one estimate of the parameter cannot be used for illustrating the magnitude of bias, so in the following a Monte-Carlo experiment is offered that proves the claim. For the moment it is merely presented as to be contrasted with FGLS estimators based on WT and WK which truly estimate β equal to 5.

Table 4. WK

Correlated Random Effects - Hausman Test
Pool: Untitled
Test cross-section random effects

Test Summary		Chi-Sq. Statistic	Chi-Sq. d.f.	Prob.
Cross-section random		0.000000	1	1.0000

* Cross-section test variance is invalid. Hausman statistic set to zero.
Cross-section random effects test comparisons:

Variable	Fixed	Random	Var(Diff.)	Prob.
X?	5.007942	5.007942	-0.038602	NA

Cross-section random effects test equation:
Dependent Variable: Y?
Method: Panel Least Squares
Date: 01/28/12 Time: 18:26
Sample: 1 10

Included observations: 10
Cross-sections included: 2
Total pool (balanced) observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	50000005	0.063682	7.85E+08	0.0000
X?	5.007942	0.005850	856.0498	0.0000

Effects Specification
Cross-section fixed (dummy variables)

R-squared	1.000000	Mean dependent var	50000057
Adjusted R-squared	1.000000	S.D. dependent var	51298887
S.E. of regression	0.075145	Akaike info criterion	-2.201304
Sum squared resid	0.095996	Schwarz criterion	-2.051944
Log likelihood	25.01304	Hannan-Quinn criter.	-2.172147
F-statistic	4.43E+18	Durbin-Watson stat	2.449182
Prob(F-statistic)	0.000000		

Table 5. WH

Correlated Random Effects - Hausman Test
Pool: Untitled
Test cross-section random effects

Test Summary	Chi-Sq. Statistic	Chi-Sq. d.f.	Prob.
Cross-section random	0.000000	1	1.0000

* Cross-section test variance is invalid. Hausman statistic set to zero.
Cross-section random effects test comparisons:

Variable	Fixed	Random	Var(Diff.)	Prob.
X?	5.007942	5.007942	-0.000000	NA

Cross-section random effects test equation:
Dependent Variable: Y?
Method: Panel Least Squares
Date: 01/28/12 Time: 18:24
Sample: 1 10
Included observations: 10
Cross-sections included: 2
Total pool (balanced) observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	50000005	0.063682	7.85E+08	0.0000
X?	5.007942	0.005850	856.0498	0.0000

Effects Specification
Cross-section fixed (dummy variables)

R-squared	1.000000	Mean dependent var	50000057
Adjusted R-squared	1.000000	S.D. dependent var	51298887
S.E. of regression	0.075145	Akaike info criterion	-2.201304
Sum squared resid	0.095996	Schwarz criterion	-2.051944
Log likelihood	25.01304	Hannan-Quinn criter.	-2.172147

F-statistic	4.43E+18	Durbin-Watson stat	2.449182
Prob(F-statistic)	0.000000		

Table 6. SA

Correlated Random Effects - Hausman Test
Pool: Untitled
Test cross-section random effects

Test Summary	Chi-Sq. Statistic	Chi-Sq. d.f.	Prob.
Cross-section random	2196988691915946900	1	0.0000

** WARNING: estimated cross-section random effects variance is zero.
Cross-section random effects test comparisons:

Variable	Fixed	Random	Var(Diff.)	Prob.
X?	5.007942	-7518791.240877	0.000026	0.0000

Cross-section random effects test equation:
Dependent Variable: Y?
Method: Panel Least Squares
Date: 01/28/12 Time: 18:27
Sample: 1 10
Included observations: 10
Cross-sections included: 2
Total pool (balanced) observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	50000005	0.063682	7.85E+08	0.0000
X?	5.007942	0.005850	856.0498	0.0000

Effects Specification
Cross-section fixed (dummy variables)

R-squared	1.000000	Mean dependent var	50000057
Adjusted R-squared	1.000000	S.D. dependent var	51298887
S.E. of regression	0.075145	Akaike info criterion	-2.201304
Sum squared resid	0.095996	Schwarz criterion	-2.051944
Log likelihood	25.01304	Hannan-Quinn criter.	-2.172147
F-statistic	4.43E+18	Durbin-Watson stat	2.449182
Prob(F-statistic)	0.000000		

iii) This bias in β and $\hat{\sigma}_u^2$ of SA estimator may inject confusing signals to the researcher about the true model. For example, Hausman test gives two polar values of 0 and 1 for the prob-values for SA and WH/WK estimators. The prob=0 for SA *wrongly* rejects the random effects model and signals in favor of fixed effects model. If follows that neglecting this fact (small n and big variance of u_i) and using SA estimator, is not only an important point for FGLS estimation of the parameters of the mean equation but also a critical point for identifying the true basic model (random or fixed effects) in the first place. This can mislead the researcher at the starting point. The computer output is presented in tables 4, 5 and 6.

iv) Some softwares introduce a RHO(=ρ) coefficient which shows the relative strength of cross-section random term (u_i). In the above simulation, a truly random effects model according to the RHO for $\sigma_u^2\left(=\frac{\sigma_u^2}{\sigma_\varepsilon^2+\sigma_u^2}\right)$ falsely appears as an absolutely non-random one, since the corresponding RHO is zero. In contrast to this false indication of RHO based on SA, the respective RHO for WH or WK indicates that the model is random effects.

Interestingly, although WH has been historically developed prior to SA and WK, in small samples (especially in terms of n) can perform better, so that in the above simulation for 1 cross-section alone, while WH is computable, this is not the case for SA or WK.

v) *Non-normal distribution of idiosyncratic error term*. Case (*i*) is valid irrespective of the variance of ε_{it}. When the number of parameters is greater or equal to the number of cross-sections, SA is not computable. Cases (*ii*) and (*iii*) follow directly from (*i*). Now, the point is that if the variance of ε_{it}, is also big and presumably ε_i is not unimodal, in terms of MSE, again WK and WH can outperform SA. Specifically, when big variance applies to idiosyncratic error term (ε_{it}) there can be large bias and inefficiency in the FGLS estimator of β based on SA estimator. This fact, with large variance of u_i, can give higher ranking to WH and WK estimators relative to SA. This result can be shown by the following simulation. Assume the true model is

$$y_{it}=\alpha+\beta x_{it}+\varepsilon_{it}+u_i$$

11 realizations for u_i are considered. The true β is 5 but as the results indicate, the bias and also the MSE of the estimates based on SA may be large and in this respect WH and WK outperform the SA. To prove this, assume ε_{it} has a bimodal distribution. Specifically, assume the idiosyncratic error term has double-sided chi-square distribution with 13 degree of freedom. (that is, in addition to the usual shape of chi-square distribution, assume it has a mirror-image in the negative quadrant). A Monte-Carlo experiment based on 100 repetitions is

designed and then the estimation is carried out. As the following table shows, in terms of MSE, the FGLS estimator of β based on SA is inferior to WK or WH.

Table 7. Histogram of FGLS estimates of β based on WK

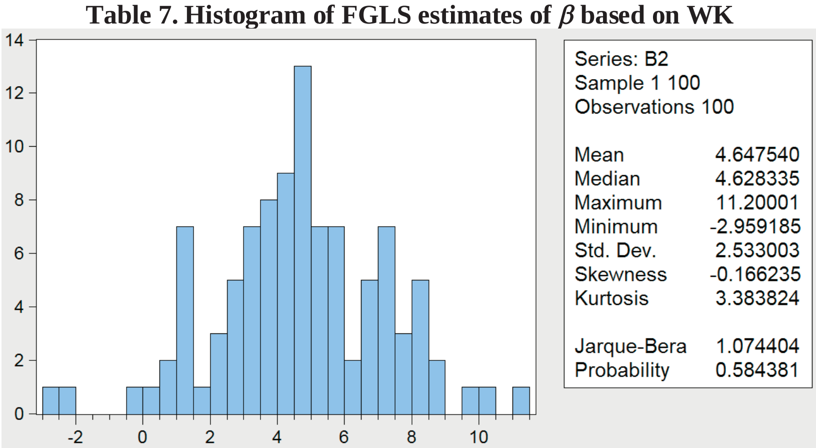
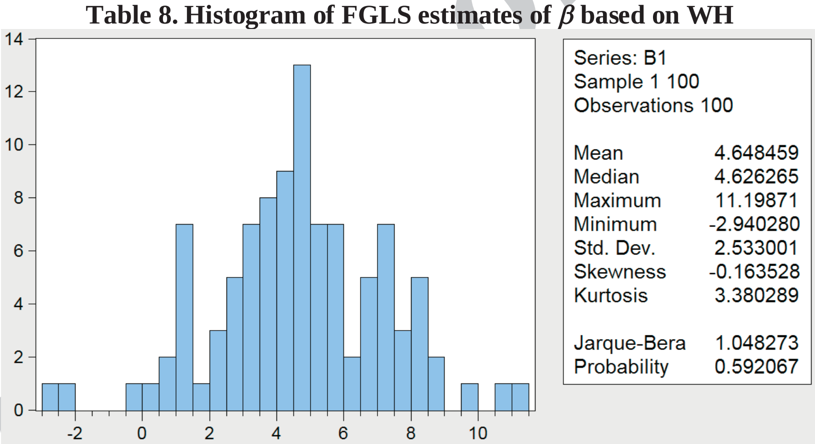


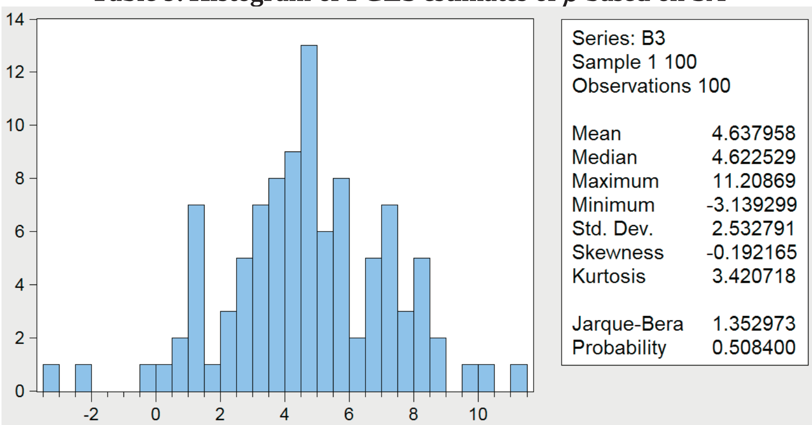
Table 8. Histogram of FGLS estimates of β based on WH



This fact gives a big advantage to HW and WK in estimating random effects model, notwithstanding the emphasis put on SA in the literature.

To sum up, small n makes SA practically unattainable and large variance of u_i or/ and ε_i makes it inferior relative to WH and WK. When the number of the parameters of the model is equal or more than cross-sections, WK or WH may appear to be the only choice.

Table 9. Histogram of FGLS estimates of β based on SA



5. Conclusion

By default, softwares estimate random effects FGLS methods by Swamy-Arora estimator of variance components. This emphasis stems from its desirable large sample properties. Studies concerning small sample properties are confined to statements about the bias of alternative estimators. In this paper, by simulation of a random effects model, cases in which alternative estimators outperform, are specified. A summary of the results are as follows:

1. When the number of cross-sections is small the SA estimator computationally breaks-down. Specifically:
   a. If the number of cross-sections is equal to the number of parameters, the SA favors the fixed-effects model wrongly. The panel estimate results in this case are dramatically false.
   b. If the number of cross-sections is smaller than the number of parameters, mathematically the SA is infeasible while WH and WK are feasible.
2. When the variance of the cross-section random is large, although computationally feasible, SA estimators gives wrong statistical signals; RHO ratios, Hausman tests are invalid.
3. The situation is much worse if the idiosyncratic random variance is large and/or is not unimodel.

Relevance of these remarks become more important when we consider that these cases, make FGLS estimates of β, σ_u^2 and σ_ε^2 so biased that the judgment as to the true model in terms of fixed or random effects, becomes blurred.

References

Amemiya, T. (1971), “The Estimation of Variances in Variance Components Model”, *International Economic Review*, Vol. 12, pp. 1-13.

Baltagi, B. H. (1981), “Pooling an Experimental Study of Alternative Testing and Estimation Procedures in Two-way Error Components Model”, *Journal of Econometrics*, Vol. 17, pp. 21-49.

Fuller, W. A. and G. E. Battese (1974), “Estimation of Linear Models with Crossed-error Structure”, *Journal of Econometrics*, Vol. 2, pp. 67-78.

Magazzini, L. and Giorgio Calzolari (2010), “Negative Variance Estimates in Panel Data Models”, Working Paper, University of Verona.

Rao, C. R. and J. Kleffe (1980), “Estimation of Variance Components”, in *Handbook of Statistics*, Krishnakumar, P. R. (ed) Vol. 1, North Holland, pp. 1-40.

Swamy, P. A. V. B. and S. S. Arora (1972), “The Exact Finite Sample Properties of Estimators of Coefficients in the Error Components Regression Models”, *Econometrica*, Vol. 40, pp. 253-260.

Wallace, T. D. and A. Hussain (1969), “The Use of Error Components Models in Combining Cross Sections with Times Series Data”, *Econometrica*, Vol. 37, pp. 55-72.

Wansbeek, T. J. and A. Kapteyn (1978), “The Separation of Individual Variation and Systematic Change in the Analysis of Panel Data”, *Annales de I’INSEE*, Vols. 30-31, pp. 659-680.