GENERALIZED ESTIMATORS OF STATIONARY RANDOM-COEFFICIENTS PANEL DATA MODELS: ASYMPTOTIC AND SMALL SAMPLE PROPERTIES

Authors: Mohamed Reda Abonazel
– Department of Applied Statistics and Econometrics,
Institute of Statistical Studies and Research, Cairo University,
Egypt (mabonazel@hotmail.com)

Abstract:

- This article provides generalized estimators for the random-coefficients panel data (RCPD) model where the errors are cross-sectional heteroskedastic and contemporaneously correlated as well as with the first-order autocorrelation of the time series errors. Of course, under the new assumptions of the error, the conventional estimators are not suitable for RCPD model. Therefore, the suitable estimator for this model and other alternative estimators have been provided and examined in this article. Furthermore, the efficiency comparisons for these estimators have been carried out in small samples and also we examine the asymptotic distributions of them. The Monte Carlo simulation study indicates that the new estimators are more efficient than the conventional estimators, especially in small samples.

Key-Words:

- *Classical pooling estimation; Contemporaneous covariance; First-order autocorrelation; Heteroskedasticity; Mean group estimation; Random coefficient regression.*

AMS Subject Classification:

- 91G70, 97K80.

1. INTRODUCTION

The econometrics literature reveals a type of data called "panel data", which refers to the pooling of observations on a cross-section of households, countries, and firms over several time periods. Pooling this data achieves a deep analysis of the data and gives a richer source of variation which allows for more efficient estimation of the parameters. With additional, more informative data, one can get more reliable estimates and test more sophisticated behavioral models with less restrictive assumptions. Also, panel data sets are more effective in identifying and estimating effects that are simply not detectable in pure cross-sectional or pure time series data. In particular, panel data sets are more effective in studying complex issues of dynamic behavior. Some of the benefits and limitations of using panel data sets are listed in Baltagi (2013) and Hsiao (2014).

The pooled least squares (classical pooling) estimator for pooled cross-sectional and time series data (panel data) models is the best linear unbiased estimator (BLUE) under the classical assumptions as in the general linear regression model.[1] An important assumption for panel data models is that the individuals in our database are drawn from a population with a common regression coefficient vector. In other words, the coefficients of a panel data model must be fixed. In fact, this assumption is not satisfied in most economic models, see, e.g., Livingston et al. (2010) and Alcacer et al. (2013). In this article, the panel data models are studied when this assumption is relaxed. In this case, the model is called "random-coefficients panel data (RCPD) model". The RCPD model has been examined by Swamy in several publications (Swamy 1970, 1973, and 1974), Rao (1982), Dielman (1992a, b), Beck and Katz (2007), Youssef and Abonazel (2009), and Mousa et al. (2011). Some statistical and econometric publications refer to this model as Swamy's model or as the random coefficient regression (RCR) model, see, e.g., Poi (2003), Abonazel (2009), and Elhorst (2014, ch.3). In RCR model, Swamy assumes that the individuals in our panel data are drawn from a population with a common regression parameter, which is a fixed component, and a random component, that will allow the coefficients to differ from unit to unit. This model has been developed by many researchers, see, e.g., Beran and Millar (1994), Chelliah (1998), Anh and Chelliah (1999), Murtazashvili and Wooldridge (2008), Cheng et al. (2013), Fu and Fu (2015), Elster and Wbbeler (2017), and Horvth and Trapani (2016).

The random-coefficients models have been applied in different fields and they constitute a unifying setup for many statistical problems. Moreover, several applications of Swamy's model have appeared in the literature of finance and economics.[2] Boot and Frankfurter (1972) used the RCR model to examine the optimal mix of short and long-term debt for firms. Feige and Swamy (1974)

[1] Dielman (1983, 1989) discussed these assumptions. In the next section in this article, we will discuss different types of classical pooling estimators under different assumptions.

[2] The RCR model has been applied also in different sciences fields, see, e.g., Bodhlyera et al. (2014).

applied this model to estimate demand equations for liquid assets, while Boness and Frankfurter (1977) used it to examine the concept of risk-classes in finance. Recently, Westerlund and Narayan (2015) used the random-coefficients approach to predictthe stock returns at the New York Stock Exchange. Swamy et al. (2015) applied a random-coefficient framework to deal with two problems frequently encountered in applied work; these problems are correcting for misspecifications in a small area level model and resolving Simpson's paradox.

Dziechciarz (1989) and Hsiao and Pesaran (2008) classified the random-coefficients models into two categories (stationary and non-stationary models), depending on the type of assumption about the coefficient variation. Stationary random-coefficients models regard the coefficients as having constant means and variance-covariances, like Swamy's (1970) model. On the other hand, the coefficients in non-stationary random-coefficients models do not have a constant mean and/or variance and can vary systematically; these models are relevant mainly for modeling the systematic structural variation in time, like the Cooley-Prescott (1973) model.[3]

The main objective of this article is to provide the researchers with general and more efficient estimators for the stationary RCPD models. To achieve this objective, we propose and examine alternative estimators of these models under an assumption that the errors are cross-sectional heteroskedastic and contemporaneously correlated as well as with the first-order autocorrelation of the time series errors.

The rest of the article is organized as follows. Section 2 presents the classical pooling (CP) estimators of fixed-coefficients models. Section 3 provides generalized least squares (GLS) estimators of the different random-coefficients models. In section 4, we examine the efficiency of these estimators, theoretically. In section 5, we discuss alternative estimators for these models. The Monte Carlo comparisons between various estimators have been carried out in section 6. Finally, section 7 offers the concluding remarks.

2. FIXED-COEFFICIENTS MODELS

Suppose the variable y for the ith cross-sectional unit at time period t is specified as a linear function of K strictly exogenous variables, x_{kit}, in the following form:

$$(2.1)\quad y_{it} = \sum_{k=1}^{K} \alpha_{ki} x_{kit} + u_{it} = \mathrm{x}_{it} \alpha_i + u_{it}, \qquad i = 1,\ 2, \ldots,\ N;\ t = 1,\ 2,\ \ldots,\ T,$$

[3] Cooley and Prescott (1973) suggested a model where coefficients vary from one time period to another on the basis of a non-stationary process. Similar models have been considered by Sant (1977) and Rausser et al. (1982).

where u_{it} denotes the random error term, x_{it} is a $1 \times K$ vector of exogenous variables, and α_i is the $K \times 1$ vector of coefficients. Stacking equation (2.1) over time, we obtain:

$$y_i = X_i \alpha_i + u_i, \tag{2.2}$$

where $y_i = (y_{i1}, \ldots, y_{iT})^{'}$, $X_i = \left(\mathrm{x}_{i1}^{'}, \ldots, \mathrm{x}_{iT}^{'}\right)^{'}$, $\alpha_i = (\alpha_{i1}, \ldots, \alpha_{iK})^{'}$, and $u_i = (u_{i1}, \ldots, u_{iT})^{'}$.

When the performance of one individual from the database is of interest, separate equation regressions can be estimated for each individual unit using the ordinary least squares (OLS) method. The OLS estimator of α_i, is given by:

$$\hat{\alpha}_i = \left(X_i^{'} X_i\right)^{-1} X_i^{'} y_i. \tag{2.3}$$

Under the following assumptions, $\hat{\alpha}_i$ is a BLUE of α_i:

Assumption 1: The errors have zero mean, i.e., $E(u_i) = 0$; $\forall\ i = 1, 2, \ldots, N$.

Assumption 2: The errors have the same variance for each individual:

$$E\left(u_i u_j^{'}\right) = \begin{cases} \sigma_u^2 I_T & if \;\; i = j \\ 0 & if \;\; i \neq j \end{cases} \qquad i, j = 1, 2, \ldots, N.$$

Assumption 3: The exogenous variables are non-stochastic, i.e., fixed in repeated samples, and hence, not correlated with the errors. Also, $rank(X_i) = K < T$; $\forall\ i = 1, 2, \ldots, N$.

These conditions are sufficient but not necessary for the optimality of the OLS estimator.[4] When OLS is not optimal, estimation can still proceed equation by equation in many cases. For example, if variance of u_i is not constant, the errors are either heteroskedastic and/or serially correlated, and the GLS method will provide relatively more efficient estimates than OLS, even if GLS was applied to each equation separately as in OLS.

Another case, If the covariances between u_i and $u_j\,(i, j = 1, 2, \ldots, N)$ do not equal to zero as in assumption (2) above, then contemporaneous correlation is present, and we have what Zellner (1962) termed as seemingly unrelated regression (SUR) equations, where the equations are related through cross-equation correlation of errors. If the $X_i\,(i = 1, 2, \ldots, N)$ matrices do not span the same column space and contemporaneous correlation exists, a relatively more efficient estimator of α_i than equation by equation OLS is the GLS estimator applied to the entire equation system, as shown in Zellner (1962).

[4] For more information about the optimality of the OLS estimators, see, e.g., Rao and Mitra (1971, ch. 8) and Srivastava and Giles (1987, pp. 17-21).

With either separate equation estimation or the SUR methodology, we obtain parameter estimates for each individual unit in the database. Now suppose it is necessary to summarize individual relationships and to draw inferences about certain population parameters. Alternatively, the process may be viewed as building a single model to describe the entire group of individuals rather than building a separate model for each. Again, assume that assumptions 1-3 are satisfied and add the following assumption:

Assumption 4: The individuals in the database are drawn from a population with a common regression parameter vector $\bar{\alpha}$, i.e., $\alpha_1 = \alpha_2 = \cdots = \alpha_N = \bar{\alpha}$.

Under this assumption, the observations for each individual can be pooled, and a single regression performed to obtain an efficient estimator of $\bar{\alpha}$. Now, the equation system is written as:

$$Y = X\bar{\alpha} + u, \tag{2.4}$$

where $Y = \left(y_1^{'}, \ \ldots, \ y_N^{'}\right)^{'}$, $X = \left(X_1^{'}, \ldots, \ X_N^{'}\right)^{'}, u = \left(u_1^{'}, \ \ldots, u_N^{'}\right)^{'}$, and $\bar{\alpha} = \left(\bar{\alpha}_1, \ \ldots, \bar{\alpha}_K\right)^{'}$ is a vector of fixed coefficients which to be estimated. We will differentiate between two cases to estimate $\bar{\alpha}$ in (2.4) based on the variance-covariance structure of u. In the first case, the errors have the same variance for each individual as given in assumption 2. In this case, the efficient and unbiased estimator of $\bar{\alpha}$ under assumptions 1-4 is:

$$\widehat{\bar{\alpha}}_{CP-OLS} = \left(X'X\right)^{-1} X'Y.$$

This estimator has been termed the classical pooling-ordinary least squares (CP-OLS) estimator. In the second case, which the errors have different variances along individuals and are contemporaneously correlated as in the SUR framework:

Assumption 5: $E\left(u_i u_j^{'}\right) = \begin{cases} \sigma_{ii} I_T \ \ if \ \ i = j \\ \sigma_{ij} I_T \ \ if \ \ i \neq j \end{cases} \qquad i, j = 1, 2, \ \ldots, \ N.$

Under assumptions 1, 3, 4 and 5, the efficient and unbiased CP estimator of $\bar{\alpha}$ is:

$$\widehat{\bar{\alpha}}_{CP-SUR} = \left[X'(\Sigma_{sur} \otimes I_T)^{-1} X\right]^{-1} \left[X'(\Sigma_{sur} \otimes I_T)^{-1} Y\right],$$

where

$$\Sigma_{sur} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_{NN} \end{pmatrix}.$$

To make this estimator ($\widehat{\bar{\alpha}}_{CP-SUR}$) a feasible, the σ_{ij} can be replaced with the following unbiased and consistent estimator:

$$\hat{\sigma}_{ij} = \frac{\hat{u}_i^{'} \hat{u}_j}{T - K}; \quad \forall \ \ i, j = 1, 2, \ \ldots, \ N, \tag{2.5}$$

where $\hat{u}_i = y_i - X_i \hat{\alpha}_i$ is the residuals vector obtained from applying OLS to equation number i.[5]

3. RANDOM-COEFFICIENTS MODELS

This section reviews the standard random-coefficients model proposed by Swamy (1970), and presents the random-coefficients model in the general case, where the errors are allowed to be cross-sectional heteroskedastic and contemporaneously correlated as well as with the first-order autocorrelation of the time series errors.

3.1. RCR model

Suppose that each regression coefficient in (2.2) is now viewed as a random variable; that is the coefficients, α_i, are viewed as invariant over time, but varying from one unit to another:

Assumption 6: (for the stationary random-coefficients approach): the coefficient vector α_i is specified as:[6] $\alpha_i = \bar{\alpha} + \mu_i$, where $\bar{\alpha}$ is a $K \times 1$ vector of constants, and μ_i is a $K \times 1$ vector of stationary random variables with zero means and constant variance-covariances:

$$E\left(\mu_i\right) = 0 \text{ and } E\left(\mu_i \mu_j^{'}\right) = \begin{cases} \Psi & if\ i = j \\ 0 & if\ i \neq j \end{cases} \qquad i, j = 1,\ 2,\ \ldots,\ N,$$

where $\Psi = diag\left\{\psi_k^2\right\}$; for $k = 1, 2, \ldots, K$, where $K < N$. Furthermore, $E\left(\mu_i \mathrm{x}_{jt}\right) = 0$ and $E\left(\mu_i u_{jt}\right) = 0 \; \forall \; i$ and j.

Also, Swamy (1970) assumed that the errors have different variances along individuals:

Assumption 7: $E\left(u_i u_j^{'}\right) = \begin{cases} \sigma_{ii} I_T & if\ \ i = j \\ 0 & if\ \ i \neq j \end{cases} \qquad i, j = 1, 2,\ \ldots,\ N.$

Under the assumption 6, the model in equation (2.2) can be rewritten as:

$$Y = X\bar{\alpha} + e; \;\; e = D\mu + u, \tag{3.1}$$

[5]The $\hat{\sigma}_{ij}$ in (2.5) are unbiased estimators because, as assumed, the number of exogenous variables of each equation is equal, i.e., $K_i = K$ for $i = 1, 2,\ \ldots,\ N$. However, in the general case, $K_i \neq K_j$, the unbiased estimator is $\hat{u}_i^{'} \hat{u}_j / \left[T - K_i - K_j + tr\left(P_{xx}\right)\right]$, where $P_{xx} = X_i \left(X_i^{'} X_i\right)^{-1} X_i^{'} X_j \left(X_j^{'} X_j\right)^{-1} X_j^{'}$. See Srivastava and Giles (1987, pp. 13–17) and Baltagi (2011, pp. 243–244).

[6]This means that the individuals in our database are drown from a population with a common regression parameter $\bar{\alpha}$, which is a fixed component, and a random component μ_i, allowed to differ from unit to unit.

where Y, X, u, and $\bar{\alpha}$ are defined in (2.4), while $\mu = \left(\mu_1', \ldots, \mu_N'\right)'$, and $D = diag\{X_i\}$; for $i = 1, 2, \ldots, N$.

The model in (3.1), under assumptions 1, 3, 6 and 7, called the 'RCR model', which was examined by Swamy (1970, 1971, 1973, and 1974), Youssef and Abonazel (2009), and Mousa et al. (2011). We will refer to assumptions 1, 3, 6 and 7 as RCR assumptions. Under these assumptions, the BLUE of $\bar{\alpha}$ in equation (3.1) is:

$$\hat{\bar{\alpha}}_{RCR} = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}Y,$$

where Ω is the variance-covariance matrix of e:

$$\Omega = (\Sigma_{rcr} \otimes I_T) + D\,(I_N \otimes \Psi)\, D',$$

where $\Sigma_{rcr} = diag\{\sigma_{ii}\}$; for $i = 1, 2, \ldots, N$. Swamy (1970) showed that the $\hat{\bar{\alpha}}_{RCR}$ estimator can be rewritten as:

$$\hat{\bar{\alpha}}_{RCR} = \left[\sum_{i=1}^{N} X_i'\left(X_i \Psi X_i' + \sigma_{ii} I_T\right)^{-1} X_i\right]^{-1} \sum_{i=1}^{N} X_i'\left(X_i \Psi X_i' + \sigma_{ii} I_T\right)^{-1} y_i.$$

The variance-covariance matrix of $\hat{\bar{\alpha}}_{RCR}$ under RCR assumptions is:

$$var\left(\hat{\bar{\alpha}}_{RCR}\right) = \left(X'\Omega^{-1}X\right)^{-1} = \left\{\sum_{i=1}^{N}\left[\Psi + \sigma_{ii}\left(X_i'X_i\right)^{-1}\right]^{-1}\right\}^{-1}.$$

To make the $\hat{\bar{\alpha}}_{RCR}$ estimator feasible, Swamy (1971) suggested using the estimator in (2.5) as an unbiased and consistent estimator of σ_{ii}, and the following unbiased estimator for Ψ:

$$\hat{\Psi} = \left[\frac{1}{N-1}\left(\sum_{i=1}^{N} \hat{\alpha}_i\, \hat{\alpha}_i' - \frac{1}{N}\sum_{i=1}^{N} \hat{\alpha}_i \sum_{i=1}^{N} \hat{\alpha}_i'\right)\right] - \left[\frac{1}{N}\sum_{i=1}^{N} \hat{\sigma}_{ii}\left(X_i'X_i\right)^{-1}\right]. \tag{3.2}$$

Swamy (1973, 1974) showed that the estimatorv $\hat{\bar{\alpha}}_{RCR}$ is consistent as both $N, T \to \infty$ and is asymptotically efficient as $T \to \infty$.[7]

It is worth noting that, just as in the error-components model, the estimator (3.2) is not necessarily non-negative definite. Mousa et al. (2011) explained that it is possible to obtain negative estimates of Swamy's estimator in (3.2) in case of small samples and if some/all coefficients are fixed. But in medium and large samples, the negative variance estimates does not appear even if all coefficients are fixed. To solve this problem, Swamy has suggested replacing (3.2) by:[8]

$$\hat{\Psi}^{+} = \frac{1}{N-1}\left(\sum_{i=1}^{N} \hat{\alpha}_i\, \hat{\alpha}_i' - \frac{1}{N}\sum_{i=1}^{N} \hat{\alpha}_i \sum_{i=1}^{N} \hat{\alpha}_i'\right).$$

[7] The statistical properties of $\hat{\bar{\alpha}}_{RCR}$ have been examined by Swamy (1971), of course, under RCR assumptions.

[8] This suggestion has been used by Stata program, specifically in xtrchh and xtrchh2 Stata's commands. See Poi (2003).

This estimator, although biased, is non-negative definite and consistent when $T \to \infty$. See Judge et al. (1985, p. 542).

3.2. Generalized RCR model

To generalize RCR model so that it would be more suitable for most economic models, we assume that the errors are cross-sectional heteroskedastic and contemporaneously correlated, as in assumption 5, as well as with the first-order autocorrelation of the time series errors. Therefore, we add the following assumption to assumption 5:

Assumption 8: $u_{it} = \rho_i u_{i,t-1} + \varepsilon_{it}$; $|\rho_i| < 1$, where ρ_i $(i = 1, 2, \ldots, N)$ are fixed first-order autocorrelation coefficients. Assume that: $E(\varepsilon_{it}) = 0$, $E(u_{i,t-1}\varepsilon_{jt}) = 0$; $\forall$ i and j, and

$$E\left(\varepsilon_i \varepsilon_j^{'}\right) = \begin{cases} \sigma_{\varepsilon_{ii}} I_T & if \;\; i = j \\ \sigma_{\varepsilon_{ij}} I_T & if \;\; i \neq j \end{cases} \qquad i, j = 1, 2, \ \ldots, \ N.$$

This means that the initial time period of the errors have the same properties as in subsequent periods, i.e., $E\left(u_{i0}^2\right) = \sigma_{\varepsilon_{ii}} / \left(1 - \rho_i^2\right)$ and $E\left(u_{i0} u_{j0}\right) = \sigma_{\varepsilon_{ij}} / \left(1 - \rho_i \rho_j\right)$ $\forall$ i and j.

We will refer to assumptions 1, 3, 5, 6, and 8 as the general RCR assumptions. Under these assumptions, the BLUE of $\bar{\alpha}$ is:

$$\widehat{\bar{\alpha}}_{GRCR} = \left(X^{'} \Omega^{*-1} X\right)^{-1} X^{'} \Omega^{*-1} Y,$$

where

(3.3)

$$\Omega^* = \begin{pmatrix} X_1 \Psi X_1^{'} + \sigma_{\varepsilon_{11}} \omega_{11} & \sigma_{\varepsilon_{12}} \omega_{12} & \cdots & \sigma_{\varepsilon_{1N}} \omega_{1N} \\ \sigma_{\varepsilon_{21}} \omega_{21} & X_2 \Psi X_2^{'} + \sigma_{\varepsilon_{22}} \omega_{22} & \cdots & \sigma_{\varepsilon_{2N}} \omega_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{\varepsilon_{N1}} \omega_{N1} & \sigma_{\varepsilon_{N2}} \omega_{N2} & \cdots & X_N \Psi X_N^{'} + \sigma_{\varepsilon_{NN}} \omega_{NN} \end{pmatrix},$$

with

$$\omega_{ij} = \frac{1}{1 - \rho_i \rho_j} \begin{pmatrix} 1 & \rho_i & \rho_i^2 & \cdots & \rho_i^{T-1} \\ \rho_j & 1 & \rho_i & \cdots & \rho_i^{T-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_j^{T-1} & \rho_j^{T-2} & \rho_j^{T-3} & \cdots & 1 \end{pmatrix}. \tag{3.4}$$

Since the elements of Ω^* are usually unknown, we develop a feasible Aitken estimator of $\bar{\alpha}$ based on consistent estimators of the elements of Ω^*:

$$\hat{\rho}_i = \frac{\sum_{t=2}^{T} \hat{u}_{it} \hat{u}_{i,t-1}}{\sum_{t=2}^{T} \hat{u}_{i,t-1}^2}, \tag{3.5}$$

where $\hat{u}_i = (\hat{u}_{i1}, \ \ldots, \hat{u}_{iT})'$ is the residuals vector obtained from applying OLS to equation number i,

$$\hat{\sigma}_{\varepsilon_{ij}} = \frac{\hat{\varepsilon}_i' \hat{\varepsilon}_j}{T-K},$$

where $\hat{\varepsilon}_i = (\hat{\varepsilon}_{i1}, \ldots, \ \hat{\varepsilon}_{iT})'$; $\hat{\varepsilon}_{i1} = \hat{u}_{i1}\sqrt{1-\hat{\rho}_i^2}$, and $\hat{\varepsilon}_{it} = \hat{u}_{it} - \hat{\rho}_i \hat{u}_{i,t-1}$ for $t = 2, \ldots, T$.

Replacing ρ_i by $\hat{\rho}_i$ in (3.4), yields consistent estimators of ω_{ij}, say $\hat{\omega}_{ij}$, which leads together with $\hat{\sigma}_{\varepsilon_{ij}}$ and $\hat{\omega}_{ij}$ to a consistent estimator of Ψ:[9]

$$\begin{aligned}\hat{\Psi}^* = {} & \tfrac{1}{N-1}\left(\sum_{i=1}^{N} \hat{\alpha}_i^* \hat{\alpha}_i^{*'} - \tfrac{1}{N}\sum_{i=1}^{N} \hat{\alpha}_i^* \sum_{i=1}^{N} \hat{\alpha}_i^{*'}\right) - \tfrac{1}{N}\sum_{i=1}^{N} \hat{\sigma}_{\varepsilon_{ii}} \left(X_i' \hat{\omega}_{ii}^{-1} X_i\right)^{-1} \\ & + \left[\begin{array}{l} \tfrac{1}{N(N-1)} \displaystyle\sum_{\substack{i \neq j \\ i,j=1}}^{N} \hat{\sigma}_{\varepsilon_{ij}} \left(X_i' \hat{\omega}_{ii}^{-1} X_i\right)^{-1} \\ X_i' \hat{\omega}_{ii}^{-1} \hat{\omega}_{ij} \hat{\omega}_{jj}^{-1} X_j \left(X_j' \hat{\omega}_{jj}^{-1} X_j\right)^{-1} \end{array}\right],\end{aligned} \tag{3.6}$$

where

$$\hat{\alpha}_i^* = \left(X_i' \hat{\omega}_{ii}^{-1} X_i\right)^{-1} X_i' \hat{\omega}_{ii}^{-1} y_i. \tag{3.7}$$

By using the consistent estimators ($\hat{\sigma}_{\varepsilon_{ij}}$, $\hat{\omega}_{ij}$, and $\hat{\Psi}^*$) in (3.3), and proceed a consistent estimator of Ω^* is obtained, say $\hat{\Omega}^*$, that leads to get the generalized RCR (GRCR) estimator of $\bar{\alpha}$:

$$\hat{\bar{\alpha}}_{GRCR} = \left(X' \hat{\Omega}^{*-1} X\right)^{-1} X' \hat{\Omega}^{*-1} Y.$$

The estimated variance-covariance matrix of $\hat{\bar{\alpha}}_{GRCR}$ is:

$$\widehat{var}\left(\hat{\bar{\alpha}}_{GRCR}\right) = \left(X' \hat{\Omega}^{*-1} X\right)^{-1}. \tag{3.8}$$

4. EFFICIENCY GAINS

In this section, we examine the efficiency gains from the use of GRCR estimator. Under the general RCR assumptions, It is easy to verify that the classical pooling estimators ($\hat{\bar{\alpha}}_{CP-OLS}$ and $\hat{\bar{\alpha}}_{CP-SUR}$) and Swamy's estimator ($\hat{\bar{\alpha}}_{RCR}$) are unbiased for $\bar{\alpha}$ and with variance-covariance matrices:

$$\begin{aligned} var\left(\hat{\bar{\alpha}}_{CP-OLS}\right) &= G_1 \Omega^* G_1'; \\ var\left(\hat{\bar{\alpha}}_{CP-SUR}\right) &= G_2 \Omega^* G_2'; \\ var\left(\hat{\bar{\alpha}}_{RCR}\right) &= G_3 \Omega^* G_3', \end{aligned}$$

[9]The estimator of ρ_i in (3.5) is consistent, but it is not unbiased. See Srivastava and Giles (1987, p. 211) for other suitable consistent estimators of ρ_i that are often used in practice.

where

$$G_1 = \left(X'X\right)^{-1}X';$$
$$G_2 = \left[X'\left(\Sigma_{sur}^{-1} \otimes I_T\right)X\right]^{-1}X'\left(\Sigma_{sur}^{-1} \otimes I_T\right); \tag{4.1}$$
$$G_3 = \left(X'\Omega^{-1}X\right)^{-1}X'\Omega^{-1}.$$

The efficiency gains, from the use of GRCR estimator, can be summarized in the following equation:

$$EG_\gamma = var\left(\widehat{\bar{\alpha}}_\gamma\right) - var\left(\widehat{\bar{\alpha}}_{GRCR}\right) = \left(G_h - G_0\right)\Omega^*\left(G_h - G_0\right)';\ \text{for } h = 1, 2, 3,$$

where the subscript γ indicates the estimator that is used (CP-OLS, CP-SUR, or RCR), $G_0 = \left(X'\Omega^{*-1}X\right)^{-1}X'\Omega^{*-1}$, and G_h (for $h = 1, 2, 3$) matrices are defined in (4.1).

Since Ω^*, Σ_{rcr}, Σ_{sur} and Ω are positive definite matrices, then EG_γ matrices are positive semi-definite matrices. In other words, the GRCR estimator is more efficient than CP-OLS, CP-SUR, and RCR estimators. These efficiency gains increase when $|\rho_i|$, $\sigma_{\varepsilon_{ij}}$, and ψ_k^2 increase. However, it is not clear to what extent these efficiency gains hold in small samples. Therefore, this will be examined in a simulation study.

5. ALTERNATIVE ESTIMATORS

A consistent estimator of $\bar{\alpha}$ can also be obtained under more general assumptions concerning α_i and the regressors. One such possible estimator is the mean group (MG) estimator, proposed by Pesaran and Smith (1995) for estimation of dynamic panel data (DPD) models with random coefficients.[10] The MG estimator is defined as the simple average of the OLS estimators:

$$\widehat{\bar{\alpha}}_{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}_i. \tag{5.1}$$

Even though the MG estimator has been used in DPD models with random coefficients, it will be used here as one of alternative estimators of static panel data models with random coefficients. Note that the simple MG estimator in (5.1) is more suitable for the RCR Model. But to make it suitable for the GRCR model, we suggest a general mean group (GMG) estimator as:

$$\widehat{\bar{\alpha}}_{GMG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}_i^*, \tag{5.2}$$

[10] For more information about the estimation methods for DPD models, see, e.g., Baltagi (2013), Abonazel (2014, 2017), Youssef et al. (2014a,b), and Youssef and Abonazel (2017).

where $\hat{\alpha}_i^*$ is defined in (3.7).

Lemma 5.1. *If the general RCR assumptions are satisfied, then* $\widehat{\bar{\alpha}}_{MG}$ *and* $\widehat{\bar{\alpha}}_{GMG}$ *are unbiased estimators of* $\bar{\alpha}$*, with the estimated variance-covariance matrices of* $\widehat{\bar{\alpha}}_{MG}$ *and* $\widehat{\bar{\alpha}}_{GMG}$ *are:*

$$\begin{aligned} \widehat{var}\left(\widehat{\bar{\alpha}}_{MG}\right) = & \frac{1}{N}\hat{\Psi}^* + \frac{1}{N^2}\sum_{i=1}^{N}\hat{\sigma}_{\varepsilon_{ii}}\left(X_i'X_i\right)^{-1}X_i'\hat{\omega}_{ii}X_i\left(X_i'X_i\right)^{-1} \\ & + \frac{1}{N^2}\sum_{\substack{i \neq j \\ i,j=1}}^{N}\hat{\sigma}_{\varepsilon_{ij}}\left(X_i'X_i\right)^{-1}X_i'\hat{\omega}_{ij}X_j\left(X_j'X_j\right)^{-1}, \end{aligned} \tag{5.3}$$

$$\widehat{var}\left(\widehat{\bar{\alpha}}_{GMG}\right) = \frac{1}{N(N-1)}\left[\begin{array}{l} \sum_{i=1}^{N}\hat{\alpha}_i^*\hat{\alpha}_i^{*\prime} - \frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}_i^*\sum_{i=1}^{N}\hat{\alpha}_i^{*\prime} \\ + \sum_{\substack{i \neq j \\ i,j=1}}^{N}\hat{\sigma}_{\varepsilon_{ij}}\left(X_i'\hat{\omega}_{ii}^{-1}X_i\right)^{-1} \\ X_i'\hat{\omega}_{ii}^{-1}\hat{\omega}_{ij}\hat{\omega}_{jj}^{-1}X_j\left(X_j'\hat{\omega}_{jj}^{-1}X_j\right)^{-1} \end{array}\right]. \tag{5.4}$$

Proof of Lemma 5.1:

a. Unbiasedness property of MG and GMG estimators:

Proof: By substituting (3.7) and (2.2) into (5.2):

$$\begin{aligned} \widehat{\bar{\alpha}}_{GMG} &= \frac{1}{N}\sum_{i=1}^{N}\left(X_i'\omega_{ii}^{-1}X_i\right)^{-1}X_i'\omega_{ii}^{-1}\left(X_i\alpha_i + u_i\right) \\ &= \frac{1}{N}\sum_{i=1}^{N}\alpha_i + \left(X_i'\omega_{ii}^{-1}X_i\right)^{-1}X_i'\omega_{ii}^{-1}u_i. \end{aligned} \tag{5.5}$$

Similarly, we can rewrite $\widehat{\bar{\alpha}}_{MG}$ in (5.1) as:

$$\widehat{\bar{\alpha}}_{MG} = \frac{1}{N}\sum_{i=1}^{N}\alpha_i + \left(X_i'X_i\right)^{-1}X_i'u_i. \tag{5.6}$$

Taking the expectation for (5.5) and (5.6), and using assumptions 1 and 6:

$$E\left(\widehat{\bar{\alpha}}_{GMG}\right) = E\left(\widehat{\bar{\alpha}}_{MG}\right) = \frac{1}{N}\sum_{i=1}^{N}\bar{\alpha} = \bar{\alpha}.$$

□

b. Derive the variance-covariance matrix of GMG:

Proof: Note first that under assumption 6, $\alpha_i = \bar{\alpha} + \mu_i$. Add $\hat{\alpha}_i^*$ to the both sides:

$$\alpha_i + \hat{\alpha}_i^* = \bar{\alpha} + \mu_i + \hat{\alpha}_i^*,$$

$$\hat{\alpha}_i^* = \bar{\alpha} + \mu_i + \hat{\alpha}_i^* - \alpha_i = \bar{\alpha} + \mu_i + \tau_i, \tag{5.7}$$

where $\tau_i = \hat{\alpha}_i^* - \alpha_i = \left(X_i'\omega_{ii}^{-1}X_i\right)^{-1} X_i'\omega_{ii}^{-1}u_i$. From (5.7):

$$\frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}_i^* = \bar{\alpha} + \frac{1}{N}\sum_{i=1}^{N}\mu_i + \frac{1}{N}\sum_{i=1}^{N}\tau_i,$$

which means that

$$\widehat{\bar{\alpha}}_{GMG} = \bar{\alpha} + \bar{\mu} + \bar{\tau}, \tag{5.8}$$

where $\bar{\mu} = \frac{1}{N}\sum\limits_{i=1}^{N}\mu_i$ and $\bar{\tau} = \frac{1}{N}\sum\limits_{i=1}^{N}\tau_i$. From (5.8) and using the general RCR assumptions:

$$\begin{aligned}
var\left(\widehat{\bar{\alpha}}_{GMG}\right) &= var\left(\bar{\mu}\right) + var\left(\bar{\tau}\right) \\
&= \frac{1}{N}\Psi + \frac{1}{N^2}\sum_{i=1}^{N}\sigma_{\varepsilon_{ii}}\left(X_i'\omega_{ii}^{-1}X_i\right)^{-1} \\
&+ \frac{1}{N^2}\sum_{\substack{i \neq j \\ i,j=1}}^{N}\sigma_{\varepsilon_{ij}}\left(X_i'\omega_{ii}^{-1}X_i\right)^{-1}X_i'\omega_{ii}^{-1}\omega_{ij}\omega_{jj}^{-1}X_j\left(X_j'\omega_{jj}^{-1}X_j\right)^{-1}.
\end{aligned}$$

Using the consistent estimators of Ψ, $\sigma_{\varepsilon_{ij}}$, and ω_{ij} defined above, then we get the formula of $\widehat{var}\left(\widehat{\bar{\alpha}}_{GMG}\right)$ as in equation (5.4). $\square$

c. Derive the variance-covariance matrix of MG:

Proof: As above, equation (2.3) can be rewritten as follows:

$$\hat{\alpha}_i = \bar{\alpha} + \mu_i + \lambda_i, \tag{5.9}$$

where $\lambda_i = \hat{\alpha}_i - \alpha_i = \left(X_i'X_i\right)^{-1}X_i'u_i$. From (5.9):

$$\frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}_i = \bar{\alpha} + \frac{1}{N}\sum_{i=1}^{N}\mu_i + \frac{1}{N}\sum_{i=1}^{N}\lambda_i,$$

which means that

$$\widehat{\bar{\alpha}}_{MG} = \bar{\alpha} + \bar{\mu} + \bar{\lambda}, \tag{5.10}$$

where $\bar{\mu} = \frac{1}{N}\sum_{i=1}^{N}\mu_i$, and $\bar{\lambda} = \frac{1}{N}\sum_{i=1}^{N}\lambda_i$. From (5.10) and using the general RCR assumptions:

$$var\left(\widehat{\bar{\alpha}}_{MG}\right) = var\left(\bar{\mu}\right) + var\left(\bar{\lambda}\right)$$
$$= \frac{1}{N}\Psi + \frac{1}{N^2}\sum_{i=1}^{N}\sigma_{\varepsilon_{ii}}\left(X_i'X_i\right)^{-1}X_i'\omega_{ii}X_i\left(X_i'X_i\right)^{-1}$$
$$+ \frac{1}{N^2}\sum_{\substack{i \neq j\\ i,j=1}}^{N}\sigma_{\varepsilon_{ij}}\left(X_i'X_i\right)^{-1}X_i'\omega_{ij}X_j\left(X_j'X_j\right)^{-1}.$$

As in the GMG estimator, and by using the consistent estimators of Ψ, $\sigma_{\varepsilon_{ij}}$, and ω_{ij}, then we get the formula of $\widehat{var}\left(\widehat{\bar{\alpha}}_{GM}\right)$ as in equation (5.3). □

It is noted from lemma 1 that the variance of the GMG estimator is less than the variance of the MG estimator when the general RCR assumptions are satisfied. In other words, the GMG estimator is more efficient than the MG estimator. But under RCR assumptions, we have:

$$var\left(\widehat{\bar{\alpha}}_{MG}\right) = var\left(\widehat{\bar{\alpha}}_{GMG}\right) = \frac{1}{N\left(N-1\right)}\left(\sum_{i=1}^{N}\alpha_i\alpha_i' - \frac{1}{N}\sum_{i=1}^{N}\alpha_i\sum_{i=1}^{N}\alpha_i'\right) = \frac{1}{N}\Psi^{+}.$$

The next lemma explains the asymptotic variances (as $T \to \infty$ with N fixed) properties of GRCR, RCR, GMG, and MG estimators. In order to justify the derivation of the asymptotic variances, we must assume the following:

Assumption 9: $\operatorname*{plim}_{T\to\infty} T^{-1}X_i'X_i$ and $\operatorname*{plim}_{T\to\infty} T^{-1}X_i'\hat{\omega}_{ii}^{-1}X_i$ are finite and positive definite for all i and for $|\rho_i| < 1$.

Lemma 5.2. *If the general RCR assumptions and assumption 9 are satisfied, then the estimated asymptotic variance-covariance matrices of GRCR, RCR, GMG, and MG estimators are equal:*

$$\operatorname*{plim}_{T\to\infty}\widehat{var}\left(\widehat{\bar{\alpha}}_{GRCR}\right) = \operatorname*{plim}_{T\to\infty}\widehat{var}\left(\widehat{\bar{\alpha}}_{RCR}\right) = \operatorname*{plim}_{T\to\infty}\widehat{var}\left(\widehat{\bar{\alpha}}_{GMG}\right)$$
$$= \operatorname*{plim}_{T\to\infty}\widehat{var}\left(\widehat{\bar{\alpha}}_{MG}\right) = N^{-1}\Psi^{+}.$$

Proof of Lemma 5.2: Following the same argument as in Parks (1967) and utilizing assumption 9, we can show that:

$$\begin{aligned}&\operatorname*{plim}_{T\to\infty}\hat{\alpha}_i = \operatorname*{plim}_{T\to\infty}\hat{\alpha}_i^{*} = \alpha_i, \ \operatorname*{plim}_{T\to\infty}\hat{\rho}_{ij} = \rho_{ij},\\ &\operatorname*{plim}_{T\to\infty}\hat{\sigma}_{\varepsilon_{ij}} = \sigma_{\varepsilon_{ij}}, \text{and} \operatorname*{plim}_{T\to\infty}\hat{\omega}_{ij} = \omega_{ij},\end{aligned} \tag{5.11}$$

and then

(5.12)
$$\begin{aligned}\underset{T\to\infty}{\operatorname{plim}} \frac{1}{T}\hat{\sigma}_{\varepsilon_{ii}} T\left(X_i'\hat{\omega}_{ii}^{-1}X_i\right)^{-1} &= \underset{T\to\infty}{\operatorname{plim}} \frac{1}{T}\hat{\sigma}_{\varepsilon_{ii}} T\left(X_i'X_i\right)^{-1}X_i'\hat{\omega}_{ii}X_i\left(X_i'X_i\right)^{-1} \\ &= \underset{T\to\infty}{\operatorname{plim}} \frac{1}{T}\hat{\sigma}_{\varepsilon_{ij}} T\left(X_i'X_i\right)^{-1}X_i'\hat{\omega}_{ij}X_j\left(X_j'X_j\right)^{-1} \\ &= \underset{T\to\infty}{\operatorname{plim}} \frac{1}{T}\hat{\sigma}_{\varepsilon_{ij}} T\left(X_i'\hat{\omega}_{ii}^{-1}X_i\right)^{-1}X_i'\hat{\omega}_{ii}^{-1}\hat{\omega}_{ij}\hat{\omega}_{jj}^{-1}X_j \\ &\quad \left(X_j'\hat{\omega}_{jj}^{-1}X_j\right)^{-1} = 0.\end{aligned}$$

Substituting (5.11) and (5.12) in (3.6):

$$\underset{T\to\infty}{\operatorname{plim}} \hat{\Psi}^* = \frac{1}{N-1}\left(\sum_{i=1}^{N}\alpha_i\alpha_i' - \frac{1}{N}\sum_{i=1}^{N}\alpha_i\sum_{i=1}^{N}\alpha_i'\right) = \Psi^+. \tag{5.13}$$

By substituting (5.11)-(5.13) into (5.3), (5.4), and (3.8):

(5.14)
$$\begin{aligned}\underset{T\to\infty}{\operatorname{plim}} \widehat{var}\left(\widehat{\bar{\alpha}}_{MG}\right) = &\ \frac{1}{N}\underset{T\to\infty}{\operatorname{plim}} \hat{\Psi}^* \\ &+ \frac{1}{N^2}\sum_{i=1}^{N}\underset{T\to\infty}{\operatorname{plim}} \frac{1}{T}\hat{\sigma}_{\varepsilon_{ii}} T\left(X_i'X_i\right)^{-1}X_i'\hat{\omega}_{ii}X_i\left(X_i'X_i\right)^{-1} \\ &+ \frac{1}{N^2}\sum_{\substack{i\neq j \\ i,j=1}}^{N}\underset{T\to\infty}{\operatorname{plim}} \frac{1}{T}\hat{\sigma}_{\varepsilon_{ij}} T\left(X_i'X_i\right)^{-1}X_i'\hat{\omega}_{ij}X_j\left(X_j'X_j\right)^{-1} \\ = &\ \frac{1}{N}\Psi^+,\end{aligned}$$

$$\begin{aligned}\underset{T\to\infty}{\operatorname{plim}} \widehat{var}\left(\widehat{\bar{\alpha}}_{GMG}\right) = &\ \tfrac{1}{N(N-1)}\underset{T\to\infty}{\operatorname{plim}}\left(\sum_{i=1}^{N}\hat{\alpha}_i^*\hat{\alpha}_i^{*'} - \tfrac{1}{N}\sum_{i=1}^{N}\hat{\alpha}_i^*\sum_{i=1}^{N}\hat{\alpha}_i^{*'}\right) \\ &+\tfrac{1}{N(N-1)}\sum_{\substack{i\neq j \\ i,j=1}}^{N}\left[\begin{array}{l}\underset{T\to\infty}{\operatorname{plim}} \tfrac{1}{T}\hat{\sigma}_{\varepsilon_{ij}} T\left(X_i'\hat{\omega}_{ii}^{-1}X_i\right)^{-1} \\ X_i'\hat{\omega}_{ii}^{-1}\hat{\omega}_{ij}\hat{\omega}_{jj}^{-1}X_j\left(X_j'\hat{\omega}_{jj}^{-1}X_j\right)^{-1}\end{array}\right] = \tfrac{1}{N}\Psi^+,\end{aligned} \tag{5.15}$$

$$\underset{T\to\infty}{\operatorname{plim}} \widehat{var}\left(\widehat{\bar{\alpha}}_{GRCR}\right) = \underset{T\to\infty}{\operatorname{plim}}\left(X'\hat{\Omega}^{*-1}X\right)^{-1} = \left[\sum_{i=1}^{N}\Psi^{+-1}\right]^{-1} = \frac{1}{N}\Psi^+. \tag{5.16}$$

Similarly, we will use the results in (5.11)-(5.13) in case of RCR estimator:

$$\begin{aligned}\underset{T\to\infty}{\operatorname{plim}} \widehat{var}\left(\widehat{\bar{\alpha}}_{RCR}\right) &= \underset{T\to\infty}{\operatorname{plim}}\left[\left(X'\hat{\Omega}^{-1}X\right)^{-1}X'\hat{\Omega}^{-1}\hat{\Omega}^*\,\hat{\Omega}^{-1}X\left(X'\hat{\Omega}^{-1}X\right)^{-1}\right] \\ &= \frac{1}{N}\Psi^+.\end{aligned} \tag{5.17}$$

From (5.14)-(5.17), we can conclude that:

$$\begin{aligned}\operatorname*{plim}_{T\to\infty} \widehat{var}\left(\widehat{\bar{\alpha}}_{GRCR}\right) &= \operatorname*{plim}_{T\to\infty} \widehat{var}\left(\widehat{\bar{\alpha}}_{RCR}\right) \\ &= \operatorname*{plim}_{T\to\infty} \widehat{var}\left(\widehat{\bar{\alpha}}_{GMG}\right) = \operatorname*{plim}_{T\to\infty} \widehat{var}\left(\widehat{\bar{\alpha}}_{MG}\right) = \frac{1}{N}\Psi^{+}.\end{aligned}$$

□

From Lemma 5.2, we can conclude that the means and the variance-covariance matrices of the limiting distributions of $\widehat{\bar{\alpha}}_{GRCR}$, $\widehat{\bar{\alpha}}_{RCR}$, $\widehat{\bar{\alpha}}_{GMG}$, and $\widehat{\bar{\alpha}}_{MG}$ are the same and are equal to $\bar{\alpha}$ and $N^{-1}\Psi$ respectively even if the errors are correlated as in assumption 8. it is not expected to increase the asymptotic efficiency of $\widehat{\bar{\alpha}}_{GRCR}$, $\widehat{\bar{\alpha}}_{RCR}$, $\widehat{\bar{\alpha}}_{GMG}$, and $\widehat{\bar{\alpha}}_{MG}$. This does not mean that the GRCR estimator cannot be more efficient than RCR, GMG, and MG in small samples when the errors are correlated as in assumption 8. This will be examined in our simulation study.

6. MONTE CARLO SIMULATION

In this section, the Monte Carlo simulation has been used for making comparisons between the behavior of the classical pooling estimators (CP-OLS and CP-SUR), random-coefficients estimators (RCR and GRCR), and mean group estimators (MG and GMG) in small and moderate samples. The program to set up the Monte Carlo simulation, written in the R language, is available upon request. Monte Carlo experiments were carried out based on the following data generating process:

$$y_{it} = \sum_{k=1}^{3} \alpha_{ki} x_{kit} + u_{it}, \qquad i = 1,\ 2, \ldots,\ N;\ t = 1,\ 2, \ldots,\ T. \tag{6.1}$$

To perform the simulation under the general RCR assumptions, the model in (6.1) was generated as follows:

1. The independent variables, $(x_{kit}; k = 1,\ 2,\ 3)$, were generated as independent standard normally distributed random variables. The values of x_{kit} were allowed to differ for each cross-sectional unit. However, once generated for all N cross-sectional units the values were held fixed over all Monte Carlo trials.
2. The errors, u_{it}, were generated as in assumption 8: $u_{it} = \rho u_{i,t-1} + \varepsilon_{it}$, where the values of $\varepsilon_i = (\varepsilon_{i1},\ \ldots, \varepsilon_{iT})'$ $\forall\, i = 1,\ 2,\ \ldots,\ N$ were generated as multivariate normally distributed with means zeros and variance-covariance

Table 1: ATSE for various estimators when $\sigma_{\varepsilon_{ii}} = 1$ and $N < T$.

$(\rho,\ \sigma_{\varepsilon_{ij}})$	(0, 0)				(0.55, 0.75)				(0.85, 0.95)			
$(N,\ T)$	(5, 8)	(5, 12)	(10, 15)	(10, 20)	(5, 8)	(5, 12)	(10, 15)	(10, 20)	(5, 8)	(5, 12)	(10, 15)	(10, 20)
						$\mu_i = 0$						
CP-OLS	0.920	0.746	0.440	0.436	0.857	0.888	0.409	0.450	1.107	1.496	0.607	0.641
CP-SUR	0.958	0.767	0.419	0.417	0.829	0.880	0.381	0.384	0.947	1.469	0.453	0.532
MG	0.947	0.765	0.470	0.469	0.886	0.910	0.442	0.468	1.133	1.475	0.608	0.636
GMG	0.702	0.556	0.369	0.375	0.638	0.662	0.289	0.305	0.644	1.098	0.302	0.291
RCR	1.012	30.746	0.517	0.497	1.064	1.130	2.241	0.726	1.365	5.960	0.856	1.326
GRCR	0.754	0.624	0.352	0.357	0.634	0.703	0.302	0.295	0.735	1.141	0.324	0.388
						$\mu_i \sim N(0,\ 5)$						
CP-OLS	4.933	4.682	2.320	2.742	2.588	2.902	2.598	2.130	3.627	5.079	2.165	2.935
CP-SUR	5.870	5.738	2.852	3.411	3.143	3.456	3.212	2.592	4.011	5.906	2.668	3.549
MG	4.057	4.112	2.086	2.494	2.173	2.478	2.352	1.888	3.094	4.040	1.938	2.626
GMG	4.057	4.110	2.084	2.494	2.176	2.479	2.348	1.879	3.052	4.024	1.908	2.606
RCR	4.053	4.114	2.083	2.493	2.632	3.304	2.352	1.888	3.287	6.422	2.052	2.648
GRCR	4.030	4.092	2.067	2.480	2.104	2.413	2.331	1.855	2.969	3.905	1.865	2.578
						$\mu_i \sim N(0,\ 25)$						
CP-OLS	7.528	7.680	7.147	6.341	8.293	8.156	6.321	6.739	7.942	7.214	4.691	6.423
CP-SUR	8.866	9.439	8.935	8.046	10.104	9.880	8.028	8.402	9.074	8.482	5.739	7.937
MG	6.272	6.549	6.324	5.597	6.879	6.650	5.541	5.917	6.442	6.083	4.118	5.672
GMG	6.271	6.548	6.324	5.597	6.881	6.650	5.538	5.913	6.422	6.078	4.103	5.662
RCR	6.271	6.548	6.324	5.597	6.885	6.657	5.541	5.917	7.546	6.098	4.122	5.686
GRCR	6.251	6.539	6.319	5.590	6.857	6.626	5.530	5.906	6.389	6.010	4.082	5.649
						$\mu_i \sim t(5)$						
CP-OLS	2.253	1.983	1.562	1.544	1.479	1.977	1.060	1.223	2.115	3.301	1.470	1.439
CP-SUR	2.626	2.419	1.925	1.912	1.694	2.266	1.275	1.454	2.403	3.903	1.717	1.643
MG	1.859	1.776	1.410	1.401	1.324	1.722	0.984	1.078	1.923	2.707	1.335	1.260
GMG	1.856	1.771	1.408	1.400	1.316	1.718	0.970	1.064	1.826	2.666	1.284	1.215
RCR	2.002	1.768	1.452	1.396	2.020	3.260	1.017	1.087	12.328	6.655	2.035	2.650
GRCR	1.788	1.727	1.377	1.375	1.215	1.655	0.926	1.019	1.786	2.552	1.221	1.155
						$\mu_i \sim t(1)$						
CP-OLS	16.112	4.096	2.732	10.189	12.490	24.982	6.424	2.837	6.685	5.668	12.763	1.786
CP-SUR	19.483	5.046	3.365	12.976	14.940	29.854	8.009	3.555	7.807	7.043	15.947	2.126
MG	11.751	3.427	2.432	9.094	9.811	19.875	5.742	2.306	5.568	4.365	11.473	1.620
GMG	11.751	3.423	2.431	9.094	9.811	19.875	5.740	2.298	5.540	4.352	11.468	1.583
RCR	11.751	3.423	2.431	9.094	9.813	19.877	5.742	2.304	5.591	7.730	11.475	1.829
GRCR	11.739	3.403	2.417	9.090	9.795	19.868	5.733	2.271	5.498	4.228	11.462	1.530

matrix:

$$\begin{pmatrix} \sigma_{\varepsilon_{ii}} & \sigma_{\varepsilon_{ij}} & \cdots & \sigma_{\varepsilon_{ij}} \\ \sigma_{\varepsilon_{ij}} & \sigma_{\varepsilon_{ii}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \sigma_{\varepsilon_{ij}} \\ \sigma_{\varepsilon_{ij}} & \cdots & \sigma_{\varepsilon_{ij}} & \sigma_{\varepsilon_{ii}} \end{pmatrix},$$

where the values of $\sigma_{\varepsilon_{ii}}$, $\sigma_{\varepsilon_{ij}}$, and ρ were chosen to be: $\sigma_{\varepsilon_{ii}} = 1$ or 100; $\sigma_{\varepsilon_{ij}} = 0, 0.75$, or 0.95, and $\rho = 0, 0.55$, or 0.85, where the values of $\sigma_{\varepsilon_{ii}}$, $\sigma_{\varepsilon_{ij}}$, and ρ are constants for all $i,\ j = 1,\ 2,\ \ldots,\ N$ in each Monte Carlo trial. The initial values of u_{it} are generated as $u_{i1} = \varepsilon_{i1}/\sqrt{1-\rho^2}\ \forall\ i = 1,\ 2,\ \ldots,\ N$. The values of errors were allowed to differ for each cross-sectional unit on a given Monte Carlo trial and were allowed to differ between trials. The errors are independent with all independent variables.

3. The coefficients, α_{ki}, were generated as in assumption 6: $\alpha_i = \bar{\alpha} + \mu_i$, where $\bar{\alpha} = (1,1,1)^{'}$, and μ_i were generated from two distributions. First, multivariate normal distribution with means zeros and variance-covariance matrix $\Psi = diag\left\{\psi_k^2\right\}; k = 1,2,3$. The values of Ψ_k^2 were chosen to be fixed for all k and equal to 5 or 25. Second, multivariate student's t distribution

Table 2: ATSE for various estimators when $\sigma_{\varepsilon_{ii}} = 1$ and $N = T$.

$(\rho,\ \sigma_{\varepsilon_{ij}})$	(0, 0)				(0.55, 0.75)				(0.85, 0.95)			
$(N,\ T)$	(5, 5)	(10, 10)	(15, 15)	(20, 20)	(5, 5)	(10, 10)	(15, 15)	(20, 20)	(5, 5)	(10, 10)	(15, 15)	(20, 20)
					$\mu_i = 0$							
CP-OLS	1.671	0.461	0.259	0.174	2.081	0.424	0.274	0.207	3.351	0.678	0.394	0.276
CP-SUR	2.387	0.550	0.299	0.178	3.340	0.478	0.291	0.182	4.301	0.716	0.293	0.192
MG	1.686	0.486	0.280	0.183	2.058	0.474	0.300	0.210	3.093	0.668	0.377	0.255
GMG	1.174	0.395	0.234	0.159	1.669	0.363	0.209	0.149	2.028	0.370	0.190	0.115
RCR	1.905	0.557	0.314	0.179	1.997	0.953	0.411	0.502	3.249	1.982	0.471	0.458
GRCR	1.294	0.320	0.173	0.102	1.678	0.264	0.151	0.093	2.480	0.380	0.145	0.094
					$\mu_i \sim\ N\,(0,\ 5)$							
CP-OLS	4.119	3.404	1.982	1.651	4.593	2.002	1.517	1.474	5.023	2.926	1.847	1.740
CP-SUR	6.478	5.521	3.511	3.097	8.141	3.313	2.735	2.737	7.176	4.951	3.313	3.368
MG	3.480	2.750	1.744	1.520	4.015	1.671	1.295	1.341	4.284	2.531	1.633	1.608
GMG	3.481	2.750	1.743	1.520	4.008	1.664	1.289	1.337	4.034	2.515	1.615	1.599
RCR	5.955	2.749	1.743	1.520	4.232	1.666	1.295	1.342	12.312	2.574	1.651	1.617
GRCR	3.400	2.727	1.730	1.513	3.826	1.622	1.266	1.328	3.913	2.463	1.591	1.590
					$\mu_i \sim\ N\,(0,\ 25)$							
CP-OLS	8.056	6.265	4.022	3.637	7.976	5.496	4.240	3.968	10.264	6.615	4.558	3.733
CP-SUR	12.776	10.403	7.168	6.869	14.233	9.622	7.606	7.540	15.004	11.368	8.361	7.229
MG	6.474	5.145	3.558	3.348	6.491	4.599	3.692	3.623	6.798	5.597	4.042	3.464
GMG	6.476	5.145	3.558	3.348	6.498	4.596	3.690	3.622	6.822	5.589	4.036	3.460
RCR	6.469	5.145	3.558	3.348	6.457	4.597	3.692	3.624	10.576	5.614	4.050	3.468
GRCR	6.412	5.134	3.552	3.345	6.399	4.581	3.683	3.618	6.534	5.566	4.027	3.456
					$\mu_i \sim\ t\,(5)$							
CP-OLS	2.017	1.444	1.054	0.818	2.719	2.306	1.452	1.202	3.512	1.374	1.130	0.866
CP-SUR	2.952	2.278	1.848	1.499	4.581	4.002	2.602	2.251	4.784	2.113	1.960	1.584
MG	1.900	1.215	0.933	0.759	2.435	1.892	1.228	1.113	3.241	1.209	1.017	0.800
GMG	1.752	1.214	0.933	0.759	2.369	1.886	1.221	1.108	2.635	1.177	0.989	0.780
RCR	2.987	1.209	0.931	0.758	2.862	1.886	1.229	1.114	11.891	1.760	1.527	0.815
GRCR	1.628	1.165	0.908	0.744	2.193	1.848	1.199	1.097	2.727	1.073	0.951	0.762
					$\mu_i \sim\ t\,(1)$							
CP-OLS	2.946	4.082	36.296	32.249	170.833	4.983	7.221	5.545	5.447	14.094	27.076	2.245
CP-SUR	4.663	6.691	70.583	64.229	291.169	8.653	13.554	10.472	7.942	25.514	54.690	4.290
MG	2.569	3.337	23.288	26.932	92.236	4.064	5.831	5.069	4.403	11.428	20.763	2.085
GMG	2.565	3.337	23.288	26.932	92.238	4.060	5.829	5.068	4.362	11.420	20.759	2.078
RCR	5.160	3.337	23.288	26.932	92.238	4.061	5.831	5.069	7.663	11.440	20.767	2.091
GRCR	2.433	3.320	23.280	26.931	92.226	4.042	5.823	5.065	4.024	11.401	20.753	2.072

with degree of freedom (df): $df = 1$ or 5. To include the case of fixed-coefficients models in our simulation study, we assume that $\mu_i = 0$.

4. The values of N and T were chosen to be 5, 8, 10, 12, 15, and 20 to represent small and moderate samples for the number of individuals and the time dimension. To compare the small and moderate samples performance for the different estimators, three different samplings schemes have been designed in our simulation, where each design contains four pairs of N and T. The first two represent small samples while the moderate samples are represented by the second two pairs. These designs have been created as follows: First, case of $N < T$, the pairs of N and T were chosen to be $(N, T) = (5, 8), (5, 12), (10, 15)$, or (10, 20). Second, case of $N = T$, the pairs are $(N, T) = (5, 5), (10, 10), (15, 15)$, or (20, 20). Third, case of $N > T$, the pairs are $(N, T) = (8, 5), (12, 5), (15, 10)$, or (20, 10).

5. All Monte Carlo experiments involved 1000 replications and all the results of all separate experiments are obtained by precisely the same series of random numbers. To raise the efficiency of the comparison between these estimators, we calculate the average of total standard errors (ATSE) for

Table 3: ATSE for various estimators when $\sigma_{\varepsilon_{ii}} = 1$ and $N > T$.

$(\rho,\ \sigma_{\varepsilon_{ij}})$	(0, 0)				(0.55, 0.75)				(0.85, 0.95)			
$(N,\ T)$	(8, 5)	(12, 5)	(15, 10)	(20, 10)	(8, 5)	(12, 5)	(15, 10)	(20, 10)	(8, 5)	(12, 5)	(15, 10)	(20, 10)
$\mu_i = 0$												
CP-OLS	1.763	3.198	0.510	0.438	1.254	1.399	0.436	0.536	1.218	1.350	0.688	0.591
CP-SUR	2.504	4.585	0.635	0.518	1.748	1.963	0.497	0.607	1.637	1.808	0.780	0.655
MG	1.856	2.927	0.576	0.475	1.434	1.455	0.501	0.618	1.528	1.523	0.830	0.631
GMG	1.288	1.767	0.452	0.391	1.017	0.995	0.350	0.417	1.014	0.982	0.468	0.433
RCR	7.356	2.702	0.567	0.573	1.353	1.333	0.693	1.625	1.490	1.468	2.432	1.605
GRCR	1.289	2.277	0.342	0.267	0.937	1.010	0.248	0.306	0.865	0.856	0.413	0.312
$\mu_i \sim N(0,\ 5)$												
CP-OLS	3.136	4.014	2.525	2.017	3.677	3.352	2.477	3.105	2.146	3.501	1.927	2.415
CP-SUR	4.590	5.845	3.576	2.888	5.279	4.824	3.485	4.396	3.080	4.935	2.687	3.393
MG	2.753	3.418	2.153	1.685	2.972	2.643	2.113	2.628	2.191	2.813	1.724	2.156
GMG	2.665	3.425	2.152	1.684	2.951	2.660	2.106	2.617	2.097	2.748	1.679	2.142
RCR	3.611	3.306	2.146	1.681	2.897	3.034	2.109	2.621	61.169	137.429	2.187	2.147
GRCR	2.400	2.982	2.103	1.636	2.774	2.399	2.066	2.572	1.852	2.550	1.532	2.075
$\mu_i \sim N(0,\ 25)$												
CP-OLS	6.919	6.434	6.179	5.259	6.442	5.639	4.972	4.460	6.279	7.428	5.480	5.366
CP-SUR	10.250	9.292	8.750	7.682	9.200	8.224	7.123	6.378	9.507	10.544	7.791	7.698
MG	5.090	5.029	5.092	4.381	4.987	4.505	4.167	3.688	5.353	5.689	4.545	4.756
GMG	5.046	5.031	5.092	4.380	4.971	4.512	4.163	3.680	5.316	5.677	4.530	4.749
RCR	4.986	4.735	5.091	4.380	4.939	4.466	4.165	3.683	5.303	6.219	4.538	4.753
GRCR	4.898	4.588	5.071	4.362	4.874	4.408	4.142	3.645	5.189	5.559	4.479	4.720
$\mu_i \sim t(5)$												
CP-OLS	1.779	2.367	1.151	1.080	1.780	2.464	1.986	1.308	2.157	2.848	1.473	1.283
CP-SUR	2.541	3.365	1.604	1.493	2.596	3.711	2.929	1.745	3.137	4.179	1.987	1.730
MG	1.839	1.989	1.010	0.943	1.647	2.276	1.603	1.074	2.109	2.401	1.260	1.467
GMG	1.577	1.974	1.008	0.942	1.563	2.245	1.586	1.076	1.730	2.362	1.235	1.255
RCR	2.573	2.327	0.991	0.960	2.785	2.945	1.591	1.097	3.523	3.020	3.322	3.509
GRCR	1.336	1.738	0.924	0.837	1.529	1.893	1.525	0.982	1.652	2.120	1.124	1.049
$\mu_i \sim t(1)$												
CP-OLS	23.572	9.953	1.708	9.638	9.612	3.030	5.400	4.609	6.932	8.340	25.666	4.259
CP-SUR	35.133	13.767	2.466	14.035	15.207	4.429	8.027	6.816	9.309	12.412	39.880	6.199
MG	17.304	6.568	1.410	6.014	7.568	2.654	4.164	3.451	4.802	6.004	16.848	3.318
GMG	17.295	6.563	1.409	6.014	7.580	2.629	4.155	3.452	4.781	5.991	16.840	3.267
RCR	17.295	6.535	1.398	6.012	7.546	2.499	4.158	3.456	6.130	5.997	16.849	4.158
GRCR	17.263	6.483	1.345	5.979	7.492	2.345	4.128	3.407	4.593	5.877	16.779	3.081

each estimator by:

$$\text{ATSE} = \frac{1}{1000} \sum_{l=1}^{1000} \left\{ trace \left[\widehat{var} \left(\widehat{\bar{\alpha}}_l \right) \right]^{0.5} \right\},$$

where $\widehat{\bar{\alpha}}_l$ is the estimated vector of $\bar{\alpha}$ in (6.1), and $\widehat{var}\left(\widehat{\bar{\alpha}}_l\right)$ is the estimated variance-covariance matrix of the estimator.

The Monte Carlo results are given in Tables 1–6. Specifically, Tables 1–3 present the ATSE values of the estimators when $\sigma_{\varepsilon_{ii}} = 1$, and in cases of $N < T, N = T$, and$N > T$, respectively. While case of $\sigma_{\varepsilon_{ii}} = 100$ is presented in Tables 4–6 in the same cases of N and T. In our simulation study, the main factors that have an effect on the ATSE values of the estimators are $N,\ T,\ \sigma_{\varepsilon_{ii}},\ \sigma_{\varepsilon_{ij}},\ \rho,\ \psi_k^2$ (for normal distribution), and *df* (for student's t distribution). From Tables 1–6, we can summarize some effects for all estimators in the following points:

- When the values of N and T are increased, the values of ATSE are decreasing for all simulation situations.

Table 4: ATSE for various estimators when $\sigma_{\varepsilon_{ii}} = 100$ and $N < T$.

$(\rho,\ \sigma_{\varepsilon_{ij}})$	(0, 0)				(0.55, 0.75)				(0.85, 0.95)			
$(N,\ T)$	(5, 8)	(5, 12)	(10, 15)	(10, 20)	(5, 8)	(5, 12)	(10, 15)	(10, 20)	(5, 8)	(5, 12)	(10, 15)	(10, 20)
						$\mu_i = 0$						
CP-OLS	2.908	2.357	1.389	1.379	2.756	2.863	1.414	1.395	3.798	5.179	2.042	2.208
CP-SUR	3.028	2.422	1.323	1.316	2.806	2.997	1.335	1.302	3.520	5.316	1.692	1.989
MG	2.993	2.419	1.486	1.483	2.830	2.984	1.492	1.503	3.850	4.907	2.010	2.292
GMG	2.221	1.759	1.168	1.187	1.975	2.180	1.027	1.004	2.132	3.466	1.022	1.191
RCR	3.199	97.225	1.634	1.570	3.205	6.691	2.576	2.846	4.711	7.169	2.708	3.170
GRCR	2.381	1.970	1.111	1.128	2.188	2.399	1.061	1.029	2.667	3.872	1.220	1.429
						$\mu_i \sim\ N(0,\ 5)$						
CP-OLS	5.096	4.872	2.481	2.890	3.298	3.570	2.732	2.260	4.432	6.390	2.479	3.180
CP-SUR	5.787	5.751	2.856	3.437	3.573	3.960	3.305	2.557	4.449	6.946	2.463	3.524
MG	4.533	4.450	2.361	2.737	3.193	3.448	2.575	2.172	4.327	5.642	2.363	3.076
GMG	4.507	4.427	2.349	2.734	2.869	3.165	2.539	2.101	3.695	5.110	2.150	2.849
RCR	11.579	5.572	2.500	2.702	3.871	8.045	3.278	3.489	7.748	9.539	5.301	22.220
GRCR	4.179	4.294	2.166	2.576	2.755	3.026	2.378	1.911	3.456	5.004	1.879	2.560
						$\mu_i \sim\ N(0,\ 25)$						
CP-OLS	7.670	7.803	7.209	6.407	8.362	8.314	6.380	6.781	7.971	7.887	4.852	6.554
CP-SUR	8.833	9.460	8.952	8.050	10.073	10.032	8.245	8.508	9.153	9.160	5.890	8.277
MG	6.570	6.760	6.431	5.714	7.118	7.016	5.653	6.018	6.812	7.017	4.338	5.913
GMG	6.556	6.749	6.426	5.713	7.116	7.013	5.625	5.991	6.658	6.996	4.240	5.795
RCR	10.949	6.908	6.423	5.706	7.103	7.629	5.647	6.008	11.120	16.814	9.260	6.478
GRCR	6.400	6.633	6.370	5.646	6.945	6.826	5.558	5.932	6.286	6.595	4.057	5.661
						$\mu_i \sim\ t(5)$						
CP-OLS	3.227	2.672	1.820	1.804	2.894	3.067	1.534	1.558	4.052	5.630	2.112	2.299
CP-SUR	3.432	2.879	1.975	1.959	3.045	3.327	1.529	1.560	3.998	6.065	1.838	2.099
MG	3.186	2.654	1.829	1.810	2.924	3.097	1.588	1.617	4.042	5.146	2.071	2.318
GMG	2.816	2.405	1.799	1.782	2.296	2.690	1.394	1.435	2.792	4.288	1.603	1.692
RCR	3.665	3.442	2.592	2.462	4.922	4.147	3.057	4.985	9.667	14.064	3.871	6.113
GRCR	2.666	2.317	1.625	1.543	2.374	2.662	1.232	1.233	3.045	4.365	1.456	1.604
						$\mu_i \sim\ t(1)$						
CP-OLS	16.193	4.345	2.882	10.228	12.527	25.028	6.481	2.957	6.842	6.962	12.819	2.363
CP-SUR	19.488	5.071	3.383	12.975	14.929	30.583	8.213	3.571	7.803	7.838	16.626	2.317
MG	11.990	3.871	2.673	9.164	9.996	19.985	5.841	2.595	6.095	5.929	11.548	2.434
GMG	11.990	3.832	2.665	9.163	9.979	19.993	5.819	2.524	5.898	5.591	11.512	1.988
RCR	11.965	4.529	2.625	9.162	9.966	19.996	5.839	3.527	13.705	59.015	11.574	14.464
GRCR	11.840	3.650	2.507	9.122	9.862	19.940	5.762	2.360	5.434	5.506	11.460	1.773

- When the value of $\sigma_{\varepsilon_{ii}}$ is increased, the values of ATSE are increasing in most situations.
- When the values of $(\rho, \sigma_{\varepsilon_{ij}})$ are increased, the values of ATSE are increasing in most situations.
- When the value of ψ_k^2 is increased, the values of ATSE are increasing for all situations.
- When the value of df is increased, the values of ATSE are decreasing for all situations.

For more deeps in simulation results, we can conclude the following results:

1. Generally, the performance of all estimators in cases of $N \leqslant T$ is better than their performance in case of $N > T$. Similarly, their performance in cases of $\sigma_{\varepsilon_{ii}} = 1$ is better than the performance in case of $\sigma_{\varepsilon_{ii}} = 100$, but not as significantly better as in N and T.

Table 5: ATSE for various estimators when $\sigma_{\varepsilon_{ii}} = 100$ and $N = T$.

$(\rho,\ \sigma_{\varepsilon_{ij}})$	(0, 0)				(0.55, 0.75)				(0.85, 0.95)			
$(N,\ T)$	(5, 5)	(10, 10)	(15, 15)	(20, 20)	(5, 5)	(10, 10)	(15, 15)	(20, 20)	(5, 5)	(10, 10)	(15, 15)	(20, 20)
					$\mu_i = 0$							
CP-OLS	5.284	1.456	0.818	0.548	6.920	1.339	0.904	0.629	11.353	2.314	1.215	0.871
CP-SUR	7.548	1.737	0.942	0.559	10.528	1.580	0.977	0.589	15.654	2.573	0.987	0.625
MG	5.331	1.537	0.886	0.577	6.606	1.417	0.998	0.658	10.554	2.362	1.238	0.839
GMG	3.712	1.250	0.741	0.503	5.470	1.105	0.693	0.466	6.959	1.419	0.602	0.410
RCR	6.023	1.759	0.990	0.564	8.315	2.026	2.034	1.388	10.978	3.817	2.088	1.241
GRCR	4.090	1.007	0.545	0.318	5.497	0.907	0.527	0.318	8.037	1.363	0.525	0.325
					$\mu_i \sim N(0,\ 5)$							
CP-OLS	5.580	3.519	2.061	1.705	7.429	2.182	1.629	1.543	10.993	3.155	1.991	1.859
CP-SUR	8.237	5.479	3.497	3.091	11.726	3.255	2.651	2.742	15.414	4.585	3.080	3.221
MG	5.622	2.996	1.876	1.592	6.993	1.987	1.522	1.438	10.338	3.017	1.864	1.733
GMG	4.959	2.994	1.876	1.591	6.571	1.968	1.459	1.406	7.682	2.893	1.712	1.649
RCR	8.572	3.064	1.861	1.588	8.773	2.645	2.696	1.435	10.818	6.531	3.172	1.779
GRCR	4.679	2.764	1.747	1.520	6.313	1.727	1.249	1.322	8.234	2.397	1.489	1.558
					$\mu_i \sim N(0,\ 25)$							
CP-OLS	8.220	6.333	4.056	3.661	9.384	5.567	4.285	3.991	12.808	6.724	4.618	3.788
CP-SUR	12.685	10.388	7.152	6.865	15.219	9.557	7.574	7.573	18.954	11.401	8.194	7.215
MG	7.404	5.282	3.620	3.380	8.388	4.740	3.779	3.657	11.236	5.845	4.138	3.523
GMG	7.257	5.281	3.620	3.380	8.438	4.728	3.754	3.645	9.858	5.787	4.073	3.482
RCR	12.035	5.272	3.618	3.380	9.526	4.731	3.774	3.658	12.921	6.137	4.153	3.545
GRCR	6.703	5.166	3.556	3.347	7.863	4.608	3.688	3.613	9.475	5.537	3.995	3.440
					$\mu_i \sim t(5)$							
CP-OLS	5.268	1.758	1.205	0.930	6.905	2.466	1.566	1.289	11.183	2.322	1.363	1.078
CP-SUR	7.487	2.302	1.826	1.505	10.462	3.902	2.518	2.232	15.445	2.648	1.486	1.354
MG	5.301	1.734	1.173	0.901	6.588	2.197	1.457	1.231	10.371	2.363	1.359	1.024
GMG	3.914	1.688	1.171	0.900	5.741	2.170	1.392	1.193	7.036	1.810	1.138	0.874
RCR	6.313	2.356	1.226	0.885	8.980	4.088	1.806	1.224	10.384	6.372	4.418	4.574
GRCR	4.238	1.313	0.937	0.764	5.796	1.894	1.179	1.094	8.124	1.489	0.823	0.688
					$\mu_i \sim t(1)$							
CP-OLS	5.492	4.176	36.310	32.254	170.969	5.046	7.246	5.564	11.208	14.166	27.093	2.332
CP-SUR	8.085	6.670	70.596	64.232	277.362	8.718	13.502	10.390	15.450	26.068	54.457	4.185
MG	5.469	3.529	23.379	26.943	92.536	4.228	5.898	5.095	10.448	11.655	20.834	2.180
GMG	4.346	3.528	23.378	26.943	92.558	4.213	5.878	5.086	7.748	11.603	20.786	2.114
RCR	7.220	3.503	23.365	26.943	92.513	4.383	5.895	5.096	13.141	12.397	20.840	2.210
GRCR	4.471	3.354	23.296	26.932	92.445	4.050	5.822	5.064	8.345	11.384	20.731	2.046

2. When $\sigma_{\varepsilon_{ij}} = \rho = \mu_i = 0$, the ATSE values of the classical pooling estimators (CP-OLS and CP-SUR) are approximately equivalent, especially when the sample size is moderate and/or $N \leqslant T$. However, the ATSE values of GMG and GRCR estimators are smaller than those of the classical pooling estimators in this situation ($\sigma_{\varepsilon_{ij}} = \rho = \mu_i = 0$) and other simulation situations (case of $\sigma_{\varepsilon_{ii}},\ \sigma_{\varepsilon_{ij}}, \rho, \psi_k^2$ are increasing, and *df* is decreasing). In other words, GMG and GRCR are more efficient than CP-OLS and CP-SUR whether the regression coefficients are fixed or random.

3. If $T \geq 15$, the values of ATSE for the MG and GMG estimators are approximately equivalent. This result is consistent with Lemma 5.2. According to our study, this case ($T \geq 15$) is achieved when the sample size is moderate in Tables 1, 2, 4, and 5. Moreover, convergence slows down if $\sigma_{\varepsilon_{ii}},\ \sigma_{\varepsilon_{ij}}$, and ρ are increased. But the situation for the RCR and GRCR estimators is different; the convergence between them is very slow even if $T = 20$. So the MG and GMG estimators are more efficient than RCR in all simulation situations.

4. When the coefficients are random (whether they are distributed as normal or student's t), the values of ATSE for GMG and GRCR are smaller than those of MG and RCR in all simulation situations (for any $N,\ T,\ \sigma_{\varepsilon_{ii}},\ \sigma_{\varepsilon_{ij}}$,

Table 6: ATSE for various estimators when $\sigma_{\varepsilon_{ii}} = 100$ and $N > T$.

$(\rho,\ \sigma_{\varepsilon_{ij}})$	(0, 0)				(0.55, 0.75)				(0.85, 0.95)			
$(N,\ T)$	(8, 5)	(12, 5)	(15, 10)	(20, 10)	(8, 5)	(12, 5)	(15, 10)	(20, 10)	(8, 5)	(12, 5)	(15, 10)	(20, 10)
						$\mu_i = 0$						
CP-OLS	5.574	3.501	1.511	1.493	5.616	4.178	1.764	1.546	8.088	9.255	2.325	2.474
CP-SUR	7.919	4.835	1.798	1.840	7.780	5.841	2.229	1.813	11.886	12.804	2.723	2.975
MG	5.868	3.453	1.659	1.676	5.678	4.306	1.908	1.629	9.127	8.473	2.678	2.773
GMG	4.073	2.490	1.349	1.337	3.643	3.717	1.515	1.219	5.788	7.373	1.382	1.581
RCR	23.253	3.498	1.759	1.808	5.403	6.417	5.387	2.286	8.172	11.799	2.744	4.156
GRCR	4.072	2.397	0.931	0.972	3.998	3.241	1.142	0.872	5.937	6.519	1.267	1.352
						$\mu_i \sim\ N(0,\ 5)$						
CP-OLS	5.574	4.258	2.867	2.692	5.221	5.014	2.744	2.396	8.256	9.261	2.333	3.037
CP-SUR	7.899	5.954	3.858	3.725	7.202	7.096	3.802	3.166	12.049	12.885	2.782	4.092
MG	5.793	3.775	2.616	2.509	5.407	4.904	2.622	2.241	9.299	8.462	2.682	3.135
GMG	4.753	3.635	2.615	2.503	4.022	4.657	2.663	2.226	6.423	7.531	2.230	2.815
RCR	7.585	5.340	2.525	2.569	25.633	6.314	8.404	2.808	10.171	10.268	15.344	8.355
GRCR	4.220	3.123	2.206	2.063	3.901	3.925	2.101	1.771	6.533	6.464	1.443	2.026
						$\mu_i \sim\ N(0,\ 25)$						
CP-OLS	7.383	6.000	5.791	4.700	6.808	7.512	4.220	6.284	7.648	11.202	4.729	4.463
CP-SUR	10.777	8.636	8.118	6.667	9.409	11.012	5.987	8.667	11.213	16.010	6.596	6.367
MG	6.876	4.940	4.816	4.146	6.287	6.642	3.722	5.162	8.635	9.623	4.346	4.168
GMG	6.442	4.902	4.815	4.143	6.205	6.532	3.765	5.156	7.205	9.360	4.171	3.961
RCR	11.741	5.730	4.792	4.090	11.299	7.379	3.776	5.160	12.146	12.980	13.643	7.505
GRCR	5.510	4.310	4.615	3.915	5.288	5.902	3.379	4.983	6.356	8.403	3.669	3.352
						$\mu_i \sim\ t(5)$						
CP-OLS	5.373	3.666	1.719	1.726	5.575	4.294	1.789	1.805	8.085	9.347	2.373	2.455
CP-SUR	7.646	5.136	2.115	2.217	7.757	5.989	2.248	2.223	11.901	13.041	2.803	2.974
MG	5.706	3.482	1.779	1.837	5.623	4.394	1.926	1.802	9.133	8.456	2.695	2.784
GMG	4.249	3.082	1.722	1.759	3.683	3.907	1.647	1.727	5.933	7.429	1.691	1.879
RCR	9.861	5.223	2.501	2.758	5.421	5.238	3.195	3.158	13.392	14.875	4.908	6.298
GRCR	3.915	2.670	1.150	1.268	4.044	3.334	1.188	1.170	6.032	6.570	1.342	1.415
						$\mu_i \sim\ t(1)$						
CP-OLS	5.821	3.703	4.328	6.252	6.016	5.931	31.442	4.149	11.344	10.999	5.576	3.013
CP-SUR	8.533	5.188	6.188	9.132	8.500	8.555	47.659	5.806	17.261	15.893	8.562	3.969
MG	5.986	3.550	3.544	5.182	5.876	5.420	21.165	3.416	11.058	9.507	4.826	3.140
GMG	4.941	3.242	3.537	5.179	5.579	5.219	21.177	3.402	8.986	9.203	4.557	2.831
RCR	8.791	13.034	13.254	5.140	7.133	6.561	21.171	3.896	13.086	12.317	10.078	10.717
GRCR	4.403	2.740	3.115	4.987	4.936	4.559	21.041	3.093	8.697	7.876	3.877	2.021

and ρ). However, the ATSE values of GRCR are smaller than those of GMG estimator in most situations, especially when the sample size is moderate. In other words, the GRCR estimator performs better than all other estimators as long as the sample size is moderate regardless of other simulation factors.

7. CONCLUSION

In this article, the classical pooling (CP-OLS and CP-SUR), random-coefficients (RCR and GRCR), and mean group (MG and GMG) estimators of stationary RCPD models were examined in different sample sizes for the case where the errors are cross-sectionally and serially correlated. Analytical efficiency comparisons for these estimators indicate that the mean group and random-coefficients estimators are equivalent when T is sufficiently large. Furthermore, the Monte Carlo simulation results show that the classical pooling estimators are absolutely not suitable for random-coefficients models. And, the MG and GMG estimators are more efficient than the RCR estimator for random- and fixed-coefficients models, especially when T is small ($T \leq 12$). But when $T \geq 20$, the MG, GMG, and GRCR estimators are approximately equivalent. However, the GRCR estimator performs better than the MG and GMG estimators in most situations, especially in moderate samples. Therefore, we conclude that the GRCR estimator is suitable to stationary RCPD models whether the coefficients are random or fixed.

REFERENCES

[1] Abonazel, M. R. (2009). *Some Properties of Random Coefficients Regression Estimators.* MSc thesis. Institute of Statistical Studies and Research. Cairo University. Egypt.

[2] Abonazel, M. R. (2014). *Some Estimation Methods for Dynamic Panel Data Models.* PhD thesis. Institute of Statistical Studies and Research. Cairo University. Egypt.

[3] Abonazel, M. R. (2017). Bias correction methods for dynamic panel data models with fixed effects, *International Journal of Applied Mathematical Research*, **6**, 2, 58–66.

[4] Alcacer, J.; Chung, W.; Hawk, A. and Pacheco-de-Almeida, G. (2013). Applying random coefficient models to strategy research: testing for firm heterogeneity, predicting firm-specific coefficients, and estimating Strategy Trade-Offs. Working Paper, No. 14-022. Harvard Business School Strategy Unit.

[5] Anh, V. and Chelliah, T. (1999). Estimated generalized least squares for random coefficient regression models, *Scandinavian Journal of Statistics*, **26**, 1, 31–46.

[6] Baltagi, B. H. (2011). *Econometrics.* 5th ed. Berlin: Springer-Verlag Berlin Heidelberg.

[7] Baltagi, B. H. (2013). *Econometric Analysis of Panel Data.* 5th ed. Chichester: John Wiley and Sons.

[8] Beck, N. and Katz, J. (2007). Random coefficient models for time-series-cross-section data: Monte Carlo experiments, *Political Analysis*, **15**, 2, 182–195.

[9] Beran, R. and Millar, P. (1994). Minimum distance estimation in random coefficient regression models, *The Annals of Statistics*, **22**, 4, 1976–1992.

[10] Bodhlyera, O.; Zewotir, T. and Ramroop, S. (2014). Random coefficient model for changes in viscosity in dissolving pulp, *Wood Research*, **59**, 4, 571–582.

[11] Boness, A. and Frankfurter, G. (1977). Evidence of non-homogeneity of capital costs within "risk-classes", *The Journal of Finance*, **32**, 3, 775–787.

[12] Boot, J. and Frankfurter, G. (1972). The dynamics of corporate debt management, decision rules, and some empirical evidence, *Journal of Financial and Quantitative Analysis*, **7**, 4, 1957–1965.

[13] Chelliah, N. (1998). A new covariance estimator in random coefficient regression model, *The Indian Journal of Statistics*, Series B, **60**, 3, 433–436.

[14] Cheng, J.; Yue, R. X. and Liu, X. (2013). Optimal designs for random coefficient regression models with heteroscedastic errors, *Communications in Statistics-Theory and Methods* **42**, 15, 2798–2809.

[15] Cooley, T. and Prescott, E. (1973). Systematic (non-random) variation models: varying parameter regression: a theory and some applications, *Annals of Economic and Social Measurement*, **2**, 4, 463-473.

[16] Dielman, T. (1983). Pooled cross-sectional and time series data: a survey of current statistical methodology, *The American Statistician*, **37**, 2, 111–122.

[17] Dielman, T. (1989). *Pooled Cross-Sectional and Time Series Data Analysis*, Marcel Dekker, New York.

[18] Dielman, T. (1992a). Misspecification in random coefficient regression models: a Monte Carlo simulation. *Statistical Papers*, **33**, 1, 241–260.

[19] Dielman, T. (1992b). Small sample properties of random coefficient regression estimators: a Monte Carlo simulation. *Communications in Statistics–Simulation and Computation*, **21**, 1, 103–132.

[20] Dziechciarz, J. (1989). *Changing and random coefficient models. A survey.* In: "Statistical Analysis and Forecasting of Economic Structural Change" (P. Hackl, Ed.), Springer Berlin Heidelberg, Berlin, 217–251.

[21] Elhorst, J. (2014). *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*, springer, Berlin.

[22] Elster, C. and Wbbeler, G. (2017). Bayesian inference using a noninformative prior for linear Gaussian random coefficient regression with inhomogeneous within-class variances, *Computational Statistics*, **32**, 1, 51–69.

[23] Feige, E. and Swamy, P. (1974). A random coefficient model of the demand for liquid assets, *Journal of Money, Credit and Banking*, **6**, 2, 241–252.

[24] Fu, K. and Fu, X. (2015). Asymptotics for the random coefficient first-order autoregressive model with possibly heavy-tailed innovations, *Journal of Computational and Applied Mathematics*, **285**, 116–124.

[25] Horvth, L. and Trapani, L. (2016). Statistical inference in a random coefficient panel model, *Journal of Econometrics*, **193**, 1, 54–75.

[26] Hsiao, C. (2014). *Analysis of Panel Data*, 3rd ed., Cambridge University Press, Cambridge.

[27] Hsiao, C. and Pesaran, M. (2008). *Random coefficient models.* In: "The Econometrics of Panel Data", Vol. 46, (L. Matyas, and P. Sevestre Eds.), Springer Berlin Heidelberg, Berlin, 185–213.

[28] Judge, G.; Griffiths, W.; Hill, R.; Ltkepohl, H. and Lee, T. (1985). *The Theory and Practice of Econometrics*, 2nd ed., Wiley, New York.

[29] Livingston, M.; Erickson, K. and Mishra, A. (2010). *Standard and Bayesian random coefficient model estimation of US Corn–Soybean farmer risk attitudes.* In: "The Economic Impact of Public Support to Agriculture", (V. Ball, R. Fanfani and L. Gutierrez Eds.), Springer, New York, 329–343.

[30] Mousa, A.; Youssef, A. and Abonazel, M. (2011). A Monte Carlo study for Swamy's estimate of random coefficient panel data model. Working paper, No. 49768. University Library of Munich, Germany.

[31] Murtazashvili, I. and Wooldridge, J. (2008). Fixed effects instrumental variables estimation in correlated random coefficient panel data models' *Journal of Econometrics*, **142**, 539–552.

[32] Parks, R. (1967). Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated, *Journal of the American Statistical Association*, **62**, 500–509.

[33] Pesaran, M. and Smith, R. (1995). Estimation of long-run relationships from dynamic heterogeneous panels, *Journal of Econometrics*, **68**, 79–114.

[34] Poi, B. (2003). From the help desk: Swamy's random-coefficients model. *The Stata Journal*, **3**, 3, 302–308.

[35] Rao, C. and Mitra, S. (1971). *Generalized Inverse of Matrices and Its Applications*, John Wiley and Sons Ltd, New York.

[36] Rao, U. (1982). A note on the unbiasedness of Swamy's estimator for the random coefficient regression model. *Journal of econometrics*, **18**, 3, 395–401.

[37] Rausser, G.; Mundlak, Y. and Johnson, S. (1982). *Structural change, updating, and forecasting.* In: "New Directions in Econometric Modeling and Forecasting US Agriculture", (G. Rausser, Ed.), North–Holland, Amsterdam.

[38] Sant, D. (1977). Generalized least squares applied to time-varying parameter models, *Annals of Economic and Social Measurement*, **6**, 3, 301–314.

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

[40] Swamy, P. (1970). Efficient inference in a random coefficient regression model, *Econometrica*, **38**, 311–323.

[41] Swamy, P. (1971). *Statistical Inference in Random Coefficient Regression Models*, Springer-Verlag, New York.

[42] Swamy, P. (1973). Criteria, constraints, and multicollinearity in random coefficient regression model, *Annals of Economic and Social Measurement*, **2**, 4, 429-450.

[43] Swamy, P. (1974). *Linear models with random coefficients.* In: "Frontiers in Econometrics", (P. Zarembka, Ed.), Academic Press, New York.

[44] Swamy, P.; Mehta, J.; Tavlas, G. and Hall, S. (2015). Two applications of the random coefficient procedure: Correcting for misspecifications in a small area level model and resolving Simpson's paradox, *Economic Modelling*, **45**, 93–98.

[45] Westerlund, J. and Narayan, P. (2015). A random coefficient approach to the predictability of stock returns in panels, *Journal of Financial Econometrics*, **13**, 3, 605–664.

[46] Youssef, A. and Abonazel, M. (2009). A comparative study for estimation parameters in panel data model. Working paper, No. 49713. University Library of Munich, Germany.

[47] Youssef, A. and Abonazel, M. (2017). Alternative GMM estimators for first-order autoregressive panel model: an improving efficiency approach, *Communications in Statistics-Simulation and Computation*, **46**, 4, 3112–3128.

[48] Youssef, A.; El-sheikh, A. and Abonazel, M. (2014a). Improving the efficiency of GMM estimators for dynamic panel models, *Far East Journal of Theoretical Statistics*, **47**, 171–189.

[49] Youssef, A.; El-sheikh, A. and Abonazel, M. (2014b). New GMM estimators for dynamic panel data models, *International Journal of Innovative Research in Science, Engineering and Technology*, **3**, 16414–16425.

[50] Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests of aggregation bias, *Journal of the American Statistical Association*, **57**, 348–368.