A GLS ESTIMATION OF THE TWO-WAY RANDOM EFFECT MODEL WITH DOUBLE AUTOCORRELATION

To my beloved Mother, KOUASSI Sally Georgette, and

the memory of my darling Father, KOUASSI Brou Bruno.

ABSTRACT

This thesis is a theoretical investigation of a frequently encountered econometric issue: the problem of autocorrelation. Under a two-way random effect context, we introduce serial correlation in the time-varying disturbances, leading to a double correlation framework.

We analyze two major situations related to the structure of the error terms. The first one considers that the time-varying disturbances follow the same correlation pattern, with the same parameters. They are allowed to exhibit series such as the autoregressive of order 1 (hereafter AR(1)) or the moving-average of order 1 (hereafter MA(1)) processes. We also examined the case of unknown correlation. A detailed generalized least squares (hereafter GLS) procedure is deduced from the spectral decomposition of the variance-covariance matrix of the composite error term. A Feasible Generalized Least Squares (hereafter FGLS) approach is derived whatever the correlation status may be.

The second error structure assumes that the time-varying disturbances can follow different correlation patterns. A general case of unknown serial correlation is considered, as well as the autoregressive and moving-average processes of order 1 models. We show that the variance-covariance matrix of the overall error term can always be written in a precise form, independently from the type of serial correlation. Once again, we deduce a GLS estimator from the inverse of this moment matrix. Underlying estimators are shown out and their asymptotic properties are studied. We find that the GLS estimator is asymptotically equivalent to a "within" estimator called the covariance estimator. Finally, a FGLS version is proposed.

RESUME

Cette thèse est une étude théorique sur l'autocorrélation, une préoccupation fréquemment rencontrée en économétrie. Dans un modèle à deux effets à erreurs composées, nous introduisons une corrélation sérielle dans les termes d'erreurs temporels, sous la forme d'une autocorrélation double.

Deux cas de figure sont analysés, selon la structure des termes d'erreurs. Dans le premier cas, nous supposons que les perturbations temporelles suivent un processus unique, avec donc les mêmes paramètres, notamment un processus autorégressif d'ordre 1 (AR(1)) ou moyenne mobile d'ordre 1 (MA(1)). Nous avons aussi examiné la situation où l'autocorrélation est de type inconnu. Une procédure par les moindres carrés généralisés (MCG) est déduite de la décomposition spectrale de la matrice de variance-covariance du terme d'erreur composite. Une version estimable est finalement proposée, quelle que soit la nature de l'autocorrélation.

Le deuxième cas de figure suppose que les perturbations temporelles suivent des processus différents. Ici également, un modèle général avec une autocorrélation de type inconnu a été considéré, en plus des processus AR(1) et MA(1). Nous montrons que la matrice des variances-covariances du terme d'erreur composite peut toujours s'écrire sous une forme précise, indépendamment de la nature de l'autocorrélation. De l'inversion de cette matrice nous déduisons un estimateur MCG, qui se présente par ailleurs comme une combinaison d'autres estimateurs MCG. Leurs propriétés asymptotiques sont étudiées et révèlent une équivalence, dans les échantillons de grande taille, entre notre estimateur MCG et un estimateur de type "intra" appelé l'estimateur de covariance. Finalement, une version estimable est proposée.

ACKNOWLEDGEMENTS

Being able to write this thesis is itself the result of daily support and assistance of numerous persons and institutions. I have been very fortunate to count on the financial support of the African Economic Research Consortium (A.E.R.C.) under the Collaborative PhD Program (C.P.P.). I am grateful to its management and executive teams for the international experience (together with other Ph.D candidates from Côte d'Ivoire and elsewhere) we accumulate by the different trips they offered to us. The core courses taken in Yaoundé (Cameroon) as well as the elective ones received in Nairobi (Kenya) have been determinant in the choice of our research area. I will always be proud to introduce myself as an AERC's "product"!

A warm thank is also due to the institutions I belong to, say the University of Abidjan-Cocody, as well as to its partners, notably the University of Yaoundé II. If the former is naturally "my University", the latter kindly collaborates in giving us core courses at a PhD level. The twelve months spent in Cameroon have been extremely valuable and no word could sufficiently capture the extent to which we have benefited from this university. May its current Dean, Pr. Georges Kobou and its whole team as well as the former one Pr. Séraphin Fouda, find here the greetings of their student and son, who would remain grateful to them forever.

I am also particularly indebted to my supervisors, Pr. Aké N'Gbo and Pr. Kaddour Hadri. The former is a model that has inspired and is still inspiring many students at the Economics Department of the University of Cocody. I am actually one of them. As a chair professor in microeconomics and in econometrics, he is a constant source of encouragement and demand for rigor. The later is a brilliant and fruitful international econometrician with publications in top journals such as Econometrica. Met in a conference in Nairobi, he has friendly accepted to supervise my dissertation. His advises and his comments on earlier drafts have been extremely useful. Unfortunately, I couldn't get the visa to join him at Durham University (United Kingdom) for a deeper and closer assistance.

This thesis could not actually be written without the support of Dr Eugene Kouassi. He taught me econometrics during the master and also during the Joint Facilities for Electives (J.F.E.) under the C.P.P. in Nairobi. He has definitely raised my interest in theoretical staff. I owe special thanks to him for his daily help and consistent comments. His door has always been opened to my questions and complaints in matters that ranged well beyond research. He helped me focus my interests and gave me the tools to pursue them. He is a second “father” for me and I am immensely thankful to him.

I cannot go ahead without thanking the members of the Jury: Pr. Aké N’Gbo, Pr. Kaddour Hadri, Dr. Eugene Kouassi, Pr. Pierre Roche Séka, Pr. Mama Ouattara, Pr. Georges Kobou and Pr. Kern O. Kymn from West-Virginia University. It was a great honor for me to be evaluated by such a prestigious committee. I owe special thanks to Pr Kern O. Kymn who has kindly accepted to assess my dissertation, under the American standards. I am also grateful to Pr. Mama Ouattara (who used to supervise my master’s dissertation) and Pr Pierre Roche Séka (the Dean of the Department of Economics of the University of Abidjan-Cocody), for their involvement in my thesis evaluation and their general concern for the well-being of the students.

I would like to thank many other people for providing me encouragements and advices. Amongst are my peers of the C.P.P., especially Komenan Aka Narcisse and Anviré Ahou-Yah Gisèle. The experiments shared together have definitely transformed our relationships from friends to brother and sister. I owe them special thanks, as well as to Baida Angèle Louise for her encouragements. I also thank other C.P.P. students, notably N’Guessan Dja Nicaise, Yéo Nahoua and Niangoran Amoa Rachelle. May all the PhD students of other African countries met under this program, especially Mugandza Tendie of Zimbabwe, find here the expression of my amity.

Finally, I thank my relatives and numerous friends who kept on telling me how proud they are. Some of them are my peers Malan Benoit and Soro Nahoua. Their moral supports have strengthened my determination. Completing my dissertation is my way of thanking them all. I am particularly very grateful to N’Dri Brou Cécile. Her assistance has been extremely valuable. Special thanks are also due to Kouassi Angora Vincent, Konan Yao Barthélémy and my beloved mother Kouassi Sally Georgette. May life bless you all.

BRIEF CONTENTS

ABSTRACT iii

RESUME iv

ACKNOWLEDGEMENTS v

BRIEF CONTENTS vii

INTRODUCTION 1

Part 1: A GLS ESTIMATION OF THE TWO-WAY RANDOM EFFECT MODEL WITH IDENTICAL TIME STRUCTURES 10

- 1.1. Identical AR(1) Error Structure 14
- 1.2. Identical MA(1) Error Structure 28
- 1.3. General Model with Identical Error Structure 46
- 1.4. FGLS Estimation of an Identical Time Structure Model 59

Part 2: A GLS ESTIMATION OF THE TWO-WAY RANDOM EFFECT MODEL WITH DIFFERENT TIME STRUCTURES 74

- 2.1. Autocorrelation and General Expression of the Variance-Covariance Matrix 75
- 2.2. GLS Estimation and Properties of the Estimators 84
- 2.3. FGLS Estimation of a Double Autocorrelation Model 118

CONCLUDING REMARKS 135

REFERENCES 138

TABLE OF CONTENTS 144

INTRODUCTION

The increasing availability of panel data has popularized the use of this particularly appealing type of datasets. This tendency is confirmed by recent papers dealing with a wide range of research questions. For instance, Bawn and Rosenbluth (2006) analyze the consequences of the number of parties in a government in 17 European countries over two decades, while Hecock (2006) examines the determinants of primary education spending with a panel of 29 states over 5 years. Likewise, Lee (2005), through a panel data approach, shows that a strong interaction between democracy and public sector size explains within-country income inequality. More challenging is the paper by Enders, Sachsida and Sandler (2006) who assess the impact of terrorism on the level of US foreign direct investments by the means of panel data models. Kelleher and Yackee (2006) and Stasavage (2005) in political science, South, Crowder and Chavez (2005) in demography, and less recently N'Gbo (1994) in agricultural economics, all these papers call for panel data frameworks.

As a definition, the term "panel data" refers to the pooling of observations on a cross-section of units (Baltagi, 2005), which are generally households, countries or firms. By surveying a number of individuals and following them over time, panel data yield a larger number of observations than one-dimensional datasets such as pure time series or pure cross-sections.

This feature is particularly helpful to developing countries which often lack reliable data. Especially, Sub-Saharan African countries, with financial shortages, often suffer from the unavailability of relevant and high-quality datasets in conducting their economic policies (Hoeffler, 2002). For those countries, panel data appear as a solution to various research and policy problems because of the greater amount of information they provide, say N-times or T-times more observations, if N and T denote the individual and time dimensions respectively. This tendency (the increasing use of panel data) still holds when dealing with integrated regions such as economic unions and monetary zones. Rather than collecting time-series on countries one by one, panel data offer the tools to study the economic questions across the countries over time through the same dataset. This has definitely prompted the popularity of the panel data in applied research, all over the word.

Furthermore, the worldwide use of panel data is also motivated by the fact that they provide such a rich environment for the development of estimation techniques and theoretical results. In fact, researchers have been able to use time-series and cross-sectional data to examine issues that could not be studied in either cross-sectional or time-series settings alone (Greene, 2008). Klevmarken (1989), Hsiao (2003) and Baltagi (2005) have listed several benefits from using panel data. One can control for individual heterogeneity by the means of time-specific or unit-specific parameters or disturbances or through some time-invariant and individual-invariant variables. Panel data give more informative data, more variability, less collinearity among the variables, more degrees of freedom and more efficiency. In studying the dynamics of adjustment within some cross-sectional units, panel data perform better than one-dimensional datasets. In addition, panel data models allow us to construct and test more complicated behavioral models and to identify and measure effects that are simply not detectable in pure cross-section or pure time-series data.

This popularity of the cross-section time series datasets keeps up-to-date the need for a precise knowledge of the appropriate econometric methods and techniques. In this thesis, we are interested in theoretical issues related to panel data approaches, especially to the error component models.

Following Balestra and Nerlove (1966), most of panel data studies model the disturbances of the regression equation using a random error component specification. Households' panels as well as macro panels with large number of units are often considered well suited to these models. The need for accounting for both individual and time-specific effects from a random errors perspective leads Wallace and Hussain (1969), Nerlove (1971), Amemiya (1971) and more recently Baltagi, Bresson and Pirotte (2005) among others, to the two-way error component model, i.e. models where the individual specificity μ_i and the time effect λ_t are taken as error terms, along with a more general disturbance v_{it} that vary with both individuals and time. The term "two-way" refers to the individual and time effects that are included while expressions like "error component" or "random effects" both describe the fact that these effects are considered as disturbances, rather than parameters. Some authors label this specification as the three error components model (see Magnus (1988) for example) stressing the presence of three error terms leading to an overall or composite disturbance u_{it}.

The classical error component model which has been extensively studied in the panel data literature assumes that the individual-specific effect is spherical, especially independently and identically normal[1] (hereafter IIN). This assumption also states the absence of spatial correlation unlike some recent studies allowing for cross-sectional dependency (see Hadri (2000) and Hadri and Larsson (2005) for instance). The error terms are all pairwise independent with zero as the mean. Moreover, this specification assumes that the correlation in the remainder disturbance is not spread among individuals, at any time period.

Another important feature that actually characterizes the classical two-way error component model is the fact that it denies the potential serial correlation in the time-varying disturbances, that is in the remainder error term (which varies with both dimensions) and in the time-specific effect. This absence of autocorrelation means that only correlation over time is due to the presence of the same individual across the panel. It is known as the equicorrelation error component model (Cameron and Trivedi, 2005). However, assuming equicorrelation results in the

[1] Normality is not actually compulsory. A simple assumption of independently and identically distributed (hereafter IID) errors is sufficient. Nonetheless, we have kept the normality hypothesis because of convergence purposes regarding the initial values in modeling the autocorrelation patterns.

same correlation coefficient for any two values u_{it} and u_{is} of the overall disturbance, no matter how far *t* is from *s*, and, more restrictively, is strictly positive (see Baltagi, 2005).

Unfortunately most of economic relationships cannot accommodate with such a limiting assumption. In the case of behavioral functions like investment or consumption, an unobserved shock this period can affect the consumer or investor decisions for at least the next few periods (see Baltagi, 2005). The time-varying disturbances (λ_t and ν_{it}) may contain their own correlation patterns, spreading shocks in the overall regression equation. As a result, the equicorrelation assumption is no longer justified.

However, serial correlation is not allowed for in the classical error component model, notably in the two-way version. Ignoring serial correlation when it is present results in consistent but inefficient estimates of the regression coefficients and biased standard errors (see Baltagi, 2005 and 2006). According to Cameron and Trivedi (2005), panel data models have errors that are usually autocorrelated over time for a given individual. As a consequence, they find that the correlation error leads to large bias in standard errors for pooled ordinary least squares (hereafter, OLS) if one does not account for the autocorrelation, and to relatively small efficiency gains as the lengths of the panel is increased.

From an empirical viewpoint, the issue of serial correlation is well known and several tests such as the Durbin-Watson one have fallen under the basic tasks performed in regression analysis. Autocorrelation in applied panel data is intuitively justified because of the time dimension. In her study on the significance of the African dummy variable through a Solow-type growth equation, Hoeffler (2002) considered a one-way error component model with a test of absence of serial correlation in the residuals. She ended up with the conclusion that the African dummy variable is not significant, and argued that earlier contradictory findings were based on misleading and careless econometric methods. Another illustrative example is provided by Cameron and Trivedi (2005) on the responsiveness of labor supply to wages, a very important question in labor economics. With data on 532 males on the 10 years from 1979 to 1988 from Ziliak (1997), they showed how important correcting for serial correlation is. They considered different panel data estimators, especially some panel-robust ones (first-differences, random effects GLS and random effects Maximum Likelihood (hereafter ML) estimators) and proved that the robust estimates performed better than the usual ones, since they controlled for heteroskedasticity and serial correlation. They even stated that correction for heteroskedasticity is usually much less important than the correction for panel correlation (see, Cameron and Trivedi, 2005, p. 712). Knowing that panel data intervene in a wide range of research areas one can definitely establish the necessity of overcoming the issue of autocorrelation in panel data when it occurs.

In their desire to account for autocorrelation when dealing with panel data, Kiefer (1980) and Bharghava et al (1982) considered the one-way fixed effects model with AR(1) remainder disturbances, while Schmidt (1983) extended the fixed effects model to cover cases with an arbitrary covariance matrix. The random effect version has also been examined. Lillard and Willis (1978) and Lillard and Weiss (1979) generalized the error component model to allow for an AR(1) process on the remainder disturbances of a one-way error component model. An extension to the MA(q) case was proposed by Baltagi and Li (1994) and a treatment of the remainder error term v_{it} following an autoregressive moving average ARMA(p,q) process was analyzed by MaCurdy (1982) and Galbraith and Zinde-Walsh (1995). Further studies including individual as well as time specificities are also available. In fact, the two-way error component model with serially correlated error terms has been considered by Revankar (1979) and Karlsson and Skoglund (2004). However, in their models, only λ_t was assumed to be serially correlated, following an AR(1) process in the former and an ARMA(p,q) series in the later. Magnus and Woodland (1988) generalized the Revankar (1979) model to a multivariate error components model with serially correlated errors and derived the corresponding ML estimator.

The new strand in investigation on the serial correlation issue in panel data is mainly directed towards non-stationary panels. This was the subject of papers by Baltagi and Kramer (1997) and Kao and Emerson (2002a, 2002b) which all include a time trend as a regressor. More recently, Baltagi, Kao and Liu (2007) examined the one-way error component model in which a regressor and the remainder disturbance are both serially correlated with the possibility to be nonstationary.

Acknowledging the need for those developments in the econometric literature, we however argue that some issues on serial correlation, especially in a two way error components framework, have to be tackled under some merely classical assumptions such as the absence of spatial correlation, the presence of nonstochastic regressors, stationarity of the errors, etc. One of these unsolved questions is the treatment of a double autocorrelation in three error components model where no more complication is added.

With panels autocorrelation can arise from two sources: (i) the autocorrelation function for the remainder error term v_{it}, and (ii) the autocorrelation function for the time-specific error term λ_t. We have labeled this situation as *double autocorrelation*. The idea of introducing a correlation in the composite terms $\left(\lambda_t + v_{it}\right)$ has firstly been considered by Nerlove (1970). However, he thereafter assumed λ_t null and his suggestion actually referred to v_{it} only. These two forms of autocorrelation can be combined to describe several cases of autocorrelation (on λ_t but not on v_{it}; or on v_{it} and not on λ_t for instance). This study focuses on the double autocorrelation case, which looks more complex than the single autocorrelation one.

The interest of a double autocorrelation in economics is straightforward. Many studies calling for panel data intend to assess individual as well as time heterogeneity, leading to a two way random effect model. The presence of the two error terms λ_t and v_{it} may hide a serial correlation. This type of autocorrelation is yet insufficiently analyzed by the econometric literature. Assuming a single correlation framework when a double one is required will yield an erroneous variance-covariance matrix, and therefore misleading *t*-statistics, inefficient estimates of the regression coefficients, biased standard errors etc. Finding a solution to such a correlation issue is a challenge that will increase the precision of estimations in panel data, for the benefit of researchers.

In this dissertation we generalize Baltagi and Li (1992a) treatment of the one-way random effect model in the presence of serial autocorrelation, and the single autocorrelation two-way approach of Revankar (1979). The particularity of this study consists in the structure of the panel data model, especially the two-way approach that we have retained, which itself requires GLS or ML rather than OLS estimation procedure. We share these authors' interest in a GLS estimator. It reasonably appears to us as the "natural" one rather than a ML estimator (or any other estimate) since autocorrelation creates a distortion in the variance-covariance matrix of the disturbances which are no longer spherical. A thorough study of the structure of this matrix of moments should yield a precise correction procedure that will remove the distortion created and bring us back to spherical disturbances. If achieved, then the GLS estimators derived would be Best Linear Unbiased Estimates (hereafter BLUE). Their consistency and efficiency would also be established (see Greene, 2008).

The choice of our modeling approach is justified by some empirical as well as theoretical considerations. Panel data are widely used in empirical research because of their time and cross-sectional dimensions. The unobserved individual effects captured by the panel models are assumed either correlated to the explanatory variables (leading to the fixed effects approach) or uncorrelated with the regressors (it is the random effect hypothesis). In the former, the differences between units are modeled as parametric shifts of the regression function, and such a model is designed for the cross-sectional units of the sample only, not for additional ones outside this sample. Keeping in mind that a sample is drawn from a population which actually may contain other units exhibiting the same characteristics as those present in that sample, the model should be considered constant over the subset of the population with the characteristics of interest. The individual specificities should then be modeled as randomly distributed across cross-sectional units, if we assume that they are randomly drawn from a large population. This is a reasonable assumption when dealing with households panels for instance. The current availability of large panel datasets leads us to build a framework allowing for large number of cross-section units and considerable time dimension, which is theoretically not well suited to the fixed effects approach. Moreover, the payoff of this form of modeling is that it greatly reduces the number of parameters to be estimated[2]. The two-way specification directly follows from the need to account for the individual and time dimensions in a more theoretically involved analysis of the autocorrelation issue.

The objective of the thesis is therefore to propose a theoretical as well as feasible GLS procedure which permits the estimation of a two-way random effects model in presence of autocorrelation. More specifically, we intend to

(i) present a theoretical GLS treatment of the autocorrelation in a two-way error components model when the time-varying disturbances follow identical correlation processes;
(ii) present a theoretical GLS treatment of the autocorrelation in a two-way error components model when the time-varying disturbances exhibit different correlation processes (real double autocorrelation case), with a special interest in the asymptotic properties of the estimators obtained;
(iii) propose a workable alternative under a general correlation pattern;
(iv) examine the applicability of the feasible version by assessing the case of AR(1) and MA(1) autocorrelations.

To achieve these goals, we allow both λ_t and ν_{it} to be serially correlated, but independently. They are assumed to either follow identical distributions (we talk about identical time structure) or exhibit different time processes in a double autocorrelation model. The investigation is conducted through two separate parts. Each part firstly considers two specific processes, say AR(1) and MA(1), then tackles a more general case in which the correlation pattern

[2] The cost of the random effects approach is the possibility of inconsistent estimates, should the assumption turn out to be inappropriate. (See Greene, 2008). In an empirical study, the researcher may perform the well-known Hausmann specification test to discriminate between the random and fixed effects models.

is undefined, and lastly suggests some estimators of the involved parameters in a FGLS perspective. Moreover, part 2 demonstrates the asymptotic properties of the GLS estimators obtained earlier.

More precisely, the remainder of the study is organized as follows: part 1 examines the case of the identical time structure shared by both λ_t and ν_{it} through four sections dealing consecutively with the AR(1) specification, the MA(1) process, the general approach with unspecified correlation pattern, and lastly a feasible version of each of the above models. The double autocorrelation structure is tackled by part 2 by the means of three sections. The first one is devoted to the derivation of a general variance-covariance matrix formula while the second section derived the subsequent GLS estimation procedure and the properties of the estimators obtained. The last section is interested in finding a FGLS procedure for the double autocorrelation situation. Some concluding remarks are then presented.

Part 1: A GLS ESTIMATION OF THE TWO-WAY RANDOM EFFECT MODEL WITH IDENTICAL TIME STRUCTURES

Any study aimed at modeling and then assessing autocorrelation is strongly affected by the pattern of that correlation. This modeling issue remains of the greatest relevance when dealing with a two-way random effect model.

Following Amemiya (1971) we consider the two-way error component model with the following disturbances:

$$u_{it} = \mu_i + \lambda_t + v_{it}\ ,\ i = 1,\ldots,N \text{ and } t = 1,\ldots,T\ , \tag{1.1}$$

where μ_i denotes the individual specificities and λ_t the time-specific effects whereas v_{it} represents the remaining errors, *i* the individuals and *t* the time periods. The assumptions related to the errors can be summarized as follows:

$$E\left(v_{it}\right) = E\left(\mu_i\right) = E\left(\lambda_t\right) = 0 \qquad \forall\ i,t \tag{1.2a}$$

$$E\left(v_{it}\mu_j\right) = E\left(v_{it}\lambda_s\right) = E\left(\mu_j\lambda_s\right) = 0 \quad \forall\ i, j, t \text{ and } s\ , \tag{1.2b}$$

$$\mathrm{Var}\left(\mu_i\right) = \sigma_\mu^2 \qquad \forall\, i = 1,\ldots,N\ , \tag{1.2c}$$

$$\mathrm{Cov}\left(\mu_i, \mu_j\right) = 0 \quad \forall\ i \neq j\ , \tag{1.2d}$$

$$\mathrm{Var}\left(v_{it}\right) = \sigma_v^2 \quad \forall\, i = 1,\ldots,N \text{ and } \forall\, t = 1,\ldots,T\ , \tag{1.2e}$$

$$\mathrm{Cov}\left(v_{it}, v_{js}\right) = 0 \qquad \forall\ i \neq j \text{ and } \forall\ t, s\ , \tag{1.2f}$$

$$\mathrm{Var}\left(\lambda_t\right) = \sigma_\lambda^2 \qquad \forall\, t = 1,\ldots,T\ , \tag{1.2g}$$

$$\mathrm{Cov}\left(\lambda_t, \lambda_s\right) = 0 \quad \forall\, t \neq s, \tag{1.2h}$$

$$\mathrm{Cov}\left(v_{it}, v_{is}\right) = 0 \quad \forall\, t \neq s. \tag{1.2i}$$

In this study, the two-way error component model considered is the following one:

$$y_{it} = \beta_0 + x_{it}\beta + u_{it},\ i = 1,\ldots,N \text{ and } t = 1,\ldots,T \tag{1.3}$$

where β_0 is the intercept and β is a $k \times 1$ vector[3] of slope coefficients, x_{it} is a $1 \times k$ row vector of explanatory variables which are uncorrelated with the disturbances u_{it} given by equation (1.1).

The general framework we have adopted is the following one. Let $r_{iv}\left(h\right) = r_v\left(h\right)$ and $r_\lambda\left(h\right)$ be the autocorrelation functions for the disturbances v_{it} and λ_t respectively. We then have,

$$\begin{cases} r_{iv}\left(h\right) = \dfrac{\mathrm{cov}\left(v_{it}, v_{i,t-h}\right)}{\sigma_v^2} = \dfrac{\gamma_{iv}\left(h\right)}{\gamma_{iv}\left(0\right)} = \dfrac{\gamma_v\left(h\right)}{\gamma_v\left(0\right)} \\ r_\lambda\left(h\right) = \dfrac{\mathrm{cov}\left(\lambda_t, \lambda_{t-h}\right)}{\sigma_\lambda^2} = \dfrac{\gamma_\lambda\left(h\right)}{\gamma_\lambda\left(0\right)} \end{cases} \tag{1.4}$$

where $h = 1,\ldots,t-1;\ t = 1,\ldots,T;\ i = 1,\ldots,N$, $\gamma_{iv}\left(h\right)$ and $\gamma_\lambda\left(h\right)$ being the autocovariance functions of v_{it} and λ_t respectively. Obviously, assumptions (1.2h) and (1.2i) no longer hold.

---

[3] Henceforth, we shall denote a matrix by a boldface capital letter as $\mathbf{\Sigma}_\lambda$ in (1.5a), while vectors and scalars are simply written with lowercase letters as β in (1.3) and μ_i in (1.1).

In vector form, we have

$$\mu = (\mu_1, \dots, \mu_N)', \ \lambda = (\lambda_1, \dots, \lambda_T)'$$

and

$$\nu = (\nu_{11}, \dots, \nu_{1T}, \dots, \nu_{N1}, \dots, \nu_{NT})' = (\nu_1 \quad \nu_2 \quad \cdots \quad \nu_N)'.$$

The variance-covariance matrices of λ_t and ν_{it} are respectively given by:

$$\mathbf{\Sigma_\lambda} = E(\lambda\lambda') = \sigma_\lambda^2 \begin{pmatrix} 1 & r_\lambda(1) & \cdots & r_\lambda(T-1) \\ r_\lambda(1) & 1 & & \vdots \\ \vdots & & \ddots & r_\lambda(1) \\ r_\lambda(T-1) & \cdots & r_\lambda(1) & 1 \end{pmatrix} = \sigma_\lambda^2 \mathbf{\Gamma_\lambda} \tag{1.5a}$$

and

$$\mathbf{\Sigma_\nu} = E(\nu\nu') = E\left[\begin{pmatrix} \nu_1 \\ \nu_2 \\ \vdots \\ \nu_N \end{pmatrix} (\nu_1' \quad \nu_2' \quad \cdots \quad \nu_N')\right] = \begin{bmatrix} E(\nu_1\nu_1') & E(\nu_1\nu_2') & \cdots & E(\nu_1\nu_N') \\ E(\nu_2\nu_1') & E(\nu_2\nu_2') & \cdots & E(\nu_2\nu_N') \\ \vdots & \vdots & \ddots & \vdots \\ E(\nu_N\nu_1') & E(\nu_N\nu_2') & \cdots & E(\nu_N\nu_N') \end{bmatrix}.$$

The elements of matrix $\mathbf{\Sigma_\nu}$ are such that

$$E(\nu_i\nu_i') = E\left[\begin{pmatrix} \nu_{i1} \\ \nu_{i2} \\ \vdots \\ \nu_{iT} \end{pmatrix} (\nu_{i1} \quad \nu_{i2} \quad \cdots \quad \nu_{iT})\right] = \begin{bmatrix} E(\nu_{i1}^2) & E(\nu_{i1}\nu_{i2}) & \cdots & E(\nu_{i1}\nu_{iT}) \\ E(\nu_{i2}\nu_{i1}) & E(\nu_{i2}^2) & \cdots & E(\nu_{i2}\nu_{iT}) \\ \vdots & \vdots & \ddots & \vdots \\ E(\nu_{iT}\nu_{i1}) & E(\nu_{iT}\nu_{i2}) & \cdots & E(\nu_{iT}^2) \end{bmatrix}$$

that is,

$$E\left(\nu_i\nu_i'\right)=\sigma_\nu^2\begin{pmatrix} 1 & r_\nu(1) & \cdots & r_\nu(T-1) \\ r_\nu(1) & 1 & & \vdots \\ \vdots & & \ddots & r_\nu(1) \\ r_\nu(T-1) & \cdots & r_\nu(1) & 1 \end{pmatrix}=\sigma_\nu^2\boldsymbol{\Gamma}_{\mathbf{v}} \qquad \forall i=1,\ldots,N.$$

and $E\left(\nu_i\nu_j'\right)=\mathbf{0} \quad \forall i\neq j$ since $E\left(\nu_{it}\nu_{js}\right)=0 \quad$ for $i\neq j, \quad \forall t,s.$

Consequently,

$$\boldsymbol{\Sigma}_{\mathbf{v}}=\sigma_\nu^2\begin{bmatrix} \boldsymbol{\Gamma}_{\mathbf{v}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma}_{\mathbf{v}} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Gamma}_{\mathbf{v}} \end{bmatrix}=\sigma_\nu^2\left(\mathbf{I}_{\mathbf{N}}\otimes\boldsymbol{\Gamma}_{\mathbf{v}}\right). \tag{1.5b}$$

In this part, we assume that both λ_t and ν_{it} follow the same process, with the same parameters. In other words, $\boldsymbol{\Gamma}_{\boldsymbol{\lambda}}=\boldsymbol{\Gamma}_{\mathbf{v}}=\boldsymbol{\Gamma}$.

This somewhat strong assumption is however very informative in the sense that it shows out all the implications of a serial correlation in a two-way error component model. Furthermore, since λ_t and ν_{it} behave identically, only one autocorrelation matrix needs to be inverted. This allows us to come out with the exact expression of the transformed data as a function of the original data.

Part 1 is organized as follows: section 1.1 presents the model in which both λ_t and ν_{it} come from the same AR(1) process, while section 1.2 deals with a model of MA(1) time errors. Next, section 1.3 generalizes this approach. Finally, section 1.4 presents a way of estimating the parameters in a FGLS framework.

1.1. Identical AR(1) Error Structure

In this section, the error terms are assumed to follow the same AR(1) process. We consider simple structures of the disturbances following an AR(1) process. Then, we explain the structure of the variance-covariance matrix of the transformed error terms and its spectral decomposition. Next, we present the GLS transformations of the original data aimed at correcting for the serial correlation in the particular context of the two-way structure. Lastly, the Best Quadratic Unbiased (hereafter, BQU) estimators are given.

1.1.1. Specification of the Model

We reconsider model (1.3), $y_{it} = \beta_0 + x_{it}\beta + u_{it}$ $i = 1,\dots,N$ and $t = 1,\dots,T$ where *i* denotes individuals and *t* time periods. The overall error term is displayed according to equation (1.1), i.e. $u_{it} = \mu_i + \lambda_t + \nu_{it}$, $i = 1,\dots,N$ and $t = 1,\dots,T$. The underlying disturbances ν_{it} and λ_t are assumed to follow the same AR(1) process. On the one hand, $\nu_{it} = \rho\nu_{i,t-1} + e_{it}$, with $|\rho| < 1$, and $e_{it} \sim IIN\left(0,\sigma_e^2\right)$ while on the other hand $\lambda_t = \rho\lambda_{t-1} + \varepsilon_t$, with $\varepsilon_t \sim IIN\left(0,\sigma_\varepsilon^2\right)$. For convergence purpose and under stationarity assumption, the initial values are defined as

$$\begin{cases} \nu_{i0} \sim N\left(0, \dfrac{\sigma_e^2}{1-\rho^2} = \sigma_\nu^2\right) \\ \lambda_0 \sim N\left(0, \dfrac{\sigma_\varepsilon^2}{1-\rho^2} = \sigma_\lambda^2\right). \end{cases}$$

This is a two-way error component model with classical individual effect, but serially correlated temporal error terms. In vector form,

$$u = \left(\mathbf{I_N} \otimes \iota_T\right)\mu + \left(\iota_N \otimes \mathbf{I_T}\right)\lambda + \nu \tag{1.6}$$

where ι_T and ι_N are vectors of one of dimension *T* and *N* respectively, I_N and I_T are identity matrices of dimension *N* and *T* respectively. Thus, u has mean zero and variance-covariance matrix

$$\mathbf{\Sigma} = E(uu') = \mathbf{\Sigma_v} + \sigma_\mu^2 \left[\mathbf{I_N} \otimes (\iota_T \iota_T') \right] + (\iota_N \iota_N') \otimes \mathbf{\Sigma_\lambda} \tag{1.7}$$

where $\mathbf{\Sigma_\lambda}$ and $\mathbf{\Sigma_v}$ are the variance-covariance matrices of the error terms vectors λ and ν respectively:

$\mathbf{\Sigma_\lambda} = E(\lambda\lambda') = \sigma_\lambda^2 \mathbf{\Gamma}$ and $\mathbf{\Sigma_v} = \sigma_\nu^2 (\mathbf{I_N} \otimes \mathbf{\Gamma})$.

In this subsection, we assume an autoregressive process of order one, leading to

$$\mathbf{\Gamma} = \begin{pmatrix} 1 & \rho & \cdots & \rho^{T-1} \\ \rho & 1 & & \vdots \\ \vdots & & \ddots & \rho \\ \rho^{T-1} & \cdots & \rho & 1 \end{pmatrix}. \tag{1.8}$$

The correlation coefficient is here:

$$\mathrm{Correl}(u_{it}, u_{js}) = \begin{cases} \left[\rho^{|t-s|} (\sigma_\lambda^2 + \sigma_\nu^2) + \sigma_\mu^2 \right] \big/ (\sigma_\mu^2 + \sigma_\lambda^2 + \sigma_\nu^2) & \text{for } i = j, t \neq s \\ \sigma_\lambda^2 \big/ (\sigma_\mu^2 + \sigma_\lambda^2 + \sigma_\nu^2) & \text{for } i \neq j, t = s \\ 1 & \text{for } i = j, t = s \\ 0 & \text{for } i \neq j, t \neq s. \end{cases} \tag{1.9}$$

Note that the correlation coefficient actually depends on the time length $|t-s|$, which is the aim of our specification of the serial correlation structure.

Finally, the variance-covariance matrix of u is

$$\mathbf{\Sigma} = E(uu') = \sigma_\nu^2 (\mathbf{I_N} \otimes \mathbf{\Gamma}) + \sigma_\mu^2 \left(\mathbf{I_N} \otimes (\iota_T \iota_T') \right) + \sigma_\lambda^2 (\iota_N \iota_N' \otimes \mathbf{\Gamma}). \tag{1.10}$$

1.1.2. Variance-Covariance of the Transformed Errors and Its Spectral Decomposition

The familiar Prais-Winsten (1954) transformation matrix of an AR(1) process is

$$\mathbf{C}=\begin{pmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & & \cdots & \vdots \\ 0 & -\rho & 1 & 0 & & \cdots & \vdots \\ 0 & 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & -\rho & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{pmatrix}. \tag{1.11}$$

The variance-covariance matrix of the transformed errors is,

$$\boldsymbol{\Sigma}^*=\left(\mathbf{I_N}\otimes\mathbf{C}\right)\boldsymbol{\Sigma}\left(\mathbf{I_N}\otimes\mathbf{C}'\right)$$

$$\boldsymbol{\Sigma}^*=\left(\mathbf{I_N}\otimes\mathbf{C}\right)\left(\mathbf{I_N}\otimes\left(\sigma_\nu^2\boldsymbol{\Gamma}\right)\right)\left(\mathbf{I_N}\otimes\mathbf{C}'\right)+\sigma_\mu^2\left(\mathbf{I_N}\otimes\mathbf{C}\right)\left(\mathbf{I_N}\otimes\iota_T\iota_T'\right)\left(\mathbf{I_N}\otimes\mathbf{C}'\right)+$$
$$\left(\mathbf{I_N}\otimes\mathbf{C}\right)\left(\iota_N\iota_N'\otimes\left(\sigma_\lambda^2\right)\boldsymbol{\Gamma}\right)\left(\mathbf{I_N}\otimes\mathbf{C}'\right)$$

$$\boldsymbol{\Sigma}^*=\left(\mathbf{I_N}\otimes\mathbf{C}\left(\sigma_\nu^2\boldsymbol{\Gamma}\right)\mathbf{C}'\right)+\sigma_\mu^2\left[\mathbf{I_N}\otimes\left(\mathbf{C}\iota_T\right)\left(\mathbf{C}\iota_T\right)'\right]+\left(\iota_N\iota_N'\otimes\mathbf{C}\left(\sigma_\lambda^2\boldsymbol{\Gamma}\right)\mathbf{C}'\right). \tag{1.12}$$

Following Baltagi and Li (1991), we set $\iota_T^\alpha=\left(\alpha,\iota_{T-1}'\right)$ where $\alpha=\sqrt{(1+\rho)/(1-\rho)}$, and $\mathbf{J}_\mathbf{T}^{\boldsymbol{\alpha}}=\left(\iota_T^\alpha\right)\left(\iota_T^\alpha\right)'$, $\mathbf{J_N}=\iota_N\iota_N'$. Moreover, we have

$$\begin{cases}\mathbf{C}\left(\sigma_\nu^2\boldsymbol{\Gamma}\right)\mathbf{C}'=\sigma_e^2\mathbf{I_T}\\ \mathbf{C}\left(\sigma_\lambda^2\boldsymbol{\Gamma}\right)\mathbf{C}'=\sigma_\varepsilon^2\mathbf{I_T}\\ \mathbf{C}\iota_T=(1-\rho)\iota_T^\alpha.\end{cases} \tag{1.13}$$

Therefore,

$$\boldsymbol{\Sigma}^*=\sigma_e^2\left(\mathbf{I_N}\otimes\mathbf{I_T}\right)+\sigma_\mu^2\left(1-\rho\right)^2\left(\mathbf{I_N}\otimes\mathbf{J}_\mathbf{T}^\alpha\right)+\sigma_\varepsilon^2\left(\mathbf{J_N}\otimes\mathbf{I_T}\right). \tag{1.14}$$

In order to get idempotent matrices we make the following transformations:

$\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} = \frac{\iota_T^{\alpha}\iota_T^{\alpha\prime}}{d_\alpha^2} = \frac{1}{d_\alpha^2}\mathbf{J}_{\mathbf{T}}^{\boldsymbol{\alpha}}$, $\bar{\mathbf{J}}_{\mathbf{N}} = \frac{\iota_N\iota_N'}{N} = \frac{1}{N}\mathbf{J}_{\mathbf{N}}$ where $d_\alpha^2 = \iota_T^{\alpha\prime}\iota_T^{\alpha} = \alpha^2 + T - 1$.

We then use Wansbeek and Kapteyn (1982, 1983) approach. In the expression of $\boldsymbol{\Sigma}^*$,

$\mathbf{I}_{\mathbf{N}}$ is replaced by $\mathbf{E}_{\mathbf{N}} + \bar{\mathbf{J}}_{\mathbf{N}}$ where $\mathbf{E}_{\mathbf{N}} = \mathbf{I}_{\mathbf{N}} - \bar{\mathbf{J}}_{\mathbf{N}}$;

$\mathbf{J}_{\mathbf{N}}$ is replaced by $\mathbf{N}\bar{\mathbf{J}}_{\mathbf{N}}$;

$\mathbf{J}_{\mathbf{T}}^{\boldsymbol{\alpha}}$ is replaced by $d_\alpha^2\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}$;

$\mathbf{I}_{\mathbf{T}}$ is replaced by $\mathbf{E}_{\mathbf{T}}^{\boldsymbol{\alpha}} + \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}$ where $\mathbf{E}_{\mathbf{T}}^{\boldsymbol{\alpha}} = \mathbf{I}_{\mathbf{T}} - \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}$.

We then obtain,

$$\boldsymbol{\Sigma}^* = \sigma_e^2\left[\left(\mathbf{E}_{\mathbf{N}} + \bar{\mathbf{J}}_{\mathbf{N}}\right)\otimes\left(\mathbf{E}_{\mathbf{T}}^{\boldsymbol{\alpha}} + \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right)\right] + d_\alpha^2(1-\rho)^2\sigma_\mu^2\left[\left(\mathbf{E}_{\mathbf{N}} + \bar{\mathbf{J}}_{\mathbf{N}}\right)\otimes\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right] + N\sigma_\varepsilon^2\left[\bar{\mathbf{J}}_{\mathbf{N}}\otimes\left(\mathbf{E}_{\mathbf{T}}^{\boldsymbol{\alpha}} + \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right)\right]$$

$$\begin{aligned}\boldsymbol{\Sigma}^* &= \sigma_e^2\left(\mathbf{E}_{\mathbf{N}}\otimes\mathbf{E}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) + \left(\sigma_e^2 + d_\alpha^2(1-\rho)^2\sigma_\mu^2\right)\left(\mathbf{E}_{\mathbf{N}}\otimes\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) + \left(\sigma_e^2 + N\sigma_\varepsilon^2\right)\left(\bar{\mathbf{J}}_{\mathbf{N}}\otimes\mathbf{E}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) + \\ &+ \left(\sigma_e^2 + d_\alpha^2(1-\rho)^2\sigma_\mu^2 + N\sigma_\varepsilon^2\right)\left(\bar{\mathbf{J}}_{\mathbf{N}}\otimes\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right).\end{aligned} \tag{1.15}$$

The spectral decomposition of $\boldsymbol{\Sigma}^*$ is,

$$\boldsymbol{\Sigma}^* = \sum_{i=1}^{4}\psi_i\mathbf{Q}_{\mathbf{i}} \tag{1.16}$$

with,

$\psi_1 = \sigma_e^2$, $\psi_2 = \sigma_e^2 + d_\alpha^2(1-\rho)^2\sigma_\mu^2$, $\psi_3 = \sigma_e^2 + N\sigma_\varepsilon^2$, $\psi_4 = \sigma_e^2 + d_\alpha^2(1-\rho)^2\sigma_\mu^2 + N\sigma_\varepsilon^2$,

$\mathbf{Q}_{\mathbf{1}} = \mathbf{E}_{\mathbf{N}}\otimes\mathbf{E}_{\mathbf{T}}^{\boldsymbol{\alpha}}$, $\mathbf{Q}_{\mathbf{2}} = \mathbf{E}_{\mathbf{N}}\otimes\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}$, $\mathbf{Q}_{\mathbf{3}} = \bar{\mathbf{J}}_{\mathbf{N}}\otimes\mathbf{E}_{\mathbf{T}}^{\boldsymbol{\alpha}}$, and $\mathbf{Q}_{\mathbf{4}} = \bar{\mathbf{J}}_{\mathbf{N}}\otimes\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}$.

1.1.3. GLS Transformation

However, at this step, the transformed composite disturbance u^* is still not spherical. This issue can be overcome by an additional GLS transformation. Following Fuller and Battese (1974), the new transformation matrix could be $\sigma_e \boldsymbol{\Sigma}^{*-1/2}$.[4]

From the spectral decomposition of $\boldsymbol{\Sigma}^*$, it follows that:

$$\sigma_e \boldsymbol{\Sigma}^{*-1/2} = \sum_{i=1}^{4} \frac{\sigma_e}{\psi_i^{1/2}} \mathbf{Q_i} = \mathbf{Q}_1 + \sum_{i=2}^{4} \frac{\sigma_e}{\psi_i^{1/2}} \mathbf{Q_i} . \tag{1.17}$$

By premultiplying the Prais-Winsten transformed observations $y^* = \left(\mathbf{I_N} \otimes \mathbf{C}\right) y$ by $\sigma_e \boldsymbol{\Sigma}^{\text{*-1/2}}$, one gets $y^{**} = \sigma_e \boldsymbol{\Sigma}^{\text{*-1/2}} y^*$.[5]

The goal of this subsection is to derive the typical elements of y^{**} in terms of y^*, i.e. y_{it}^{**} as a function of y_{it}^*.

From the spectral decomposition in equation (1.16), it follows

---

[4] As in any GLS approach, the same transformation is applied to each column vector of the matrix **X.** Therefore any result obtained with vector *y* is totally and identically relevant for the columns of **X**. This will remain true all along the thesis.

[5] We then have $E\left(u^{**}u^{**\prime}\right) = \sigma_e^2 \boldsymbol{\Sigma}^{\text{*-1/2}} E\left(u^{**}u^{**\prime}\right) \boldsymbol{\Sigma}^{\text{*-1/2}} = \sigma_e^2 \mathbf{I_{NT}}$. Any constant could be used instead of σ_e^2, making u^{**} immediately spherical. We have retained σ_e^2 so that $\frac{\sigma_e}{\psi_i^{1/2}} \mathbf{Q_i}$ collapses into $\mathbf{Q_i}$.

$$\sigma_e \mathbf{\Sigma}^{*-1/2} y^* = \sum_{i=1}^{4} \frac{\sigma_e}{\psi_i^{1/2}} \mathbf{Q_i} y^* = \mathbf{Q}_1 y^* + \sum_{i=2}^{4} \frac{\sigma_e}{\psi_i^{1/2}} \mathbf{Q_i} y^* . \tag{1.18}$$

Firstly, we deal with the product $\mathbf{Q}_1 y^*$. We can write,

$$\begin{aligned}\mathbf{Q_1} y^* &= \left(\mathbf{E_N} \otimes \mathbf{E_T^\alpha}\right) y^* = \left[\left(\mathbf{I_N} - \overline{\mathbf{J}}_\mathbf{N}\right) \otimes \left(\mathbf{I_T} - \overline{\mathbf{J}}_\mathbf{T}^\alpha\right)\right] y^* \\ &= \left(\mathbf{I_N} \otimes \mathbf{I_T}\right) y^* - \underbrace{\left(\mathbf{I_N} \otimes \overline{\mathbf{J}}_\mathbf{T}^\alpha\right) y^*}_{(A)} - \underbrace{\left(\overline{\mathbf{J}}_\mathbf{N} \otimes \mathbf{I_T}\right) y^*}_{(B)} + \underbrace{\left(\overline{\mathbf{J}}_\mathbf{N} \otimes \overline{\mathbf{J}}_\mathbf{T}^\alpha\right) y^*}_{(C)}\end{aligned}$$

where (*A*), (*B*) and (*C*) have to be determined explicitly. We have

$$(A) = \left(\mathbf{I_N} \otimes \overline{\mathbf{J}}_\mathbf{T}^\alpha\right) y^* = \mathrm{diag}\left[\overline{\mathbf{J}}_\mathbf{T}^\alpha\right] \begin{bmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_N^* \end{bmatrix} = \begin{bmatrix} \overline{\mathbf{J}}_\mathbf{T}^\alpha y_1^* \\ \overline{\mathbf{J}}_\mathbf{T}^\alpha y_2^* \\ \vdots \\ \overline{\mathbf{J}}_\mathbf{T}^\alpha y_N^* \end{bmatrix} \text{ where}$$

$$\overline{\mathbf{J}}_\mathbf{T}^\alpha y_i^* = \frac{1}{d_\alpha^2} \left[\begin{pmatrix} \alpha \\ 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} \alpha & 1 & \cdots & 1 \end{pmatrix} \right] y_i^* = \frac{1}{d_\alpha^2} \begin{bmatrix} \alpha^2 & \alpha & \cdots & \alpha \\ \alpha & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_{i1}^* \\ y_{i2}^* \\ \vdots \\ y_{iT}^* \end{bmatrix} = \frac{1}{d_\alpha^2} \begin{bmatrix} \alpha\left(\alpha y_{i1}^* + \sum_{t=2}^{T} y_{it}^*\right) \\ \alpha y_{i1}^* + \sum_{t=2}^{T} y_{it}^* \\ \vdots \\ \alpha y_{i1}^* + \sum_{t=2}^{T} y_{it}^* \end{bmatrix} = \begin{bmatrix} \alpha b_i \\ b_i \\ \vdots \\ b_i \end{bmatrix}.$$

with $b_i = \dfrac{1}{d_\alpha^2}\left(\alpha y_{i1}^* + \sum_{t=2}^{T} y_{it}^*\right) \quad \forall i = 1, \ldots, N.$

Hence, $(A) = \begin{pmatrix} \alpha b_1 & \cdots & b_1 & \vdots & \cdots & \vdots & \alpha b_N & \cdots & b_N \end{pmatrix}'$.

Likewise, one has

$$(B) = \left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \mathbf{I}_{\mathbf{T}}\right) y^* = \frac{1}{N}\begin{bmatrix} \mathbf{I}_{\mathbf{T}} & \cdots & \mathbf{I}_{\mathbf{T}} \\ \vdots & \ddots & \vdots \\ \mathbf{I}_{\mathbf{T}} & \cdots & \mathbf{I}_{\mathbf{T}} \end{bmatrix}\begin{bmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_N^* \end{bmatrix} = \frac{1}{N}\begin{bmatrix} \sum_{i=1}^{N} y_1^* \\ \sum_{i=1}^{N} y_2^* \\ \vdots \\ \sum_{i=1}^{N} y_N^* \end{bmatrix} \text{ where}$$

$$\frac{1}{N}\sum_{i=1}^{N} y_i^* = \frac{1}{N}\left[\begin{pmatrix} y_{11}^* \\ y_{12}^* \\ \vdots \\ y_{1T}^* \end{pmatrix} + \begin{pmatrix} y_{21}^* \\ y_{22}^* \\ \vdots \\ y_{2T}^* \end{pmatrix} + \cdots + \begin{pmatrix} y_{N1}^* \\ y_{N2}^* \\ \vdots \\ y_{NT}^* \end{pmatrix}\right] = \frac{1}{N}\begin{pmatrix} \sum_{i=1}^{N} y_{i1}^* \\ \sum_{i=1}^{N} y_{i2}^* \\ \vdots \\ \sum_{i=1}^{N} y_{iT}^* \end{pmatrix} = \begin{pmatrix} \bar{y}_{\bullet 1}^* \\ \bar{y}_{\bullet 2}^* \\ \vdots \\ \bar{y}_{\bullet T}^* \end{pmatrix}.$$

Hence, $(B) = \left(\bar{y}_{\bullet 1}^* \quad \cdots \quad \bar{y}_{\bullet T}^* \quad \vdots \quad \cdots \quad \vdots \quad \bar{y}_{\bullet 1}^* \quad \cdots \quad \bar{y}_{\bullet T}^*\right)'$.

The last term (C) of $\mathbf{Q}_1 y^*$ can also be expanded as,

$$(C) = \left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^* = \frac{1}{N}\begin{bmatrix} \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} & \cdots & \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} \\ \vdots & \ddots & \vdots \\ \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} & \cdots & \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} \end{bmatrix}\begin{bmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_N^* \end{bmatrix} = \frac{1}{N}\begin{bmatrix} \sum_{i=1}^{N} \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_i^* \\ \sum_{i=1}^{N} \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_i^* \\ \vdots \\ \sum_{i=1}^{N} \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_i^* \end{bmatrix} \text{ where}$$

$$\frac{1}{N}\sum_{i=1}^{N}\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}y_i^{*}=\frac{1}{N}\left[\begin{pmatrix}\alpha b_1\\ b_1\\ \vdots\\ b_1\end{pmatrix}+\begin{pmatrix}\alpha b_2\\ b_2\\ \vdots\\ b_2\end{pmatrix}+\cdots+\begin{pmatrix}\alpha b_N\\ b_N\\ \vdots\\ b_N\end{pmatrix}\right]=\frac{1}{N}\begin{pmatrix}\alpha\sum_{i=1}^{N}b_i\\ \sum_{i=1}^{N}b_i\\ \vdots\\ \sum_{i=1}^{N}b_i\end{pmatrix}=\begin{pmatrix}\alpha b\\ b\\ \vdots\\ b\end{pmatrix}\qquad\text{with}\qquad b=\frac{1}{N}\sum_{i=1}^{N}b_i\ .$$

$$(C)=\begin{pmatrix}\alpha b & \cdots & b & \vdots & \cdots & \vdots & \alpha b & \cdots & b\end{pmatrix}'.$$

From (A), (B), and (C), one gets

$$\mathbf{Q_1}y^{*}=\begin{pmatrix}y_{11}^{*}\\ \vdots\\ y_{1T}^{*}\\ \cdots\cdots\\ \vdots\\ \cdots\cdots\\ y_{N1}^{*}\\ \vdots\\ y_{NT}^{*}\end{pmatrix}-\begin{pmatrix}\alpha b_1\\ \vdots\\ b_1\\ \cdots\cdots\\ \vdots\\ \cdots\cdots\\ \alpha b_N\\ \vdots\\ b_N\end{pmatrix}-\begin{pmatrix}\bar{y}_{\bullet1}^{*}\\ \vdots\\ \bar{y}_{\bullet T}^{*}\\ \cdots\cdots\\ \vdots\\ \cdots\cdots\\ \bar{y}_{\bullet1}^{*}\\ \vdots\\ \bar{y}_{\bullet T}^{*}\end{pmatrix}+\begin{pmatrix}\alpha b\\ \vdots\\ b\\ \cdots\cdots\\ \vdots\\ \cdots\cdots\\ \alpha b\\ \vdots\\ b\end{pmatrix}=\begin{pmatrix}\left(y_{11}^{*}-\bar{y}_{\bullet1}^{*}\right)-\alpha\left(b_1-b\right)\\ \left(y_{12}^{*}-\bar{y}_{\bullet2}^{*}\right)-\left(b_1-b\right)\\ \vdots\\ \left(y_{1T}^{*}-\bar{y}_{\bullet T}^{*}\right)-\left(b_1-b\right)\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \vdots\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \left(y_{N1}^{*}-\bar{y}_{\bullet1}^{*}\right)-\alpha\left(b_N-b\right)\\ \left(y_{N2}^{*}-\bar{y}_{\bullet2}^{*}\right)-\left(b_N-b\right)\\ \vdots\\ \left(y_{NT}^{*}-\bar{y}_{\bullet T}^{*}\right)-\left(b_N-b\right)\end{pmatrix}\qquad(1.19)$$

The second product of interest is $\mathbf{Q}_2y^{*}$. Its expression at the observation level can also be determined through a similar procedure. We have,

$$\mathbf{Q}_2y^{*}=\left(\mathbf{E_N}\otimes\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right)y^{*}=\left[\left(\mathbf{I_N}-\bar{\mathbf{J}}_{\mathbf{N}}\right)\otimes\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right]y^{*}=\underbrace{\left(\mathbf{I_N}\otimes\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right)y^{*}}_{(A)}-\underbrace{\left(\bar{\mathbf{J}}_{\mathbf{N}}\otimes\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right)y^{*}}_{(C)}$$

$$\mathbf{Q}_2 y^* = \begin{pmatrix} \alpha b_1 \\ \vdots \\ b_1 \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha b_N \\ \vdots \\ b_N \end{pmatrix} - \begin{pmatrix} \alpha b \\ \vdots \\ b \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha b \\ \vdots \\ b \end{pmatrix} = \begin{pmatrix} \alpha\left(b_1 - b\right) \\ \vdots \\ b_1 - b \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha\left(b_N - b\right) \\ \vdots \\ b_N - b \end{pmatrix}. \tag{1.20}$$

The third product $\mathbf{Q}_3 y^*$ can be rewritten as,

$$\mathbf{Q}_3 y^* = \left(\bar{\mathbf{J}}_\mathbf{N} \otimes \mathbf{E}_\mathbf{T}^{\boldsymbol{\alpha}}\right) y^* = \left[\bar{\mathbf{J}}_\mathbf{N} \otimes \left(\mathbf{I}_\mathbf{T} - \bar{\mathbf{J}}_\mathbf{T}^{\boldsymbol{\alpha}}\right)\right] y^* = \underbrace{\left(\bar{\mathbf{J}}_\mathbf{N} \otimes \mathbf{I}_\mathbf{T}\right) y^*}_{(B)} - \underbrace{\left(\bar{\mathbf{J}}_\mathbf{N} \otimes \bar{\mathbf{J}}_\mathbf{T}^{\boldsymbol{\alpha}}\right) y^*}_{(C)}$$

$$\mathbf{Q}_3 y^* = \begin{pmatrix} \bar{y}_{\bullet 1}^* \\ \vdots \\ \bar{y}_{\bullet T}^* \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \bar{y}_{\bullet 1}^* \\ \vdots \\ \bar{y}_{\bullet T}^* \end{pmatrix} - \begin{pmatrix} \alpha b \\ \vdots \\ b \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha b \\ \vdots \\ b \end{pmatrix} = \begin{pmatrix} \bar{y}_{\bullet 1}^* - \alpha b \\ \bar{y}_{\bullet 2}^* - b \\ \vdots \\ \bar{y}_{\bullet T}^* - b \\ \cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots \\ \bar{y}_{\bullet 1}^* - \alpha b \\ \bar{y}_{\bullet 2}^* - b \\ \vdots \\ \bar{y}_{\bullet T}^* - b \end{pmatrix}. \tag{1.21}$$

The last product $\mathbf{Q}_4 y^*$ is straightforward,

$$\mathbf{Q}_4 y^* = \left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^* = (C) = \begin{pmatrix} \alpha b \\ \vdots \\ b \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha b \\ \vdots \\ b \end{pmatrix}. \tag{1.22}$$

As a consequence, the equation (1.18) becomes

$$\sigma_e \mathbf{\Sigma}^{*-1/2} y^* = \begin{pmatrix} \left(y_{11}^* - \bar{y}_{\bullet 1}^*\right) + \alpha\left(b_1 - b\right) \\ \left(y_{12}^* - \bar{y}_{\bullet 2}^*\right) + \left(b_1 - b\right) \\ \vdots \\ \left(y_{1T}^* - \bar{y}_{\bullet T}^*\right) + \left(b_1 - b\right) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \left(y_{N1}^* - \bar{y}_{\bullet 1}^*\right) + \alpha\left(b_N - b\right) \\ \left(y_{N2}^* - \bar{y}_{\bullet 2}^*\right) + \left(b_N - b\right) \\ \vdots \\ \left(y_{NT}^* - \bar{y}_{\bullet T}^*\right) + \left(b_N - b\right) \end{pmatrix} + \frac{\sigma_e}{\psi_2^{1/2}} \begin{pmatrix} \alpha\left(b_1 - b\right) \\ b_1 - b \\ \vdots \\ b_1 - b \\ \cdots\cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha\left(b_N - b\right) \\ b_N - b \\ \vdots \\ b_N - b \end{pmatrix} + \frac{\sigma_e}{\psi_3^{1/2}} \begin{pmatrix} \bar{y}_{\bullet 1}^* - \alpha b \\ \bar{y}_{\bullet 2}^* - b \\ \vdots \\ \bar{y}_{\bullet T}^* - b \\ \cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots \\ \bar{y}_{\bullet 1}^* - \alpha b \\ \bar{y}_{\bullet 2}^* - b \\ \vdots \\ \bar{y}_{\bullet T}^* - b \end{pmatrix} + \frac{\sigma_e}{\psi_4^{1/2}} \begin{pmatrix} \alpha b \\ \vdots \\ b \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha b \\ \vdots \\ b \end{pmatrix}$$

$$
\sigma_e \mathbf{\Sigma}^{*-1/2} y^* = \begin{pmatrix}
y_{11}^* - \left(1-\frac{\sigma_e}{\psi_3^{1/2}}\right)\bar{y}_{\bullet 1}^* - \alpha\left(1-\frac{\sigma_e}{\psi_2^{1/2}}\right)b_1 + \left(\left(1-\frac{\sigma_e}{\psi_2^{1/2}}\right)+\left(1-\frac{\sigma_e}{\psi_3^{1/2}}\right)+\frac{\sigma_e}{\psi_4^{1/2}}-1\right)\alpha b \\
y_{12}^* - \left(1-\frac{\sigma_e}{\psi_3^{1/2}}\right)\bar{y}_{\bullet 1}^* - \left(1-\frac{\sigma_e}{\psi_2^{1/2}}\right)b_1 + \left(\left(1-\frac{\sigma_e}{\psi_2^{1/2}}\right)+\left(1-\frac{\sigma_e}{\psi_3^{1/2}}\right)+\frac{\sigma_e}{\psi_4^{1/2}}-1\right)b \\
\vdots \\
y_{1T}^* - \left(1-\frac{\sigma_e}{\psi_3^{1/2}}\right)\bar{y}_{\bullet T}^* - \left(1-\frac{\sigma_e}{\psi_2^{1/2}}\right)b_1 + \left(\left(1-\frac{\sigma_e}{\psi_2^{1/2}}\right)+\left(1-\frac{\sigma_e}{\psi_3^{1/2}}\right)+\frac{\sigma_e}{\psi_4^{1/2}}-1\right)b \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\vdots \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
y_{N1}^* - \left(1-\frac{\sigma_e}{\psi_3^{1/2}}\right)\bar{y}_{\bullet 1}^* - \alpha\left(1-\frac{\sigma_e}{\psi_2^{1/2}}\right)b_N + \left(\left(1-\frac{\sigma_e}{\psi_2^{1/2}}\right)+\left(1-\frac{\sigma_e}{\psi_3^{1/2}}\right)+\frac{\sigma_e}{\psi_4^{1/2}}-1\right)\alpha b \\
y_{N2}^* - \left(1-\frac{\sigma_e}{\psi_3^{1/2}}\right)\bar{y}_{\bullet 2}^* - \left(1-\frac{\sigma_e}{\psi_2^{1/2}}\right)b_N + \left(\left(1-\frac{\sigma_e}{\psi_2^{1/2}}\right)+\left(1-\frac{\sigma_e}{\psi_3^{1/2}}\right)+\frac{\sigma_e}{\psi_4^{1/2}}-1\right)b \\
\vdots \\
y_{NT}^* - \left(1-\frac{\sigma_e}{\psi_3^{1/2}}\right)\bar{y}_{\bullet T}^* - \left(1-\frac{\sigma_e}{\psi_2^{1/2}}\right)b_N + \left(\left(1-\frac{\sigma_e}{\psi_2^{1/2}}\right)+\left(1-\frac{\sigma_e}{\psi_3^{1/2}}\right)+\frac{\sigma_e}{\psi_4^{1/2}}-1\right)b
\end{pmatrix}
$$

$$y^{**} = \sigma_e \mathbf{\Sigma}^{*-1/2} y^* = \begin{pmatrix} y_{11}^* - \theta_2 \bar{y}_{\bullet 1}^* - \alpha\theta_1 b_1 + \left(\theta_1 + \theta_2 + \dfrac{\sigma_e}{\psi_4^{1/2}} - 1\right)\alpha b \\ y_{12}^* - \theta_2 \bar{y}_{\bullet 2}^* - \theta_1 b_1 + \left(\theta_1 + \theta_2 + \dfrac{\sigma_e}{\psi_4^{1/2}} - 1\right) b \\ \vdots \\ y_{1T}^* - \theta_2 \bar{y}_{\bullet T}^* - \theta_1 b_1 + \left(\theta_1 + \theta_2 + \dfrac{\sigma_e}{\psi_4^{1/2}} - 1\right) b \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ y_{N1}^* - \theta_2 \bar{y}_{\bullet 1}^* - \alpha\theta_1 b_N + \left(\theta_1 + \theta_2 + \dfrac{\sigma_e}{\psi_4^{1/2}} - 1\right)\alpha b \\ y_{N2}^* - \theta_2 \bar{y}_{\bullet 2}^* - \theta_1 b_N + \left(\theta_1 + \theta_2 + \dfrac{\sigma_e}{\psi_4^{1/2}} - 1\right) b \\ \vdots \\ y_{NT}^* - \theta_2 \bar{y}_{\bullet T}^* - \theta_1 b_N + \left(\theta_1 + \theta_2 + \dfrac{\sigma_e}{\psi_4^{1/2}} - 1\right) b \end{pmatrix} = \begin{pmatrix} y_{11}^* - \alpha\theta_1 b_1 - \theta_2 \bar{y}_{\bullet 1}^* + \theta_3 \alpha b \\ y_{12}^* - \theta_1 b_1 - \theta_2 \bar{y}_{\bullet 2}^* + \theta_3 b \\ \vdots \\ y_{1T}^* - \theta_1 b_1 - \theta_2 \bar{y}_{\bullet T}^* + \theta_3 b \\ \cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots \\ y_{N1}^* - \alpha\theta_1 b_N - \theta_2 \bar{y}_{\bullet 1}^* + \theta_3 \alpha b \\ y_{N2}^* - \theta_1 b_N - \theta_2 \bar{y}_{\bullet 2}^* + \theta_3 b \\ \vdots \\ y_{NT}^* - \theta_1 b_N - \theta_2 \bar{y}_{\bullet T}^* + \theta_3 b \end{pmatrix}.$$

Thus, the typical element of the vector y^{**} is given by

$$y_{it}^{**} = \begin{cases} y_{i1}^* - \theta_1 \alpha b_i - \theta_2 \bar{y}_{\bullet 1}^* + \theta_3 \alpha b & \text{if} \quad t = 1 \\ y_{it}^* - \theta_1 b_i - \theta_2 \bar{y}_{\bullet t}^* + \theta_3 b & \text{if} \quad t = 2, \ldots, T \end{cases} \tag{1.23}$$

where

$$\theta_1 = 1 - \frac{\sigma_e}{\psi_2^{1/2}}, \; \theta_2 = 1 - \frac{\sigma_e}{\psi_3^{1/2}}, \; \theta_3 = \theta_1 + \theta_2 + \frac{\sigma_e}{\psi_4^{1/2}} - 1, \; b = \sum_{i=1}^{N} b_i, \tag{1.24}$$

and

$$b_i = \frac{1}{d_\alpha^2}\left(\alpha y_{i1}^* + \sum_{t=2}^{T} y_{it}^*\right). \tag{1.25}$$

The b_is are weighted averages of Prais-Winsten transformed observations with a special weight α / d_α^2 given to the first observation.

Likewise the one-way AR(1) serially correlated error component model (Baltagi, 2005), the two-way model can be estimated through two steps: firstly, by applying the Prais-Winsten transformation as it is usually done in the time series literature, and lastly by subtracting a pseudo-average from these transformed data.

The estimator associated to this last transformation is defined by

$$\eta_{GLS} = \left(\mathbf{X}^{**\prime}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}y^{**}. \tag{1.26}$$

Moreover, it is possible to reduce this procedure to a one-step one, since y^{**} can be directly expressed in terms of y:

$$y^{**} = \sigma_e \boldsymbol{\Sigma}^{*-1/2}\left(\mathbf{I_N} \otimes \mathbf{C}\right) y .$$

It is also possible to derive their typical elements. Knowing that

$$\forall i = 1,\dots,N \qquad y_{it}^{*} = \begin{cases} \sqrt{1-\rho^2}\, y_{i1} & if\ t = 1 \\ y_{it} - \rho y_{i,t-1} & if\ t = 2,\dots,T, \end{cases}$$

one can express y_{it}^{**} in terms of y_{it}. We have:

$$y_{it}^{**} = \begin{cases} \sqrt{1-\rho^2}\, y_{i1} - \theta_1 \alpha b_i - \dfrac{\theta_2}{N}\sum_{i=1}^{N}\left(\sqrt{1-\rho^2}\, y_{i1}\right) + \theta_3 \alpha b & if\ t = 1 \\ y_{it} - \rho y_{i,t-1} - \theta_1 b_i - \dfrac{\theta_2}{N}\sum_{i=1}^{N}\left(y_{it} - \rho y_{i,t-1}\right) + \theta_3 b & if\ t = 2,\dots,T. \end{cases}$$

$$y_{it}^{**} = \begin{cases} \sqrt{1-\rho^2}\, y_{i1} - \theta_1 \alpha b_i - \theta_2 \sqrt{1-\rho^2} \left(\dfrac{1}{N} \sum_{i=1}^{N} y_{i1} \right) + \theta_3 \alpha b & \text{if } t = 1 \\ y_{it} - \rho y_{i,t-1} - \theta_1 b_i - \theta_2 \left(\dfrac{1}{N} \sum_{i=1}^{N} y_{it} - \rho \dfrac{1}{N} \sum_{i=1}^{N} y_{i,t-1} \right) + \theta_3 b & \text{if } t = 2,\ldots,T. \end{cases}$$

$$y_{it}^{**} = \begin{cases} \sqrt{1-\rho^2} \left(y_{i1} - \theta_2 \bar{y}_{\bullet 1} \right) - \theta_1 \alpha b_i + \theta_3 \alpha b & \text{if } t = 1 \\ y_{it} - \rho y_{i,t-1} - \theta_1 b_i - \theta_2 \left(\bar{y}_{\bullet t} - \rho \bar{y}_{\bullet, t-1} \right) + \theta_3 b & \text{if } t = 2,\ldots,T. \end{cases}$$

$$y_{it}^{**} = \begin{cases} \sqrt{1-\rho^2} \left(y_{i1} - \theta_2 \bar{y}_{\bullet 1} \right) - \theta_1 \alpha b_i + \theta_3 \alpha b & \text{if } t = 1 \\ \left(y_{it} - \theta_2 \bar{y}_{\bullet t} \right) - \rho \left(y_{i,t-1} - \theta_2 \bar{y}_{\bullet, t-1} \right) - \theta_1 b_i + \theta_3 b & \text{if } t = 2,\ldots,T. \end{cases} \tag{1.27}$$

We shall mention that b_i s can now be seen as a weighted average of the original observations:

$$b_i = \frac{1}{d_\alpha^2} \left(\alpha y_{i1}^* + \sum_{t=2}^{T} y_{it}^* \right) = \frac{1}{d_\alpha^2} \left(\alpha \sqrt{1-\rho^2}\, y_{i1} + \sum_{t=2}^{T} \left(y_{it} - \rho y_{i,t-1} \right) \right)$$

$$b_i = \frac{1}{d_\alpha^2} \left(\sqrt{\frac{1+\rho}{1-\rho}} \sqrt{1-\rho^2}\, y_{i1} + \sum_{t=2}^{T} y_{it} - \rho \sum_{t=2}^{T} y_{i,t-1} \right) = \frac{1}{d_\alpha^2} \left[(1+\rho) y_{i1} + \left(\sum_{t=2}^{T-1} y_{it} + y_{iT} \right) - \rho \left(\sum_{t=2}^{T-1} y_{it} + y_{i1} \right) \right]$$

$$b_i = \frac{1}{d_\alpha^2} \left[(1+\rho-\rho) y_{i1} + \sum_{t=2}^{T-1} y_{it} - \rho \sum_{t=2}^{T-1} y_{it} + y_{iT} \right]$$

$$b_i = \frac{1}{d_\alpha^2} \left(y_{i1} + (1-\rho) \sum_{t=2}^{T-1} y_{it} + y_{iT} \right) \tag{1.28}$$

where b_is are now seen as weighted average of the original observations.

1.1.4. BQU Estimates

The BQU estimates of the variance components arise naturally from the spectral decomposition of the variance-covariance matrix $\mathbf{\Sigma^*}$ (see Baltagi (2005), p. 36). We have:

$E\left(\mathbf{Q_i}u^*\right)=\mathbf{Q_i}E\left(u^*\right)=\mathbf{Q_i}\left(\mathbf{I_N}\otimes\mathbf{C}\right)E\left(u\right)=0$ and

$$\mathrm{Var}\left(\mathbf{Q_i}u^*\right)=\mathbf{Q_i}E\left(u^*u^{*\prime}\right)\mathbf{Q_i}=\mathbf{Q_i}\mathbf{\Sigma^*}\mathbf{Q_i}=\mathbf{Q_i}\left(\psi_i\mathbf{Q_i}+\sum_{j\neq i}\psi_j\mathbf{Q_j}\right)\mathbf{Q_i}=\psi_i\mathbf{Q_i}\,.$$

In other words,

$$\mathbf{Q_i}u^*\sim(0,\psi_i\mathbf{Q_i}),\ \text{for } i=1,\ldots,4. \tag{1.29}$$

The best quadratic unbiased estimator of ψ_i is equal to $\hat{\psi}_i=\dfrac{u^{*\prime}\mathbf{Q_i}u^*}{\mathrm{trace}(\mathbf{Q_i})}$, for all *i*.

$$\text{Thus, }\begin{cases}\hat{\sigma}_e^2=\dfrac{u^{*\prime}\left(\mathbf{E_N}\otimes\mathbf{E_T^{\alpha}}\right)u^*}{\left(N-1\right)\left(T-1\right)}\\[2ex] \hat{\sigma}_e^2+\hat{d}_\alpha^2(1-\hat{\rho})^2\hat{\sigma}_\mu^2=\dfrac{u^{*\prime}\left(\mathbf{E_N}\otimes\mathbf{\bar{J}_T^{\alpha}}\right)u^*}{\left(N-1\right)}\\[2ex] \hat{\sigma}_e^2+N\hat{\sigma}_\varepsilon^2=\dfrac{u^{*\prime}\left(\mathbf{\bar{J}_N}\otimes\mathbf{E_T^{\alpha}}\right)u^*}{\left(T-1\right)}\\[2ex] \hat{\sigma}_e^2+\hat{d}_\alpha^2(1-\hat{\rho})^2\hat{\sigma}_\mu^2+N\hat{\sigma}_\varepsilon^2=u^{*\prime}\left(\mathbf{\bar{J}_N}\otimes\mathbf{\bar{J}_T^{\alpha}}\right)u^*.\end{cases} \tag{1.30}$$

1.2. Identical MA(1) Error Structure

We present a treatment of the two-way random effect model in the presence of MA(1) serial correlation. We consider simple structures of the disturbances following an MA(1) process, as in Baltagi and Li (1992b). The structure of the variance-covariance matrix of the transformed error terms and its spectral decomposition are then investigated. Next, we present the resulting GLS transformations and lastly the BQU estimators.

1.2.1. Specification of the Model

Here, we set $\nu_{it} = e_{it} - \theta e_{i,t-1}$, with $|\theta| < 1$, and $e_{it} \sim IIN\left(0, \sigma_e^2\right)$ while $\lambda_t = \varepsilon_t - \theta \varepsilon_{t-1}$, for $\varepsilon_t \sim IIN\left(0, \sigma_\varepsilon^2\right)$. The individual-specific effect is spherical, i.e. $\mu_i \sim IIN\left(0, \sigma_\mu^2\right)$. For convergence purpose and assuming stationarity, the initial values are defined as

$$\begin{cases} \nu_{i0} \sim N\left(0, \sigma_\nu^2 = \left(1+\theta^2\right)\sigma_e^2\right) \\ \lambda_0 \sim N\left(0, \sigma_\lambda^2 = \left(1+\theta^2\right)\sigma_\varepsilon^2\right). \end{cases}$$

Once again, the overall disturbance u has mean zero and variance-covariance matrix

$$\boldsymbol{\Sigma} = E\left(uu'\right) = \boldsymbol{\Sigma}_{\mathbf{v}} + \sigma_\mu^2\left[\mathbf{I_N} \otimes \left(\iota_T \iota_T'\right)\right] + \left(\iota_N \iota_N'\right) \otimes \boldsymbol{\Sigma}_{\boldsymbol{\lambda}}$$

as in equation (1.7). However, $\boldsymbol{\Sigma}_{\mathbf{v}}$ and $\boldsymbol{\Sigma}_\lambda$ are defined differently:

$$\boldsymbol{\Sigma}_{\mathbf{v}} = \sigma_e^2 \begin{bmatrix} \boldsymbol{\Gamma} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Gamma} \end{bmatrix} = \sigma_e^2\left(\mathbf{I_N} \otimes \boldsymbol{\Gamma}\right) \tag{1.31}$$

and

$$\boldsymbol{\Sigma}_\lambda = \sigma_\varepsilon^2 \begin{bmatrix} 1+\theta^2 & -\theta & 0 & 0 & \cdots & 0 \\ -\theta & 1+\theta^2 & -\theta & 0 & \cdots & 0 \\ 0 & -\theta & 1+\theta^2 & -\theta & \ddots & \vdots \\ 0 & 0 & -\theta & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & -\theta \\ 0 & 0 & \cdots & 0 & -\theta & 1+\theta^2 \end{bmatrix} = \sigma_\varepsilon^2 \boldsymbol{\Gamma} \tag{1.32}$$

with, of course,

$$\mathbf{\Gamma}=\begin{bmatrix}1+\theta^2 & -\theta & 0 & 0 & \cdots & 0\\ -\theta & 1+\theta^2 & -\theta & 0 & \cdots & 0\\ 0 & -\theta & 1+\theta^2 & -\theta & \ddots & \vdots\\ 0 & 0 & -\theta & \ddots & \ddots & 0\\ \vdots & \vdots & \ddots & \ddots & \ddots & -\theta\\ 0 & 0 & \cdots & 0 & -\theta & 1+\theta^2\end{bmatrix}=\text{Toeplitz}\left(1+\theta^2,-\theta,0,\ldots,0\right). \tag{1.33}$$

Under this MA(1) framework, the correlation coefficient is now:

$$\text{Correl}\left(u_{it},u_{js}\right)=\begin{cases}\left[\sigma_\mu^2-\theta\left(\sigma_e^2+\sigma_\varepsilon^2\right)\right]\Big/\left[\sigma_\mu^2+\left(1+\theta^2\right)\left(\sigma_e^2+\sigma_\varepsilon^2\right)\right] & \text{for } \; i=j,\left|t-s\right|=1\\ \sigma_\mu^2\Big/\left[\sigma_\mu^2+\left(1+\theta^2\right)\left(\sigma_e^2+\sigma_\varepsilon^2\right)\right] & \text{for } \; i=j,\left|t-s\right|\geq 2\\ 1 & \text{for } \; i=j,t=s\\ 0 & \text{for } \; i\neq j,\forall t,s.\end{cases} \tag{1.34}$$

Once again, the equicorrelation has been removed and the correlation coefficient actually depends on the time length $\left|t-s\right|$.

1.2.2. Variance-Covariance of the Transformed Errors and Its Spectral Decomposition

Balestra (1980) suggest the following transformation matrix $\mathbf{C}$ for an MA(1) process: $\mathbf{C}=\mathbf{D}^{-1/2}\mathbf{P}$; where

$$\mathbf{P}=\begin{bmatrix}1 & 0 & 0 & 0 & \cdots & 0\\ \theta & a_1 & 0 & 0 & \cdots & 0\\ \theta^2 & a_1\theta & a_2 & 0 & \cdots & 0\\ \theta^3 & a_1\theta^2 & a_2\theta & a_3 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ \theta^{T-1} & a_1\theta^{T-2} & a_2\theta^{T-3} & a_3\theta^{T-4} & \cdots & a_{T-1}\end{bmatrix} \tag{1.35}$$

with

$a_p = \sum_{k=0}^{p} \theta^{2k} \quad p = 0,1,\ldots,T$; and $\mathbf{D} = \mathrm{diag}\left(a_0 a_1, a_1 a_2, a_2 a_3, \cdots, a_{T-1} a_T\right)$.

Hence,

$$\mathbf{C} = \begin{bmatrix} \left(\frac{a_0}{a_1}\right)^{1/2} & 0 & 0 & 0 & \cdots & 0 \\ \theta \frac{a_0}{\left(a_1 a_2\right)^{1/2}} & \left(\frac{a_1}{a_2}\right)^{1/2} & 0 & 0 & \cdots & 0 \\ \theta^2 \frac{a_0}{\left(a_2 a_3\right)^{1/2}} & \frac{a_1 \theta}{\left(a_2 a_3\right)^{1/2}} & \left(\frac{a_2}{a_3}\right)^{1/2} & 0 & \cdots & 0 \\ \theta^3 \frac{a_0}{\left(a_3 a_4\right)^{1/2}} & \frac{a_1 \theta^2}{\left(a_3 a_4\right)^{1/2}} & \frac{a_2 \theta}{\left(a_3 a_4\right)^{1/2}} & \left(\frac{a_3}{a_4}\right)^{1/2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \theta^{T-1} \frac{a_0}{\left(a_{T-1} a_T\right)^{1/2}} & \frac{a_1 \theta^{T-2}}{\left(a_{T-1} a_T\right)^{1/2}} & \frac{a_2 \theta^{T-3}}{\left(a_{T-1} a_T\right)^{1/2}} & \frac{a_3 \theta^{T-4}}{\left(a_{T-1} a_T\right)^{1/2}} & \cdots & \left(\frac{a_{T-1}}{a_T}\right)^{1/2} \end{bmatrix}. \tag{1.36}$$

This matrix is such that $\mathbf{C\Gamma C'} = \mathbf{I_T}$. The transformed error vector is

$$u^* = \begin{bmatrix} u_1^* \\ u_2^* \\ \vdots \\ u_N^* \end{bmatrix} = \begin{bmatrix} \mathbf{C}u_1 \\ \mathbf{C}u_2 \\ \vdots \\ \mathbf{C}u_N \end{bmatrix} = \begin{bmatrix} \mathbf{C} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix} = \left(\mathbf{I_N} \otimes \mathbf{C}\right) u .$$

The variance-covariance matrix of the transformed errors is,

$$\mathbf{\Sigma}^{*}=\left(\mathbf{I_N}\otimes\mathbf{C}\right)\mathbf{\Sigma}\left(\mathbf{I_N}\otimes\mathbf{C}'\right)$$

$$\mathbf{\Sigma}^{*}=\left(\mathbf{I_N}\otimes\mathbf{C}\right)\left[\mathbf{I_N}\otimes\left(\sigma_e^2\mathbf{\Gamma}\right)\right]\left(\mathbf{I_N}\otimes\mathbf{C}'\right)+\sigma_\mu^2\left(\mathbf{I_N}\otimes\mathbf{C}\right)\left[\mathbf{I_N}\otimes\left(\iota_T\iota_T'\right)\right]\left(\mathbf{I_N}\otimes\mathbf{C}'\right)+$$
$$\left(\mathbf{I_N}\otimes\mathbf{C}\right)\left[\left(\iota_N\iota_N'\right)\otimes\left(\sigma_\varepsilon^2\right)\mathbf{\Gamma}\right]\left(\mathbf{I_N}\otimes\mathbf{C}'\right)$$

$$\mathbf{\Sigma}^{*}=\sigma_e^2\left(\mathbf{I_N}\otimes\mathbf{C\Gamma C}'\right)+\sigma_\mu^2\left[\mathbf{I_N}\otimes\left(\mathbf{C}\iota_T\right)\left(\mathbf{C}\iota_T\right)'\right]+\sigma_\varepsilon^2\left[\left(\iota_N\iota_N'\right)\otimes\mathbf{C\Gamma C}'\right]$$

$$\mathbf{\Sigma}^{*}=\sigma_e^2\left(\mathbf{I_N}\otimes\mathbf{I_T}\right)+\sigma_\mu^2\left[\mathbf{I_N}\otimes\left(\iota_T^\alpha\iota_T^{\alpha\prime}\right)\right]+\sigma_\varepsilon^2\left[\left(\iota_N\iota_N'\right)\otimes\mathbf{I_T}\right] \tag{1.37}$$

where $\iota_T^\alpha=\mathbf{C}\iota_T=\left(\alpha_1\quad\alpha_2\quad\cdots\quad\alpha_T\right)'$ with[6] $\alpha_t=\dfrac{1}{\left(a_{t-1}a_t\right)^{1/2}}\sum_{s=1}^{t}\theta^{t-s}a_{s-1}\quad t=1,\ldots,T$ or $\alpha_t=\sum_{s=1}^{t}\alpha_{ts}$ if we set $\alpha_{ts}=\dfrac{\theta^{t-s}a_{s-1}}{\left(a_{t-1}a_t\right)^{1/2}}\quad s=1,\ldots,t$ and $t=1,\ldots,T$.

Likewise in the AR(1) case, we define $\mathbf{J_T^\alpha}=\iota_T^\alpha\iota_T^{\alpha\prime}$, $\mathbf{J_N}=\iota_N\iota_N'$, $\mathbf{\bar{J}_T^\alpha}=\dfrac{1}{d_\alpha^2}\mathbf{J_T^\alpha}$, and $\mathbf{\bar{J}_N}=\dfrac{1}{N}\mathbf{J_N}$ with $d_\alpha^2=\iota_T^{\alpha'}\iota_T^\alpha=\sum_{t=1}^{T}\alpha_t=\sum_{t=1}^{T}\dfrac{1}{\left(a_{t-1}a_t\right)^{1/2}}\sum_{s=1}^{t}\theta^{t-s}a_{s-1}$. Therefore,

$$\mathbf{\Sigma}^{*}=\sigma_e^2\left(\mathbf{I_N}\otimes\mathbf{I_T}\right)+d_\alpha^2\sigma_\mu^2\left(\mathbf{I_N}\otimes\mathbf{\bar{J}_T^\alpha}\right)+N\sigma_\varepsilon^2\left(\mathbf{\bar{J}_N}\otimes\mathbf{I_T}\right). \tag{1.38}$$

---

[6] For all processes, apart from AR(1), we shall set $\iota_T^\alpha=\mathbf{C}\iota_T$, where $\mathbf{C}$ denotes the correlation correction matrix. However, for the AR(1) series, we specifically set $\iota_T^\alpha=\dfrac{1}{1-\rho}\mathbf{C}\iota_T$ following Baltagi and Li (1991) and Baltagi (2005). This slight difference in the definition has no consequence on the results since the two ι_T^α are proportional. It is actually useful in the sense that ρ appears in the coefficient, which substantially simplifies the analysis.

Again, the Wansbeek and Kapteyn (1982, 1983) approach requires the following substitutions in the expression of $\boldsymbol{\Sigma}^*$:

$\mathbf{I_N}$ is replaced by $\mathbf{E_N} + \bar{\mathbf{J}}_\mathbf{N}$ where $\mathbf{E_N} = \mathbf{I_N} - \bar{\mathbf{J}}_\mathbf{N}$;

$\mathbf{J_N}$ is replaced by $\mathbf{N}\bar{\mathbf{J}}_\mathbf{N}$;

$\mathbf{J}_\mathbf{T}^{\boldsymbol{\alpha}}$ is replaced by $d_\alpha^2 \bar{\mathbf{J}}_\mathbf{T}^{\boldsymbol{\alpha}}$;

$\mathbf{I_T}$ is replaced by $\mathbf{E}_\mathbf{T}^{\boldsymbol{\alpha}} + \bar{\mathbf{J}}_\mathbf{T}^{\boldsymbol{\alpha}}$ where $\mathbf{E}_\mathbf{T}^{\boldsymbol{\alpha}} = \mathbf{I_T} - \bar{\mathbf{J}}_\mathbf{T}^{\boldsymbol{\alpha}}$.

We then get,

$$\boldsymbol{\Sigma}^* = \sigma_e^2\left[\left(\mathbf{E_N} + \bar{\mathbf{J}}_\mathbf{N}\right)\otimes\left(\mathbf{E}_\mathbf{T}^{\boldsymbol{\alpha}} + \bar{\mathbf{J}}_\mathbf{T}^{\boldsymbol{\alpha}}\right)\right] + d_\alpha^2\sigma_\mu^2\left[\left(\mathbf{E_N} + \bar{\mathbf{J}}_\mathbf{N}\right)\otimes\bar{\mathbf{J}}_\mathbf{T}^{\boldsymbol{\alpha}}\right] + N\sigma_\varepsilon^2\left[\bar{\mathbf{J}}_\mathbf{N}\otimes\left(\mathbf{E}_\mathbf{T}^{\boldsymbol{\alpha}} + \bar{\mathbf{J}}_\mathbf{T}^{\boldsymbol{\alpha}}\right)\right]$$

$$\begin{aligned}\boldsymbol{\Sigma}^* &= \sigma_e^2\left(\mathbf{E_N}\otimes\mathbf{E}_\mathbf{T}^{\boldsymbol{\alpha}}\right) + \left(\sigma_e^2 + d_\alpha^2\sigma_\mu^2\right)\left(\mathbf{E_N}\otimes\bar{\mathbf{J}}_\mathbf{T}^{\boldsymbol{\alpha}}\right) + \left(\sigma_e^2 + N\sigma_\varepsilon^2\right)\left(\bar{\mathbf{J}}_\mathbf{N}\otimes\mathbf{E}_\mathbf{T}^{\boldsymbol{\alpha}}\right) + \\ &\left(\sigma_e^2 + d_\alpha^2\sigma_\mu^2 + N\sigma_\varepsilon^2\right)\left(\bar{\mathbf{J}}_\mathbf{N}\otimes\bar{\mathbf{J}}_\mathbf{T}^{\boldsymbol{\alpha}}\right).\end{aligned} \tag{1.39}$$

The spectral decomposition of $\boldsymbol{\Sigma}^*$ is summarized by equation (1.16):

$\boldsymbol{\Sigma}^* = \sum_{i=1}^{4}\psi_i\mathbf{Q_i}$ with

$\psi_1 = \sigma_e^2$, $\psi_2 = \sigma_e^2 + d_\alpha^2\sigma_\mu^2$, $\psi_3 = \sigma_e^2 + N\sigma_\varepsilon^2$, and $\psi_4 = \sigma_e^2 + d_\alpha^2\sigma_\mu^2 + N\sigma_\varepsilon^2$.

The $\mathbf{Q}_i$s have the same definitions as in subsection 1.1.2.

1.2.3. GLS Transformation

Unfortunately, the overall disturbance u^* is not spherical at this step, as in subsection (1.1.3). The suggestion of Fuller and Battese (1974) still works. We intend to find the typical elements of y^{**} in terms of y^*.

From the spectral decomposition given by equation (1.16) we again deduce equation (1.18)

$$\sigma_e \boldsymbol{\Sigma}^{*-1/2} y^* = \sum_{i=1}^{4} \frac{\sigma_e}{\psi_i^{1/2}} \mathbf{Q_i} y^* = \mathbf{Q}_1 y^* + \sum_{i=2}^{4} \frac{\sigma_e}{\psi_i^{1/2}} \mathbf{Q_i} y^* .$$

Likewise the AR(1) model, we consider the products $\mathbf{Q_i} y^*$, $i = 1, \ldots, 4$ one by one.

Firstly,

$$\begin{aligned}
\mathbf{Q_1} y^* &= \left(\mathbf{E_N} \otimes \mathbf{E_T^{\alpha}}\right) y^* = \left[\left(\mathbf{I_N} - \bar{\mathbf{J}}_{\mathbf{N}}\right) \otimes \left(\mathbf{I_T} - \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right)\right] y^* \\
&= \left(\mathbf{I_N} \otimes \mathbf{I_T}\right) y^* - \underbrace{\left(\mathbf{I_N} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^*}_{(A')} - \underbrace{\left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \mathbf{I_T}\right) y^*}_{(B')} + \underbrace{\left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^*}_{(C')}
\end{aligned}$$

where (*A'*), (*B'*) and (*C'*) will be determined explicitly. We have

$$(A') = \left(\mathbf{I_N} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^* = \operatorname{diag}\left[\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right] \begin{bmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_N^* \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_1^* \\ \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_2^* \\ \vdots \\ \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_N^* \end{bmatrix} \text{ where}$$

$$\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_i^* = \frac{1}{d_\alpha^2}\left[\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\\ \vdots\\ \alpha_T\end{pmatrix}\begin{pmatrix}\alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_T\end{pmatrix}\right] y_i^* = \frac{1}{d_\alpha^2}\begin{bmatrix}\alpha_1^2 & \alpha_1\alpha_2 & \alpha_1\alpha_3 & \cdots & \alpha_1\alpha_T\\ \alpha_2\alpha_1 & \alpha_2^2 & \alpha_2\alpha_3 & \cdots & \alpha_2\alpha_T\\ \alpha_3\alpha_1 & \alpha_2 & \alpha_3^2 & \cdots & \alpha_3\alpha_T\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \alpha_T\alpha_1 & \alpha_T\alpha_2 & \alpha_T\alpha_3 & \cdots & \alpha_T^2\end{bmatrix}\begin{bmatrix}y_{i1}^*\\ y_{i2}^*\\ \vdots\\ y_{iT}^*\end{bmatrix}$$

i.e.,

$$\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_i^* = \frac{1}{d_\alpha^2}\begin{pmatrix}\alpha_1\sum_{t=1}^{T}\alpha_t y_{it}^*\\ \alpha_2\sum_{t=1}^{T}\alpha_t y_{it}^*\\ \vdots\\ \alpha_T\sum_{t=1}^{T}\alpha_t y_{it}^*\end{pmatrix} = \begin{pmatrix}\alpha_1 b_i\\ \alpha_2 b_i\\ \vdots\\ \alpha_T b_i\end{pmatrix} \text{ with } b_i = \frac{1}{d_\alpha^2}\sum_{t=1}^{T}\alpha_t y_{it}^* \qquad \forall i = 1,\ldots,N.$$

Consequently, $(A') = \begin{pmatrix}\alpha_1 b_1 & \alpha_2 b_1 & \cdots & \alpha_T b_1 & \vdots & \cdots & \vdots & \alpha_1 b_N & \alpha_2 b_N & \cdots & \alpha_T b_N\end{pmatrix}'$.

$$(B') = \left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \mathbf{I}_{\mathbf{T}}\right) y^* = \frac{1}{N}\begin{bmatrix}\mathbf{I}_{\mathbf{T}} & \cdots & \mathbf{I}_{\mathbf{T}}\\ \vdots & \ddots & \vdots\\ \mathbf{I}_{\mathbf{T}} & \cdots & \mathbf{I}_{\mathbf{T}}\end{bmatrix}\begin{bmatrix}y_1^*\\ y_2^*\\ \vdots\\ y_N^*\end{bmatrix} = \frac{1}{N}\begin{bmatrix}\sum_{i=1}^{N} y_1^*\\ \sum_{i=1}^{N} y_2^*\\ \vdots\\ \sum_{i=1}^{N} y_N^*\end{bmatrix} \text{ where}$$

$$\frac{1}{N}\sum_{i=1}^{N} y_i^* = \frac{1}{N}\left[\begin{pmatrix} y_{11}^* \\ y_{12}^* \\ \vdots \\ y_{1T}^* \end{pmatrix} + \begin{pmatrix} y_{21}^* \\ y_{22}^* \\ \vdots \\ y_{2T}^* \end{pmatrix} + \cdots + \begin{pmatrix} y_{N1}^* \\ y_{N2}^* \\ \vdots \\ y_{NT}^* \end{pmatrix}\right] = \frac{1}{N}\begin{pmatrix} \sum_{i=1}^{N} y_{i1}^* \\ \sum_{i=1}^{N} y_{i2}^* \\ \vdots \\ \sum_{i=1}^{N} y_{iT}^* \end{pmatrix} = \begin{pmatrix} \overline{y}_{\bullet 1}^* \\ \overline{y}_{\bullet 2}^* \\ \vdots \\ \overline{y}_{\bullet T}^* \end{pmatrix}.$$

Hence, $\left(B'\right) = \left(\overline{y}_{\bullet 1}^* \quad \overline{y}_{\bullet 2}^* \quad \cdots \quad \overline{y}_{\bullet T}^* \quad \vdots \quad \cdots \quad \vdots \quad \overline{y}_{\bullet 1}^* \quad \overline{y}_{\bullet 2}^* \quad \cdots \quad \overline{y}_{\bullet T}^*\right)' = (B)$.

$$\left(C'\right) = \left(\overline{\mathbf{J}}_{\mathbf{N}} \otimes \overline{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^* = \frac{1}{N}\begin{bmatrix} \overline{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} & \cdots & \overline{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} \\ \vdots & \ddots & \vdots \\ \overline{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} & \cdots & \overline{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} \end{bmatrix}\begin{bmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_N^* \end{bmatrix} = \frac{1}{N}\begin{bmatrix} \sum_{i=1}^{N} \overline{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_i^* \\ \sum_{i=1}^{N} \overline{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_i^* \\ \vdots \\ \sum_{i=1}^{N} \overline{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_i^* \end{bmatrix} \text{ where}$$

$$\frac{1}{N}\sum_{i=1}^{N} \overline{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_i^* = \frac{1}{N}\left[\begin{pmatrix} \alpha_1 b_1 \\ \alpha_2 b_1 \\ \vdots \\ \alpha_T b_1 \end{pmatrix} + \begin{pmatrix} \alpha_1 b_2 \\ \alpha_2 b_2 \\ \vdots \\ \alpha_T b_2 \end{pmatrix} + \cdots + \begin{pmatrix} \alpha_1 b_N \\ \alpha_2 b_N \\ \vdots \\ \alpha_T b_N \end{pmatrix}\right] = \frac{1}{N}\begin{pmatrix} \alpha_1 \sum_{i=1}^{N} b_i \\ \alpha_2 \sum_{i=1}^{N} b_i \\ \vdots \\ \alpha_T \sum_{i=1}^{N} b_i \end{pmatrix} = \begin{pmatrix} \alpha_1 b \\ \alpha_2 b \\ \vdots \\ \alpha_T b \end{pmatrix}$$

with $b = \frac{1}{N}\sum_{i=1}^{N} b_i$.

It comes that $(C') = \begin{pmatrix} \alpha_1 b \\ \alpha_2 b \\ \vdots \\ \alpha_T b \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 b \\ \alpha_2 b \\ \vdots \\ \alpha_T b \end{pmatrix}$

From (*A'*), (*B'*), and (*C'*), one gets

$$\mathbf{Q_1} y^* = \begin{pmatrix} y_{11}^* \\ y_{12}^* \\ \vdots \\ y_{1T}^* \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ y_{N1}^* \\ y_{N2}^* \\ \vdots \\ y_{NT}^* \end{pmatrix} - \begin{pmatrix} \alpha_1 b_1 \\ \alpha_2 b_1 \\ \vdots \\ \alpha_T b_1 \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 b_N \\ \alpha_2 b_N \\ \vdots \\ \alpha_T b_N \end{pmatrix} - \begin{pmatrix} \bar{y}_{\bullet 1}^* \\ \bar{y}_{\bullet 2}^* \\ \vdots \\ \bar{y}_{\bullet T}^* \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \bar{y}_{\bullet 1}^* \\ \bar{y}_{\bullet 2}^* \\ \vdots \\ \bar{y}_{\bullet T}^* \end{pmatrix} + \begin{pmatrix} \alpha_1 b \\ \alpha_2 b \\ \vdots \\ \alpha_T b \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 b \\ \alpha_2 b \\ \vdots \\ \alpha_T b \end{pmatrix} = \begin{pmatrix} \left(y_{11}^* - \bar{y}_{\bullet 1}^*\right) - \alpha_1 \left(b_1 - b\right) \\ \left(y_{12}^* - \bar{y}_{\bullet 2}^*\right) - \alpha_2 \left(b_1 - b\right) \\ \vdots \\ \left(y_{1T}^* - \bar{y}_{\bullet T}^*\right) - \alpha_T \left(b_1 - b\right) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \left(y_{N1}^* - \bar{y}_{\bullet 1}^*\right) - \alpha_1 \left(b_N - b\right) \\ \left(y_{N2}^* - \bar{y}_{\bullet 2}^*\right) - \alpha_2 \left(b_N - b\right) \\ \vdots \\ \left(y_{NT}^* - \bar{y}_{\bullet T}^*\right) - \alpha_T \left(b_N - b\right) \end{pmatrix} \qquad (1.40)$$

Secondly,

$$\mathbf{Q}_2 y^* = \left(\mathbf{E_N} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^* = \left[\left(\mathbf{I_N} - \bar{\mathbf{J}}_{\mathbf{N}}\right) \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right] y^* = \underbrace{\left(\mathbf{I_N} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^*}_{(A')} - \underbrace{\left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^*}_{(C')}$$

$$\mathbf{Q}_2 y^* = \begin{pmatrix} \alpha_1 b_1 \\ \alpha_2 b_1 \\ \vdots \\ \alpha_T b_1 \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 b_N \\ \alpha_2 b_N \\ \vdots \\ \alpha_T b_N \end{pmatrix} - \begin{pmatrix} \alpha_1 b \\ \alpha_2 b \\ \vdots \\ \alpha_T b \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 b \\ \alpha_2 b \\ \vdots \\ \alpha_T b \end{pmatrix} = \begin{pmatrix} \alpha_1 \left(b_1 - b\right) \\ \alpha_2 \left(b_1 - b\right) \\ \vdots \\ \alpha_T \left(b_1 - b\right) \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 \left(b_N - b\right) \\ \alpha_2 \left(b_N - b\right) \\ \vdots \\ \alpha_T \left(b_N - b\right) \end{pmatrix}. \tag{1.41}$$

Thirdly,

$$\mathbf{Q}_3 y^* = \left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \mathbf{E}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^* = \left[\bar{\mathbf{J}}_{\mathbf{N}} \otimes \left(\mathbf{I_T} - \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right)\right] y^* = \underbrace{\left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \mathbf{I_T}\right) y^*}_{(B')} - \underbrace{\left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^*}_{(C')}$$

$$\mathbf{Q}_3 y^* = \begin{pmatrix} \bar{y}^*_{\bullet 1} \\ \bar{y}^*_{\bullet 2} \\ \vdots \\ \bar{y}^*_{\bullet T} \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \bar{y}^*_{\bullet 1} \\ \bar{y}^*_{\bullet 2} \\ \vdots \\ \bar{y}^*_{\bullet T} \end{pmatrix} - \begin{pmatrix} \alpha_1 b \\ \alpha_2 b \\ \vdots \\ \alpha_T b \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 b \\ \alpha_2 b \\ \vdots \\ \alpha_T b \end{pmatrix} = \begin{pmatrix} \bar{y}^*_{\bullet 1} - \alpha_1 b \\ \bar{y}^*_{\bullet 2} - \alpha_2 b \\ \vdots \\ \bar{y}^*_{\bullet T} - \alpha_T b \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \bar{y}^*_{\bullet 1} - \alpha_1 b \\ \bar{y}^*_{\bullet 2} - \alpha_2 b \\ \vdots \\ \bar{y}^*_{\bullet T} - \alpha_T b \end{pmatrix}. \tag{1.42}$$

Lastly,

$$\mathbf{Q}_4 y^* = \left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \bar{\mathbf{J}}^{\boldsymbol{\alpha}}_{\mathbf{T}} \right) y^* = \left(C' \right) = \begin{pmatrix} \alpha_1 b \\ \alpha_2 b \\ \vdots \\ \alpha_T b \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 b \\ \alpha_2 b \\ \vdots \\ \alpha_T b \end{pmatrix}. \tag{1.43}$$

Here, equation (1.18) becomes

$$\sigma_e \boldsymbol{\Sigma}^{*-1/2} y^* = \begin{pmatrix} \left(y_{11}^* - \bar{y}_{\bullet 1}^*\right) - \alpha_1\left(b_1 - b\right) \\ \left(y_{12}^* - \bar{y}_{\bullet 2}^*\right) - \alpha_2\left(b_1 - b\right) \\ \vdots \\ \left(y_{1T}^* - \bar{y}_{\bullet T}^*\right) - \alpha_T\left(b_1 - b\right) \\ \cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots \\ \left(y_{N1}^* - \bar{y}_{\bullet 1}^*\right) - \alpha_1\left(b_N - b\right) \\ \left(y_{N2}^* - \bar{y}_{\bullet 2}^*\right) - \alpha_2\left(b_N - b\right) \\ \vdots \\ \left(y_{NT}^* - \bar{y}_{\bullet T}^*\right) - \alpha_T\left(b_N - b\right) \end{pmatrix} + \frac{\sigma_e}{\psi_2^{1/2}} \begin{pmatrix} \alpha_1\left(b_1 - b\right) \\ \alpha_2\left(b_1 - b\right) \\ \vdots \\ \alpha_T\left(b_1 - b\right) \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1\left(b_N - b\right) \\ \alpha_2\left(b_N - b\right) \\ \vdots \\ \alpha_T\left(b_N - b\right) \end{pmatrix} + \frac{\sigma_e}{\psi_3^{1/2}} \begin{pmatrix} \bar{y}_{\bullet 1}^* - \alpha_1 b \\ \bar{y}_{\bullet 2}^* - \alpha_2 b \\ \vdots \\ \bar{y}_{\bullet T}^* - \alpha_T b \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \bar{y}_{\bullet 1}^* - \alpha_1 b \\ \bar{y}_{\bullet 2}^* - \alpha_2 b \\ \vdots \\ \bar{y}_{\bullet T}^* - \alpha_T b \end{pmatrix} + \frac{\sigma_e}{\psi_4^{1/2}} \begin{pmatrix} \alpha_1 b \\ \alpha_2 b \\ \vdots \\ \alpha_T b \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 b \\ \alpha_2 b \\ \vdots \\ \alpha_T b \end{pmatrix}$$

$$\sigma_e \boldsymbol{\Sigma}^{*-1/2} y^* = \begin{pmatrix} y_{11}^* - \left(1 - \frac{\sigma_e}{\psi_3^{1/2}}\right)\bar{y}_{\bullet 1}^* - \left(1 - \frac{\sigma_e}{\psi_2^{1/2}}\right)\alpha_1 b_1 + \left[\left(1 - \frac{\sigma_e}{\psi_2^{1/2}}\right) + \left(1 - \frac{\sigma_e}{\psi_3^{1/2}}\right) + \frac{\sigma_e}{\psi_4^{1/2}} - 1\right]\alpha_1 b \\ y_{12}^* - \left(1 - \frac{\sigma_e}{\psi_3^{1/2}}\right)\bar{y}_{\bullet 2}^* - \left(1 - \frac{\sigma_e}{\psi_2^{1/2}}\right)\alpha_2 b_1 + \left[\left(1 - \frac{\sigma_e}{\psi_2^{1/2}}\right) + \left(1 - \frac{\sigma_e}{\psi_3^{1/2}}\right) + \frac{\sigma_e}{\psi_4^{1/2}} - 1\right]\alpha_2 b \\ \vdots \\ y_{1T}^* - \left(1 - \frac{\sigma_e}{\psi_3^{1/2}}\right)\bar{y}_{\bullet T}^* - \left(1 - \frac{\sigma_e}{\psi_2^{1/2}}\right)\alpha_T b_1 + \left[\left(1 - \frac{\sigma_e}{\psi_2^{1/2}}\right) + \left(1 - \frac{\sigma_e}{\psi_3^{1/2}}\right) + \frac{\sigma_e}{\psi_4^{1/2}} - 1\right]\alpha_T b \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ y_{N1}^* - \left(1 - \frac{\sigma_e}{\psi_3^{1/2}}\right)\bar{y}_{\bullet 1}^* - \left(1 - \frac{\sigma_e}{\psi_2^{1/2}}\right)\alpha_1 b_N + \left[\left(1 - \frac{\sigma_e}{\psi_2^{1/2}}\right) + \left(1 - \frac{\sigma_e}{\psi_3^{1/2}}\right) + \frac{\sigma_e}{\psi_4^{1/2}} - 1\right]\alpha_1 b \\ y_{N2}^* - \left(1 - \frac{\sigma_e}{\psi_3^{1/2}}\right)\bar{y}_{\bullet 2}^* - \left(1 - \frac{\sigma_e}{\psi_2^{1/2}}\right)\alpha_2 b_N + \left[\left(1 - \frac{\sigma_e}{\psi_2^{1/2}}\right) + \left(1 - \frac{\sigma_e}{\psi_3^{1/2}}\right) + \frac{\sigma_e}{\psi_4^{1/2}} - 1\right]\alpha_2 b \\ \vdots \\ y_{NT}^* - \left(1 - \frac{\sigma_e}{\psi_3^{1/2}}\right)\bar{y}_{\bullet T}^* - \left(1 - \frac{\sigma_e}{\psi_2^{1/2}}\right)\alpha_T b_N + \left[\left(1 - \frac{\sigma_e}{\psi_2^{1/2}}\right) + \left(1 - \frac{\sigma_e}{\psi_3^{1/2}}\right) + \frac{\sigma_e}{\psi_4^{1/2}} - 1\right]\alpha_T b \end{pmatrix}$$

$$\sigma_e \boldsymbol{\Sigma}^{*-1/2} y^* = \begin{pmatrix} y_{11}^* - \theta_1 \bar{y}_{\bullet 1}^* - \theta_1 \alpha_1 b_1 + \left(\theta_1 + \theta_2 + \dfrac{\sigma_e}{\psi_4^{1/2}} - 1 \right) \alpha_1 b \\ y_{12}^* - \theta_2 \bar{y}_{\bullet 2}^* - \theta_1 \alpha_2 b_1 + \left(\theta_1 + \theta_2 + \dfrac{\sigma_e}{\psi_4^{1/2}} - 1 \right) \alpha_2 b \\ \vdots \\ y_{1T}^* - \theta_2 \bar{y}_{\bullet T}^* - \theta_2 \alpha_T b_1 + \left(\theta_1 + \theta_2 + \dfrac{\sigma_e}{\psi_4^{1/2}} - 1 \right) \alpha_T b \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ y_{N1}^* - \theta_2 \bar{y}_{\bullet 1}^* - \theta_1 \alpha_1 b_N + \left(\theta_1 + \theta_2 + \dfrac{\sigma_e}{\psi_4^{1/2}} - 1 \right) \alpha_1 b \\ y_{N2}^* - \theta_2 \bar{y}_{\bullet 2}^* - \theta_1 \alpha_2 b_N + \left(\theta_1 + \theta_2 + \dfrac{\sigma_e}{\psi_4^{1/2}} - 1 \right) \alpha_2 b \\ \vdots \\ y_{NT}^* - \theta_2 \bar{y}_{\bullet T}^* - \theta_1 \alpha_T b_N + \left(\theta_1 + \theta_2 + \dfrac{\sigma_e}{\psi_4^{1/2}} - 1 \right) \alpha_T b \end{pmatrix} = \begin{pmatrix} y_{11}^* - \theta_2 \bar{y}_{\bullet 1}^* - \theta_1 \alpha_1 b_1 + \theta_3 \alpha_1 b \\ y_{12}^* - \theta_2 \bar{y}_{\bullet 2}^* - \theta_1 \alpha_2 b_1 + \theta_3 \alpha_2 b \\ \vdots \\ y_{1T}^* - \theta_2 \bar{y}_{\bullet T}^* - \theta_1 \alpha_T b_1 + \theta_3 \alpha_T b \\ \cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots \\ y_{N1}^* - \theta_2 \bar{y}_{\bullet 1}^* - \theta_1 \alpha_1 b_N + \theta_3 \alpha_1 b \\ y_{N2}^* - \theta_2 \bar{y}_{\bullet 2}^* - \theta_1 \alpha_2 b_N + \theta_3 \alpha_2 b \\ \vdots \\ y_{NT}^* - \theta_2 \bar{y}_{\bullet T}^* - \theta_1 \alpha_T b_N + \theta_3 \alpha_T b \end{pmatrix}$$

with $\theta_1 = 1 - \dfrac{\sigma_e}{\psi_2^{1/2}}$, $\theta_2 = 1 - \dfrac{\sigma_e}{\psi_3^{1/2}}$, $\theta_3 = \theta_1 + \theta_2 + \dfrac{\sigma_e}{\psi_4^{1/2}} - 1$.

Finally, $y^{**} = \sigma_e \boldsymbol{\Sigma}^{*-1/2} y^*$ with the following typical elements:

$$y_{it}^{**} = y_{it}^* - \theta_2 \bar{y}_{\bullet t}^* - \theta_1 \alpha_t b_i + \theta_3 \alpha_t b \qquad i = 1, \ldots, N \quad t = 1, \ldots, T. \tag{1.44}$$

where

$$b_i = \frac{1}{d_\alpha^2} \sum_{t=1}^{T} \alpha_t y_{it}^* \qquad \forall i = 1, \ldots, N. \tag{1.45}$$

The other parameters have kept their definitions of subsection 1.1.3. The estimator associated to this last transformation is once again defined by equation (1.26). Here, the b_is are weighted averages of MA(1) corrected observations with weights α_t / d_α^2 changing from one observation to another.

Likewise the MA(1) serially correlated one-way error component model (see Baltagi, 2005), and the above AR(1) two-way correlated model, the MA(1) two-way model can also be estimated through two steps: (i) one uses the Balestra transformation to correct for serial correlation, and (ii) one subtracts a pseudo-average from these transformed data.

It is again possible to reduce this procedure to a one-step one, since y^{**} can be directly expressed in terms of *y*: $y^{**} = \sigma_e \mathbf{\Sigma}^{*-1/2} \left(\mathbf{I_N} \otimes \mathbf{C} \right) y$. We are then interested in expressing y_{it}^{**} as a function of y_{it}. We therefore need to find out the relationship between y_{it}^{*} and y_{it}.

In contrast to the AR(1) model, the link between y_{it}^{*} and y_{it} is not so obvious.

We have

$$y^{*} = \left(\mathbf{I_N} \otimes \mathbf{C} \right) y = \begin{pmatrix} \mathbf{C} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} \mathbf{C} y_1 \\ \mathbf{C} y_2 \\ \vdots \\ \mathbf{C} y_N \end{pmatrix}$$

with

$$y_i^* = \mathbf{C}y_i = \begin{bmatrix} \left(\frac{a_0}{a_1}\right)^{1/2} & 0 & 0 & 0 & \cdots & 0 \\ \theta\frac{a_0}{(a_1a_2)^{1/2}} & \left(\frac{a_1}{a_2}\right)^{1/2} & 0 & 0 & \cdots & 0 \\ \theta^2\frac{a_0}{(a_2a_3)^{1/2}} & \frac{a_1\theta}{(a_2a_3)^{1/2}} & \left(\frac{a_2}{a_3}\right)^{1/2} & 0 & \cdots & 0 \\ \theta^3\frac{a_0}{(a_3a_4)^{1/2}} & \frac{a_1\theta^2}{(a_3a_4)^{1/2}} & \frac{a_2\theta}{(a_3a_4)^{1/2}} & \left(\frac{a_3}{a_4}\right)^{1/2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \theta^{T-1}\frac{a_0}{(a_{T-1}a_T)^{1/2}} & \frac{a_1\theta^{T-2}}{(a_{T-1}a_T)^{1/2}} & \frac{a_2\theta^{T-3}}{(a_{T-1}a_T)^{1/2}} & \frac{a_3\theta^{T-4}}{(a_{T-1}a_T)^{1/2}} & \cdots & \left(\frac{a_{T-1}}{a_T}\right)^{1/2} \end{bmatrix} \begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ y_{i4} \\ \vdots \\ y_{iT} \end{pmatrix}$$

$$y_i^* = \begin{bmatrix} \left(\frac{a_0}{a_1}\right)^{1/2} y_{i1} \\ \frac{a_0\theta}{(a_1a_2)^{1/2}} y_{i1} + \left(\frac{a_1}{a_2}\right)^{1/2} y_{i2} \\ \frac{a_0\theta^2}{(a_2a_3)^{1/2}} y_{i1} + \frac{a_1\theta}{(a_2a_3)^{1/2}} y_{i2} + \left(\frac{a_2}{a_3}\right)^{1/2} y_{i3} \\ \frac{a_0\theta^3}{(a_3a_4)^{1/2}} y_{i1} + \frac{a_1\theta^2}{(a_3a_4)^{1/2}} y_{i2} + \frac{a_2\theta}{(a_3a_4)^{1/2}} y_{i3} + \left(\frac{a_3}{a_4}\right)^{1/2} y_{i4} \\ \vdots \\ \frac{a_0\theta^{T-1}}{(a_{T-1}a_T)^{1/2}} y_{i1} + \frac{a_1\theta^{T-2}}{(a_{T-1}a_T)^{1/2}} y_{i2} + \frac{a_2\theta^{T-3}}{(a_{T-1}a_T)^{1/2}} y_{i3} + \cdots + \frac{a_{T-2}\theta}{(a_{T-1}a_T)^{1/2}} y_{i,T-1} + \left(\frac{a_{T-1}}{a_T}\right)^{1/2} y_{iT} \end{bmatrix}$$

$$y_i^* = \begin{pmatrix} \frac{1}{(a_0 a_1)^{1/2}} \sum_{s=1}^{1} \theta^{1-s} a_{s-1} y_{is} \\ \frac{1}{(a_1 a_2)^{1/2}} \sum_{s=1}^{2} \theta^{2-s} a_{s-1} y_{is} \\ \vdots \\ \frac{1}{(a_{t-1} a_t)^{1/2}} \sum_{s=1}^{t} \theta^{t-s} a_{s-1} y_{is} \\ \vdots \\ \frac{1}{(a_{T-1} a_T)^{1/2}} \sum_{s=1}^{T} \theta^{T-s} a_{s-1} y_{is} \end{pmatrix}. \tag{1.46}$$

At the observation level, we obtain

$$y_{it}^* = \frac{1}{(a_{t-1} a_t)^{1/2}} \sum_{s=1}^{t} \theta^{t-s} a_{s-1} y_{is} = \sum_{s=1}^{t} \alpha_{ts} y_{is} \quad \forall i, t. \tag{1.47}$$

We are now able to express y_{it}^{**} in terms of y_{it}.

$$y_{it}^{**} = y_{it}^* - \theta_2 \bar{y}_{\bullet t}^* - \theta_1 \alpha_t b_i + \theta_3 \alpha_t b = y_{it}^* - \theta_2 \frac{1}{N} \sum_{1=1}^{N} y_{it}^* - \theta_1 \alpha_t b_i + \theta_3 \alpha_t b$$

$$y_{it}^{**} = \sum_{s=1}^{t} \alpha_{ts} y_{is} - \theta_2 \frac{1}{N} \sum_{i=1}^{N} \sum_{s=1}^{t} \alpha_{ts} y_{is} - \theta_1 \alpha_t b_i + \theta_3 \alpha_t b$$

$$y_{it}^{**} = \sum_{s=1}^{t} \alpha_{ts} y_{is} - \theta_2 \sum_{s=1}^{t} \alpha_{ts} \left(\frac{1}{N} \sum_{i=1}^{N} y_{is} \right) - \theta_1 \alpha_t b_i + \theta_3 \alpha_t b$$

$$y_{it}^{**} = \sum_{s=1}^{t} \alpha_{ts} y_{is} - \theta_2 \sum_{s=1}^{t} \alpha_{ts} \bar{y}_{\bullet s} - \theta_1 \sum_{s=1}^{t} \alpha_{ts} b_i + \theta_3 \sum_{s=1}^{t} \alpha_{ts} b.$$

Thus, the typical elements of $y^{**} = \sigma_e \mathbf{\Sigma}^{*-1/2} \left(\mathbf{I_N} \otimes \mathbf{C} \right) y$ are given by,

$$y_{it}^{**} = \sum_{s=1}^{t} \alpha_{ts} \left(y_{is} - \theta_2 \bar{y}_{\bullet s} - \theta_1 b_i + \theta_3 b \right). \tag{1.48}$$

The b_is are also related to the original data. In fact, we observe that:

$$b_i = \frac{1}{d_\alpha^2} \sum_{t=1}^{T} \alpha_t y_{it}^{*} = \frac{1}{d_\alpha^2} \sum_{t=1}^{T} \alpha_t \left(\sum_{s=1}^{t} \alpha_{ts} y_{is} \right)$$
$$b_i = \frac{1}{d_\alpha^2} \sum_{t=1}^{T} \sum_{s=1}^{t} \alpha_t \alpha_{ts} y_{is}.$$

Hence, $b_i = \dfrac{1}{d_\alpha^2} \sum_{t=1}^{T} \sum_{s=1}^{t} \pi_{ts} y_{is} \quad \forall i = 1, \ldots, N$, (1.49)

with

$$\pi_{ts} = \alpha_t \alpha_{ts} \quad s = 1, \ldots, t \text{ and } t = 1, \ldots, T. \tag{1.50}$$

1.2.4. BQU Estimates

The BQU estimates of the variance components arise from equation (1.29):

$\mathbf{Q_i} u^* \sim (0, \psi_i \mathbf{Q_i})$, for $i = 1, \cdots, 4.$

The BQU estimator of ψ_i is given by $\hat{\psi}_i = \dfrac{u^{*\prime} \mathbf{Q_i} u^*}{\text{trace}(\mathbf{Q_i})}$, for all *i*.

Thus, $$\begin{cases} \hat{\sigma}_e^2 = \dfrac{u^{*\prime}\left(\mathbf{E_N} \otimes \mathbf{E_T^{\alpha}}\right)u^*}{(N-1)(T-1)} \\ \hat{\sigma}_e^2 + \hat{d}_\alpha^2\hat{\sigma}_\mu^2 = \dfrac{u^{*\prime}\left(\mathbf{E_N} \otimes \overline{\mathbf{J}}_{\mathbf{T}}^{\alpha}\right)u^*}{(N-1)} \\ \hat{\sigma}_e^2 + N\hat{\sigma}_\varepsilon^2 = \dfrac{u^{*\prime}\left(\overline{\mathbf{J}}_{\mathbf{N}} \otimes \mathbf{E_T^{\alpha}}\right)u^*}{(T-1)} \\ \hat{\sigma}_e^2 + \hat{d}_\alpha^2\hat{\sigma}_\mu^2 + N\hat{\sigma}_\varepsilon^2 = u^{*\prime}\left(\overline{\mathbf{J}}_{\mathbf{N}} \otimes \overline{\mathbf{J}}_{\mathbf{T}}^{\alpha}\right)u^*. \end{cases} \tag{1.51}$$

1.3. General Model with Identical Error Structure

In this section, the general model is assessed. The autocorrelation pattern is not defined, but the time-varying disturbances are produced by the same time series process. We first consider the specification of the model; afterward we derive the spectral decomposition of the variance-covariance matrix of the composite error. Next, the resulting GLS transformations are presented, followed by the BQU estimates of the variances involved.

1.3.1. Specification of the Model

We reconsider model (1.3) $y_{it} = \beta_0 + x_{it}\beta + u_{it}$, $i = 1,\ldots,N$ and $t = 1,\ldots,T$ with the composite error $u_{it} = \mu_i + \lambda_t + \nu_{it}$, $i = 1,\ldots,N$ and $t = 1,\ldots,T$. In vector form, the two-way error term is given by equation (1.6): $u = \left(\mathbf{I_N} \otimes \iota_T\right)\mu + \left(\iota_N \otimes \mathbf{I_T}\right)\lambda + \nu$.

We are dealing with a general framework where the time-varying disturbances ν_{it} and λ_t follow the same stationary process, the only requirement being that their variance-covariance matrices should be written as

$$E\left(\lambda\lambda'\right) = \sigma_\lambda^2\mathbf{\Gamma},\ E\left(\nu_i\nu_i'\right) = \sigma_\nu^2\mathbf{\Gamma} \qquad \forall i = 1,\ldots,N. \tag{1.52}$$

In fact, their variance covariance-matrix, at the individual level, should be proportional to a real-valued symmetric positive definite matrix $\mathbf{\Gamma}$ [7]. This is likely to be matched by all classical time series processes, i.e. autoregressive as well as moving-average and the mixed ones. We still assume that $E\left(\nu_{it}\nu_{js}\right)=0$ for $i \neq j, \quad \forall t,s$ so that $E\left(\nu_i\nu_j'\right)=\mathbf{0} \quad \forall i \neq j$. Thus, u has mean zero and variance-covariance matrix

$$\mathbf{\Sigma}=E\left(uu^{'}\right)=\mathbf{\Sigma_v}+\sigma_\mu^2\left(\mathbf{I_N}\otimes\left(\iota_T\iota_T^{'}\right)\right)+\left(\iota_N\iota_N^{'}\right)\otimes\mathbf{\Sigma_\lambda}$$

as in equation (1.7). $\mathbf{\Sigma_v}$ and $\mathbf{\Sigma_\lambda}$ are such that

$$\mathbf{\Sigma_\lambda}=\sigma_\lambda^2\mathbf{\Gamma} \tag{1.53}$$

and

$$\mathbf{\Sigma_v}=E\left(\nu\nu'\right)=\sigma_\nu^2\begin{bmatrix}\mathbf{\Gamma} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{\Gamma} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Gamma}\end{bmatrix}=\sigma_\nu^2\left(\mathbf{I_N}\otimes\mathbf{\Gamma}\right). \tag{1.54}$$

Hence, the variance-covariance matrix of the composite error terms vector *u* can be written as

---

[7]If we follow Baltagi and Li (1994) and Baltagi (2005), we would have assumed that ν_{it} and λ_t follow a q-stationary process: $E\left(\nu_{it}\nu_{i,t-s}\right)=\begin{cases}\varphi_s & 0\leq|s|\leq q\\ 0 & \text{otherwise,}\end{cases}$ and $E\left(\lambda_t\lambda_{t-s}\right)=\begin{cases}\varphi_s & 0\leq|s|\leq q\\ 0 & \text{otherwise.}\end{cases}$. There is in fact no need to assume a q-dependent process in our general model. A stationary process is simply sufficient. These authors were interested in building a framework for a general MA(q) model, which is not our goal here.

$$\mathbf{\Sigma} = \sigma_\mu^2 \left[\mathbf{I_N} \otimes \left(\iota_T \iota_T' \right) \right] + \sigma_\lambda^2 \left(\iota_N \iota_N' \right) \otimes \mathbf{\Gamma} + \sigma_\nu^2 \left(\mathbf{I_N} \otimes \mathbf{\Gamma} \right). \tag{1.55}$$

1.3.2. Variance-Covariance of the Transformed Errors and Its Spectral Decomposition

Since $\mathbf{\Gamma}$ is a real symmetric positive-definite matrix, there exists a matrix $\mathbf{C}$ such that $\mathbf{C\Gamma C'} = \mathbf{I_T}$. In order to correct for serial correlation, this matrix $\mathbf{C}$ will help us transforming the composite error into a new one denoted by u^*. The transformed error vector is

$$u^* = \begin{bmatrix} u_1^* \\ u_2^* \\ \vdots \\ u_N^* \end{bmatrix} = \begin{bmatrix} \mathbf{C}u_1 \\ \mathbf{C}u_2 \\ \vdots \\ \mathbf{C}u_N \end{bmatrix} = \begin{bmatrix} \mathbf{C} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix} = \left(\mathbf{I_N} \otimes \mathbf{C} \right) u.$$

The variance-covariance matrix of the transformed errors is

$$\mathbf{\Sigma^*} = \left(\mathbf{I_N} \otimes \mathbf{C} \right) \mathbf{\Sigma} \left(\mathbf{I_N} \otimes \mathbf{C}' \right)$$

$$\begin{aligned} \mathbf{\Sigma^*} = {} & \left(\mathbf{I_N} \otimes \mathbf{C} \right) \left\{ \sigma_\mu^2 \left[\mathbf{I_N} \otimes \left(\iota_T \iota_T' \right) \right] \right\} \left(\mathbf{I_N} \otimes \mathbf{C}' \right) + \\ & \left(\mathbf{I_N} \otimes \mathbf{C} \right) \left[\sigma_\lambda^2 \left(\iota_N \iota_N' \right) \otimes \mathbf{\Gamma} \right] \left(\mathbf{I_N} \otimes \mathbf{C}' \right) + \left(\mathbf{I_N} \otimes \mathbf{C} \right) \left[\sigma_\nu^2 \left(\mathbf{I_N} \otimes \mathbf{\Gamma} \right) \right] \left(\mathbf{I_N} \otimes \mathbf{C}' \right) \end{aligned}$$

$$\mathbf{\Sigma^*} = \sigma_\mu^2 \left[\mathbf{I_N} \otimes \left(\mathbf{C} \iota_T \right) \left(\mathbf{C} \iota_T \right)' \right] + \sigma_\lambda^2 \left[\left(\iota_N \iota_N' \right) \otimes \mathbf{C\Gamma C'} \right] + \sigma_\nu^2 \left(\mathbf{I_N} \otimes \mathbf{C\Gamma C'} \right)$$

$$\mathbf{\Sigma^*} = \sigma_\mu^2 \left[\mathbf{I_N} \otimes \left(\iota_T^\alpha \iota_T^{\alpha'} \right) \right] + \sigma_\lambda^2 \left[\left(\iota_N \iota_N' \right) \otimes \mathbf{I_T} \right] + \sigma_\nu^2 \left(\mathbf{I_N} \otimes \mathbf{I_T} \right) \tag{1.56}$$

where $\iota_T^\alpha = \mathbf{C} \iota_T = \left(\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_T \right)'$ is a $T \times 1$ vector of constants depending on the serial correlation process specified.

Likewise in the previous models, we define $\mathbf{J_T^{\alpha}} = \iota_T^{\alpha}\iota_T^{\alpha'}$, $\mathbf{J_N} = \iota_N \iota_N'$, $\mathbf{\bar{J}_T^{\alpha}} = \frac{1}{d_\alpha^2}\mathbf{J_T^{\alpha}}$, and $\mathbf{\bar{J}_N} = \frac{1}{N}\mathbf{J_N}$ with $d_\alpha^2 = \iota_T^{\alpha}\iota_T^{\alpha'} = \sum_{t=1}^{T}\alpha_t^2$. Therefore, we can write

$$\mathbf{\Sigma^*} = d_\alpha^2\sigma_\mu^2\left(\mathbf{I_N}\otimes\mathbf{\bar{J}_T^{\alpha}}\right) + N\sigma_\lambda^2\left(\mathbf{\bar{J}_N}\otimes\mathbf{I_T}\right) + \sigma_\nu^2\left(\mathbf{I_N}\otimes\mathbf{I_T}\right). \tag{1.57}$$

We then use Wansbeek and Kapteyn (1982, 1983) method:

$\mathbf{I_N}$ is replaced by $\mathbf{E_N} + \mathbf{\bar{J}_N}$ where $\mathbf{E_N} = \mathbf{I_N} - \mathbf{\bar{J}_N}$ by definition;

$\mathbf{J_N}$ is replaced by $\mathbf{N\bar{J}_N}$;

$\mathbf{J_T^{\alpha}}$ is replaced by $d_\alpha^2\mathbf{\bar{J}_T^{\alpha}}$;

$\mathbf{I_T}$ is replaced by $\mathbf{E_T^{\alpha}} + \mathbf{\bar{J}_T^{\alpha}}$ where $\mathbf{E_T^{\alpha}} = \mathbf{I_T} - \mathbf{\bar{J}_T^{\alpha}}$ by definition.

We get,

$$\mathbf{\Sigma^*} = d_\alpha^2\sigma_\mu^2\left[\left(\mathbf{E_N}+\mathbf{\bar{J}_N}\right)\otimes\mathbf{\bar{J}_T^{\alpha}}\right] + N\sigma_\lambda^2\left[\mathbf{\bar{J}_N}\otimes\left(\mathbf{E_T^{\alpha}}+\mathbf{\bar{J}_T^{\alpha}}\right)\right] + \sigma_\nu^2\left[\left(\mathbf{E_N}+\mathbf{\bar{J}_N}\right)\otimes\left(\mathbf{E_T^{\alpha}}+\mathbf{\bar{J}_T^{\alpha}}\right)\right]$$

$$\begin{aligned}\mathbf{\Sigma^*} = {} & \sigma_\nu^2\left(\mathbf{E_N}\otimes\mathbf{E_T^{\alpha}}\right) + \left(\sigma_\nu^2 + d_\alpha^2\sigma_\mu^2\right)\left(\mathbf{E_N}\otimes\mathbf{\bar{J}_T^{\alpha}}\right) + \left(\sigma_\nu^2 + N\sigma_\lambda^2\right)\left(\mathbf{\bar{J}_N}\otimes\mathbf{E_T^{\alpha}}\right) + \\ & \left(\sigma_\nu^2 + d_\alpha^2\sigma_\mu^2 + N\sigma_\lambda^2\right)\left(\mathbf{\bar{J}_N}\otimes\mathbf{\bar{J}_T^{\alpha}}\right).\end{aligned} \tag{1.58}$$

Once again, the spectral decomposition of $\mathbf{\Sigma^*}$ is summarized by equation (1.16):

$\mathbf{\Sigma^*} = \sum_{i=1}^{4}\psi_i\mathbf{Q_i}$ with

$$\psi_1 = \sigma_\nu^2, \quad \psi_2 = \sigma_\nu^2 + d_\alpha^2\sigma_\mu^2, \quad \psi_3 = \sigma_\nu^2 + N\sigma_\lambda^2, \quad \psi_4 = \sigma_\nu^2 + d_\alpha^2\sigma_\mu^2 + N\sigma_\lambda^2. \tag{1.59}$$

$\mathbf{Q}_i$s have the same definitions as in subsection 1.1.2.

1.3.3. GLS Transformation

As expected, the overall disturbance vector u^* is still non spherical. As in subsections 1.1.3 and 1.2.3, another transformation matrix is used, say $\sigma_\nu \mathbf{\Sigma}^{*-1/2}$. The typical elements of $y^{**} = \sigma_\nu \mathbf{\Sigma}^{*-1/2} y^*$ can also be obtained. We will proceed as in the above subsections.

From the spectral decomposition appearing in subsection 1.3.2, it comes

$$\sigma_\nu \mathbf{\Sigma}^{*-1/2}\mathbf{y}^* = \sum_{i=1}^{4} \frac{\sigma_\nu}{\psi_i^{1/2}} \mathbf{Q}_k \mathbf{y}^* = \mathbf{Q}_1 \mathbf{y}^* + \sum_{i=2}^{4} \frac{\sigma_\nu}{\psi_i^{1/2}} \mathbf{Q}_k \mathbf{y}^* \tag{1.60}$$

which is similar to equation (1.18).

Firstly, one has

$$\begin{aligned}
\mathbf{Q_1} y^* &= \left(\mathbf{E_N} \otimes \mathbf{E_T^{\alpha}}\right) y^* = \left[\left(\mathbf{I_N} - \bar{\mathbf{J}}_\mathbf{N}\right) \otimes \left(\mathbf{I_T} - \bar{\mathbf{J}}_\mathbf{T}^{\alpha}\right)\right] y^* \\
&= \left(\mathbf{I_N} \otimes \mathbf{I_T}\right) y^* - \underbrace{\left(\mathbf{I_N} \otimes \bar{\mathbf{J}}_\mathbf{T}^{\alpha}\right) y^*}_{(A")} - \underbrace{\left(\bar{\mathbf{J}}_\mathbf{N} \otimes \mathbf{I_T}\right) y^*}_{(B")} + \underbrace{\left(\bar{\mathbf{J}}_\mathbf{N} \otimes \bar{\mathbf{J}}_\mathbf{T}^{\alpha}\right) y^*}_{(C")}
\end{aligned}$$

where (*A''*), (*B''*) and (*C''*) have to be determined explicitly.

We have

$$(A") = \left(\mathbf{I_N} \otimes \bar{\mathbf{J}}_\mathbf{T}^{\alpha}\right) y^* = \operatorname{diag}\left[\bar{\mathbf{J}}_\mathbf{T}^{\alpha}\right] \begin{bmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_N^* \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{J}}_\mathbf{T}^{\alpha} y_1^* \\ \bar{\mathbf{J}}_\mathbf{T}^{\alpha} y_2^* \\ \vdots \\ \bar{\mathbf{J}}_\mathbf{T}^{\alpha} y_N^* \end{bmatrix}$$

where $\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_i^* = \dfrac{1}{d_\alpha^2}\left[\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\\ \vdots\\ \alpha_T\end{pmatrix}\begin{pmatrix}\alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_T\end{pmatrix}\right]\begin{bmatrix}y_{i1}^*\\ y_{i2}^*\\ \vdots\\ y_{iT}^*\end{bmatrix} = \dfrac{1}{d_\alpha^2}\sum_{t=1}^{T}\alpha_t y_{it}^*\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\\ \vdots\\ \alpha_T\end{pmatrix}$

i.e.

$$\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}} y_i^* = \frac{1}{d_\alpha^2}\begin{pmatrix}\alpha_1\sum_{t=1}^{T}\alpha_t y_{it}^*\\ \alpha_2\sum_{t=1}^{T}\alpha_t y_{it}^*\\ \vdots\\ \alpha_T\sum_{t=1}^{T}\alpha_t y_{it}^*\end{pmatrix} = \begin{pmatrix}\alpha_1 h_i\\ \alpha_2 h_i\\ \vdots\\ \alpha_T h_i\end{pmatrix} \text{ with } h_i = \frac{1}{d_\alpha^2}\sum_{t=1}^{T}\alpha_t y_{it}^* \qquad \forall i = 1,\ldots,N.$$

Consequently, $(A") = \begin{pmatrix}\alpha_1 h_1 & \alpha_2 h_1 & \cdots & \alpha_T h_1 & \vdots & \cdots & \vdots & \alpha_1 h_N & \alpha_2 h_N & \cdots & \alpha_T h_N\end{pmatrix}'$.

$$(B") = \left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \mathbf{I}_{\mathbf{T}}\right) y^* = \frac{1}{N}\begin{bmatrix}\mathbf{I}_{\mathbf{T}} & \cdots & \mathbf{I}_{\mathbf{T}}\\ \vdots & \ddots & \vdots\\ \mathbf{I}_{\mathbf{T}} & \cdots & \mathbf{I}_{\mathbf{T}}\end{bmatrix}\begin{bmatrix}y_1^*\\ y_2^*\\ \vdots\\ y_N^*\end{bmatrix} = \frac{1}{N}\begin{bmatrix}\sum_{i=1}^{N} y_1^*\\ \sum_{i=1}^{N} y_2^*\\ \vdots\\ \sum_{i=1}^{N} y_N^*\end{bmatrix} \text{ where}$$

$$\frac{1}{N}\sum_{i=1}^{N} y_i^* = \frac{1}{N}\left[\begin{pmatrix}y_{11}^*\\ y_{12}^*\\ \vdots\\ y_{1T}^*\end{pmatrix} + \begin{pmatrix}y_{21}^*\\ y_{22}^*\\ \vdots\\ y_{2T}^*\end{pmatrix} + \cdots + \begin{pmatrix}y_{N1}^*\\ y_{N2}^*\\ \vdots\\ y_{NT}^*\end{pmatrix}\right] = \frac{1}{N}\begin{pmatrix}\sum_{i=1}^{N} y_{i1}^*\\ \sum_{i=1}^{N} y_{i2}^*\\ \vdots\\ \sum_{i=1}^{N} y_{iT}^*\end{pmatrix} = \begin{pmatrix}\bar{y}_{\bullet 1}^*\\ \bar{y}_{\bullet 2}^*\\ \vdots\\ \bar{y}_{\bullet T}^*\end{pmatrix}.$$

Hence, $\left(B''\right)=\left(\bar{y}^{*}_{\bullet 1} \quad \cdots \quad \bar{y}^{*}_{\bullet T} \quad \vdots \quad \cdots \quad \vdots \quad \bar{y}^{*}_{\bullet 1} \quad \cdots \quad \bar{y}^{*}_{\bullet T}\right)'=(B')=(B)$.

$$\left(C''\right)=\left(\bar{\mathbf{J}}_{\mathbf{N}}\otimes\bar{\mathbf{J}}^{\boldsymbol{\alpha}}_{\mathbf{T}}\right)y^{*}=\frac{1}{N}\begin{bmatrix}\bar{\mathbf{J}}^{\boldsymbol{\alpha}}_{\mathbf{T}} & \cdots & \bar{\mathbf{J}}^{\boldsymbol{\alpha}}_{\mathbf{T}}\\ \vdots & \ddots & \vdots\\ \bar{\mathbf{J}}^{\boldsymbol{\alpha}}_{\mathbf{T}} & \cdots & \bar{\mathbf{J}}^{\boldsymbol{\alpha}}_{\mathbf{T}}\end{bmatrix}\begin{bmatrix}y^{*}_{1}\\ y^{*}_{2}\\ \vdots\\ y^{*}_{N}\end{bmatrix}=\frac{1}{N}\begin{bmatrix}\sum_{i=1}^{N}\bar{\mathbf{J}}^{\boldsymbol{\alpha}}_{\mathbf{T}}y^{*}_{i}\\ \sum_{i=1}^{N}\bar{\mathbf{J}}^{\boldsymbol{\alpha}}_{\mathbf{T}}y^{*}_{i}\\ \vdots\\ \sum_{i=1}^{N}\bar{\mathbf{J}}^{\boldsymbol{\alpha}}_{\mathbf{T}}y^{*}_{i}\end{bmatrix} \text{ where}$$

$$\frac{1}{N}\sum_{i=1}^{N}\bar{\mathbf{J}}^{\boldsymbol{\alpha}}_{\mathbf{T}}y^{*}_{i}=\frac{1}{N}\left[\begin{pmatrix}\alpha_1 h_1\\ \alpha_2 h_1\\ \vdots\\ \alpha_T h_1\end{pmatrix}+\begin{pmatrix}\alpha_1 h_2\\ \alpha_2 h_2\\ \vdots\\ \alpha_T h_2\end{pmatrix}+\cdots+\begin{pmatrix}\alpha_1 h_N\\ \alpha_2 h_N\\ \vdots\\ \alpha_T h_N\end{pmatrix}\right]=\frac{1}{N}\begin{pmatrix}\alpha_1\sum_{i=1}^{N}h_i\\ \alpha_2\sum_{i=1}^{N}h_i\\ \vdots\\ \alpha_T\sum_{i=1}^{N}h_i\end{pmatrix}=\begin{pmatrix}\alpha_1 h\\ \alpha_2 h\\ \vdots\\ \alpha_T h\end{pmatrix}$$

with $h=\frac{1}{N}\sum_{i=1}^{N}h_i$.

$$\left(C''\right)=\begin{pmatrix}\alpha_1 h\\ \alpha_2 h\\ \vdots\\ \alpha_T h\\ \cdots\cdots\\ \vdots\\ \cdots\cdots\\ \alpha_1 h\\ \alpha_2 h\\ \vdots\\ \alpha_T h\end{pmatrix}.$$

From (*A''*), (*B''*), and (*C''*), one gets

$$\mathbf{Q_1}y^* = \begin{pmatrix} y_{11}^* \\ y_{12}^* \\ \vdots \\ y_{1T}^* \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ y_{N1}^* \\ y_{N2}^* \\ \vdots \\ y_{NT}^* \end{pmatrix} - \begin{pmatrix} \alpha_1 h_1 \\ \alpha_2 h_1 \\ \vdots \\ \alpha_T h_1 \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 h_N \\ \alpha_2 h_N \\ \vdots \\ \alpha_T h_N \end{pmatrix} - \begin{pmatrix} \bar{y}_{\bullet 1}^* \\ \bar{y}_{\bullet 2}^* \\ \vdots \\ \bar{y}_{\bullet T}^* \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \bar{y}_{\bullet 1}^* \\ \bar{y}_{\bullet 2}^* \\ \vdots \\ \bar{y}_{\bullet T}^* \end{pmatrix} + \begin{pmatrix} \alpha_1 h \\ \alpha_2 h \\ \vdots \\ \alpha_T h \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 h \\ \alpha_2 h \\ \vdots \\ \alpha_T h \end{pmatrix} = \begin{pmatrix} \left(y_{11}^* - \bar{y}_{\bullet 1}^*\right) - \alpha_1\left(h_1 - h\right) \\ \left(y_{12}^* - \bar{y}_{\bullet 2}^*\right) - \alpha_2\left(h_1 - h\right) \\ \vdots \\ \left(y_{1T}^* - \bar{y}_{\bullet T}^*\right) - \alpha_T\left(h_1 - h\right) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \left(y_{N1}^* - \bar{y}_{\bullet 1}^*\right) - \alpha_1\left(h_N - h\right) \\ \left(y_{N2}^* - \bar{y}_{\bullet 2}^*\right) - \alpha_2\left(h_N - h\right) \\ \vdots \\ \left(y_{NT}^* - \bar{y}_{\bullet T}^*\right) - \alpha_T\left(h_N - h\right) \end{pmatrix} \qquad (1.61)$$

Secondly,

$$\mathbf{Q}_2 y^* = \left(\mathbf{E_N} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^* = \left[\left(\mathbf{I_N} - \bar{\mathbf{J}}_{\mathbf{N}}\right) \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right] y^* = \underbrace{\left(\mathbf{I_N} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^*}_{(A'')} - \underbrace{\left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^*}_{(C'')}$$

$$\mathbf{Q}_2 y^* = \begin{pmatrix} \alpha_1 h_1 \\ \alpha_2 h_1 \\ \vdots \\ \alpha_T h_1 \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 h_N \\ \alpha_2 h_N \\ \vdots \\ \alpha_T h_N \end{pmatrix} - \begin{pmatrix} \alpha_1 h \\ \alpha_2 h \\ \vdots \\ \alpha_T h \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 h \\ \alpha_2 h \\ \vdots \\ \alpha_T h \end{pmatrix} = \begin{pmatrix} \alpha_1\left(h_1 - h\right) \\ \alpha_2\left(h_1 - h\right) \\ \vdots \\ \alpha_T\left(h_1 - h\right) \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1\left(h_N - h\right) \\ \alpha_2\left(h_N - h\right) \\ \vdots \\ \alpha_T\left(h_N - h\right) \end{pmatrix}. \qquad (1.62)$$

Thirdly,

$$\mathbf{Q}_3 y^* = \left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \mathbf{E}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^* = \left[\bar{\mathbf{J}}_{\mathbf{N}} \otimes \left(\mathbf{I}_{\mathbf{T}} - \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right)\right] y^* = \underbrace{\left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \mathbf{I}_{\mathbf{T}}\right) y^*}_{(B")} - \underbrace{\left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^*}_{(C")}$$

$$\mathbf{Q}_3 y^* = \begin{pmatrix} \bar{y}^*_{\bullet 1} \\ \bar{y}^*_{\bullet 2} \\ \vdots \\ \bar{y}^*_{\bullet T} \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \bar{y}^*_{\bullet 1} \\ \bar{y}^*_{\bullet 2} \\ \vdots \\ \bar{y}^*_{\bullet T} \end{pmatrix} - \begin{pmatrix} \alpha_1 h \\ \alpha_2 h \\ \vdots \\ \alpha_T h \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 h \\ \alpha_2 h \\ \vdots \\ \alpha_T h \end{pmatrix} = \begin{pmatrix} \bar{y}^*_{\bullet 1} - \alpha_1 h \\ \bar{y}^*_{\bullet 2} - \alpha_2 h \\ \vdots \\ \bar{y}^*_{\bullet T} - \alpha_T h \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \bar{y}^*_{\bullet 1} - \alpha_1 h \\ \bar{y}^*_{\bullet 2} - \alpha_2 h \\ \vdots \\ \bar{y}^*_{\bullet T} - \alpha_T h \end{pmatrix}. \tag{1.63}$$

Lastly,

$$\mathbf{Q}_4 y^* = \left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right) y^* = \left(C"\right) = \begin{pmatrix} \alpha_1 h \\ \alpha_2 h \\ \vdots \\ \alpha_T h \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 h \\ \alpha_2 h \\ \vdots \\ \alpha_T h \end{pmatrix}. \tag{1.64}$$

Equation (1.60) becomes

$$\sigma_\nu \mathbf{\Sigma}^{*-1/2} y^* = \begin{pmatrix} \left(y_{11}^* - \bar{y}_{\bullet 1}^*\right) - \alpha_1\left(h_1 - h\right) \\ \left(y_{12}^* - \bar{y}_{\bullet 2}^*\right) - \alpha_2\left(h_1 - h\right) \\ \vdots \\ \left(y_{1T}^* - \bar{y}_{\bullet T}^*\right) - \alpha_T\left(h_1 - h\right) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \left(y_{N1}^* - \bar{y}_{\bullet 1}^*\right) - \alpha_1\left(h_N - h\right) \\ \left(y_{N2}^* - \bar{y}_{\bullet 2}^*\right) - \alpha_2\left(h_N - h\right) \\ \vdots \\ \left(y_{NT}^* - \bar{y}_{\bullet T}^*\right) - \alpha_T\left(h_N - h\right) \end{pmatrix} + \frac{\sigma_\nu}{\psi_2^{1/2}} \begin{pmatrix} \alpha_1\left(h_1 - h\right) \\ \alpha_2\left(h_1 - h\right) \\ \vdots \\ \alpha_T\left(h_1 - h\right) \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1\left(h_N - h\right) \\ \alpha_2\left(h_N - h\right) \\ \vdots \\ \alpha_T\left(h_N - h\right) \end{pmatrix} + \frac{\sigma_\nu}{\psi_3^{1/2}} \begin{pmatrix} \bar{y}_{\bullet 1}^* - \alpha_1 h \\ \bar{y}_{\bullet 2}^* - \alpha_2 h \\ \vdots \\ \bar{y}_{\bullet T}^* - \alpha_T h \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \bar{y}_{\bullet 1}^* - \alpha_1 h \\ \bar{y}_{\bullet 2}^* - \alpha_2 h \\ \vdots \\ \bar{y}_{\bullet T}^* - \alpha_T h \end{pmatrix} + \frac{\sigma_\nu}{\psi_4^{1/2}} \begin{pmatrix} \alpha_1 h \\ \alpha_2 h \\ \vdots \\ \alpha_T h \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \alpha_1 h \\ \alpha_2 h \\ \vdots \\ \alpha_T h \end{pmatrix}$$

$$\sigma_\nu \mathbf{\Sigma}^{*-1/2} y^* = \begin{pmatrix} y_{11}^* - \left(1-\frac{\sigma_\nu}{\psi_3^{1/2}}\right)\bar{y}_{\bullet 1}^* - \left(1-\frac{\sigma_\nu}{\psi_2^{1/2}}\right)\alpha_1 h_1 + \left[\left(1-\frac{\sigma_\nu}{\psi_2^{1/2}}\right)+\left(1-\frac{\sigma_\nu}{\psi_3^{1/2}}\right)+\frac{\sigma_\nu}{\psi_4^{1/2}}-1\right]\alpha_1 h \\ y_{12}^* - \left(1-\frac{\sigma_\nu}{\psi_3^{1/2}}\right)\bar{y}_{\bullet 2}^* - \left(1-\frac{\sigma_\nu}{\psi_2^{1/2}}\right)\alpha_2 h_1 + \left[\left(1-\frac{\sigma_\nu}{\psi_2^{1/2}}\right)+\left(1-\frac{\sigma_\nu}{\psi_3^{1/2}}\right)+\frac{\sigma_\nu}{\psi_4^{1/2}}-1\right]\alpha_2 h \\ \vdots \\ y_{1T}^* - \left(1-\frac{\sigma_\nu}{\psi_3^{1/2}}\right)\bar{y}_{\bullet T}^* - \left(1-\frac{\sigma_\nu}{\psi_2^{1/2}}\right)\alpha_T h_1 + \left[\left(1-\frac{\sigma_\nu}{\psi_2^{1/2}}\right)+\left(1-\frac{\sigma_\nu}{\psi_3^{1/2}}\right)+\frac{\sigma_\nu}{\psi_4^{1/2}}-1\right]\alpha_T h \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_{N1}^* - \left(1-\frac{\sigma_\nu}{\psi_3^{1/2}}\right)\bar{y}_{\bullet 1}^* - \left(1-\frac{\sigma_\nu}{\psi_2^{1/2}}\right)\alpha_1 h_N + \left[\left(1-\frac{\sigma_\nu}{\psi_2^{1/2}}\right)+\left(1-\frac{\sigma_\nu}{\psi_3^{1/2}}\right)+\frac{\sigma_\nu}{\psi_4^{1/2}}-1\right]\alpha_1 h \\ y_{N2}^* - \left(1-\frac{\sigma_\nu}{\psi_3^{1/2}}\right)\bar{y}_{\bullet 2}^* - \left(1-\frac{\sigma_\nu}{\psi_2^{1/2}}\right)\alpha_2 h_N + \left[\left(1-\frac{\sigma_\nu}{\psi_2^{1/2}}\right)+\left(1-\frac{\sigma_\nu}{\psi_3^{1/2}}\right)+\frac{\sigma_\nu}{\psi_4^{1/2}}-1\right]\alpha_2 h \\ \vdots \\ y_{NT}^* - \left(1-\frac{\sigma_\nu}{\psi_3^{1/2}}\right)\bar{y}_{\bullet T}^* - \left(1-\frac{\sigma_\nu}{\psi_2^{1/2}}\right)\alpha_T h_N + \left[\left(1-\frac{\sigma_\nu}{\psi_2^{1/2}}\right)+\left(1-\frac{\sigma_\nu}{\psi_3^{1/2}}\right)+\frac{\sigma_\nu}{\psi_4^{1/2}}-1\right]\alpha_T h \end{pmatrix}$$

$$\sigma_\nu \boldsymbol{\Sigma}^{*-1/2} y^* = \begin{pmatrix} y_{11}^* - \theta_1 \bar{y}_{\bullet 1}^* - \theta_1 \alpha_1 h_1 + \left(\theta_1 + \theta_2 + \dfrac{\sigma_\nu}{\psi_4^{1/2}} - 1 \right) \alpha_1 h \\ y_{12}^* - \theta_2 \bar{y}_{\bullet 2}^* - \theta_1 \alpha_2 h_1 + \left(\theta_1 + \theta_2 + \dfrac{\sigma_\nu}{\psi_4^{1/2}} - 1 \right) \alpha_2 h \\ \vdots \\ y_{1T}^* - \theta_2 \bar{y}_{\bullet T}^* - \theta_2 \alpha_T h_1 + \left(\theta_1 + \theta_2 + \dfrac{\sigma_\nu}{\psi_4^{1/2}} - 1 \right) \alpha_T h \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_{N1}^* - \theta_2 \bar{y}_{\bullet 1}^* - \theta_1 \alpha_1 h_N + \left(\theta_1 + \theta_2 + \dfrac{\sigma_\nu}{\psi_4^{1/2}} - 1 \right) \alpha_1 h \\ y_{N2}^* - \theta_2 \bar{y}_{\bullet 2}^* - \theta_1 \alpha_2 h_N + \left(\theta_1 + \theta_2 + \dfrac{\sigma_\nu}{\psi_4^{1/2}} - 1 \right) \alpha_2 h \\ \vdots \\ y_{NT}^* - \theta_2 \bar{y}_{\bullet T}^* - \theta_1 \alpha_T h_N + \left(\theta_1 + \theta_2 + \dfrac{\sigma_\nu}{\psi_4^{1/2}} - 1 \right) \alpha_T h \end{pmatrix} = \begin{pmatrix} y_{11}^* - \theta_2 \bar{y}_{\bullet 1}^* - \theta_1 \alpha_1 h_1 + \theta_3 \alpha_1 h \\ y_{12}^* - \theta_2 \bar{y}_{\bullet 2}^* - \theta_1 \alpha_2 h_1 + \theta_3 \alpha_2 h \\ \vdots \\ y_{1T}^* - \theta_2 \bar{y}_{\bullet T}^* - \theta_1 \alpha_T h_1 + \theta_3 \alpha_T h \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ y_{N1}^* - \theta_2 \bar{y}_{\bullet 1}^* - \theta_1 \alpha_1 h_N + \theta_3 \alpha_1 h \\ y_{N2}^* - \theta_2 \bar{y}_{\bullet 2}^* - \theta_1 \alpha_2 h_N + \theta_3 \alpha_2 h \\ \vdots \\ y_{NT}^* - \theta_2 \bar{y}_{\bullet T}^* - \theta_1 \alpha_T h_N + \theta_3 \alpha_T h \end{pmatrix}.$$

Finally, The typical elements of $y^{**} = \sigma_\nu \boldsymbol{\Sigma}^{*-1/2} y^*$ are given as follows,

$$y_{it}^{**} = y_{it}^* - \theta_2 \bar{y}_{\bullet t}^* - \theta_1 \alpha_t h_i + \theta_3 \alpha_t h \qquad i = 1, \ldots, N \quad t = 1, \ldots, T \; ^{8} \tag{1.65}$$

with

---

[8] This expression is similar to equation (1.38), its counterpart in the MA(1) identical time structure model. Here, h_i is used instead of b_i, in order for us to stress the fact that unlike the later, the former cannot be expressed in terms of the original data y_{it} unless the correlation pattern is clearly specified. Another difference is the presence of σ_ν instead of σ_e.

$$\theta_1 = 1 - \frac{\sigma_\nu}{\psi_2^{1/2}},\ \theta_2 = 1 - \frac{\sigma_\nu}{\psi_3^{1/2}},\ \theta_3 = \theta_1 + \theta_2 + \frac{\sigma_\nu}{\psi_4^{1/2}} - 1,\ h_i = \frac{1}{d_\alpha^2}\sum_{t=1}^{T}\alpha_t y_{it}^* \qquad \forall i = 1,\dots,N \qquad (1.66)$$

and $h = \sum_{i=1}^{N} h_i$.

This is a general version of the Fuller and Battese (1974) transformations, extended to the case of a two-way error component model with a general but identical serial correlation structure. The estimator associated to this last transformation is still given by equation (1.26).

Likewise the previous models, the general two-way model with identical serial correlation pattern should be estimated through two steps, by: (i) applying the Balestra transformation to correct for serial correlation, and (ii) subtracting a pseudo-average from these transformed data.

It is also possible to solve it through a one-step procedure by expressing y^{**} in terms of *y*. However, the general and unspecified correlation pattern does not permit its determination. The precise matrix **C** needs to be known first.

1.3.4. BQU Estimates

The BQU estimates of the variance components are still derived from equation (1.29):

$\mathbf{Q_i} u^* \sim (0, \psi_i \mathbf{Q_i})$, for $i = 1,\dots,4.$

The BQU estimator of ψ_i is equal to $\hat{\psi}_i = \dfrac{u^{*\prime}\mathbf{Q}_i u^*}{\text{trace}(\mathbf{Q}_i)},\quad i = 1,\dots,4.$

Thus, $$\begin{cases} \hat{\sigma}_{\nu}^{2} = \dfrac{u^{*\prime}\left(\mathbf{E_N} \otimes \mathbf{E_T^{\alpha}}\right)u^{*}}{(N-1)(T-1)} \\ \hat{\sigma}_{\nu}^{2} + \hat{d}_{\alpha}^{2}\hat{\sigma}_{\mu}^{2} = \dfrac{u^{*\prime}\left(\mathbf{E_N} \otimes \overline{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right)u^{*}}{(N-1)} \\ \hat{\sigma}_{\nu}^{2} + N\hat{\sigma}_{\lambda}^{2} = \dfrac{u^{*\prime}\left(\overline{\mathbf{J}}_{\mathbf{N}} \otimes \mathbf{E_T^{\alpha}}\right)u^{*}}{(T-1)} \\ \hat{\sigma}_{\nu}^{2} + \hat{d}_{\alpha}^{2}\hat{\sigma}_{\mu}^{2} + N\hat{\sigma}_{\lambda}^{2} = u^{*\prime}\left(\overline{\mathbf{J}}_{\mathbf{N}} \otimes \overline{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right)u^{*}. \end{cases} \tag{1.67}$$

1.4. FGLS Estimation of an Identical Time Structure Model

Since the beginning of part 1, the variance-covariance matrix has been assumed to be known. In fact, one of the major difficulties in practicing econometrics is that most of the parameters of a model have to be determined. It is often the case of the error terms variance-covariance matrix, with all the involved parameters. Therefore, a FGLS approach is welcome, in order to overcome this issue of unknown relevant parameters. This section is devoted to this objective. The cases of AR(1) and MA(1) error structures are solved in the first two subsections. Afterward, a feasible treatment of the general model is proposed. Lastly, the approach of Swamy and Arora (1972) is presented as an alternative to our method.

1.4.1. AR(1) Model

From a practical point of view, we face several unknown parameters: $\rho, \sigma_e, \sigma_\varepsilon, \sigma_\mu, \sigma_e, \gamma_2, \gamma_3$ and γ_4. We first need an estimate of the AR(1) parameter ρ. Following the recommendation of Baltagi and Li (1997) in the one-way serially correlated model, an estimator of ρ based on the autocovariance function $\gamma(s) = E\left(u_{it}u_{i,t-s}\right)$ will be derived. From $u_{it} = \mu_i + \lambda_t + \nu_{it}$, $\nu_{it} = \rho\nu_{i,t-1} + e_{it}$, and $\lambda_t = \rho\lambda_{t-1} + \varepsilon_t$, it comes :

$$\gamma(s) = E\left[(\mu_i + \lambda_t + \nu_{it})(\mu_i + \lambda_{t-s} + \nu_{i,t-s})\right]$$

$$\gamma(s) = E\left(\mu_i^2\right) + E\left[\mu_i\left(\lambda_{t-s} + \nu_{i,t-s}\right)\right] + E\left[\lambda_t\left(\mu_i + \nu_{i,t-s}\right)\right] + E\left(\lambda_t\lambda_{t-s}\right) + E\left[\nu_{it}\left(\mu_i + \lambda_{t-s}\right)\right] + E\left(\nu_{it}\nu_{i,t-s}\right)$$

$$\gamma(s) = \sigma_\mu^2 + \rho^s\left(\sigma_\lambda^2 + \sigma_\nu^2\right). \tag{1.68}$$

In particular,

$$\gamma(0) = \sigma_\mu^2 + \sigma_\lambda^2 + \sigma_\nu^2,\ \gamma(1) = \sigma_\mu^2 + \rho\left(\sigma_\lambda^2 + \sigma_\nu^2\right), \text{ and } \gamma(2) = \sigma_\mu^2 + \rho^2\left(\sigma_\lambda^2 + \sigma_\nu^2\right).$$

One deduces:

$$\rho = \frac{\gamma(0)-\gamma(2)}{\gamma(0)-\gamma(1)} - 1 = \frac{\gamma(1)-\gamma(2)}{\gamma(0)-\gamma(1)}, \tag{1.69}$$

which is exactly the same relation obtained by Baltagi and Li (1997) in the one-way case. Therefore, their estimator $\hat{\rho}$ remains relevant to our current study. $\hat{\rho}$, which is consistent for large N, is given by

$$\hat{\rho} = \frac{\hat{\gamma}(1)-\hat{\gamma}(2)}{\hat{\gamma}(0)-\hat{\gamma}(1)} \tag{1.70}$$

where $\hat{\gamma}(s) = \frac{1}{N(T-s)}\sum_{i=1}^{N}\sum_{t=s+1}^{T}\hat{u}_{it}\hat{u}_{i,t-s}$ and $\hat{u}_{it}$ denotes the OLS residuals of model (1.3).

Once the AR(1) parameter is known, we deduce

$$\hat{\alpha} = \sqrt{(1+\hat{\rho})/(1-\hat{\rho})}, \tag{1.71}$$

and

$$\hat{d}_\alpha^2 = \hat{\alpha}^2 + T - 1. \tag{1.72}$$

The vector ι_T^α and matrices $\bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}$ and $\mathbf{E}_{\mathbf{T}}^{\boldsymbol{\alpha}}$ [9] are now fully known and can be used to get the BQU estimates of some variance components:

$$\begin{cases} \hat{\sigma}_e^2 = \dfrac{\hat{u}^{*\prime}\left(\mathbf{E_N} \otimes \mathbf{E}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right)\hat{u}^*}{(N-1)(T-1)} \\ \hat{\sigma}_\mu^2 = \dfrac{1}{\hat{d}_\alpha^2(1-\hat{\rho})^2}\left(\dfrac{\hat{u}^{*\prime}\left(\mathbf{E_N} \otimes \bar{\mathbf{J}}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right)\hat{u}^*}{(N-1)} - \hat{\sigma}_e^2\right) \\ \hat{\sigma}_\varepsilon^2 = \dfrac{1}{N}\left(\dfrac{\hat{u}^{*\prime}\left(\bar{\mathbf{J}}_{\mathbf{N}} \otimes \mathbf{E}_{\mathbf{T}}^{\boldsymbol{\alpha}}\right)\hat{u}^*}{(T-1)} - \hat{\sigma}_e^2\right) \end{cases} \tag{1.73}$$

where $\hat{u}^*$ are OLS residuals from the regression of y^* on $\mathbf{X}^*$. We also deduce $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_\lambda^2$ as

$$\hat{\sigma}_\nu^2 = \frac{\hat{\sigma}_e^2}{1-\hat{\rho}^2} \tag{1.74}$$

and

$$\hat{\sigma}_\lambda^2 = \frac{\hat{\sigma}_\varepsilon^2}{1-\hat{\rho}^2} \tag{1.75}$$

respectively. Hence, the estimates of the ψ_i s and θ_i s are obtained and the GLS transformations proposed in subsection 1.1.3 can now be performed. One gets the estimated coefficients and

[9] One can write $\left(\hat{\bar{\mathbf{J}}}_{\mathbf{T}}^{\alpha}\right)$ and $\left(\hat{\mathbf{E}}_{\mathbf{T}}^{\alpha}\right)$ to stress the fact that they are actually estimates of the idempotent matrices $\bar{\mathbf{J}}_{\mathbf{T}}^{\alpha}$ and $\mathbf{E}_{\mathbf{T}}^{\alpha}$ respectively, since α is substituted by its estimation $\hat{\alpha}$ in their definitions.

completes the estimation of our two-way serially correlated AR(1) error component model with identical error structure.

1.4.2. MA(1) Model

If the AR(1) FGLS approach was a straight application of the GLS transformations developed earlier, the FGLS counterpart of the MA(1) model is not that intuitive. In fact, the estimation process employed in the AR(1) case relies on the availability of an acceptable estimate of the AR(1) parameter ρ. Hence a feasible version of the correction matrix of the autocorrelation, i.e. matrix $\mathbf{C}$, was determined. Thereafter, all the parameters involved in the GLS transformations were estimated.

Unfortunately, in the MA(1) case, a direct determination results in a nonlinear estimation of the MA(1) parameter θ. Following Baltagi (2005) in the one-way serially correlated model, an estimator of θ based on the autocovariance function $\gamma(s) = E\left(u_{it}u_{i,t-s}\right), \quad s = 0,\ldots,t-1$ can be derived. From $u_{it} = \mu_i + \lambda_t + \nu_{it}$, $\nu_{it} = e_{it} - \theta e_{i,t-1}$ and $\lambda_t = \varepsilon_t - \theta\varepsilon_{t-1}$, it comes :

$$\gamma(s) = E\left[\left(\mu_i + \lambda_t + \nu_{it}\right)\left(\mu_i + \lambda_{t-s} + \nu_{i,t-s}\right)\right].$$

$$\gamma(s) = \sigma_\mu^2 + \rho_s\left(\sigma_\varepsilon^2 + \sigma_e^2\right) \tag{1.76}$$

with

$$\rho_s = \begin{cases} 1+\theta^2 & \text{if} \quad s = 0 \\ -\theta & \text{if} \quad s = 1 \\ 0 & \text{if} \quad s > 1. \end{cases} \tag{1.77}$$

We deduce that, for some $j > 1$,

$$\gamma(1) - \gamma(j) = \sigma_\mu^2 - \theta\left(\sigma_\varepsilon^2 + \sigma_e^2\right) - \sigma_\mu^2 = -\theta\left(\sigma_\varepsilon^2 + \sigma_e^2\right)$$

and

$$\gamma(0)-\gamma(j)=\sigma_\mu^2+\left(1+\theta^2\right)\left(\sigma_\varepsilon^2+\sigma_e^2\right)-\sigma_\mu^2=\left(1+\theta^2\right)\left(\sigma_\varepsilon^2+\sigma_e^2\right).$$

If we set $r=\dfrac{-\theta}{1+\theta^2}$, we then deduce

$$r=f(\theta)=\frac{\gamma(1)-\gamma(j)}{\gamma(0)-\gamma(j)} \tag{1.78}$$

which can be estimated by $\hat{r}=\dfrac{\hat{\gamma}(1)-\hat{\gamma}(j)}{\hat{\gamma}(0)-\hat{\gamma}(j)}$, knowing that $\hat{\gamma}(s)=\dfrac{1}{N(T-s)}\sum_{i=1}^{N}\sum_{t=s+1}^{T}\hat{u}_{it}\hat{u}_{i,t-s}$ where $\hat{u}_{it}$ s denote the OLS residuals of the initial model. It comes that

$$\hat{\theta}=f^{-1}(\hat{r})=\frac{-1\pm\sqrt{\Delta}}{2\hat{r}}. \tag{1.79}$$

The estimator $\hat{\theta}$ is then obtained provided $\Delta=1-4\hat{r}^2\geq 0$ and it must satisfy the invertibility condition $\left|\hat{\theta}\right|<1$. There is no warranty about the success of this nonlinear method since these conditions may not hold.

As a solution, Baltagi and Li (1994) propose a simple approach requiring only linear least squares for a one-way autocorrelated error component model with the remainder disturbances following a general MA(q) process. Their method consists in estimating the autocovariances of the composite error term which are obtained from linear least squares instead of finding the MA(q) parameters which require nonlinear least squares. We adapt their approach to our two-way error component model with identical MA(1) error structure. The theoretical model developed in section 1.2 has to be slightly modified. The variance-covariance matrix of the composite error terms vector *u* is now written as

$$\boldsymbol{\Sigma}=\sigma_\mu^2\left(\mathbf{I_N}\otimes\iota_T\iota_T'\right)+\sigma^2\left[\iota_N\iota_N'\otimes\left(\sigma_\nu^2\boldsymbol{\Gamma}\right)\right]+\mathbf{I_N}\otimes\left(\sigma_\nu^2\boldsymbol{\Gamma}\right) \tag{1.80}$$

where $\sigma^2 = \dfrac{\sigma_\lambda^2}{\sigma_\nu^2}$, and $\mathbf{\Gamma} = \text{Toeplitz}\left(1, r, 0, \ldots, 0\right)$ with $r = \dfrac{-\theta}{1+\theta^2}$. In fact, we have:

$$E\left(\nu_i \nu_i'\right) = \frac{\sigma_\nu^2}{1+\theta^2}\begin{bmatrix} 1+\theta^2 & -\theta & 0 & 0 & \cdots & 0 \\ -\theta & 1+\theta^2 & -\theta & 0 & \cdots & 0 \\ 0 & -\theta & 1+\theta^2 & -\theta & \ddots & \vdots \\ 0 & 0 & -\theta & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & -\theta \\ 0 & 0 & \cdots & 0 & -\theta & 1+\theta^2 \end{bmatrix}$$

or

$$E\left(\nu_i \nu_i'\right) = \sigma_\nu^2 \begin{bmatrix} 1 & \frac{-\theta}{1+\theta^2} & 0 & 0 & \cdots & 0 \\ \frac{-\theta}{1+\theta^2} & 1 & \frac{-\theta}{1+\theta^2} & 0 & \cdots & 0 \\ 0 & \frac{-\theta}{1+\theta^2} & 1 & \frac{-\theta}{1+\theta^2} & \ddots & \vdots \\ 0 & 0 & \frac{-\theta}{1+\theta^2} & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \frac{-\theta}{1+\theta^2} \\ 0 & 0 & \cdots & 0 & \frac{-\theta}{1+\theta^2} & 1 \end{bmatrix}.$$

The matrix $\mathbf{C} = \mathbf{D}^{-1/2}\mathbf{P}$ defined in section 1.2 is no longer appropriate. Baltagi and Li (1994) suggest a standard orthogonalizing algorithm for the general MA(q) process. This algorithm is summarized for the MA(1) case in Baltagi (2005). Let $\mathbf{C_T}$ denote the matrix correcting the correlation in ν_{it}, i.e. a matrix such that $\mathbf{C_T} E\left(\nu_i \nu_i'\right)\mathbf{C_T'} = \mathbf{C_T}\left(\sigma_\nu^2 \mathbf{\Gamma}\right)\mathbf{C_T'} = \mathbf{I_T}$. The transformation of the composite error term by $\mathbf{I_N} \otimes \mathbf{C_T}$ yields the same formulas as in the general model of section 1.3. Keeping the same notations, we have

$$\mathbf{\Sigma}^{*}=\sigma_{\mu}^{2}\left[\mathbf{I_N}\otimes(\mathbf{C_T}\iota_T)(\mathbf{C_T}\iota_T)'\right]+\sigma^{2}\left[(\iota_N\iota_N')\otimes\mathbf{C_T}\left(\sigma_{\nu}^{2}\mathbf{\Gamma}\right)\mathbf{C_T'}\right]+\left(\mathbf{I_N}\otimes\mathbf{C_T}\left(\sigma_{\nu}^{2}\mathbf{\Gamma}\right)\mathbf{C_T'}\right)$$

i.e. $\mathbf{\Sigma}^{*}=\sigma_{\mu}^{2}\left(\mathbf{I_N}\otimes\iota_T^{\alpha}\iota_T^{\alpha\prime}\right)+\sigma^{2}\left(\iota_N\iota_N'\otimes\mathbf{I_T}\right)+\left(\mathbf{I_N}\otimes\mathbf{I_T}\right)$

In fact, this equation is another version of equation (1.56) obtained under the general identical time structure model. The parameters σ_{λ}^{2} and σ_{ν}^{2} have been substituted with σ^{2} and 1 respectively. As a consequence, one gets another version of equation (1.58):

$$\begin{aligned}\mathbf{\Sigma}^{*}=&\left(\mathbf{E_N}\otimes\mathbf{E_T^{\alpha}}\right)+\left(1+d_{\alpha}^{2}\sigma_{\mu}^{2}\right)\left(\mathbf{E_N}\otimes\mathbf{\bar{J}_T^{\alpha}}\right)+\left(1+N\sigma^{2}\right)\left(\mathbf{\bar{J}_N}\otimes\mathbf{E_T^{\alpha}}\right)+\\&\left(1+d_{\alpha}^{2}\sigma_{\mu}^{2}+N\sigma^{2}\right)\left(\mathbf{\bar{J}_N}\otimes\mathbf{\bar{J}_T^{\alpha}}\right).\end{aligned}\tag{1.81}$$

The spectral decomposition of $\mathbf{\Sigma}^{*}$ is summarized by equation (1.16), i.e. $\mathbf{\Sigma}^{*}=\sum_{i=1}^{4}\psi_i\mathbf{Q_i}$ with the ψ_i s given by:

$$\psi_1=1,\quad \psi_2=1+d_{\alpha}^{2}\sigma_{\mu}^{2},\quad \psi_3=1+N\sigma^{2},\quad \psi_4=1+d_{\alpha}^{2}\sigma_{\mu}^{2}+N\sigma^{2}.$$

The final transformation applied is $y^{**}=\sigma_{\nu}\mathbf{\Sigma}^{*-1/2}y^{*}$.[10] The typical elements are defined by equations (1.65) and (1.66). In order to implement these last transformations, we firstly consider the autocovariance function of the composite error term *u*. We can write it as

[10] Here, one can easily set $\sigma_{\nu}=1$ and consider $y^{**}=\sigma_{\nu}\mathbf{\Sigma}^{*-1/2}y^{*}$, leading to $E\left(u^{**}u^{**\prime}\right)=\mathbf{I_{NT}}$. However, we have kept $y^{**}=\sigma_{\nu}\mathbf{\Sigma}^{*-1/2}y^{*}$ so that $E\left(u^{**}u^{**\prime}\right)=\sigma_{\nu}^{2}\mathbf{I_{NT}}$.

$$\gamma(s) = E\left(u_{it}u_{i,t-s}\right) = \gamma_\nu(s) + \sigma_\mu^2 + \gamma_\lambda(s)$$

where $\gamma_\nu(s)$ and $\gamma_\lambda(s)$ are the autocovariance functions of the error terms λ_t and ν_{it} respectively.

$$\begin{cases} \gamma(0) = \sigma_\mu^2 + \gamma_\nu(0) + \gamma_\lambda(0) = \sigma_\mu^2 + \sigma_\nu^2 + \sigma_\lambda^2 \\ \gamma(1) = \sigma_\mu^2 - \gamma_\nu(1) - \gamma_\lambda(1) = \sigma_\mu^2 + r\left[\gamma_\nu(0) + \gamma_\lambda(0)\right] = \sigma_\mu^2 + r\left(\sigma_\nu^2 + \sigma_\lambda^2\right) \\ \gamma(s) = \sigma_\mu^2 \quad \text{for} \quad s \geq 2. \end{cases} \tag{1.82}$$

Hence, we deduce, for some $j \geq 2$, the estimates of σ_μ^2 and *r*:

$$\hat{\sigma}_\mu^2 = \hat{\gamma}(j), \tag{1.83}$$

and

$$\hat{r} = \frac{\hat{\gamma}(1) - \hat{\gamma}(j)}{\hat{\gamma}(0) - \hat{\gamma}(j)}. \tag{1.84}$$

Note that, for $s = 0,\ldots,t-1$, $\hat{\gamma}(s) = \dfrac{1}{N(T-s)}\sum_{i=1}^{N}\sum_{t=s+1}^{T}\hat{u}_{it}\hat{u}_{i,t-s}$ where $\hat{u}_{it}$ s denote the OLS residuals of the initial model.

A remaining issue is about discriminating between $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_\lambda^2$. This can be solved by applying the within transformation to the initial model (subsection 2.3.1 presents this approach which is very helpful in the autocorrelation model with different error structures). The original model is transformed by $\mathbf{E_N} \otimes \mathbf{I_T}$. The autocovariance function of the resulting composite error term $\tilde{u}_{it} = \left(\mathbf{E_N} \otimes \mathbf{I_T}\right)u_{it} = \tilde{\mu}_i + \tilde{\nu}_{it}$ is given by

$$\tilde{\gamma}(h) = E\left(\tilde{u}_{it}\tilde{u}_{i,t-h}\right) = \frac{N-1}{N}\left[\sigma_\mu^2 + \gamma_\nu(h)\right].$$

As a consequence,

$$\hat{\sigma}_\nu^2 = \hat{\gamma}_\nu(0) = \frac{N}{N-1}\hat{\hat{\gamma}}(0) - \hat{\sigma}_\mu^2$$

where $\hat{\hat{\gamma}}(s) = \frac{1}{N(T-s)}\sum_{i=1}^{N}\sum_{t=s+1}^{T}\hat{\hat{u}}_{it}\hat{\hat{u}}_{i,t-s}$, $s = 1,\ldots,t-1$ is the empirical autocovariance function and $\hat{\hat{u}}_{it}$ s are the OLS residuals of the within equation. Hence, one deduces $\hat{\sigma}_\lambda^2$ as

$$\hat{\sigma}_\lambda^2 = \hat{\gamma}(0) - \hat{\sigma}_\nu^2 - \hat{\sigma}_\mu^2. \tag{1.85}$$

Thereafter, we get $\hat{\sigma}^2$ as

$$\hat{\sigma}^2 = \frac{\hat{\sigma}_\lambda^2}{\hat{\sigma}_\nu^2}. \tag{1.86}$$

Following Baltagi and Li (1994) and Baltagi (2005), we shall point out two important steps:

Step 1: Compute $y_{i1}^* = \frac{y_{i1}}{\sqrt{g_1}}$ and $y_{it}^* = \frac{y_{it} - \frac{\hat{r}y_{i,t-1}^*}{\sqrt{g_{t-1}}}}{\sqrt{g_t}}$ for $t = 2,\ldots,T$ where $g_1 = 1$ and $g_t = 1 - \frac{\hat{r}^2}{g_{t-1}}$ for $t = 2,\ldots,T$.

Step 2: Compute $y^{**} = \sigma_\nu \mathbf{\Sigma}^{*-1/2} y^*$ using the fact that $\iota_T^\alpha = \mathbf{C}\iota_T = (\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_T)'$ with $\alpha_1 = 1$ and $\alpha_t = \frac{1 - \frac{\hat{r}}{\sqrt{g_{t-1}}}}{\sqrt{g_t}}$.

Since the estimates $\hat{r}$ and $\hat{\sigma}_\nu^2$ have already been determined, we can go through these two steps. We also have the estimates of d_α^2 and of the ψ_i s from which the $\hat{\theta}_i$ s are deduced:

$$\hat{d}_\alpha^2 = \sum_{t=1}^{T} \hat{\alpha}_t^2\,, \quad \hat{\psi}_1 = \psi_1 = 1, \quad \hat{\psi}_2 = 1 + \hat{d}_\alpha^2 \hat{\sigma}_\mu^2, \quad \hat{\psi}_3 = 1 + N\hat{\sigma}^2, \quad \hat{\psi}_4 = 1 + \hat{d}_\alpha^2 \hat{\sigma}_\mu^2 + N\hat{\sigma}^2, \quad \hat{\theta}_1 = 1 - \frac{\hat{\sigma}_\nu}{\hat{\psi}_2^{1/2}}\,,$$

$$\hat{\theta}_2 = 1 - \frac{\hat{\sigma}_\nu}{\hat{\psi}_3^{1/2}}\,,\ \hat{\theta}_3 = \hat{\theta}_1 + \hat{\theta}_2 + \frac{\hat{\sigma}_\nu}{\hat{\psi}_4^{1/2}} - 1,\ \hat{h}_i = \frac{1}{\hat{d}_\alpha^2} \sum_{t=1}^{T} \hat{\alpha}_t y_{it}^* \qquad \forall i = 1,\ldots,N \text{ and } \hat{h} = \sum_{i=1}^{N} \hat{h}_i\,.$$

1.4.3. General Model

Along with the theoretical GLS approach presented in the section 1.3, we are interested in the feasible version of our general two-way error component model with identical error structure. The degree of sophistication in implementing such a FGLS method strongly relies on how "simple" the correction matrix $\mathbf{C}$ might be. Indeed, estimates of the serial correlation parameters are needed. Baltagi and Li (1994) argued that, for the AR(p) model these parameters are easily obtainable while the estimation method is more involved when dealing with the MA(q) process. Even in time series literature, moving-averages processes are known to be more complex. Additionally, for the error component models in which time-varying disturbances carry serial correlation, the difficulty is likely to be strengthened. The algorithm of Baltagi and Li (1994) for inverting the MA(q) process is therefore welcome. Galbraith and Zinde-Walsh (1995) propose a generalization to the ARMA(p,q) and then apply it to a one-way error component model.

Assuming that the correction matrix $\mathbf{C}$ of section 1.3 is known, all the idempotent matrices $\mathbf{Q}_\mathbf{i}$s are immediately obtainable. The BQU estimators of ψ_is are then derived as $\hat{\psi}_i = \dfrac{\hat{u}^{*\prime}\mathbf{Q}_i\hat{u}^*}{\text{trace}(\mathbf{Q}_i)}$, $i = 1,\ldots,4.$ $\hat{u}^*$ is the vector of OLS residuals from the regression of y^* on $\mathbf{X}^*$. In particular, one gets $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_\lambda^2$:

$$\begin{cases} \hat{\sigma}_{\nu}^{2} = \dfrac{\hat{u}^{*\prime}\left(\mathbf{E_N} \otimes \mathbf{E_T^{\alpha}}\right)\hat{u}^{*}}{\text{trace}\left(\mathbf{E_N} \otimes \mathbf{E_T^{\alpha}}\right)} \\ \hat{\sigma}_{\lambda}^{2} = \dfrac{1}{N}\left(\dfrac{\hat{u}^{*\prime}\left(\mathbf{\bar{J}_N} \otimes \mathbf{E_T^{\alpha}}\right)\hat{u}^{*}}{\text{trace}\left(\mathbf{\bar{J}_N} \otimes \mathbf{E_T^{\alpha}}\right)} - \hat{\sigma}_{\nu}^{2}\right). \end{cases} \tag{1.87}$$

Distinguishing between $\hat{d}_{\alpha}^{2}$ and $\hat{\sigma}_{\mu}^{2}$ can be done through the autocovariance functions of the original overall disturbances u. We have

$$\gamma(s) = E\left(u_{it}u_{i,t-s}\right) = \sigma_{\mu}^{2} + E\left(\nu_{it}\nu_{i,t-s}\right) + E\left(\lambda_{t}\lambda_{t-s}\right) \text{ or } \gamma(s) = \sigma_{\mu}^{2} + \gamma_{\nu}(s) + \gamma_{\lambda}(s)$$

for $s = 0,1,\dots,t-1$. As a consequence, one can obtain the estimate of σ_{μ}^{2}:

$$\hat{\sigma}_{\mu}^{2} = \hat{\gamma}(0) - \hat{\gamma}_{\lambda}(0) - \hat{\gamma}_{\nu}(0) = \hat{\gamma}(0) - \hat{\sigma}_{\lambda}^{2} - \hat{\sigma}_{\nu}^{2}. \tag{1.82}$$

The empirical autocovariance function of u is given by $\hat{\gamma}(s) = \dfrac{1}{N(T-s)}\sum_{i=1}^{N}\sum_{t=s+1}^{T}\hat{u}_{it}\hat{u}_{i,t-s}$ with $\hat{u}_{it}$ denoting the OLS residuals of the regression of y on $\mathbf{X}$. Hence,

$$\hat{d}_{\alpha}^{2} = \frac{1}{\hat{\sigma}_{\mu}^{2}}\left(\frac{\hat{u}^{**\prime}\left(\mathbf{E_N} \otimes \mathbf{\bar{J}_T^{\nu}}\right)\hat{u}^{**}}{\text{trace}\left(\mathbf{E_N} \otimes \mathbf{\bar{J}_T^{\nu}}\right)} - \hat{\sigma}_{\nu}^{2}\right). \tag{1.83}$$

It is worth mentioning that another estimator of d_{α}^{2} is obtainable[11].

[11] These two estimators can be statistically different. A comparison of their asymptotic and small sample properties will be welcome in further papers. An investigation by Monte Carlo experiments will also be useful regarding the issue of negative variances which

From $\mathbf{C}\iota_T = (\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_T)'$, the values of the α_t s are known.

Then one obtains $\hat{d}_\alpha^2 = \sum_{t=1}^{T} \hat{\alpha}_t^2$. Finally, from the knowledge of the $\hat{\psi}_i$ s and of $\hat{\sigma}_\nu^2$, we get the estimates of the θ_i s, h_i s and of h.

The GLS transformation $y_{it}^{**} = y_{it}^{*} - \theta_2 \bar{y}_{\bullet t}^{*} - \theta_1 \alpha_t h_i + \theta_3 \alpha_t h \quad i = 1,\ldots,N \quad t = 1,\ldots,T$ can now be performed, the GLS estimator associated to this last transformation being the one defined by equation (1.26).

1.4.4. The Swamy and Arora Method and other FGLS Approaches

The BQU estimates of the characteristic roots of the variance-covariance matrix are given, throughout part 1, by equation (1.29): $\hat{\psi}_i = \dfrac{u^{*\prime}\mathbf{Q}_i u^*}{\text{trace}(\mathbf{Q}_i)}$, $i = 1,\ldots,4$, where $u^* = (\mathbf{I_N} \otimes \mathbf{C})u$ is a vector of transformed errors.

Generally speaking, in a non-correlated two-way random effect model (see Baltagi, 2005), these BQU estimates are obtained from $\dfrac{u'\mathbf{Q}_i u}{\text{trace}(\mathbf{Q}_i)}$, $i = 1,\ldots,4$, u being the true disturbance vector of the regression of y on $\mathbf{X}$, and $\mathbf{Q_i}$ s some idempotent matrices. One can then replace vector u which is unknown, by OLS residuals, as suggested by Wallace and Hussain (1969), or by the within residuals (Amemiya 1971). However, the formers lead to asymptotically inefficient estimates of the variances with limiting distributions that are different from the ones obtained with the true disturbances. They result in biased standard errors and t-statistics (Amemiya, 1971). In

is likely to appear in the two-way error component models. Once again, we shall restate our desire to tackle these small sample properties questions in the future.

contrast, the within estimators are unbiased and asymptotically efficient and have the same asymptotic distributions as that knowing the true disturbances (see Amemiya (1971), Prucha (1984) and Baltagi (2005)). Since the within regression uses only part of the available data, Swamy and Arora (1972) suggest a FGLS estimator in three steps, each one consisting in transforming the model by an idempotent matrix involved in the spectral decomposition and then computing the estimates of some variance components (the root associated to this idempotent matrix).

In our case, we suggest replacing u^* by the OLS residuals of the transformed equation $y^* = \mathbf{X}^*\eta + u^*$, say $\hat{u}^*$. By this proposition, we follow Baltagi and Li (1991), Baltagi and Li (1994) and Baltagi (2005). This issue of the choice of the residuals is not that relevant here since we are not dealing with simple OLS residuals, but OLS residuals from a transformed model, that is with FGLS residuals.

It is well-known that true GLS estimators are BLUE. However, the variance components are usually not known and have to be estimated. Baltagi (1981) performed a Monte Carlo study on a simple regression equation with two-way error component disturbances and then studied the properties of several FGLS estimators corresponding to the methods developed by Wallace and Hussain (1969), Amemiya (1971), Swamy and Arora (1972), Rao (1972), Fuller and Battese (1974) and Nerlove (1971). They found that all the FGLS estimators considered are asymptotically efficient. It is consistent with Swamy and Arora (1972) and Prucha (1984) findings on the existence of a family of asymptotically efficient two-stage FGLS estimators of the regression coefficients, even though their variance estimation methods differ. This leaves undecided the question of which estimator is the best to use.

We shortly present the crux of Swamy and Arora (1972)'s method. As an example, we consider the AR(1) model of section 1.1. Their method suggests running three least squares regressions by transforming the data by some $\boldsymbol{Q_i}$s matrices.

The first one consists in transforming the Prais-Winsten data by $\mathbf{Q_1} = \mathbf{E_N} \otimes \mathbf{E_T^{\alpha}}$. It yields an estimate of σ_e^2:

$$\hat{\hat{\psi}}_1 = \hat{\hat{\sigma}}_e^2 = \left[y^{*\prime}\mathbf{Q_1}y^* - y^{*\prime}\mathbf{Q_1}X^*\left(X^{*\prime}\mathbf{Q_1}X^*\right)^{-1}X^{*\prime}\mathbf{Q_1}y^* \right] \Big/ \left[(N-1)(T-1)-K\right]. \tag{1.84}$$

The second regression transforms the Prais-Winsten data by $\mathbf{Q_2} = \mathbf{E_N} \otimes \mathbf{\bar{J}_T^{\alpha}}$ and suggests an estimate of $\psi_2 = \sigma_e^2 + d_\alpha^2(1-\rho)^2\sigma_\mu^2$:

$$\hat{\hat{\psi}}_2 = \left[y^{*\prime}\mathbf{Q_2}y^* - y^{*\prime}\mathbf{Q_2}X^*\left(X^{*\prime}\mathbf{Q_2}X^*\right)^{-1}X^{*\prime}\mathbf{Q_2}y^* \right] \Big/ \left[(N-1)-K\right] \tag{1.85}$$

from which one gets

$$\hat{\hat{\sigma}}_\mu^2 = \left(\hat{\hat{\psi}}_2 - \hat{\hat{\sigma}}_e^2\right) \Big/ \left[\hat{d}_\alpha^2(1-\hat{\rho})^2\right] \tag{1.86}$$

The third regression uses $\mathbf{Q_3} = \mathbf{\bar{J}_N} \otimes \mathbf{E_T^{\alpha}}$ to obtain an estimate of $\psi_3 = \sigma_e^2 + N\sigma_\varepsilon^2$:

$$\hat{\hat{\psi}}_3 = \left[y^{*\prime}\mathbf{Q_3}y^* - y^{*\prime}\mathbf{Q_3}X^*\left(X^{*\prime}\mathbf{Q_3}X^*\right)^{-1}X^{*\prime}\mathbf{Q_3}y^* \right] \Big/ \left[(T-1)-K\right] \tag{1.87}$$

from which it comes

$$\hat{\hat{\sigma}}_\varepsilon^2 = \left(\hat{\hat{\psi}}_3 - \hat{\hat{\sigma}}_e^2\right) \Big/ N. \tag{1.88}$$

The parameter $\hat{\hat{\psi}}_4$ is deduced as

$$\hat{\hat{\psi}}_4 = \hat{\hat{\sigma}}_e^2 + \hat{d}_\alpha^2(1-\hat{\rho})^2\hat{\hat{\sigma}}_\mu^2 + N\hat{\hat{\sigma}}_\varepsilon^2. \tag{1.89}$$

Unfortunately, Swamy and Arora (1972) found that their FGLS estimates in a two-way random effect model are less efficient than the OLS estimates if σ_μ^2 and σ_λ^2 are small. They are also less efficient than the within estimates when σ_μ^2 and σ_λ^2 are large. Baltagi (2005) stresses the fact that the later result is amazing since the within estimator uses only part of the available data while the Swamy and Arora estimator are based on all the available data. These are the reasons why we didn't use their estimates in our FGLS subsections. However, we find that it was worth mentioning the existence of such methods in determining the estimates of some variance components in error component models.

As a conclusion, part 1 has developed a GLS way of treating the identical correlation structure in a two-way error component model. The FGLS counterparts are deduced in several cases, according to the different correlation series considered, including the AR(1) and MA(1) processes. Henceforth, how would the treatment differ if the time-varying error terms were allowed to exhibit different correlation patterns? This is the purpose of the next part.

Part 2: A GLS ESTIMATION OF THE TWO-WAY RANDOM EFFECT MODEL WITH DIFFERENT TIME STRUCTURES

The time structure of the error terms has been assumed to be identical in the previous part. In other words, the time-varying components of the two-way disturbances λ_t and ν_{it}, were following the same processes.

However, this assumption is in fact a strong one. What can justify such a hypothesis when dealing with real-world data? Outside theoretical conceptualization purposes, actual data are not likely to display such correlation figures. As components of the same error term, one can expect λ_t and ν_{it} to follow similar processes, but not necessarily the same one. Therefore, the next level of modeling the correlation pattern consists in allowing the parameters of the time series to be different, even though the processes are of the same type. This part investigates the consequence of this hypothesis, notably for AR(1) and MA(1) processes.

Lastly, the time-varying error terms λ_t and ν_{it} are considered from a more general perspective. They may follow any time series process, even of different types. This assumption which is free of restrictions is also examined here.

Beyond the evolution of the correlation pattern, from specific features to a general framework, part 2 shows that the variance-covariance matrix of the overall disturbance takes a particular form, whatever the structure of the serial correlation might be. Section 2.1 is devoted to this objective. Hence, a GLS estimation based on the inversion of this variance-covariance matrix is suggested in section 2.2, with a special interest in the asymptotic properties of the regression estimates. Finally, section 2.3 is aimed at finding some feasible estimates for the parameters involved in the model when the variance-covariance matrix of the composite disturbances is unknown.

2.1. Autocorrelation and General Expression of the Variance-Covariance Matrix

In this part, the error terms carrying the serial correlation are allowed to follow several processes. Under this double autocorrelation error structure, the second order moments of the disturbances are more complex than those encountered in part 1 which had assumed an identical correlation framework. Fortunately, the resulting variance-covariance matrices can all be written in a specific form. It is the purpose of this section.

Subsection 2.1.1 and subsection 2.1.2 deal with the AR(1) and MA(1) time structures while subsection 2.1.3 tackles the more general case.

2.1.1. Double AR(1) Error Structure

Equation (1.3), $y_{it} = \beta_0 + x_{it}\beta + u_{it}$ $i = 1,\ldots,N$ and $t = 1,\ldots,T$ is still considered as the general regression model. In matrix form, we write $y = \mathbf{X}\eta + u$. In the overall error term given by equation (1.1), i.e. $u_{it} = \mu_i + \lambda_t + \nu_{it}$, $i = 1,\ldots,N$ and $t = 1,\ldots,T$, the underlying disturbances ν_{it} and λ_t are assumed to follow different AR(1) processes. On the one hand, $\nu_{it} = \rho_\nu \nu_{i,t-1} + e_{it}$ with $|\rho_\nu| < 1$, and $e_{it} \sim IIN\left(0, \sigma_e^2\right)$ and on the other hand $\lambda_t = \rho_\lambda \lambda_{t-1} + \varepsilon_t$ with $|\rho_\lambda| < 1$, $\rho_\nu \neq \rho_\lambda$ and $\varepsilon_t \sim IIN\left(0, \sigma_\varepsilon^2\right)$. For convergence purpose and under stationarity assumption, the initial values are defined as

$$\begin{cases} \nu_{i0} \sim N\left(0, \dfrac{\sigma_e^2}{1-\rho_\nu^2} = \sigma_\nu^2\right) \\ \lambda_0 \sim N\left(0, \dfrac{\sigma_\varepsilon^2}{1-\rho_\lambda^2} = \sigma_\lambda^2\right). \end{cases}$$

Let $\mathbf{C_1}$ and $\mathbf{C}_2$ denoting the following $(T-1)\times T$ and $(T-2)\times(T-1)$ matrices:

$$\mathbf{C_1} = \begin{pmatrix} -\rho_v & 1 & 0 & \cdots & 0 \\ 0 & -\rho_v & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -\rho_v & 1 \end{pmatrix}, \text{ and } \mathbf{C_2} = \begin{pmatrix} -\rho_\lambda & 1 & 0 & \cdots & 0 \\ 0 & -\rho_\lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -\rho_\lambda & 1 \end{pmatrix}. \tag{2.1}$$

Then, we have

$$\mathbf{C}_2\mathbf{C}_1 = \begin{pmatrix} \rho_\lambda\rho_v & -(\rho_\lambda+\rho_v) & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \rho_\lambda\rho_v & -(\rho_\lambda+\rho_v) & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \rho_\lambda\rho_v & -(\rho_\lambda+\rho_v) & 1 \end{pmatrix} \tag{2.2}$$

leading to

$$\mathbf{C}_2\mathbf{C}_1 v_i = \begin{pmatrix} e_{i3} - \rho_\lambda e_{i2} \\ \vdots \\ e_{iT} - \rho_\lambda e_{i,T-1} \end{pmatrix} = v_i^* \text{ and } \mathbf{C}_2\mathbf{C}_1 \lambda = \begin{pmatrix} \varepsilon_{i3} - \rho_v \varepsilon_{i2} \\ \vdots \\ \varepsilon_{iT} - \rho_v \varepsilon_{i,T-1} \end{pmatrix} = \lambda^*. \tag{2.3}$$

The transformed errors v_i^* and λ^* are now following two different MA(1) processes, of parameters ρ_λ and ρ_v respectively. Thus, by applying the appropriate transformation matrices, the autoregressive error structure can be changed into a moving average one. The only cost is the loss of the first two pseudo-differences, which has no serious consequence for a long time dimension.

Premultiplying the regression model (1.3) by $\mathbf{I_N} \otimes \mathbf{C_2 C_1}$ yields:

$$y^* = (\mathbf{I_N} \otimes \mathbf{C}_2\mathbf{C}_1)\, y = \mathbf{X}^*\eta + u^* \tag{2.4}$$

A typical element of this transformation is

$$y_{it}^* = y_{it} - (\rho_v + \rho_\lambda)\, y_{i,t-1} + \rho_v\rho_\lambda y_{i,t-2} \quad t = 3,\dots,T\,. \tag{2.5}$$

The transformed overall error term is given by

$$u^* = \left(\mathbf{I_N} \otimes \mathbf{C}_2\mathbf{C}_1\right)\nu + \mu \otimes \left(\mathbf{C}_2\mathbf{C}_1\iota_T\right) + i_N \otimes \left(\mathbf{C}_2\mathbf{C}_1\lambda\right)$$

$$u^* = v^* + \mu \otimes \left(\mathbf{C}_2\mathbf{C}_1\iota_T\right) + \iota_N \otimes \lambda^* . \tag{2.6}$$

where $v^* = \left(\mathbf{I_N} \otimes \mathbf{C}_2\mathbf{C}_1\right)\nu$ and $\lambda^* = \mathbf{C}_2\mathbf{C}_1\lambda$.

The subsequent variance-covariance matrix is

$$\mathbf{\Sigma^*} = \sigma_e^2\left(\mathbf{I_N} \otimes \mathbf{\Gamma_\lambda}\right) + \sigma_\mu^2\mathbf{I_N} \otimes \left(\mathbf{C}_2\mathbf{C}_1\iota_T\iota_T'\mathbf{C}_1'\mathbf{C}_2'\right) + \sigma_\varepsilon^2\left(\iota_N\iota_N'\right) \otimes \mathbf{\Gamma_v} \tag{2.7}$$

where

$$\mathbf{\Gamma_\lambda} = \begin{pmatrix} 1+\rho_\lambda^2 & -\rho_\lambda & 0 & \cdots & 0 \\ -\rho_\lambda & 1+\rho_\lambda^2 & -\rho_\lambda & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\rho_\lambda \\ 0 & \cdots & 0 & -\rho_\lambda & 1+\rho_\lambda^2 \end{pmatrix} = \mathbf{\Gamma}\left(\rho_\lambda\right)$$

and

$$\mathbf{\Gamma_v} = \begin{pmatrix} 1+\rho_v^2 & -\rho_v & 0 & \cdots & 0 \\ -\rho_v & 1+\rho_v^2 & -\rho_v & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\rho_v \\ 0 & \cdots & 0 & -\rho_v & 1+\rho_v^2 \end{pmatrix} = \mathbf{\Gamma}\left(\rho_v\right)$$ are positive-definite matrices of order *T-2*

where $\mathbf{\Gamma}(\ )$ is defined by

$$\mathbf{\Gamma}(x)=\begin{pmatrix} 1+x^2 & -x & 0 & \cdots & 0 \\ -x & 1+x^2 & -x & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -x \\ 0 & \cdots & 0 & -x & 1+x^2 \end{pmatrix}=\text{Toeplitz}\left(1+x^2,-x,0,\ldots,0\right) \tag{2.8}$$

for any real number *x*.

We are facing typical covariance matrices (aside a multiplicative factor) encountered in first order moving-average models. The exact inverse of such matrices can be found in Pesaran (1974) and Revankar (1979). At the opposite of the one suggested by Balestra (1980) and used in subsection 1.2.2, the inverse proposed by Pesaran is independent from the parameters ρ_ν and ρ_λ. This is the main reason why it is used here, in the straight line of Revankar (1979).

Let $\mathbf{P}'$ be the orthogonal matrix whose *t*-th row $C_t'(x)$ is the *t*-th eigenvector of $\mathbf{\Gamma}(x)$ corresponding to the eigenvalues $d_t(x)$. We have, assuming that $\mathbf{\Gamma}(x)$ is of order *T*:

$$C_t'(x)=\sqrt{\frac{2}{T+1}}\left[\sin\left(\frac{t\pi}{T+1}\right),\sin\left(\frac{2t\pi}{T+1}\right),\ldots,\sin\left(\frac{Tt\pi}{T+1}\right)\right] \tag{2.9}$$

$$\mathbf{P}'\mathbf{\Gamma}(x)\mathbf{P}=\mathbf{D}(x)=\text{diag}\left(d_1(x),\ldots,d_T(x)\right) \tag{2.10}$$

with

$$d_t(x)=1+x^2-2x\cos\left(\frac{t\pi}{T+1}\right). \tag{2.11}$$

In our case, $\mathbf{\Gamma}(x)$ is of order *T-2* and *x* is set to ρ_ν and ρ_λ successively. We obtain:

$$C'_t = \sqrt{\frac{2}{T-1}}\left[\sin\left(\frac{t\pi}{T-1}\right),\sin\left(\frac{2t\pi}{T-1}\right),\ldots,\sin\left(\frac{t(T-2)\pi}{T-1}\right)\right] \quad t=1,\ldots,T-2$$

and

$$\begin{cases}\mathbf{P\Gamma_{\lambda}P'} = \mathbf{D}\\ \mathbf{P\Gamma_{v}P'} = \mathbf{\Lambda}\end{cases} \tag{2.12}$$

where

$$\begin{cases}\mathbf{D} = \operatorname{diag}\left(d_1,\cdots,d_{T-2}\right) \text{ where } d_{\mathrm{t}} = 1+\rho_{\lambda}^2 - 2\rho_{\lambda}\cos\left(\dfrac{t\pi}{T-1}\right)\\ \mathbf{\Lambda} = \operatorname{diag}\left(\Lambda_1,\cdots,\Lambda_{T-2}\right) \text{ where } \Lambda_{\mathrm{t}} = 1+\rho_{v}^2 - 2\rho_{v}\cos\left(\dfrac{t\pi}{T-1}\right).\end{cases}$$

Premultiplying model (2.4) by matrix $\mathbf{I_N} \otimes \mathbf{P}$ leads to

$$y^{**} = \left(\mathbf{I_N} \otimes \mathbf{P}\right) y^{*} = \mathbf{X^{**}}\eta + u^{**}.$$

The resulting variance-covariance matrix is

$$\mathbf{\Sigma^{**}} = \left(\mathbf{I_N} \otimes \mathbf{P}\right)\mathbf{\Sigma^{*}}\left(\mathbf{I_N} \otimes \mathbf{P'}\right)$$

$$\mathbf{\Sigma^{**}} = \sigma_1^2\left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_2^2\left(\mathbf{I_N} \otimes \iota_T^{\lambda}\iota_T^{\lambda\prime}\right) + \sigma_3^2\left(\iota_N\iota'_N \otimes \mathbf{\Lambda}\right). \tag{2.13}$$

where $\iota_T^{\lambda} = \mathbf{PC}_2\mathbf{C}_1\iota_T$, $\sigma_1^2 = \sigma_e^2$, $\sigma_2^2 = \sigma_{\mu}^2$, and $\sigma_3^2 = \sigma_{\varepsilon}^2$. [12]

[12] Another expression of the variance-covariance matrix $\mathbf{\Sigma^{**}}$ is available : $\mathbf{\Sigma^{**}} = \sigma_e^2\left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_{\mu}^2\left(1-\rho_v\right)^2\left(1-\rho_{\lambda}\right)^2\left(\mathbf{I_N} \otimes \mathbf{P}\iota_{T-2}\iota'_{T-2}\mathbf{P'}\right) + \sigma_{\varepsilon}^2\left(\iota_N\iota'_N\right) \otimes \mathbf{\Lambda}$, or

2.1.2. Double MA(1) Error Structure

The time-varying disturbances ν_{it} and λ_t are assumed to follow different MA(1) processes. We set $\nu_{it} = e_{it} - \rho_\nu e_{i,t-1}$, with $|\rho_\nu| < 1$, and $e_{it} \sim IIN\left(0, \sigma_e^2\right)$ while $\lambda_t = \varepsilon_t - \rho_\lambda \varepsilon_{t-1}$ for $\varepsilon_t \sim IIN\left(0, \sigma_\varepsilon^2\right)$, $\rho_\nu \neq \rho_\lambda$ and $|\rho_\lambda| < 1$. The individual-specific effect is spherical: $\mu_i \sim IIN\left(0, \sigma_\mu^2\right)$. For convergence purpose and assuming stationarity, the initial values are defined as

$$\begin{cases} \nu_{i0} \sim N\left(0, \sigma_\nu^2 = \left(1+\rho_\nu^2\right)\sigma_e^2\right) \\ \lambda_0 \sim N\left(0, \sigma_\lambda^2 = \left(1+\rho_\lambda^2\right)\sigma_\varepsilon^2\right). \end{cases}.$$

The variance-covariance matrix of model (1.3) is given by

$$\mathbf{\Sigma} = \sigma_e^2\left(\mathbf{I_N} \otimes \mathbf{\Gamma_v}\right) + \sigma_\mu^2\left(\mathbf{I_N} \otimes \iota_T \iota_T'\right) + \sigma_\varepsilon^2\left(\iota_N \iota_N' \otimes \mathbf{\Gamma_\lambda}\right). \tag{2.14}$$

$\mathbf{\Gamma_\lambda}$ and $\mathbf{\Gamma_v}$ are defined as in the previous subsection but are now of order *T*.

Let $\mathbf{P}'$ be the Pesaran orthogonal matrix whose *t*-th row is given by:

$$\mathbf{L_t} = \sqrt{\frac{2}{T+1}}\left[\sin\left(\frac{t\pi}{T+1}\right), \sin\left(\frac{2t\pi}{T+1}\right), \ldots, \sin\left(\frac{Tt\pi}{T+1}\right)\right].$$

We have

$\mathbf{\Sigma}^{**} = \sigma_1^2\left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_2^2\left(1-\rho_\nu\right)^2\left(1-\rho_\lambda\right)^2\left(\mathbf{I_N} \otimes \iota_{T-2}^{\lambda} \iota_{T-2}^{\lambda}{}'\right) + \sigma_3^2\left(\iota_N \iota_N' \otimes \mathbf{\Lambda}\right)$ based on the fact that $\mathbf{C}_2\mathbf{C}_1\iota_T = \left(1-\rho_\nu\right)\left(1-\rho_\lambda\right)\iota_{T-2}$ and with $\iota_{T-2}^{\lambda} = P\iota_{T-2}$. However, (2.13) is still more general regarding our objective.

$$\begin{cases} \mathbf{P}\boldsymbol{\Gamma}_{\lambda}\mathbf{P}' = \boldsymbol{\Lambda} \\ \mathbf{P}\boldsymbol{\Gamma}_{\mathbf{v}}\mathbf{P}' = \mathbf{D} \end{cases} \tag{2.15}$$

where

$$\begin{cases} \boldsymbol{\Lambda} = \text{diag}\left(\Lambda_1, \cdots, \Lambda_T\right) \text{ where } \Lambda_{\text{t}} = 1 + \rho_v^2 - 2\rho_v \cos\left(\dfrac{t\pi}{T+1}\right) \\ \mathbf{D} = \text{diag}\left(d_1, \cdots, d_T\right) \text{ where } d_{\text{t}} = 1 + \rho_\lambda^2 - 2\rho_\lambda \cos\left(\dfrac{t\pi}{T+1}\right). \end{cases}$$

Premultiplying model (1.3) by $\mathbf{I_N} \otimes \mathbf{P}$ yields

$$y^* = \left(\mathbf{I_N} \otimes \mathbf{P}\right) y = \mathbf{X}^* \eta + u^* .$$

The variance-covariance matrix of u^* is

$$\boldsymbol{\Sigma}^* = \left(\mathbf{I_N} \otimes \mathbf{P}\right) \boldsymbol{\Sigma} \left(\mathbf{I_N} \otimes \mathbf{P}'\right) = \sigma_e^2 \left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_\mu^2 \left(\mathbf{I_N} \otimes \mathbf{P} \iota_T \iota_T' \mathbf{P}'\right) + \sigma_\varepsilon^2 \left(\iota_N \iota_N' \otimes \boldsymbol{\Lambda}\right)$$

$$\boldsymbol{\Sigma}^* = \sigma_1^2 \left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_2^2 \left(\mathbf{I_N} \otimes \iota_T^\lambda \iota_T^{\lambda\prime}\right) + \sigma_3^2 \left(\iota_N \iota_N' \otimes \boldsymbol{\Lambda}\right). \tag{2.16}$$

where $\iota_T^\lambda = \mathbf{P} \iota_T$, $\sigma_1^2 = \sigma_e^2$, $\sigma_2^2 = \sigma_\mu^2$, and $\sigma_3^2 = \sigma_\varepsilon^2$.

2.1.3. General Double Error Structure

The difference with the AR(1) and the MA(1) double structure consists in the fact that here we don't know the explicit form of the transformation matrix $\mathbf{P}$. However, we are still able to express the variance-covariance matrix of the error terms with a formula similar to those obtained in the above subsections.

From equations (1.1), (1.5a) and (1.5b) and under model (1.3), the covariance matrix of the composite disturbances vector *u* is given by:

$$\mathbf{\Sigma} = \sigma_v^2 \left(\mathbf{I_N} \otimes \mathbf{\Gamma_v}\right) + \sigma_\mu^2 \left(\mathbf{I_N} \otimes \iota_T \iota_T^{'}\right) + \sigma_\lambda^2 \left(\iota_N \iota_N^{'} \otimes \mathbf{\Gamma}_\lambda\right).$$

Let $\mathbf{P}_\lambda$ denote the matrix such that $\mathbf{P}_\lambda \mathbf{\Gamma}_\lambda \mathbf{P}_\lambda' = \mathbf{I_T}$. This matrix does exist for $\mathbf{\Gamma}_\lambda$ is a positive-definite matrix.

Model (1.3) can then be transformed by $\mathbf{I_N} \otimes \mathbf{P}_\lambda$, and one gets

$$y^* = \left(\mathbf{I_N} \otimes \mathbf{P}_\lambda\right) y = \mathbf{X}^* \eta + u^*.$$

It results in the following variance-covariance matrix:

$$\mathbf{\Sigma}^* = \left(\mathbf{I_N} \otimes \mathbf{P}_\lambda\right) \mathbf{\Sigma} \left(\mathbf{I_N} \otimes \mathbf{P}_\lambda'\right) = \sigma_v^2 \left(\mathbf{I_N} \otimes \mathbf{P}_\lambda \mathbf{\Gamma_v} \mathbf{P}_\lambda'\right) + \sigma_\mu^2 \left(\mathbf{I_N} \otimes \mathbf{P}_\lambda \iota_T \iota_T' \mathbf{P}_\lambda'\right) + \sigma_\lambda^2 \left(\iota_N \iota_N^{'} \otimes \mathbf{I_T}\right). \quad (2.17)$$

Let $\mathbf{P}'$ be an orthogonal matrix and $\mathbf{D} = \mathrm{diag}\left(d_1, \cdots, d_T\right)$ a diagonal matrix such that $\mathbf{P}'\left(\mathbf{P}_\lambda \mathbf{\Gamma_v} \mathbf{P}_\lambda'\right)\mathbf{P} = \mathbf{D}$. Applying a second transformation by $\mathbf{I_N} \otimes \mathbf{P}'$ yields

$$y^{**} = \left(\mathbf{I_N} \otimes \mathbf{P}'\right) y^* = \mathbf{X}^{**} \eta + u^{**}. \quad (2.18)$$

The variance-covariance matrix is written as

$$\mathbf{\Sigma}^{**} = \left(\mathbf{I_N} \otimes \mathbf{P}'\right) \mathbf{\Sigma}^* \left(\mathbf{I_N} \otimes \mathbf{P}\right)$$

$$\mathbf{\Sigma}^{**} = \sigma_v^2 \left(\mathbf{I_N} \otimes \mathbf{P}'\left(\mathbf{P}_\lambda \mathbf{\Gamma_v} \mathbf{P}_\lambda'\right)\mathbf{P}\right) + \sigma_\mu^2 \left(\mathbf{I_N} \otimes \mathbf{P}'\mathbf{P}_\lambda \iota_T \iota_T' \mathbf{P}_\lambda' \mathbf{P}\right) + \sigma_\lambda^2 \left(\iota_N \iota_N^{'} \otimes \mathbf{P}'\mathbf{P}\right)$$

$$\mathbf{\Sigma}^{**} = \sigma_1^2 \left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_2^2 \left(\mathbf{I_N} \otimes \iota_T^\lambda \iota_T^{\lambda\prime}\right) + \sigma_3^2 \left(\iota_N \iota_N^{'} \otimes \mathbf{\Lambda}\right) \quad (2.19)$$

where $\iota_T^{\lambda} = \mathbf{P}'\mathbf{P}_{\lambda}\iota_T$, $\mathbf{\Lambda} = \mathbf{P}'\mathbf{P}$, $\sigma_1^2 = \sigma_\nu^2$, $\sigma_2^2 = \sigma_\mu^2$, and $\sigma_3^2 = \sigma_\lambda^2$. Here, because of the choice of matrices $\mathbf{P}_\lambda$ and $\mathbf{P}$, we end up with $\mathbf{\Lambda} = \mathbf{I_T}$ since $\mathbf{P}$ is an orthogonal matrix. Generally speaking, $\mathbf{\Lambda}$ and $\mathbf{D}$ will be diagonal matrices [13]as in the AR(1) and MA(1) cases seen earlier.

Moreover, we implement the use of a GLS approach by introducing a scalar factor σ^2 which we define as $\sigma^2 = \sigma_\nu^2 + \sigma_\mu^2 + \sigma_\lambda^2$. As a consequence, it appears appealing to reset the σ_i^2 s as

$$\sigma_1^2 = \frac{\sigma_\nu^2}{\sigma^2},\ \sigma_2^2 = \frac{\sigma_\mu^2}{\sigma^2},\ \text{and}\ \sigma_3^2 = \frac{\sigma_\lambda^2}{\sigma^2} = 1 - \sigma_1^2 - \sigma_2^2. \tag{2.20}$$

This scaling process is similar to the one used in Revankar (1979). The error terms u^{**} are thus endowed with a new definition of their variance-covariance matrix which is now proportional to the previous one:

$$E\left(u^{**}u^{**\prime}\right) = \sigma^2\mathbf{\Sigma}^{**} \tag{2.21}$$

with $\mathbf{\Sigma}^{**}$ given by equation (2.19). The GLS approach can therefore be employed since the variance-covariance matrix of the disturbances is expressed as a factor times a positive definite matrix $\mathbf{\Sigma}^{**}$. Hereafter, we shall refer to this matrix $\mathbf{\Sigma}^{**}$ as the variance-covariance matrix of the disturbances. The reader should remember that there is an underground scalar factor.

[13] Replacing $\mathbf{\Lambda}$ by $\mathbf{I_T}$ in equation (2.19) could have been considered as the general formula. However, taking $\mathbf{\Lambda}$ as a diagonal matrix is more general and more informative than setting $\mathbf{\Lambda} = \mathbf{I_T}$, i.e. equal to the identity matrix.

Finally, we have established that, given any pattern of serial correlation, the variance-covariance matrix of the overall disturbances vector can be written as

$$\mathbf{\Sigma}^{**} = \sigma_1^2\left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_2^2\left(\mathbf{I_N} \otimes \iota_T^{\lambda}\iota_T^{\lambda'}\right) + \sigma_3^2\left(\iota_N\iota_N^{'} \otimes \mathbf{\Lambda}\right)$$

where $\mathbf{D}$ and $\mathbf{\Lambda}$ are diagonal matrices and $\iota_T^{\lambda} = \mathbf{M}\iota_T$, $\mathbf{M}$ being a matrix.

2.2. GLS Estimation and Properties of the Estimators

The previous section has derived a unique formula for the variance-covariance matrix of the overall disturbances, whatever the correlation pattern might be. This result enables us to derive a GLS estimation based on the inverse of the general variance-covariance matrix. Subsection 2.2.1 presents this inverse while subsection 2.2.2 interprets the subsequent GLS estimator. Lastly, subsection 2.2.3 provides some asymptotic properties of this GLS estimator.

2.2.1. Inverse of the Variance- Covariance Matrix of the Overall Errors

The inverse of $\mathbf{\Sigma}^{**}$ is obtained using the method of Revankar (1979, pp 156-159). Although this author considered an autocorrelation of order 1 on the error term λ_t, he encountered a similar covariance matrix in which $\mathbf{\Lambda}$ was equal to the identity matrix of order *T*.

We have established that $\mathbf{\Sigma}^{**} = \sigma_1^2\left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_2^2\left(\mathbf{I_N} \otimes \iota_T^{\lambda}\iota_T^{\lambda'}\right) + \sigma_3^2\left(\iota_N\iota_N^{'} \otimes \mathbf{\Lambda}\right)$, with $\mathbf{D} = \operatorname{diag}\left(d_1,\cdots,d_T\right)$, $\mathbf{\Lambda} = \operatorname{diag}\left(\Lambda_1,\cdots,\Lambda_T\right)$. By setting

$$\mathbf{G} = \sigma_1^2\left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_3^2\left[\left(\iota_N\iota_N^{'}\right) \otimes \mathbf{\Lambda}\right] \tag{2.22}$$

we can rewrite our variance-covariance matrix as:

$$\mathbf{\Sigma}^{**} = \mathbf{G} + \sigma_2^2\left(\mathbf{I_N} \otimes \iota_T^{\lambda}\right)\left(\mathbf{I_N} \otimes \iota_T^{\lambda'}\right) = \mathbf{G} + \sigma_2^2\mathbf{J}\mathbf{J}' \tag{2.23}$$

where $\mathbf{J} = \left(\mathbf{I_N} \otimes \iota_T^{\lambda}\right)$.

By the means of an updating formula (see Greene, 2008), we deduce the formula of the inverse of $\boldsymbol{\Sigma}^{**}$:

$$\left(\boldsymbol{\Sigma}^{**}\right)^{-1} = \mathbf{G^{-1}} - \mathbf{G^{-1}J}\left(\mathbf{J'G^{-1}J} + \frac{1}{\sigma_2^2}\mathbf{I_N}\right)^{-1}\mathbf{J'G^{-1}}. \tag{2.24}$$

We need to obtain $\mathbf{G^{-1}}$ and the inverse of the bracketed expression. On the one hand,

$$\mathbf{G} = \sigma_1^2\left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_3^2\left(\left(\iota_N \iota_N^{'}\right) \otimes \boldsymbol{\Lambda}\right) = \left(\mathbf{I_N} \otimes \mathbf{D^{1/2}}\right)\left[\sigma_1^2\mathbf{I_{NT}} + \sigma_3^2\left(\iota_N \iota_N^{'}\right) \otimes \boldsymbol{\Lambda}\mathbf{D^{-1}}\right]\left(\mathbf{I_N} \otimes \mathbf{D^{1/2}}\right). \tag{2.25}$$

Let $\mathbf{H}$ denote the matrix $\sigma_1^2\mathbf{I_{NT}} + \sigma_3^2\left(\iota_N \iota_N^{'}\right) \otimes \boldsymbol{\Lambda}\mathbf{D^{-1}}$. At this step, the inverse of $\mathbf{H}$ is required.

Let $\mathbf{C}' = \left(i_N / \sqrt{N}, \mathbf{C}_a\right)'$ be a $N \times N$ orthogonal matrix. Then,

$$\left(\mathbf{C}' \otimes \mathbf{I_T}\right)\mathbf{H}\left(\mathbf{C} \otimes \mathbf{I_T}\right) = \left(\mathbf{C}' \otimes \mathbf{I_T}\right)\left(\sigma_1^2\mathbf{I_{NT}} + \sigma_3^2\left(\iota_N \iota_N^{'}\right) \otimes \boldsymbol{\Lambda}\mathbf{D^{-1}}\right)\left(\mathbf{C} \otimes \mathbf{I_T}\right)$$

$$\left(\mathbf{C}' \otimes \mathbf{I_T}\right)\mathbf{H}\left(\mathbf{C} \otimes \mathbf{I_T}\right) = \sigma_1^2 I_{NT} + \sigma_3^2\left(\mathbf{C}'\iota_N \iota_N^{'}\mathbf{C}\right) \otimes \boldsymbol{\Lambda}\mathbf{D^{-1}} = \sigma_1^2\left(\mathbf{I_N} \otimes \mathbf{I_T}\right) + \sigma_3^2\mathbf{B}$$

$$\left(\mathbf{C}' \otimes \mathbf{I_T}\right)\mathbf{H}\left(\mathbf{C} \otimes \mathbf{I_T}\right) = \mathrm{diag}\left(\sigma_1^2 + \frac{N\Lambda_1\sigma_3^2}{d_1}, \cdots, \sigma_1^2 + \frac{N\Lambda_T\sigma_3^2}{d_T}, \sigma_1^2, \cdots, \sigma_1^2\right)' \tag{2.26}$$

with

$$\mathbf{B} = \begin{pmatrix} \mathbf{N}\boldsymbol{\Lambda}\mathbf{D^{-1}} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}. \tag{2.27}$$

It is worth mentioning that $\mathbf{C}'_a i_N = \mathbf{C}'_a \frac{1}{\sqrt{N}} \iota_N = 0$ for $\mathbf{C_a}$ and $\frac{1}{\sqrt{N}} \iota_N$ are different columns of the same diagonal matrix. It is therefore obvious that $\mathbf{H}$ has been diagonalized through equation (2.26).

As a consequence, the inverse of $\mathbf{H}$ is given by

$$\mathbf{H^{-1}} = \left(\mathbf{C} \otimes \mathbf{I_T}\right) \operatorname{diag}\left(\frac{d_1}{d_1\sigma_1^2 + N\Lambda_1\sigma_3^2}, \cdots, \frac{d_T}{d_T\sigma_1^2 + N\Lambda_T\sigma_3^2}, \frac{1}{\sigma_1^2}, \cdots, \frac{1}{\sigma_1^2} \right) \left(\mathbf{C}' \otimes \mathbf{I_T}\right)$$

$$\mathbf{H^{-1}} = \frac{1}{\sigma_1^2}\left(\mathbf{C}_a\mathbf{C}'_a\right) \otimes \mathbf{I_T} + \left(\frac{1}{\sqrt{N}} \iota_N \otimes \mathbf{I_T} \right) \mathbf{A} \left(\frac{1}{\sqrt{N}} \iota'_N \otimes \mathbf{I_T} \right) \tag{2.28}$$

where

$$\mathbf{A} = \operatorname{diag}\left(\frac{d_1}{d_1\sigma_1^2 + N\Lambda_1\sigma_3^2}, \cdots, \frac{d_T}{d_T\sigma_1^2 + N\Lambda_T\sigma_3^2} \right). \tag{2.29}$$

Since

$$\mathbf{C}'_a \iota_N = 0 \tag{2.30}$$

and

$$\mathbf{C}'_a\mathbf{C}_a = \mathbf{I_{N-1}}, \tag{2.31}$$

we have

$$\mathbf{C}_a\mathbf{C}'_a = \mathbf{I_N} - \frac{1}{N}\iota_N\iota'_N = \mathbf{E_N}. \tag{2.32}$$

Therefore

$$\mathbf{H}^{-1} = \frac{1}{\sigma_1^2}\mathbf{E_N} \otimes \mathbf{I_T} + \left(\frac{1}{\sqrt{N}}\iota_N \otimes \mathbf{I_T}\right)\mathbf{A}\left(\frac{1}{\sqrt{N}}\iota_N' \otimes \mathbf{I_T}\right). \tag{2.33}$$

It follows that

$$\mathbf{G}^{-1} = \left(\mathbf{I_N} \otimes \mathbf{D}^{-1/2}\right)\left[\frac{1}{\sigma_1^2}\mathbf{E_N} \otimes \mathbf{I_T} + \left(\frac{1}{\sqrt{N}}\iota_N \otimes \mathbf{I_T}\right)\mathbf{A}\left(\frac{1}{\sqrt{N}}\iota_N' \otimes \mathbf{I_T}\right)\right]\left(\mathbf{I_N} \otimes \mathbf{D}^{-1/2}\right) \tag{2.34}$$

$$\mathbf{G}^{-1} = \frac{1}{\sigma_1^2}\mathbf{E_N} \otimes \mathbf{D}^{-1} + \left(\frac{1}{\sqrt{N}}\iota_N \otimes \mathbf{D}^{-1/2}\right)\mathbf{A}\left(\frac{1}{\sqrt{N}}\iota_N' \otimes \mathbf{D}^{-1/2}\right) = \frac{1}{\sigma_1^2}\mathbf{E_N} \otimes \mathbf{D}^{-1} + \left(\frac{1}{N}\iota_N\iota_N' \otimes \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}\right)$$

$$\mathbf{G}^{-1} = \frac{1}{\sigma_1^2}\left[\mathbf{E_N} \otimes \mathbf{D}^{-1} + \left(\frac{1}{N^2}\iota_N\iota_N' \otimes \mathbf{S}\right)\right] \tag{2.35}$$

in which

$$\mathbf{S} = \operatorname{diag}\left(s_1, \cdots, s_T\right) \tag{2.36}$$

with

$$s_t = \frac{N\sigma_1^2}{d_t\sigma_1^2 + N\Lambda_t\sigma_3^2}, \quad t = 1, \ldots, T. \tag{2.37}$$

On the other hand, the matrix $\left(\mathbf{J}'\mathbf{G}^{-1}\mathbf{J} + \frac{1}{\sigma_2^2}\mathbf{I_N}\right)^{-1}$ has to be determined. We have:

$$\mathbf{J}'\mathbf{G}^{-1}\mathbf{J} = \frac{1}{\sigma_1^2}\left(\mathbf{I_N} \otimes \iota_T^{\lambda\prime}\right)\left[\mathbf{E_N} \otimes \mathbf{D}^{-1} + \left(\frac{1}{N^2}\iota_N\iota_N' \otimes \mathbf{S}\right)\right]\left(\mathbf{I_N} \otimes \iota_T^{\lambda}\right) \tag{2.38}$$

$$\mathbf{J}'\mathbf{G}^{-1}\mathbf{J} = \frac{1}{\sigma_1^2}\left[\mathbf{E_N} \otimes \iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda} + \frac{1}{N^2}\iota_N\iota_N' \otimes \iota_T^{\lambda\prime}\mathbf{S}\iota_T^{\lambda}\right] = \frac{1}{\sigma_1^2}\left[\left(\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}\right)\mathbf{E_N} + \left(\iota_T^{\lambda\prime}\mathbf{S}\iota_T^{\lambda}\right)\frac{1}{N^2}\iota_N\iota_N'\right] \tag{2.39}$$

Hence,

$$\mathbf{J}'\mathbf{G}^{-1}\mathbf{J}+\frac{1}{\sigma_2^2}\mathbf{I_N}=\frac{1}{\sigma_1^2}\left[\left(\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}\right)\mathbf{E_N}+\left(\iota_T^{\lambda\prime}\mathbf{S}\iota_T^{\lambda}\right)\frac{1}{N^2}\iota_N\iota_N'\right]+\frac{1}{\sigma_2^2}\mathbf{I_N}$$

$$\mathbf{J}'\mathbf{G}^{-1}\mathbf{J}+\frac{1}{\sigma_2^2}\mathbf{I_N}=\frac{1}{\sigma_1^2}\left[\left(\mathbf{I_N}-\frac{1}{N}\iota_N\iota_N'\right)\left(\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}\right)+\left(\iota_T^{\lambda\prime}\mathbf{S}\iota_T^{\lambda}\right)\frac{1}{N^2}\iota_N\iota_N'+\frac{\sigma_1^2}{\sigma_2^2}\mathbf{I_N}\right]$$

$$\mathbf{J}'\mathbf{G}^{-1}\mathbf{J}+\frac{1}{\sigma_2^2}\mathbf{I_N}=\frac{1}{\sigma_1^2}\left[\mathbf{I_N}\left(\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}+\frac{\sigma_1^2}{\sigma_2^2}\right)+\iota_N\iota_N'\left(\frac{\iota_T^{\lambda\prime}\mathbf{S}\iota_T^{\lambda}}{N^2}-\frac{\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}}{N}\right)\right]=a\mathbf{I_N}+b\iota_N\iota_N' \tag{2.40}$$

where

$$a=\frac{\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}}{\sigma_1^2}+\frac{1}{\sigma_2^2} \tag{2.41}$$

and

$$b=\frac{\iota_T^{\lambda\prime}\mathbf{S}\iota_T^{\lambda}}{\sigma_1^2N^2}-\frac{\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}}{\sigma_1^2N}. \tag{2.42}$$

Knowing that $\left(a\mathbf{I_N}+b\iota_N\iota_N'\right)^{-1}=\frac{1}{a}\left(\mathbf{I_N}-\frac{b}{a+bN}\iota_N\iota_N'\right)$, we deduce $\left(\mathbf{J}'\mathbf{G}^{-1}\mathbf{J}+\frac{1}{\sigma_2^2}\mathbf{I_N}\right)^{-1}$.

We are now interested in the expression $\mathbf{G}^{-1}\mathbf{J}\left(\mathbf{J}'\mathbf{G}^{-1}\mathbf{J}+\frac{1}{\sigma_2^2}\mathbf{I_N}\right)^{-1}\mathbf{J}'\mathbf{G}^{-1}$. We have:

$$\mathbf{G}^{-1}\mathbf{J}\left(\mathbf{J}'\mathbf{G}^{-1}\mathbf{J}+\frac{1}{\sigma_2^2}\mathbf{I_N}\right)^{-1}\mathbf{J}'\mathbf{G}^{-1}=\frac{1}{a}\mathbf{G}^{-1}\mathbf{J}\left(\mathbf{I_N}-\frac{b}{a+bN}\iota_N\iota_N'\right)\mathbf{J}'\mathbf{G}^{-1}$$

$$\mathbf{G}^{-1}\mathbf{J}\left(\mathbf{J}'\mathbf{G}^{-1}\mathbf{J}+\frac{1}{\sigma_2^2}\mathbf{I_N}\right)^{-1}\mathbf{J}'\mathbf{G}^{-1}=\frac{1}{a}\left(\mathbf{G}^{-1}\mathbf{J}\mathbf{J}'\mathbf{G}^{-1}-\frac{b}{a+bN}\mathbf{G}^{-1}\mathbf{J}\iota_N\iota_N'\mathbf{J}'\mathbf{G}^{-1}\right). \quad (2.43)$$

From the definitions of the matrices $\mathbf{G}$ and $\mathbf{J}$, we can write:

$$\mathbf{G}^{-1}\mathbf{J}=\frac{1}{\sigma_1^2}\left[\mathbf{E_N}\otimes\mathbf{D}^{-1}\iota_T^{\lambda}+\left(\frac{1}{N^2}\iota_N\iota_N'\otimes\mathbf{S}\iota_T^{\lambda}\right)\right], \quad (2.44)$$

and

$$\mathbf{G}^{-1}\mathbf{J}\iota_N=\frac{\iota_N}{N\sigma_1^2}\otimes\mathbf{S}\iota_T^{\lambda}, \quad (2.45)$$

so that

$$\mathbf{G}^{-1}\mathbf{J}\mathbf{J}'\mathbf{G}^{-1}=\frac{1}{\sigma_1^4}\left[\left(\mathbf{E_N}\otimes\mathbf{D}^{-1}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{D}^{-1}\right)+\left(\frac{1}{N^3}\iota_N i_N'\otimes\mathbf{S}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{S}\right)\right], \quad (2.46)$$

and lastly

$$\mathbf{G}^{-1}\mathbf{J}i_N i_N'\mathbf{J}'\mathbf{G}^{-1}=\frac{1}{N^2\sigma_1^4}\iota_N\iota_N'\otimes\mathbf{S}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{S}. \quad (2.47)$$

It comes that

$$\mathbf{G}^{-1}\mathbf{J}\left(\mathbf{J}'\mathbf{G}^{-1}\mathbf{J}+\frac{1}{\sigma_2^2}\mathbf{I_N}\right)^{-1}\mathbf{J}'\mathbf{G}^{-1}=\frac{1}{a\sigma_1^4}\left[\mathbf{E_N}\otimes\mathbf{D}^{-1}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{D}^{-1}+\left(\frac{1}{N^3}\iota_N\iota_N'\otimes\mathbf{S}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{S}\right)-\frac{b}{N^2(a+bN)}\iota_N\iota_N'\otimes\mathbf{S}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{S}\right]$$

$$\mathbf{G}^{-1}\mathbf{J}\left(\mathbf{J}'\mathbf{G}^{-1}\mathbf{J}+\frac{1}{\sigma_2^2}\mathbf{I_N}\right)^{-1}\mathbf{J}'\mathbf{G}^{-1}=\frac{1}{a\sigma_1^4}\left[\mathbf{E_N}\otimes\left(\mathbf{D}^{-1}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{D}^{-1}\right)+\frac{1}{N^2}\left(\frac{1}{N}-\frac{b}{(a+bN)}\right)\left(i_N i_N'\otimes\mathbf{S}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{S}\right)\right]$$

$$\mathbf{G}^{-1}\mathbf{J}\left(\mathbf{J}'\mathbf{G}^{-1}\mathbf{J}+\frac{1}{\sigma_2^2}\mathbf{I_N}\right)^{-1}\mathbf{J}'\mathbf{G}^{-1}=\frac{1}{a\sigma_1^4}\mathbf{E_N}\otimes\left(\mathbf{D}^{-1}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{D}^{-1}\right)+\frac{1}{a\sigma_1^4N^3}\left(\frac{a}{a+bN}\right)\left(i_N i_N'\otimes\mathbf{S}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{S}\right). \tag{2.48}$$

Finally, the inverse of $\mathbf{\Sigma}^{**}$ can be derived:

$$\left(\mathbf{\Sigma}^{**}\right)^{-1}=\mathbf{G}^{-1}-\mathbf{G}^{-1}\mathbf{J}\left(\mathbf{J}'\mathbf{G}^{-1}\mathbf{J}+\frac{1}{\sigma_2^2}\mathbf{I_N}\right)^{-1}\mathbf{J}'\mathbf{G}^{-1} \tag{2.49}$$

$$\left(\mathbf{\Sigma}^{**}\right)^{-1}=\frac{1}{\sigma_1^2}\mathbf{E_N}\otimes\mathbf{D}^{-1}+\left(\frac{1}{\sigma_1^2N^2}\iota_N\iota_N'\otimes\mathbf{S}\right)-\frac{1}{a\sigma_1^4}\mathbf{E_N}\otimes\left(\mathbf{D}^{-1}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{D}^{-1}\right)-\frac{1}{a\sigma_1^4N^3}\left(\frac{a}{a+bN}\right)\left(\iota_N\iota_N'\otimes\mathbf{S}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{S}\right)$$

$$\left(\mathbf{\Sigma}^{**}\right)^{-1}=\mathbf{E_N}\otimes\left(\frac{1}{\sigma_1^2}\mathbf{D}^{-1}-\frac{1}{a\sigma_1^4}\mathbf{D}^{-1}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{D}^{-1}\right)+\mathbf{J_N}\otimes\left(\frac{1}{\sigma_1^2N^2}\mathbf{S}-\frac{1}{\sigma_1^4N^3\left(a+bN\right)}\mathbf{S}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{S}\right) \tag{2.50}$$

with $\mathbf{J_N}=\iota_N\iota_N'$.

We set

$$\mathbf{K_T}=\mathbf{D}^{-1}-\mathbf{L_T} \tag{2.51}$$

and

$$\mathbf{L_T}=\frac{1}{\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}}\mathbf{D}^{-1}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{D}^{-1}. \tag{2.52}$$

We then have

$$\left(\mathbf{\Sigma}^{**}\right)^{-1}=\mathbf{E_N}\otimes\left(\frac{1}{\sigma_1^2}\mathbf{D}^{-1}-\frac{1}{a\sigma_1^4}\mathbf{D}^{-1}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{D}^{-1}\right)+\mathbf{J_N}\otimes\left(\frac{1}{\sigma_1^2N^2}\mathbf{S}-\frac{1}{\sigma_1^4N^3\left(a+bN\right)}\mathbf{S}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{S}\right)$$

$$\left(\boldsymbol{\Sigma}^{**}\right)^{\mathbf{-1}} = \frac{1}{\sigma_1^2}\mathbf{E_N} \otimes \left(\mathbf{D^{-1}} - \mathbf{L_T} + \mathbf{L_T} - \frac{1}{a\sigma_1^2}\mathbf{D^{-1}}\iota_T^\lambda \iota_T^{\lambda\prime}\mathbf{D^{-1}}\right) + \mathbf{J_N} \otimes \left(\frac{1}{\sigma_1^2 N^2}\mathbf{S} - \frac{1}{\sigma_1^4 N^3\left(a+bN\right)}\mathbf{S}\iota_T^\lambda \iota_T^{\lambda\prime}\mathbf{S}\right)$$

$$\left(\boldsymbol{\Sigma}^{**}\right)^{\mathbf{-1}} = \frac{1}{\sigma_1^2}\left(\mathbf{E_N} \otimes \mathbf{K_T}\right) + \frac{1}{\sigma_1^2}\mathbf{E_N} \otimes \left(\mathbf{L_T} - \frac{\iota_T^{\lambda\prime}\mathbf{D^{-1}}\iota_T^\lambda}{a\sigma_1^2}\mathbf{L_T}\right) + \mathbf{J_N} \otimes \left(\frac{1}{\sigma_1^2 N^2}\mathbf{S} - \frac{1}{\sigma_1^4 N^3\left(a+bN\right)}\mathbf{S}\iota_T^\lambda \iota_T^{\lambda\prime}\mathbf{S}\right)$$

$$\left(\boldsymbol{\Sigma}^{**}\right)^{\mathbf{-1}} = \frac{1}{\sigma_1^2}\left(\mathbf{E_N} \otimes \mathbf{K_T}\right) + \frac{1}{\sigma_1^2}\left(1 - \frac{\left(\iota_T^{\lambda\prime}\mathbf{D^{-1}}\iota_T^\lambda\right)\sigma_2^2}{\left(\iota_T^{\lambda\prime}\mathbf{D^{-1}}\iota_T^\lambda\right)\sigma_2^2 + \sigma_1^2}\right)\left(\mathbf{E_N} \otimes \mathbf{L_T}\right) + \mathbf{J_N} \otimes \left(\frac{1}{\sigma_1^2 N^2}\mathbf{S} - \frac{1}{\sigma_1^4 N^3\left(a+bN\right)}\mathbf{S}\iota_T^\lambda \iota_T^{\lambda\prime}\mathbf{S}\right)$$

$$\left(\boldsymbol{\Sigma}^{**}\right)^{\mathbf{-1}} = \frac{1}{\sigma_1^2}\left(\mathbf{E_N} \otimes \mathbf{K_T}\right) + \frac{1}{\left(\iota_T^{\lambda\prime}\mathbf{D^{-1}}\iota_T^\lambda\right)\sigma_2^2 + \sigma_1^2}\left(\mathbf{E_N} \otimes \mathbf{L_T}\right) + \frac{1}{N^2}\left(\mathbf{J_N} \otimes \mathbf{S_T}\right) \tag{2.53}$$

with

$$\mathbf{S_T} = \left(\frac{1}{\sigma_1^2}\mathbf{S} - \frac{1}{\sigma_1^4 N\left(a+bN\right)}\mathbf{S}\iota_T^\lambda \iota_T^{\lambda\prime}\mathbf{S}\right) = \left(\frac{1}{\sigma_1^2}\mathbf{S} - \frac{1}{\dfrac{\sigma_1^4 N + \sigma_1^2\sigma_2^2\left(\iota_T^{\lambda\prime}\mathbf{S}\iota_T^\lambda\right)}{\sigma_2^2}}\mathbf{S}\iota_T^\lambda \iota_T^{\lambda\prime}\mathbf{S}\right) \text{, i.e.}$$

$$\mathbf{S_T} = \left(\frac{1}{\sigma_1^2}\mathbf{S} - \frac{\sigma_2^2}{\sigma_1^4 N + \sigma_1^2\sigma_2^2\left(\iota_T^{\lambda\prime}\mathbf{S}\iota_T^\lambda\right)}\mathbf{S}\iota_T^\lambda \iota_T^{\lambda\prime}\mathbf{S}\right). \tag{2.54}$$

To put it in a nutshell, $\left(\boldsymbol{\Sigma}^{**}\right)^{-1}$ is given by:

$$\left(\boldsymbol{\Sigma}^{**}\right)^{-1} = \frac{1}{\sigma_1^2}\left(\mathbf{E_N} \otimes \mathbf{K_T}\right) + \frac{1}{d}\left(\mathbf{E_N} \otimes \mathbf{L_T}\right) + \frac{1}{N^2}\left(\mathbf{J_N} \otimes \mathbf{S_T}\right) \tag{2.55}$$

where

$$\begin{cases} d = \left(\iota_T^{\lambda\prime}\mathbf{D^{-1}}\iota_T^{\lambda}\right)\sigma_2^2 + \sigma_1^2 \\ \mathbf{J_N} = \iota_N\iota_N' \quad \text{and} \quad \mathbf{E_N} = \mathbf{I_N} - \dfrac{1}{N}\mathbf{J_N} \\ \mathbf{K_T} = \mathbf{D^{-1}} - \mathbf{L_T} \quad \text{with} \quad \mathbf{L_T} = \dfrac{1}{\iota_T^{\lambda\prime}\mathbf{D^{-1}}\iota_T^{\lambda}}\mathbf{D^{-1}}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{D^{-1}} \\ \mathbf{S_T} = \left(\dfrac{1}{\sigma_1^2}\mathbf{S} - \dfrac{\sigma_2^2}{\sigma_1^4 N + \sigma_1^2\sigma_2^2\left(\iota_T^{\lambda\prime}\mathbf{S}\iota_T^{\lambda}\right)}\mathbf{S}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{S}\right), \ \mathbf{S} = \operatorname{diag}\left(s_1,\cdots,s_T\right), \ \ s_t = \dfrac{N\sigma_1^2}{d_t\sigma_1^2 + N\Lambda_t\sigma_3^2}. \end{cases} \tag{2.56}$$

The resulting GLS estimator is written as[14]:

$$\eta_{GLS} = \left(\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}y^{**}. \tag{2.57}$$

2.2.2. Interpretation of the GLS Estimator

In classical two-way regression models, Swamy and Arora (1972) and Nerlove (1971) provide an interpretation of the GLS estimator, which is appealing in view of the sources of variation in sample data. In the straight line of their work, the GLS estimator may be viewed as obtained by pooling three uncorrelated estimators: the covariance estimator (or within estimator), the between-time estimator and the within-individual estimator. As explained by Revankar (1979), these estimators are derived from orthogonal transformations. Since the inverse of the covariance matrix of the errors is known, we can determine the appropriate transformation matrices to get our three estimators of interest. They are the same as those suggested by Revankar (1979). We have

---

[14] Throughout this dissertation, we didn't emphasize the implications of the absence or presence of a constant term. Contrary to the presentation of Revankar (1979), we instead focus on the coefficient vector η which may contain an intercept or not, the first column of $\mathbf{X}$ being a vector of one if necessary. The reader can even consider η as the slope vector. See ,for example, Revankar (1979), Greene (2008), Baltagi (2005) for the consequences of including an intercept.

1. The covariance estimator $\eta_c = \left(\mathbf{X}^{**\prime}\mathbf{A_1}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A_1}y^{**}$ where $\mathbf{A}_1 = \mathbf{E_N}\otimes\mathbf{K_T}$.

2. The between-time estimator $\eta_B = \left(\mathbf{X}^{**\prime}\mathbf{A_2}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A_2}y^{**}$ with $\mathbf{A_2} = \frac{1}{N^2}\mathbf{J_N}\otimes\mathbf{S_T}$.

3. The within-individual estimator $\eta_T = \left(\mathbf{X}^{**\prime}\mathbf{A_3}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A_3}y^{**}$ where $\mathbf{A_3} = \mathbf{E_N}\otimes\mathbf{L_T}$.

These estimators are obtained from some transformations of the regression model (2.18):

$$y^{**} = \mathbf{X}^{**}\eta + u^{**}.$$

It is premultiplied by three matrices which are $\mathbf{M}_1 = \mathbf{E_N}\otimes\mathbf{K_T}$, $\mathbf{M_2} = \frac{1}{N}\iota_N'\otimes\mathbf{I_T}$ and $\mathbf{M_3} = \mathbf{E_N}\otimes\mathbf{L_T}$, respectively. The remainder of this subsection is aimed at giving details on the derivations of the covariance, between-individual and within-time estimators.

Firstly, we premultiply equation (2.18) by $\mathbf{M}_1 = \mathbf{E_N}\otimes\mathbf{K_T}$, leading to the following model:

$$\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)y^{**} = \left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}\eta + \left(\mathbf{E_N}\otimes\mathbf{K_T}\right)u^{**}. \tag{2.58}$$

The error variance covariance matrix of the errors in this new model is given by

$$\mathbf{\Sigma_1} = \left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{\Sigma}^{**}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)$$

$$\mathbf{\Sigma_1} = \left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\left[\sigma_1^2\left(\mathbf{I_N}\otimes\mathbf{D}\right) + \sigma_2^2\left(\mathbf{I_N}\otimes\left(\iota_T^\lambda\iota_T^{\lambda\prime}\right)\right) + \sigma_3^2\left(\left(\iota_N\iota_N'\right)\otimes\mathbf{\Lambda}\right)\right]\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)$$

$$\mathbf{\Sigma_1} = \sigma_1^2\left(\mathbf{E_N}\otimes\mathbf{K_T}\mathbf{D}\mathbf{K_T}\right) \tag{2.59}$$

since

$$\mathbf{K_T}\iota_T^\lambda = \mathbf{E_N}\iota_N = 0. \tag{2.60}$$

Recalling the definitions of matrices $\mathbf{K_T}$ and $\mathbf{L_T}$, we point out that

$$\mathbf{L_T D L_T} = \frac{1}{\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}}\left(\mathbf{D}^{-1}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{D}^{-1}\right)\mathbf{D}\frac{1}{\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}}\left(\mathbf{D}^{-1}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{D}^{-1}\right) = \frac{1}{\left(\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}\right)^2}\mathbf{D}^{-1}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{D}^{-1}$$

$$\mathbf{L_T D L_T} = \frac{1}{\left(\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}\right)^2}\mathbf{D}^{-1}\iota_T^{\lambda}\left(\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}\right)\iota_T^{\lambda\prime}\mathbf{D}^{-1} = \mathbf{L_T} \tag{2.61}$$

and $\mathbf{K_T D K_T = (D^{-1} - L_T)D(D^{-1} - L_T) = D^{-1} - L_T - L_T + L_T D L_T = D^{-1} - L_T - L_T + L_T}$, i.e.

$$\mathbf{K_T D K_T = K_T}. \tag{2.62}$$

As a consequence it comes that

$$\mathbf{\Sigma_1} = \sigma_1^2\left(\mathbf{E_N}\otimes\mathbf{K_T}\right) \tag{2.63}$$

and that a g-inverse of $\mathbf{K_T}$ is $\mathbf{D}$.

Also, since $\mathbf{E_N}$ is an idempotent matrix, its Moore-Penrose inverse is itself. Thus, a g-inverse of the covariance matrix $\mathbf{\Sigma_1}$ is simply $\frac{1}{\sigma_1^2}\left(\mathbf{E_N}\otimes\mathbf{D}\right)$.

The GLS estimator deduced from this equation, is then

$$\eta_c = \left[\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\left[\frac{1}{\sigma_1^2}\mathbf{E_N}\otimes\mathbf{D}\right]\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}\right]^{-1}\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\left[\frac{1}{\sigma_1^2}\mathbf{E_N}\otimes\mathbf{D}\right]\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)y^{**}$$

$$\eta_c = \left(\mathbf{X}^{**\prime}\mathbf{A_1}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A_1}y^{**} \tag{2.64}$$

with

$\mathbf{A_1} = \left(\mathbf{E_N} \otimes \mathbf{K_T}\right)\left(\mathbf{E_N} \otimes \mathbf{D}\right)\left(\mathbf{E_N} \otimes \mathbf{K_T}\right) = \mathbf{E_N} \otimes \mathbf{K_T}$, i.e

$$\mathbf{A_1} = \mathbf{E_N} \otimes \mathbf{K_T} = \mathbf{M_1}. \tag{2.65}$$

Thus, when the model (2.18) is premultiplied by $\mathbf{M_1} = \mathbf{E_N} \otimes \mathbf{K_T}$, the transformation annihilates the individual- and time- effects, and also eliminates the first column of the matrix of explanatory variables. This then amounts to covariance estimation of the slope vector. It is equivalent to the within estimator in the classical two-way error components model. (See Revankar, 1979, Greene, 2008 and Baltagi, 2005).

Secondly, we premultiply the model (2.18) by $\mathbf{M_2} = \frac{1}{N}\iota_N' \otimes \mathbf{I_T}$. The transformed model is

$$\left(\frac{\iota_N'}{N} \otimes \mathbf{I_T}\right) y^{**} = \left(\frac{\iota_N'}{N} \otimes \mathbf{I_T}\right)\mathbf{X^{**}}\eta + \left(\frac{\iota_N'}{N} \otimes \mathbf{I_T}\right) u^{**}. \tag{2.66}$$

The error term variance covariance matrix of the model is then

$$\mathbf{\Sigma_2} = \left(\frac{\iota_N'}{N} \otimes \mathbf{I_T}\right)\mathbf{\Sigma^{**}}\left(\frac{\iota_N}{N} \otimes \mathbf{I_T}\right)$$

$$\mathbf{\Sigma_2} = \left(\frac{\iota_N'}{N} \otimes \mathbf{I_T}\right)\left\{\sigma_1^2\left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_2^2\left[\mathbf{I_N} \otimes \left(\iota_T^\lambda \iota_T^{\lambda\prime}\right)\right] + \sigma_3^2\left[\left(\iota_N \iota_N'\right) \otimes \mathbf{\Lambda}\right]\right\}\left(\frac{\iota_N}{N} \otimes \mathbf{I_T}\right)$$

$$\mathbf{\Sigma_2} = \left(\frac{\sigma_1^2}{N}\mathbf{D} + \frac{\sigma_2^2}{N}\left(\iota_T^\lambda \iota_T^{\lambda\prime}\right)\right) + \sigma_3^2\mathbf{\Lambda} = \mathrm{diag}\left(\sigma_1^2\frac{d_t}{N} + \sigma_3^2\Lambda_t\right) + \frac{\sigma_2^2}{N}\left(\iota_T^\lambda \iota_T^{\lambda\prime}\right) = \sigma_1^2\mathbf{S^{-1}} + \frac{\sigma_2^2}{N}\left(\iota_T^\lambda \iota_T^{\lambda\prime}\right) \tag{2.67}$$

where

$$\mathbf{S} = \mathrm{diag}\left(\frac{N\sigma_1^2}{d_t\sigma_1^2 + N\Lambda_t\sigma_3^2}\right). \tag{2.68}$$

Once again, the use of an updating formula yields

$$\left[\sigma_1^2\mathbf{S}^{-1}+\frac{\sigma_2^2}{N}\left(\iota_T^{\lambda}\iota_T^{\lambda\prime}\right)\right]^{-1}=\frac{1}{\sigma_1^2}\left(\mathbf{S}-\frac{\dfrac{\sigma_2^2}{N\sigma_1^2}}{1+\dfrac{\sigma_2^2}{N\sigma_1^2}\iota_T^{\lambda\prime}\mathbf{S}\mathbf{i}_T^{\lambda}}\mathbf{S}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{S}\right)$$

$$\left[\sigma_1^2\mathbf{S}^{-1}+\frac{\sigma_2^2}{N}\left(\iota_T^{\lambda}\iota_T^{\lambda\prime}\right)\right]^{-1}=\left(\frac{1}{\sigma_1^2}\mathbf{S}-\frac{\sigma_2^2}{N\sigma_1^4+\sigma_1^2\sigma_2^2\iota_T^{\lambda\prime}\mathbf{S}\iota_T^{\lambda}}\mathbf{S}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{S}\right)$$

$$\left[\sigma_1^2\mathbf{S}^{-1}+\frac{\sigma_2^2}{N}\left(\iota_T^{\lambda}\iota_T^{\lambda\prime}\right)\right]^{-1}=\mathbf{S_T} \tag{2.69}$$

that is,

$$\mathbf{\Sigma_2^{-1}}=\mathbf{S_T}. \tag{2.70}$$

The estimator η_B of this model is defined as

$$\eta_B=\left[\mathbf{X}^{**\prime}\left(\frac{\iota_N}{N}\otimes\mathbf{I_T}\right)\mathbf{\Sigma_2^{-1}}\left(\frac{\iota_N}{N}\otimes\mathbf{I_T}\right)\mathbf{X}^{**}\right]^{-1}\mathbf{X}^{**\prime}\left(\frac{\iota_N}{N}\otimes\mathbf{I_T}\right)\mathbf{\Sigma_2^{-1}}\left(\frac{\iota_N}{N}\otimes\mathbf{I_T}\right)y^{**} \tag{2.71}$$

with

$$\left(\frac{\iota_N}{N}\otimes\mathbf{I_T}\right)\mathbf{\Sigma_2^{-1}}\left(\frac{\iota_N'}{N}\otimes\mathbf{I_T}\right)=\frac{\iota_N\iota_N'}{N^2}\otimes\mathbf{\Sigma_2^{-1}}=\frac{\iota_N\iota_N'}{N^2}\otimes\mathbf{S_T}=\frac{1}{N^2}\left(\mathbf{J_N}\otimes\mathbf{S_T}\right). \tag{2.72}$$

Hence,

$$\eta_B=\left(\mathbf{X}^{**\prime}\mathbf{A_2}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A_2}y^{**} \tag{2.73}$$

with

$$\mathbf{A_2} = \frac{\iota_N \iota_N'}{N^2} \otimes \mathbf{S_T} = \frac{1}{N^2} \mathbf{J_N} \otimes \mathbf{S_T}. \tag{2.74}$$

It is worth mentioning that the matrices $\mathbf{M_2}$ and $\mathbf{A_2}$ are different. This is due to the fact $\mathbf{M_2}$ is not an idempotent matrix. Finally, it appears that the second transformation is equivalent to averaging individual equations for each time period. It leads to the between-time estimator, as defined by Revankar (1979).

Lastly, we consider the third transformation defined by $\mathbf{M_3} = \left(\mathbf{E_N} \otimes \mathbf{L_T}\right)$. The new variance covariance matrix is

$$\mathbf{\Sigma_3} = \left(\mathbf{E_N} \otimes \mathbf{L_T}\right)\mathbf{\Sigma}^{**}\left(\mathbf{E_N} \otimes \mathbf{L_T}\right)$$

$$\mathbf{\Sigma_3} = \left(\mathbf{E_N} \otimes \mathbf{L_T}\right)\left[\sigma_1^2\left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_2^2\left(\mathbf{I_N} \otimes \left(\iota_T^{\lambda}\iota_T^{\lambda\prime}\right)\right) + \sigma_3^2\left(\left(\iota_N \iota_N'\right) \otimes \mathbf{\Lambda}\right)\right]\left(\mathbf{E_N} \otimes \mathbf{L_T}\right)$$

$$\mathbf{\Sigma_3} = \sigma_1^2\left(\mathbf{E_N} \otimes \mathbf{L_T D L_T}\right) + \sigma_2^2\left(\mathbf{E_N} \otimes \left(\mathbf{L_T}\iota_T^{\lambda}\iota_T^{\lambda\prime}\mathbf{L_T}\right)\right) = \sigma_1^2\left(\mathbf{E_N} \otimes \mathbf{L_T}\right) + \sigma_2^2\left(\iota_T^{\lambda\prime}\mathbf{D^{-1}}\iota_T^{\lambda}\right)\left(\mathbf{E_N} \otimes \mathbf{L_T}\right)$$

$$\mathbf{\Sigma_3} = \left[\sigma_1^2 + \sigma_2^2\left(\iota_T^{\lambda\prime}\mathbf{D^{-1}}\iota_T^{\lambda}\right)\right]\left(\mathbf{E_N} \otimes \mathbf{L_T}\right) = d\left(\mathbf{E_N} \otimes \mathbf{L_T}\right) \tag{2.75}$$

with

$$d = \sigma_1^2 + \sigma_2^2\left(\iota_T^{\lambda\prime}\mathbf{D^{-1}}\iota_T^{\lambda}\right).$$

According to equation (2.61),

$$\mathbf{L_T D L_T} = \mathbf{L_T}$$

Hence, a g-inverse of $\mathbf{L_T}$ is $\mathbf{D}$. We already know that $\mathbf{E_N}$ is its own Moore-Penrose inverse. Therefore, a g-inverse of $\mathbf{\Sigma_3}$ is $\frac{1}{d}(\mathbf{E_N} \otimes \mathbf{L_T})$.

The estimator η_T of this last model is such that:

$$\eta_T = \left(\mathbf{X}^{**\prime}\mathbf{A_3}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A_3}y^{**} \tag{2.76}$$

with

$$\mathbf{A_3} = \left(\mathbf{E_N} \otimes \mathbf{L_T}\right)\left(\mathbf{E_N} \otimes \mathbf{D}\right)\left(\mathbf{E_N} \otimes \mathbf{L_T}\right) = \left(\mathbf{E_N} \otimes \mathbf{L_T}\right)$$

$$\mathbf{A_3} = \left(\mathbf{E_N} \otimes \mathbf{L_T}\right) = \mathbf{M_3}. \tag{2.77}$$

Thus, we shall mention that the presence of the idempotent matrix $\mathbf{E_N}$ in $\mathbf{M_3}$ indicates that this transformation wipes out the constant term as well as the time specific error term λ_t, along with the findings of Revankar (1979). However, we disagree with Revankar (1979) about the name given to this last estimator: he called it "between-individual" estimate. We argue that the idempotent matrix $\mathbf{E_N}$, is a within operator rather than a between one. Moreover, a between transformation naturally keeps the constant term as in the case of $\mathbf{M_2}$ while a within one removes it. In addition, $\mathbf{M_3}$ differs from the classical within (or covariance) operator by the fact that it annihilates the time-specific effect, but not the individual-heterogeneity disturbance. In order to stress the presence of individual specificities in the resulting model under a within transformation, the GLS estimator obtained has been called a *within-individual* estimator.

Another important feature of the three estimators η_C, η_B, and η_T has to be pointed out. We find that the GLS estimator η_{GLS} is a weighted average of these estimators. It is similar to the findings of Maddala (1971a, 1971b). Let us show it.

From equation (2.57), the GLS estimator η_{GLS} is defined by

$$\eta_{GLS} = \left(\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}y^{**}$$

or

$$\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}y^{**} = \left(\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}\right)\eta_{GLS} \tag{2.78}$$

with

$$\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}y^{**} = \frac{1}{\sigma_1^2}\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)y^{**} + \frac{1}{d}\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)y^{**} + \mathbf{X}^{**\prime}\left[\frac{1}{N^2}\left(\mathbf{J_N}\otimes\mathbf{S_T}\right)\right]y^{**}$$

i.e.

$$\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}y^{**} = \frac{1}{\sigma_1^2}\mathbf{X}^{**\prime}\mathbf{A_1}y^{**} + \mathbf{X}^{**\prime}\mathbf{A_2}y^{**} + \frac{1}{d}\mathbf{X}^{**\prime}\mathbf{A_3}y^{**}. \tag{2.79}$$

But from their definitions, the estimators η_c, η_B and η_T are respectively such that

$$\mathbf{X}^{**\prime}\mathbf{A_1}y^{**} = \left(\mathbf{X}^{**\prime}\mathbf{A_1}\mathbf{X}^{**}\right)\eta_C, \tag{2.80}$$

$$\mathbf{X}^{**\prime}\mathbf{A_2}y^{**} = \left(\mathbf{X}^{**\prime}\mathbf{A_2}\mathbf{X}^{**}\right)\eta_B \tag{2.81}$$

and

$$\mathbf{X}^{**\prime}\mathbf{A_3}y^{**} = \left(\mathbf{X}^{**\prime}\mathbf{A_3}\mathbf{X}^{**}\right)\eta_T. \tag{2.82}$$

Therefore,

$$\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1} y^{**} = \frac{1}{\sigma_1^2}\left(\mathbf{X}^{**\prime}\mathbf{A_1}\mathbf{X}^{**}\right)\eta_C + \left(\mathbf{X}^{**\prime}\mathbf{A_2}\mathbf{X}^{**}\right)\eta_B + \frac{1}{d}\left(\mathbf{X}^{**\prime}\mathbf{A_3}\mathbf{X}^{**}\right)\eta_T . \tag{2.83}$$

In other words,

$$\left(\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}\right)\eta_{GLS} = \frac{1}{\sigma_1^2}\left(\mathbf{X}^{**\prime}\mathbf{A_1}\mathbf{X}^{**}\right)\eta_C + \left(\mathbf{X}^{**\prime}\mathbf{A_2}\mathbf{X}^{**}\right)\eta_B + \frac{1}{d}\left(\mathbf{X}^{**\prime}\mathbf{A_3}\mathbf{X}^{**}\right)\eta_T \tag{2.84}$$

$$\begin{aligned}\eta_{GLS} = \frac{1}{\sigma_1^2}&\left(\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}\right)^{-1}\left(\mathbf{X}^{**\prime}\mathbf{A_1}\mathbf{X}^{**}\right)\eta_C + \left(\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}\right)^{-1}\left(\mathbf{X}^{**\prime}\mathbf{A_2}\mathbf{X}^{**}\right)\eta_B + \\ &\frac{1}{d}\left(\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}\right)^{-1}\left(\mathbf{X}^{**\prime}\mathbf{A_3}\mathbf{X}^{**}\right)\eta_T .\end{aligned} \tag{2.85}$$

Thus,

$$\eta_{GLS} = \mathbf{F_C}\eta_C + \mathbf{F_B}\eta_B + \mathbf{F_T}\eta_T \tag{2.86}$$

with

$$\mathbf{F_C} = \frac{1}{\sigma_1^2}\left(\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A_1}\mathbf{X}^{**}, \tag{2.87}$$

$$\mathbf{F_B} = \left(\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_2\mathbf{X}^{**}, \tag{2.88}$$

and

$$\mathbf{F_T} = \frac{1}{d}\left(\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_3\mathbf{X}^{**} = \mathbf{I} - \mathbf{F_C} - \mathbf{F_B} . \tag{2.89}$$

Finally, the GLS estimator η_{GLS} is the weighted average of η_C, η_B, and η_T, where the weights are given by the matrices $\mathbf{F_C}$, $\mathbf{F_B}$, and $\mathbf{F_T}$.

To complete the interpretation of the double autocorrelation GLS estimator η_{GLS}, we show that the three estimators η_C, η_B, and η_T are mutually uncorrelated.

By definition, these estimates are such that :

$$\begin{cases} \eta_c = \left(\mathbf{X}^{**\prime}\mathbf{A}_1\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_1 y^{**} = \eta + \left(\mathbf{X}^{**\prime}\mathbf{A}_1\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_1 u^{**} \\ \eta_B = \left(\mathbf{X}^{**\prime}\mathbf{A}_2\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_2 y^{**} = \eta + \left(\mathbf{X}^{**\prime}\mathbf{A}_2\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_2 u^{**} \\ \eta_T = \left(\mathbf{X}^{**\prime}\mathbf{A}_3\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_3 y^{**} = \eta + \left(\mathbf{X}^{**\prime}\mathbf{A}_3\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_3 u^{**}. \end{cases}$$

It follows that

$$\begin{cases} \operatorname{cov}(\eta_c,\eta_B) = \operatorname{cov}\left[\left(\mathbf{X}^{**\prime}\mathbf{A}_1\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_1 y^{**}, \left(\mathbf{X}^{**\prime}\mathbf{A}_2\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_2 y^{**}\right] \\ \operatorname{cov}(\eta_c,\eta_T) = \operatorname{cov}\left[\left(\mathbf{X}^{**\prime}\mathbf{A}_1\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_1 y^{**}, \left(\mathbf{X}^{**\prime}\mathbf{A}_3\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_3 y^{**}\right] \\ \operatorname{cov}(\eta_B,\eta_T) = \operatorname{cov}\left[\left(\mathbf{X}^{**\prime}\mathbf{A}_2\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_2 y^{**}, \left(\mathbf{X}^{**\prime}\mathbf{A}_3\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_3 y^{**}\right]. \end{cases}$$

In other words,

$$\begin{cases} \operatorname{cov}(\eta_c,\eta_B) = \left(\mathbf{X}^{**\prime}\mathbf{A}_1\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_1\left[\operatorname{var}\left(y^{**}\right)\right]\mathbf{A}_2\mathbf{X}^{**}\left(\mathbf{X}^{**\prime}\mathbf{A}_2\mathbf{X}^{**}\right)^{-1} \\ \operatorname{cov}(\eta_c,\eta_T) = \left(X^{**\prime}A_1X^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_1\left[\operatorname{var}\left(y^{**}\right)\right]\mathbf{A}_3\mathbf{X}^{**}\left(\mathbf{X}^{**\prime}\mathbf{A}_3\mathbf{X}^{**}\right)^{-1} \\ \operatorname{cov}(\eta_B,\eta_T) = \left(X^{**\prime}A_2X^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_2\left[\operatorname{var}\left(y^{**}\right)\right]\mathbf{A}_3\mathbf{X}^{**}\left(\mathbf{X}^{**\prime}\mathbf{A}_3\mathbf{X}^{**}\right)^{-1} \end{cases}$$

or

$$\begin{cases} \operatorname{cov}(\eta_c, \eta_B) = \left(\mathbf{X}^{**\prime}\mathbf{A_1}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\left(\mathbf{A_1}\boldsymbol{\Sigma}^{**}\mathbf{A_2}\right)\mathbf{X}^{**}\left(\mathbf{X}^{**\prime}\mathbf{A_2}\mathbf{X}^{**}\right)^{-1} \\ \operatorname{cov}(\eta_c, \eta_T) = \left(\mathbf{X}^{**\prime}\mathbf{A_1}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\left(\mathbf{A_1}\boldsymbol{\Sigma}^{**}\mathbf{A_3}\right)\mathbf{X}^{**}\left(\mathbf{X}^{**\prime}\mathbf{A_3}\mathbf{X}^{**}\right)^{-1} \\ \operatorname{cov}(\eta_B, \eta_T) = \left(\mathbf{X}^{**\prime}\mathbf{A_2}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\left(\mathbf{A_2}\boldsymbol{\Sigma}^{**}\mathbf{A_3}\right)\mathbf{X}^{**}\left(\mathbf{X}^{**\prime}\mathbf{A_3}\mathbf{X}^{**}\right)^{-1}. \end{cases}$$

Firstly,

$$\mathbf{A_1}\boldsymbol{\Sigma}^{**}\mathbf{A_2} = \left(\mathbf{E_N} \otimes \mathbf{K_T}\right)\left[\sigma_1^2\left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_2^2\left(\mathbf{I_N} \otimes \left(\iota_T^{\lambda}\iota_T^{\lambda\prime}\right)\right) + \sigma_3^2\left(\left(\iota_N\iota_N'\right) \otimes \boldsymbol{\Lambda}\right)\right]\left(\frac{1}{N^2}\mathbf{J_N} \otimes \mathbf{S_T}\right)$$

$$\mathbf{A_1}\boldsymbol{\Sigma}^{**}\mathbf{A_2} = \frac{\sigma_1^2}{N^2}\left(\mathbf{E_N}\mathbf{J_N} \otimes \mathbf{K_T}\mathbf{D}\mathbf{S_T}\right) = \mathbf{0} \tag{2.90}$$

since

$$\mathbf{K_T}\iota_T^{\lambda} = \mathbf{E_N}\iota_N = 0. \tag{2.91}$$

Secondly,

$$\mathbf{A_1}\boldsymbol{\Sigma}^{**}\mathbf{A_3} = \left(\mathbf{E_N} \otimes \mathbf{K_T}\right)\left[\sigma_1^2\left(\mathbf{I_N} \otimes \mathbf{D}\right) + \sigma_2^2\left(\mathbf{I_N} \otimes \left(\iota_T^{\lambda}\iota_T^{\lambda\prime}\right)\right) + \sigma_3^2\left(\left(\iota_N\iota_N'\right) \otimes \boldsymbol{\Lambda}\right)\right]\left(\mathbf{E_N} \otimes \mathbf{L_T}\right)$$

$$\mathbf{A_1}\boldsymbol{\Sigma}^{**}\mathbf{A_3} = \sigma_1^2\left(\mathbf{E_N} \otimes \mathbf{K_T}\mathbf{D}\mathbf{L_T}\right) \tag{2.92}$$

Because $\mathbf{K_T}\iota_T^{\lambda} = \mathbf{E_N}\iota_N = 0$.

We also have, by definition, $\mathbf{K_T} = \mathbf{D}^{-1} - \mathbf{L_T}$ and we have already shown that $\mathbf{L_T}\mathbf{D}\mathbf{L_T} = \mathbf{L_T}$. Hence,

$$\mathbf{K_T}\mathbf{D}\mathbf{L_T} = \mathbf{L_T} - \mathbf{L_T}\mathbf{D}\mathbf{L_T} = \mathbf{L_T} - \mathbf{L_T} = \mathbf{0}. \tag{2.93}$$

Thus, $\mathbf{A_1}\boldsymbol{\Sigma}^{**}\mathbf{A_3} = \mathbf{0}$. (2.94)

Thirdly,

$$\mathbf{A}_2\boldsymbol{\Sigma}^{**}\mathbf{A_3} = \left(\frac{1}{N^2}\mathbf{J_N}\otimes\mathbf{S_T}\right)\left[\sigma_1^2\left(\mathbf{I_N}\otimes\mathbf{D}\right)+\sigma_2^2\left(\mathbf{I_N}\otimes\left(\iota_T^{\lambda}\iota_T^{\lambda\prime}\right)\right)+\sigma_3^2\left(\mathbf{J_N}\otimes\boldsymbol{\Lambda}\right)\right]\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)$$

$$\mathbf{A}_2\boldsymbol{\Sigma}^{**}\mathbf{A_3} = \mathbf{0} \tag{2.95}$$

since $\mathbf{J_N}\mathbf{E_N} = \mathbf{0}$.

Finally, we find that

$$\mathbf{A_1}\boldsymbol{\Sigma}^{**}\mathbf{A_2} = \mathbf{A_1}\boldsymbol{\Sigma}^{**}\mathbf{A_3} = \mathbf{A_2}\boldsymbol{\Sigma}^{**}\mathbf{A_3} = \mathbf{0}.$$

As a result,

$$\operatorname{cov}\left(\eta_c,\eta_B\right) = \operatorname{cov}\left(\eta_c,\eta_T\right) = \operatorname{cov}\left(\eta_B,\eta_T\right) = 0. \tag{2.96}$$

Moreover, as explained by Revankar (1979), the fact that

$$\operatorname{rank}\left(\mathbf{M_1}\right)+\operatorname{rank}\left(\mathbf{M_2}\right)+\operatorname{rank}\left(\mathbf{M_3}\right) = \left(N-1\right)\left(T-1\right)+T+N-1 = NT \tag{2.97}$$

gives evidence on the use of all the available information from the sample. The estimators η_c, η_B, and η_T together use up the entire set of information to build our GLS estimator η_{GLS} with no lost at all.

2.2.3. Asymptotic Properties of the GLS Estimator

We assume, for convergence purpose (see Revankar, 1979, and Wallace and Hussain, 1969), that the x_{it} s are weakly non-stochastic, i.e. do not repeat in repeated samples.[15] Here we show that the GLS and our estimators of the coefficient vector, say η_{GLS}, η_C, η_B, and η_T are all consistent and asymptotically normally distributed. In particular, we show that η_{GLS} and η_C are asymptotically equivalent. It is a result similar to the one obtained in the classical two-way error component model (see Amemiya, 1971).

We first need to set few assumptions. We assume that the following matrices exist and are positive definite with finite elements:

(a1) $$\text{plim}\left(\frac{\mathbf{X}^{**\prime}\mathbf{A}_1\mathbf{X}^{**}}{NT}\right) = \text{plim}\left(\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}}{NT}\right)$$ (for the first transformation)

(b1) $$\text{plim}\left(\frac{\mathbf{X}^{**\prime}\mathbf{A}_2\mathbf{X}^{**}}{T}\right) = \text{plim}\left(\frac{\mathbf{X}^{**\prime}\left(\frac{1}{N^2}\mathbf{J_N}\otimes\mathbf{S_T}\right)\mathbf{X}^{**}}{T}\right)$$ (for the second transformation)

(c1) $$\text{plim}\left(\frac{\mathbf{X}^{**\prime}\mathbf{A}_3\mathbf{X}^{**}}{NT}\right) = \text{plim}\left(\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**}}{NT}\right)$$ (for the third transformation).

Furthermore, in the straight line of Revankar's (1979) approach, we also assume that

(a2) $$\text{plim}\left(\frac{\mathbf{X}^{**\prime}\mathbf{A}_1 u^{**}}{NT}\right) = \text{plim}\left(\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)u^{**}}{NT}\right) = 0$$

[15] One can also consider some general requirements, such as those of Gordin (1969), say summability of autocovariances, asymptotic uncorrelatedness, and asymptotic negligibility of innovations (see Greene, 2008).

(b2) $\text{plim} = \left(\dfrac{\mathbf{X}^{**\prime} \mathbf{A_2} u^{**}}{T} \right) = \text{plim} \left(\dfrac{\mathbf{X}^{**\prime} \left(\dfrac{1}{N^2} \mathbf{J_N} \otimes \mathbf{S_T} \right) u^{**}}{T} \right) = 0$

(c2) $\text{plim} = \left(\dfrac{\mathbf{X}^{**\prime} \mathbf{A_3} u^{**}}{NT} \right) = \text{plim} \left(\dfrac{\mathbf{X}^{**\prime} \left(\mathbf{E_N} \otimes \mathbf{L_T} \right) u^{**}}{NT} \right) = 0$.

We finally state that $\lim\limits_{T \to \infty} \left(\iota_T^{\lambda\prime} \mathbf{D}^{\mathbf{-1}} \iota_T^{\lambda} \right) = \infty$, so that the quantity $d = \left(\iota_T^{\lambda\prime} \mathbf{D}^{\mathbf{-1}} \iota_T^{\lambda} \right) \sigma_2^2 + \sigma_1^2$ remains infinite as $T \to \infty$. The limits and probabilities are taken as $N \to \infty$ and $T \to \infty$. It is worth mentioning that from all these assumptions we deduce the unbiasedness, consistency as well as the asymptotic normality of all the three GLS estimators, say η_C, η_B, and η_T.

Firstly, under the **M1**-transformation, we have:

$$\left(\mathbf{E_N} \otimes \mathbf{K_T} \right) y^{**} = \left(\mathbf{E_N} \otimes \mathbf{K_T} \right) \mathbf{X}^{**} \eta + \left(\mathbf{E_N} \otimes \mathbf{K_T} \right) u^{**}.$$

We start by analyzing some properties of this model. The mean of the **M1**-model's error term is

$$E\left[\left(\mathbf{E_N} \otimes \mathbf{K_T} \right) u^{**} \right] = \left(\mathbf{E_N} \otimes \mathbf{K_T} \right) E\left(u^{**} \right) = 0. \tag{2.98}$$

Moreover its variance is given by

$$\text{Var}\left[\left(\mathbf{E_N} \otimes \mathbf{K_T} \right) u^{**} \right] = V\left[\left(\mathbf{E_N} \otimes \mathbf{K_T} \right) u^{**} \right] = \left(\mathbf{E_N} \otimes \mathbf{K_T} \right) \mathbf{\Sigma}^{**} \left(\mathbf{E_N} \otimes \mathbf{K_T} \right) = \sigma_1^2 \left(\mathbf{E_N} \otimes \mathbf{K_T} \right) \tag{2.99}$$

and its inverse is equal to $\dfrac{1}{\sigma_1^2} \left(\mathbf{E_N} \otimes \mathbf{D} \right)$, as shown earlier. In addition, assumption (a2) establishes the lack of correlation between the explanatory variables and the error term. Lastly, in the straight line of the approach of Judge and al (1985) in their version of the Mann and Wald (1943) theorem, we show that the quantity $\frac{\mathbf{X}^{**'}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\left\{V\left[\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)u^{**}\right]\right\}^{-1}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}}{NT}$ converges in probability to a finite and positive definite matrix. We have,

$$\text{plim}\left(\frac{\mathbf{X}^{**'}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\left\{V\left[\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)u^{**}\right]\right\}^{-1}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}}{NT}\right)=\text{plim}\left(\frac{\mathbf{X}^{**'}\mathbf{A_1}\left[\frac{1}{\sigma_1^2}\left(\mathbf{E_N}\otimes\mathbf{D}\right).\right]\mathbf{A_1}\mathbf{X}^{**}}{NT}\right)$$

$$\text{plim}\left(\frac{\mathbf{X}^{**'}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\left\{V\left[\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)u^{**}\right]\right\}^{-1}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}}{NT}\right)=\frac{1}{\sigma_1^2}\text{plim}\left(\frac{\mathbf{X}^{**'}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}}{NT}\right).$$

(2.100)

This expression is positive definite according to assumption (a1).

Unfortunately, the conditions for the use of Mann and Wald (1943) theorem are not fully met, especially the need for $\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)u^{**}$ to be spherical, i.e. $V\left[\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)u^{**}\right]$ being proportional to the identity matrix. We can nonetheless use a more classical approach.

From the above properties, we firstly deduce the consistency of the covariance estimate η_C. We write:

$$\eta_C=\left(\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)y^{**}$$

$$\eta_C = \left(\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}\eta + \left(\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)u^{**}$$

$$\eta_C = \eta + \left(\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)u^{**}. \tag{2.101}$$

Hence,

$$\text{plim}\left(\eta_C - \eta\right) = \text{plim}\left(\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}}{NT}\right)^{-1} \bullet \text{plim}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)u^{**}}{NT} \tag{2.102}$$

Making use of assumptions (a1) and (a2), we obtain $\text{plim}\left(\eta_C - \eta\right) = 0$, i.e.

$$\text{plim}\left(\eta_C\right) = \eta. \tag{2.103}$$

The estimator η_C is therefore consistent.

We also deduce the asymptotic normality of η_C based on Gordin's (1969) central limit theorem (see Greene, 2008) for the time dimension. All along this subsection, following Revankar (1979), we will consider the "usual" assumptions regarding the elements of u^{**}, as stated in Theil (1971, p. 398) and Wallace and Hussain (1969, p. 64), which ensures the asymptotic normality. [16]

[16] Another strand of literature is the multi-index asymptotic theory of Phillips and Moon (1999, 2000). However, it proves more useful when dealing with nonstationary panels, which is not our case. Here, we use more classical approaches in order to stress the similarities as well as the differences with earlier works on close issues, notably with the paper of Revankar (1979).

From equations (2.98) and (2.99), it comes that

$$\frac{\mathbf{X}^{**\prime}(\mathbf{E_N} \otimes \mathbf{K_T})u^{**}}{\sqrt{NT}} \xrightarrow{d} N\left(0, \sigma_1^2 \text{plim} \frac{\mathbf{X}^{**\prime}(\mathbf{E_N} \otimes \mathbf{K_T})\mathbf{X}^{**}}{NT}\right). \tag{2.104}$$

Moreover, we have

$$\sqrt{NT}(\eta_C - \eta) = \left(\frac{\mathbf{X}^{**\prime}(\mathbf{E_N} \otimes \mathbf{K_T})\mathbf{X}^{**}}{NT}\right)^{-1} \frac{\mathbf{X}^{**\prime}(\mathbf{E_N} \otimes \mathbf{K_T})u^{**}}{\sqrt{NT}} \tag{2.105}$$

from which we deduce that

$$\sqrt{NT}(\eta_C - \eta) \xrightarrow{d} N(a_1, b_1) \tag{2.106}$$

with

$$a_1 = \text{plim}\left(\frac{\mathbf{X}^{**\prime}(\mathbf{E_N} \otimes \mathbf{K_T})\mathbf{X}^{**}}{NT}\right)^{-1} .\mathbf{0} = \mathbf{0} \tag{2.107}$$

and

$$b_1 = \sigma_1^2 \left(\text{plim} \frac{\mathbf{X}^{**\prime}(\mathbf{E_N} \otimes \mathbf{K_T})\mathbf{X}^{**}}{NT}\right)^{-1} \left[\text{plim} \frac{\mathbf{X}^{**\prime}(\mathbf{E_N} \otimes \mathbf{K_T})\mathbf{X}^{**}}{NT}\right] \left(\text{plim} \frac{\mathbf{X}^{**\prime}(\mathbf{E_N} \otimes \mathbf{K_T})\mathbf{X}^{**}}{NT}\right)^{-1}$$

$$b_1 = \sigma_1^2 \left(\text{plim} \frac{\mathbf{X}^{**\prime}(\mathbf{E_N} \otimes \mathbf{K_T})\mathbf{X}^{**}}{NT}\right)^{-1}. \tag{2.108}$$

Thus, we deduce the asymptotic normality of the covariance estimator η_C:

$$\eta_C \xrightarrow{a} N\left(\eta, \frac{1}{NT\sigma_1^2}\left(\text{plim}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}}{NT}\right)^{-1}\right). \tag{2.109}$$

Secondly, under the **M2-**transformation, we get:

$$\left(\frac{i_N^{'}}{N}\otimes\mathbf{I_T}\right)y^{**} = \left(\frac{i_N^{'}}{N}\otimes\mathbf{I_T}\right)\mathbf{X}^{**}\eta + \left(\frac{i_N^{'}}{N}\otimes\mathbf{I_T}\right)u^{**}.$$

We firstly analyze some simple properties of this model error terms. Again, the mean of the **M2**-model's error term vanishes:

$$E\left[\left(\frac{i_N^{'}}{N}\otimes\mathbf{I_T}\right)u^{**}\right] = \left(\frac{i_N^{'}}{N}\otimes\mathbf{I_T}\right)E\left(u^{**}\right) = 0. \tag{2.110}$$

The variance of this error term is written as

$$V\left[\left(\frac{i_N^{'}}{N}\otimes\mathbf{I_T}\right)u^{**}\right] = \sigma_1^2\mathbf{S^{-1}} + \frac{\sigma_2^2}{N}\left(i_T^{\lambda}i_T^{\lambda'}\right). \tag{2.111}$$

Its inverse is $\left(\sigma_1^2\mathbf{S^{-1}} + \frac{\sigma_2^2}{N}\left(i_T^{\lambda}i_T^{\lambda'}\right)\right)^{-1} = \mathbf{S_T}$. Once again, assumption (b2) states the absence of correlation between regressors and disturbances under the **M2** transformation. Moreover,

$$\left(\frac{\iota_N'}{N}\otimes\mathbf{I_T}\right)\left(V\left(\frac{\iota_N'}{N}\otimes\mathbf{I_T}\right)u^{**}\right)^{-1}\left(\frac{\iota_N}{N}\otimes\mathbf{I_T}\right)=\left(\frac{\iota_N'}{N}\otimes\mathbf{I_T}\right)\left(\sigma_1^2\mathbf{S^{-1}}+\frac{\sigma_2^2}{N}\left(i_T^{\lambda}i_T^{\lambda'}\right)\right)^{-1}\left(\frac{\iota_N'}{N}\otimes\mathbf{I_T}\right)=\frac{\iota_N\iota_N'}{N^2}\otimes\mathbf{S_T}.$$

So that

$$\text{plim}\left(\frac{\mathbf{X}^{**'}\left(\frac{\iota_N'}{N}\otimes\mathbf{I_T}\right)\left(V\left(\left(\frac{\iota_N'}{N}\otimes\mathbf{I_T}\right)u^{**}\right)\right)^{-1}\left(\frac{\iota_N}{N}\otimes I_T\right)\mathbf{X}^{**}}{T}\right)=\text{plim}\left(\frac{\mathbf{X}^{**'}\left(\frac{\iota_N\iota_N'}{N^2}\otimes\mathbf{S_T}\right)\mathbf{X}^{**}}{T}\right)$$

$$\text{plim}\left(\frac{\mathbf{X}^{**'}\left(\frac{\iota_N'}{N}\otimes\mathbf{I_T}\right)\left(V\left(\left(\frac{\iota_N'}{N}\otimes\mathbf{I_T}\right)u^{**}\right)\right)^{-1}\left(\frac{\iota_N}{N}\otimes I_T\right)\mathbf{X}^{**}}{T}\right)=\text{plim}\left(\frac{\mathbf{X}^{**'}\left(\frac{\mathbf{J_N}}{N^2}\otimes S_T\right)\mathbf{X}^{**}}{T}\right)$$

(2.112)

which is positive definite by assumption (b1). Once again, the conditions for the use of Mann and Wald (1943) theorem are not fully met. We can nevertheless deduce some asymptotical properties of η_B.

Firstly, its consistency is immediately derived:

$$\eta_B=\left(\mathbf{X}^{**\prime}\mathbf{A_2}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A_2}y^{**}\ \text{with}\ \mathbf{A_2}=\frac{1}{N^2}\mathbf{J_N}\otimes\mathbf{S_T}$$

$$\eta_B=\left(\mathbf{X}^{**\prime}\mathbf{A_2}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A_2}\left(\mathbf{X}^{**}\eta+u^{**}\right)$$

$$\eta_B = \eta + \left(\mathbf{X}^{**\prime}\mathbf{A_2}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A_2}u^{**}. \tag{2.113}$$

Hence,

$$\text{plim}\left(\eta_B - \eta\right) = \text{plim}\left(\frac{\mathbf{X}^{**\prime}\left(\frac{1}{N^2}\mathbf{J_N}\otimes\mathbf{S_T}\right)\mathbf{X}^{**}}{T}\right)^{-1} \bullet \text{plim}\frac{\mathbf{X}^{**\prime}\left(\frac{1}{N^2}\mathbf{J_N}\otimes\mathbf{S_T}\right)u^{**}}{T} \tag{2.114}$$

Assumptions (b1) and (b2) imply that

$$\text{plim}\left(\eta_B - \eta\right) = 0,$$

i.e.

$$\text{plim}\left(\eta_B\right) = \eta. \tag{2.115}$$

The estimator η_B is then consistent.

We secondly deduce its asymptotical normality from equations (2.110) and (2.111). We deduce that

$$\frac{\mathbf{X}^{**\prime}\left(\frac{1}{N^2}\mathbf{J_N}\otimes\mathbf{S_T}\right)u^{**}}{\sqrt{T}} \xrightarrow{d} N\left(0, \text{plim}\frac{\mathbf{X}^{**\prime}\left(\frac{1}{N^2}\mathbf{J_N}\otimes\mathbf{S_T}\right)\mathbf{X}^{**}}{T}\right), \tag{2.116}$$

In addition, we have

$$\sqrt{T}\left(\eta_B-\eta\right)=\left(\frac{\mathbf{X}^{**\prime}\left(\frac{\mathbf{J_N}}{N^2}\otimes\mathbf{S_T}\right)\mathbf{X}^{**}}{T}\right)^{-1}\frac{\mathbf{X}^{**\prime}\left(\frac{\mathbf{J_N}}{N^2}\otimes\mathbf{S_T}\right)u^{**}}{\sqrt{T}} \tag{2.117}$$

from which we deduce that

$$\sqrt{T}\left(\eta_B-\eta\right)\xrightarrow{d}N\left(a_2,b_2\right) \tag{2.118}$$

with

$$a_2=\text{plim}\left(\frac{\mathbf{X}^{**\prime}\left(\frac{\mathbf{J_N}}{N^2}\otimes\mathbf{S_T}\right)\mathbf{X}^{**}}{T}\right)^{-1}.\mathbf{0}=\mathbf{0} \tag{2.119}$$

and

$$b_2=\left(\text{plim}\frac{\mathbf{X}^{**\prime}\left(\frac{\mathbf{J_N}}{N^2}\otimes\mathbf{S_T}\right)\mathbf{X}^{**}}{T}\right)^{-1}\left[\text{plim}\frac{\mathbf{X}^{**\prime}\left(\frac{\mathbf{J_N}}{N^2}\otimes\mathbf{S_T}\right)\mathbf{X}^{**}}{T}\right]\left(\text{plim}\frac{\mathbf{X}^{**\prime}\left(\frac{\mathbf{J_N}}{N^2}\otimes\mathbf{S_T}\right)\mathbf{X}^{**}}{T}\right)^{-1}$$

i.e.

$$b_2=\left(\text{plim}\frac{\mathbf{X}^{**\prime}\left(\frac{\mathbf{J_N}}{N^2}\otimes\mathbf{S_T}\right)\mathbf{X}^{**}}{T}\right)^{-1}. \tag{2.120}$$

Thus,

$$\eta_B \xrightarrow{a} N\left(\eta, \frac{1}{T}\left(\operatorname{plim}\frac{\mathbf{X}^{**\prime}\left(\frac{1}{N^2}\mathbf{J_N}\otimes\mathbf{S_T}\right)\mathbf{X}^{**}}{T}\right)^{-1}\right). \tag{2.121}$$

Lastly, under the **M3**-transformation, we obtain:

$$\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)y^{**} = \left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**}\eta + \left(\mathbf{E_N}\otimes\mathbf{L_T}\right)u^{**}.$$

Once again, the mean of the new error term is equal to zero:

$$E\left[\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)u^{**}\right] = \left(\mathbf{E_N}\otimes\mathbf{L_T}\right)E\left(u^{**}\right) = 0\,. \tag{2.122}$$

The variance of $\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)u^{**}$ is obtained as:

$$\operatorname{Var}\left[\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)u^{**}\right] = V\left[\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)u^{**}\right] = d\left(\mathbf{E_N}\otimes\mathbf{L_T}\right). \tag{2.123}$$

The inverse of this matrix is $\frac{1}{d}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)$. There is no correlation between the explanatory variables and the error term. Furthermore, we have

$$\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\left\{V\left[\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)u^{**}\right]\right\}^{-1}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**} = \mathbf{X}^{**\prime}\mathbf{A}_3\left[\frac{1}{d}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\right]\mathbf{A}_3\mathbf{X}^{**}, \tag{2.124}$$

leading to

$$\text{plim}\left(\frac{\mathbf{X}^{**'}(\mathbf{E_N}\otimes\mathbf{L_T})\left\{V\left[(\mathbf{E_N}\otimes\mathbf{L_T})u^{**}\right]\right\}^{-1}(\mathbf{E_N}\otimes\mathbf{L_T})\mathbf{X}^{**}}{NT}\right)=\frac{1}{d}\text{plim}\left(\frac{\mathbf{X}^{**'}(\mathbf{E_N}\otimes\mathbf{L_T})\mathbf{X}^{**}}{NT}\right).$$

(2.125)

This last expression is positive definite according to assumption (c1).

We can now determine the asymptotical properties of η_T even though the Mann and Wald (1943) theorem cannot be used. We start by establishing its consistency.

$\eta_T=\left(\mathbf{X}^{**\prime}\mathbf{A}_3\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_3 y^{**}$ with $\mathbf{A_3}=\mathbf{E_N}\otimes\mathbf{L_T}$

$$\eta_T=\left(\mathbf{X}^{**\prime}\mathbf{A}_3\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_3\left(\mathbf{X}^{**}\eta+u^{**}\right)$$

$$\eta_T=\eta+\left(\mathbf{X}^{**\prime}\mathbf{A}_3\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\mathbf{A}_3 u^{**}. \tag{2.126}$$

Hence,

$$\text{plim}(\eta_T-\eta)=\text{plim}\left(\frac{\mathbf{X}^{**\prime}(\mathbf{E_N}\otimes\mathbf{L_T})\mathbf{X}^{**}}{T}\right)^{-1}\bullet\text{plim}\frac{\mathbf{X}^{**\prime}(\mathbf{E_N}\otimes\mathbf{L_T})u^{**}}{T} \tag{2.127}$$

The assumptions (c1) and (c2) imply that

$$\text{plim}(\eta_T-\eta)=0, \tag{2.128}$$

i.e.

$$\text{plim}\left(\eta_T\right) = \eta . \tag{2.129}$$

Thus, η_T is a consistent estimator of η.

Next, we deduce its asymptotical normality from equations (2.122) and (2.123). We find that $\frac{1}{d}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)u^{**}}{\sqrt{NT}}$ converges in distribution to a normal variable with mean 0 and variance $\frac{1}{d}\text{plim}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**}}{NT}$, one gets

$$\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)u^{**}}{\sqrt{NT}} \xrightarrow{d} N\left(0, d.\text{plim}\left(\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**}}{NT}\right)\right). \tag{2.130}$$

Furthermore, we have

$$\sqrt{NT}\left(\eta_T - \eta\right) = \left(\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**}}{NT}\right)^{-1}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)u^{**}}{\sqrt{NT}} \tag{2.131}$$

from which we deduce that

$$\sqrt{NT}\left(\eta_T - \eta\right) \xrightarrow{d} N\left(a_3, b_3\right) \tag{2.132}$$

with $a_3 = \text{plim}\left(\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**}}{NT}\right)^{-1}.\mathbf{0} = \mathbf{0}$ (2.133)

and

$$b_3 = d\left(\operatorname{plim}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**}}{NT}\right)^{-1}\left[\operatorname{plim}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**}}{NT}\right]\left(\operatorname{plim}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**}}{NT}\right)^{-1}$$

$$b_3 = d\left(\operatorname{plim}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**}}{NT}\right)^{-1}. \tag{2.134}$$

Thus,.

$$\eta_T \xrightarrow{a} N\left(\eta, \frac{d}{NT}\left(\operatorname{plim}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**}}{NT}\right)^{-1}\right). \tag{2.135}$$

We are now capable of investigating the asymptotic properties of the coefficient GLS estimator η_{GLS}. We show that it is equivalent to the covariance estimator η_C and then deduce its normality. We have:

$$\eta_{GLS} - \eta = \left(\mathbf{X}^{**\prime}\left(\boldsymbol{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\left(\boldsymbol{\Sigma}^{**}\right)^{-1}u^{**}$$

$$\sqrt{NT}\left(\eta_{GLS}-\eta\right) = \left(\frac{\mathbf{X}^{**\prime}\left(\boldsymbol{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}}{NT}\right)^{-1}\frac{\mathbf{X}^{**\prime}\left(\boldsymbol{\Sigma}^{**}\right)^{-1}u^{**}}{\sqrt{NT}}. \tag{2.136}$$

On the one hand, we have:

$$\frac{\mathbf{X}^{**\prime}\left(\boldsymbol{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}}{NT} = \frac{\mathbf{X}^{**\prime}\left(\frac{1}{\sigma_1^2}\mathbf{E_N}\otimes\mathbf{K_T} + \frac{1}{d}\mathbf{E_N}\otimes\mathbf{L_T} + \frac{1}{N^2}\mathbf{J_N}\otimes\mathbf{S_T}\right)X^{**}}{NT}, \text{ or}$$

$$\frac{\mathbf{X}^{**\prime}\left(\boldsymbol{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}}{NT}=\frac{1}{\sigma_1^2}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}}{NT}+\frac{1}{d}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**}}{NT}+\frac{1}{N}\frac{\mathbf{X}^{**\prime}\left(\mathbf{J_N}\otimes\mathbf{S_T}\right)\mathbf{X}^{**}}{N^2T}$$

where $d=\left(\iota_T^{\lambda\prime}\mathbf{D}^{-1}\iota_T^{\lambda}\right)\sigma_2^2+\sigma_1^2\to\infty$ when $T\to\infty$ Therefore, from assumption (a1), one finds

$\frac{1}{d}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)\mathbf{X}^{**}}{NT}\to\mathbf{0}$ when $N,T\to\infty$. Likewise, assumption (a2) leads up to $\frac{1}{N}\frac{\mathbf{X}^{**\prime}\left(\mathbf{J_N}\otimes\mathbf{S_T}\right)\mathbf{X}^{**}}{N^2T}\to\mathbf{0}$ when N and $T\to\infty$. Hence,

$$\operatorname{plim}\frac{\mathbf{X}^{**\prime}\left(\boldsymbol{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}}{NT}=\frac{1}{\sigma_1^2}\operatorname{plim}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}}{NT}. \tag{2.137}$$

On the other hand, we have :

$$\frac{\mathbf{X}^{**\prime}\left(\boldsymbol{\Sigma}^{**}\right)^{-1}u^{**}}{\sqrt{NT}}=\frac{\mathbf{X}^{**\prime}\left[\frac{1}{\sigma_1^2}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)+\frac{1}{d}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)+\frac{1}{N^2}\left(\mathbf{J_N}\otimes\mathbf{S_T}\right)\right]u^{**}}{\sqrt{NT}}$$

$$\frac{\mathbf{X}^{**\prime}\left(\boldsymbol{\Sigma}^{**}\right)^{-1}u^{**}}{\sqrt{NT}}=\frac{1}{\sigma_1^2}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)u^{**}}{\sqrt{NT}}+\frac{1}{d}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)u^{**}}{\sqrt{NT}}+\frac{1}{N^2}\frac{\mathbf{X}^{**\prime}\left(\mathbf{J_N}\otimes\mathbf{S_T}\right)u^{**}}{\sqrt{NT}}.$$

Under the transformations M_1 and M_2, one has:

$$\operatorname{plim}\frac{1}{d}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{L_T}\right)u^{**}}{\sqrt{NT}}=\operatorname{plim}\frac{1}{N^2}\frac{\mathbf{X}^{**\prime}\left(\mathbf{J_N}\otimes\mathbf{S_T}\right)u^{**}}{\sqrt{NT}}=0, \tag{2.138}$$

leading to

$$\operatorname{plim}\frac{\mathbf{X}^{**\prime}\left(\boldsymbol{\Sigma}^{**}\right)^{-1}u^{**}}{\sqrt{NT}}=\frac{1}{\sigma_1^2}\operatorname{plim}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)u^{**}}{\sqrt{NT}}. \tag{2.139}$$

Hence, we can write

$$\text{plim}\left[\sqrt{NT}\left(\eta_{GLS}-\eta\right)\right]=\text{plim}\left[\left(\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}}{NT}\right)^{-1}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)u^{**}}{\sqrt{NT}}\right]$$

i.e. $\text{plim}\left[\sqrt{NT}\left(\eta_{GLS}-\eta\right)\right]=\text{plim}\left[\sqrt{NT}\left(\eta_{c}-\eta\right)\right].$ (2.140)

Finally, $\sqrt{NT}\left(\eta_{GLS}-\eta\right)$ has the same limiting distribution as $\sqrt{NT}\left(\eta_{c}-\eta\right)$.

This shows the asymptotic equivalence of these two estimators. We deduce that

$$\eta_{GLS}\xrightarrow{a}N\left(\eta,\frac{1}{NT\sigma_1^2}\left(\text{plim}\frac{\mathbf{X}^{**\prime}\left(\mathbf{E_N}\otimes\mathbf{K_T}\right)\mathbf{X}^{**}}{NT}\right)^{-1}\right). \tag{2.141}$$

Thus, the GLS estimator η_{GLS} suggested under the double autocorrelation error structure has the desired asymptotic properties.

2.3. FGLS Estimation of a Double Autocorrelation Model

The variance-covariance matrix, which has so far been assumed to be known, is actually unknown, as well as all the parameters involved in its determination. Therefore, a FGLS approach is required. It is the objective of this section. The first subsection exhibits a well-known method which is presented here as a principle: the within approach. It consists in removing the time specific effect to obtain a one-way error component model where only v_{it} carries the serial correlation. Afterward, applications to AR(1) and MA(1) processes are made in the last two subsections.

2.3.1. Principle: The Within Approach

In getting an operational version the general approach retained here is the "within" one. Its principle is quite simple: we wipe out the time specific effect to obtain a one-way error component model where only v_{it} carries the serial correlation. Afterward, the resulting variance-covariance matrix of the new error term $\tilde{u}$ is derived by the use of a correction matrix of the covariance matrix of $\tilde{v}_{it}$. Hence, some BQU estimation formulas and the autocovariance functions of $\tilde{u}$ and $\tilde{v}_{it}$ as well as of u, v_{it}, and λ_t help us coming out with the estimates of some unknown parameters.

We transform the initial model by: $\left(\mathbf{I_N} - \frac{1}{N}\iota_N\iota_N'\right)\otimes\mathbf{I_T} = \mathbf{E_N}\otimes\mathbf{I_T}$ which is the "deviations from the individual mean" operator (Greene, 2008). It yields the following model:

$$\tilde{y} = \tilde{\mathbf{X}}\eta + \tilde{u}\,. \tag{2.142}$$

At the individual level, we can write

$$\tilde{u}_{it} = \left(\mu_i - \frac{1}{N}\sum_{i}^{N}\mu_i\right) + \left(v_{it} - \frac{1}{N}\sum_{i}^{N}v_{it}\right) = \tilde{\mu}_i + \tilde{v}_{it} \tag{2.143a}$$

with

$$\tilde{\mu}_i = \left(\mu_i - \frac{1}{N}\sum_{i}^{N}\mu_i\right) \tag{2.143b}$$

and

$$\tilde{v}_{it} = \left(v_{it} - \frac{1}{N}\sum_{i}^{N}v_{it}\right) \tag{2.143c}$$

Hence, one gets

$$\tilde{u} = \left(\mathbf{E_N} \otimes \mathbf{I_T}\right) u = \left(\mathbf{E_N} \otimes \mathbf{I_T}\right)\left(\mu \otimes \iota_T\right) + \left(\mathbf{E_N} \otimes \mathbf{I_T}\right)\nu = \left(\mathbf{E_N}\mu \otimes \iota_T\right) + \left(\mathbf{E_N} \otimes \mathbf{I_T}\right)\nu = \tilde{\mu} \otimes \iota_T + \tilde{\nu}.$$

In this expression, $\tilde{\mu}$ and $\tilde{\nu}$ are such that $E\left(\tilde{\mu}\tilde{\mu}'\right) = E\left(\mathbf{E_N}\mu\mu'\mathbf{E_N}\right) = \sigma_\mu^2 \mathbf{E_N}$ and $E\left(\tilde{\nu}\tilde{\nu}'\right) = E\left[\left(\mathbf{E_N} \otimes \mathbf{I_T}\right)\nu\nu'\left(\mathbf{E_N} \otimes \mathbf{I_T}\right)\right] = \left(\mathbf{E_N} \otimes \mathbf{I_T}\right)\left[\sigma_\nu^2\left(\mathbf{I_N} \otimes \mathbf{\Gamma_v}\right)\right]\left(\mathbf{E_N} \otimes \mathbf{I_T}\right) = \sigma_\nu^2\left(\mathbf{E_N} \otimes \mathbf{\Gamma_v}\right)$.

In contrast to their untransformed counterparts μ_i s, the $\tilde{\mu}_i$ s are not spherical, as a consequence of taking the deviations from the individual mean. Although this could appear as an additional and superfluous obstacle to the correlation correction, it is actually with no consequence.

The variance-covariance matrix of the disturbances vector $\tilde{u}$ is therefore given by

$$\tilde{\mathbf{\Sigma}} = E\left(\tilde{u}\tilde{u}'\right) = \sigma_\mu^2\left(\mathbf{E_N} \otimes \iota_T\iota_T'\right) + \sigma_\nu^2\left(\mathbf{E_N} \otimes \mathbf{\Gamma_v}\right). \tag{2.144}$$

Let $\mathbf{C_v}$ be a matrix such that $\mathbf{C_v}\mathbf{\Gamma_v}\mathbf{C_v}' = \mathbf{I_T}$. This matrix exists since $\mathbf{\Gamma_v} = \dfrac{1}{\sigma_\nu^2} E\left(\nu_i\nu_i'\right) \quad \forall i = 1,\ldots,N$ is a positive definite matrix. We then apply the transformation matrix $\mathbf{I_N} \otimes \mathbf{C_v}$ to our model which becomes:

$$y^* = \left(\mathbf{I_N} \otimes \mathbf{C_v}\right)\tilde{y} = \mathbf{X}^*\eta + u^* \tag{2.145}$$

where $u^* = \left(\mathbf{I_N} \otimes \mathbf{C_v}\right)\tilde{u}$ and $\mathbf{X}^* = \left(\mathbf{I_N} \otimes \mathbf{C_v}\right)\tilde{\mathbf{X}}$. The resulting within-type estimator is defined as

$$\eta_W = \left(\mathbf{X}^*\mathbf{X}^{*'}\right)^{-1}\mathbf{X}^{*'}y^*. \tag{2.146}$$

Hence, the variance-covariance matrix of the new error term u^* is obtained as follows:

$$\mathbf{\Sigma^*} = E\left(u^*u^{*'}\right) = \left(\mathbf{I_N} \otimes \mathbf{C_v}\right)\left[E\left(\tilde{u}\tilde{u}'\right)\right]\left(\mathbf{I_N} \otimes \mathbf{C_v}'\right)$$

$$\mathbf{\Sigma^*} = \left(\mathbf{I_N} \otimes \mathbf{C_v}\right)\left[\sigma_\mu^2\left(\mathbf{E_N} \otimes \iota_T\iota_T'\right) + \sigma_\nu^2\left(\mathbf{E_N} \otimes \mathbf{\Gamma_v}\right)\right]\left(\mathbf{I_N} \otimes \mathbf{C_v}'\right).$$

If we set $\iota_T^{\nu} = \mathbf{C_v}\iota_T = \begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_T\end{pmatrix}'$, $d_\nu^2 = \iota_T^{\nu\prime}\iota_T^{\nu} = \sum_{t=1}^{T}\alpha_t^2$ and $\mathbf{\bar{J}_T^v} = \frac{1}{d_\nu^2}\iota_T^{\nu}\iota_T^{\nu\prime}$, then $\mathbf{C_v}\iota_T\iota_T'\mathbf{C_v}'$ can be written as $\iota_T^{\nu}\iota_T^{\nu\prime} = d_\nu^2\mathbf{\bar{J}_T^v}$. Therefore,

$$\mathbf{\Sigma^*} = \sigma_\mu^2 d_\nu^2\left(\mathbf{E_N}\otimes\mathbf{\bar{J}_T^v}\right) + \sigma_\nu^2\left(\mathbf{E_N}\otimes\mathbf{I_T}\right). \tag{2.147}$$

Once again, we use the Wansbeek and Kapteyn (1982, 1983) trick and get the spectral decomposition of $\mathbf{\Sigma^*}$:

$$\mathbf{\Sigma^*} = \left(\sigma_\mu^2 d_\nu^2 + \sigma_\nu^2\right)\left(\mathbf{E_N}\otimes\mathbf{\bar{J}_T^v}\right) + \sigma_\nu^2\left(\mathbf{E_N}\otimes\mathbf{E_T^v}\right) \tag{2.148}$$

with $\mathbf{E_T^v} = \mathbf{I_T} - \mathbf{\bar{J}_T^v}$.

Providing the knowledge and the tractability of the matrix $\mathbf{C_v}$, the idempotent matrices $\mathbf{E_N}\otimes\mathbf{\bar{J}_T^v}$ and $\mathbf{E_N}\otimes\mathbf{E_T^v}$ lead to the following BQU estimates:

$$\begin{cases} \hat{\sigma}_\nu^2 = \dfrac{\hat{u}^{*\prime}\left(\mathbf{E_N}\otimes\mathbf{E}_\mathbf{T}^{\nu}\right)\hat{u}^*}{\text{trace}\left(\mathbf{E_N}\otimes\mathbf{E}_\mathbf{T}^{\nu}\right)} \\ \hat{\sigma}_\nu^2 + \hat{d}_\nu^2\hat{\sigma}_\mu^2 = \dfrac{\hat{u}^{*\prime}\left(\mathbf{E_N}\otimes\mathbf{\bar{J}}_\mathbf{T}^{\nu}\right)\hat{u}^*}{\text{trace}\left(\mathbf{E_N}\otimes\mathbf{\bar{J}}_\mathbf{T}^{\nu}\right)}. \end{cases} \tag{2.149}$$

$\hat{u}^*$ is the vector of OLS residuals from equation (2.145) . It is a vector of residuals obtained after applying a within transformation, and is therefore a within residuals vector itself. According to Amemiya (1971) and Baltagi (2005), such residuals yield efficient and unbiased estimates of the variances.

Distinguishing between $\hat{d}_\nu^2$ and $\hat{\sigma}_\mu^2$ can be done through the autocovariance functions of $\tilde{u}$, i.e.

$$\tilde{\gamma}(h) = E\left(\tilde{u}_{it}\tilde{u}_{i,t-h}\right) = \frac{N-1}{N}\left[\sigma_\mu^2 + E\left(\nu_{it}\nu_{i,t-h}\right)\right] = \frac{N-1}{N}\left[\sigma_\mu^2 + \gamma_\nu(h)\right] \text{ for } h = 0,1,\ldots,t-1. \quad (2.150)$$

In fact, $\tilde{\gamma}(0) = \frac{N-1}{N}\left(\sigma_\mu^2 + \sigma_\nu^2\right)$ from which we deduce that

$$\hat{\sigma}_\mu^2 = \frac{N}{N-1}\hat{\tilde{\gamma}}(0) - \hat{\sigma}_\nu^2 . \quad (2.151)$$

Note that $\hat{\tilde{\gamma}}(h) = \frac{1}{N(T-h)}\sum_{i=1}^{N}\sum_{t=h+1}^{T}\hat{\tilde{u}}_{it}\hat{\tilde{u}}_{i,t-h}$ is the empirical autocovariance function and $\hat{\tilde{u}}_{it}$ s are the OLS residuals of the within equation (2.142). Hence,

$$\hat{d}_\nu^2 = \frac{1}{\hat{\sigma}_\mu^2}\left(\frac{\hat{u}^{*'}\left(\mathbf{E_N} \otimes \overline{\mathbf{J}}_\mathbf{T}^\nu\right)\hat{u}^*}{\operatorname{trace}\left(\mathbf{E_N} \otimes \overline{\mathbf{J}}_\mathbf{T}^\nu\right)} - \hat{\sigma}_\nu^2\right). \quad (2.152)$$

Likewise, we can make use of the autocovariance function of the original overall disturbances u. We have

$$\gamma(h) = E\left(u_{it}u_{i,t-h}\right) = \sigma_\mu^2 + E\left(\nu_{it}\nu_{i,t-h}\right) + E\left(\lambda_t\lambda_{t-h}\right)$$

or

$$\gamma(h) = \frac{N}{N-1}\tilde{\gamma}(h) + \gamma_\lambda(h) \text{ for } h = 0,1,\ldots,t. \quad (2.153)$$

As a consequence, one can obtain the estimate of σ_λ^2 as

$$\hat{\sigma}_\lambda^2 = \hat{\gamma}(0) - \frac{N}{N-1}\hat{\tilde{\gamma}}(0) \text{ or } \hat{\sigma}_\lambda^2 = \hat{\gamma}(0) - \hat{\sigma}_\mu^2 - \hat{\sigma}_\nu^2 . \quad (2.154)$$

The empirical autocovariance function of u is given by $\hat{\gamma}(h) = \frac{1}{N(T-h)} \sum_{i=1}^{N} \sum_{t=h+1}^{T} \hat{u}_{it} \hat{u}_{i,t-h}$ with $\hat{u}_{it}$ denoting the OLS residuals of $y = \mathbf{X}\eta + u$.

For any other parameter remaining (for instance α_t, $t = 1,\ldots,T$), the knowledge of the exact process followed by the time-varying error terms is required.

One can go one step further and suggest transforming the model by $\sigma_\nu \mathbf{\Sigma}^{*-1/2}$. A within- -type GLS estimator of the coefficient vector, say η_{WGLS} can be obtained by applying OLS on the transformed model $y^{**} = \sigma_\nu \mathbf{\Sigma}^{*-1/2} y^*$. It is similar to the correlation-correction GLS estimator obtained under the identical time structure. In this subsection, it is quite different. We can derive the resulting typical elements. We have

$$\sigma_\nu \mathbf{\Sigma}^{*-1/2} y^* = \sum_{k=1}^{2} \frac{\sigma_\nu}{\tilde{\psi}_k^{1/2}} \tilde{\mathbf{Q}}_k y^* = \tilde{\mathbf{Q}}_1 y^* + \frac{\sigma_\nu}{\tilde{\psi}_2^{1/2}} \tilde{\mathbf{Q}}_2 y^* \tag{2.155}$$

where $\tilde{\mathbf{Q}}_\mathbf{1} = \mathbf{E_N} \otimes \mathbf{E_T^v}$, $\tilde{\mathbf{Q}}_\mathbf{2} = \mathbf{E_N} \otimes \bar{\mathbf{J}}_\mathbf{T}^\mathbf{v}$, $\tilde{\psi}_1 = \sigma_\nu^2$ and $\tilde{\psi}_2 = \sigma_\mu^2 d_\nu^2 + \sigma_\nu^2$.

Since matrices $\bar{\mathbf{J}}_\mathbf{T}^\mathbf{v}$ and $\mathbf{E_T^v}$ are extremely similar to $\bar{\mathbf{J}}_\mathbf{T}^{\boldsymbol{\alpha}}$ and $\mathbf{E}_\mathbf{T}^{\boldsymbol{\alpha}}$ respectively, aside from a constant ($\frac{1}{d_\nu^2}$ instead of $\frac{1}{d_\alpha^2}$), the expressions of $\tilde{\mathbf{Q}}_\mathbf{1} y^*$ and $\tilde{\mathbf{Q}}_2 y^*$ are respectively identical to $\mathbf{Q}_1 y^*$ and $\mathbf{Q}_2 y^*$ which have been established in subsection 1.3.3 for instance (we just need to substitute d_ν^2 to d_α^2). The typical elements of $\tilde{\mathbf{Q}}_1 y^*$ and $\tilde{\mathbf{Q}}_2 y^*$ are therefore given by $\left(y_{1t}^* - \bar{y}_{\bullet t}^*\right) - \alpha_t \left(\tilde{h}_i - \tilde{h}\right)$ and $\alpha_t \left(\tilde{h}_i - \tilde{h}\right)$ respectively. We then deduce the typical elements of $y^{**} = \sigma_\nu \mathbf{\Sigma}^{*-1/2} y^*$ as

$$y_{it}^{**} = y_{it}^* - \bar{y}_{\bullet t}^* - \tilde{\theta} \alpha_t \left(\tilde{h}_i - \tilde{h}\right) \qquad i = 1,\ldots,N \quad t = 1,\ldots,T \tag{2.156}$$

with $\tilde{\theta} = 1 - \frac{\sigma_\nu}{\tilde{\psi}_2^{1/2}}$, $\tilde{\psi}_2 = \sigma_\mu^2 d_\nu^2 + \sigma_\nu^2$, $\tilde{h}_i = \frac{1}{d_\nu^2}\sum_{t=1}^{T}\alpha_t y_{it}^* \quad \forall i = 1,\dots,N$, and $\tilde{h} = \frac{1}{N}\sum_{i=1}^{N}\tilde{h}_i$.

All the included parameters are obtainable from the above results, notably $\tilde{\theta}$, $\tilde{\psi}_2$, $\tilde{h}_i$ and $\tilde{h}$. The within-type estimator η_{WGLS} is finally derived.

2.3.2. Feasible Double AR(1) Model

We assume that $\nu_{it} = \rho_\nu \nu_{i,t-1} + e_{it}$ and $\lambda_t = \rho\lambda_{t-1} + \varepsilon_t$ where

$$\begin{cases} e_{it} \sim IIN\left(0, \sigma_e^2\right) \\ \varepsilon_t \sim IIN\left(0, \sigma_\varepsilon^2\right) \end{cases}.$$

Following the within principle, we can write the within error term u^* vector as

$$\tilde{u} = \left(\mathbf{E_N} \otimes \mathbf{I_T}\right)u = \left(\mathbf{E_N} \otimes \mathbf{I_T}\right)\left(\mu \otimes \iota_T\right) + \left(\mathbf{E_N} \otimes \mathbf{I_T}\right)\nu = \left(\mathbf{E_N}\mu \otimes \iota_T\right) + \left(\mathbf{E_N} \otimes \mathbf{I_T}\right)\nu = \tilde{\mu} \otimes \iota_T + \tilde{\nu}.$$

and then deduce its variance-covariance matrix:

$$E\left(\tilde{u}\tilde{u}'\right) = E\left[\left(\tilde{\mu} \otimes \iota_T + \tilde{\nu}\right)\left(\tilde{\mu} \otimes \iota_T + \tilde{\nu}\right)'\right] = E\left(\tilde{\mu}\tilde{\mu}'\right) \otimes \iota_T\iota_T' + E\left(\tilde{\nu}\tilde{\nu}'\right)$$

where

$E\left(\tilde{\mu}\tilde{\mu}'\right) = \sigma_\mu^2 \mathbf{E_N}$ and $E\left(\tilde{\nu}\tilde{\nu}'\right) = \sigma_\nu^2\left(\mathbf{E_N} \otimes \mathbf{\Gamma_v}\right)$ with

$$\mathbf{\Gamma_v} = \frac{1}{\sigma_\nu^2}E\left(\nu_i\nu_i'\right) \quad \forall i = 1,\dots,N.$$

Hence,

$\tilde{\boldsymbol{\Sigma}} = E\left(\tilde{u}\tilde{u}'\right) = \sigma_\mu^2\left(\mathbf{E_N} \otimes \iota_T \iota_T^{'}\right) + \sigma_\nu^2\left(\mathbf{E_N} \otimes \boldsymbol{\Gamma_v}\right)$.

Since v_{it} follows an AR(1) process of parameter ρ_v, we define the matrix $\mathbf{C_v}$ as the familiar Prais-Winsten (1954) transformation matrix:

$$\mathbf{C_v} = \begin{pmatrix} \sqrt{1-\rho_\nu^2} & 0 & 0 & \cdots & \cdots & 0 & 0 \\ -\rho_\nu & 1 & 0 & \cdots & & \cdots & \vdots \\ 0 & -\rho_\nu & 1 & 0 & & \cdots & \vdots \\ 0 & 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & -\rho_\nu & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\rho_\nu & 1 \end{pmatrix}.$$

This matrix is such that $\mathbf{C_v}\left(\sigma_\nu^2 \boldsymbol{\Gamma_v}\right)\mathbf{C}_\nu' = \left(1-\rho_\nu^2\right)\sigma_\nu^2 \mathbf{I_T} = \sigma_e^2 \mathbf{I_T}$. The resulting GLS estimator is once again given by equation (2.146), where y^*, u^* and $\mathbf{X}^*$ have kept their definitions of subsection 2.3.1:

$\eta_W = \left(\mathbf{X}^* \mathbf{X}^{*'}\right)^{-1} \mathbf{X}^{*'} y^*$.

The covariance matrix of $u^* = \left(\mathbf{I_N} \otimes \mathbf{C_v}\right)\tilde{u}$ is :

$$\boldsymbol{\Sigma}^* = \left(\sigma_\mu^2 d_\nu^2 + \sigma_e^2\right)\left(\mathbf{E_N} \otimes \bar{\mathbf{J}}_\mathbf{T}^\mathbf{v}\right) + \sigma_e^2\left(\mathbf{E_N} \otimes \mathbf{E}_\mathbf{T}^\mathbf{v}\right) \tag{2.157}$$

where,

$$\iota_T^\nu = \mathbf{C_v}\iota_T = \left(1-\rho_\nu\right)\begin{pmatrix} \sqrt{\dfrac{1+\rho_\nu}{1-\rho_\nu}} & 1 & \cdots & 1 \end{pmatrix}' = \left(1-\rho_\nu\right)\begin{pmatrix}\alpha & 1 & \cdots & 1\end{pmatrix}' = \left(1-\rho_\nu\right)\iota_T^\alpha,$$

$$d_\nu^2 = \iota_T^{\nu'}\iota_T^\nu = \left(1-\rho_\nu\right)^2 \iota_T^{\alpha'}\iota_T^\alpha = \left(1-\rho_\nu\right)^2\left[\alpha^2 + \left(T-1\right)\right] = \left(1-\rho_\nu\right)^2 d_\alpha^2.$$

Thus,

$$\mathbf{\Sigma}^{*} = \left(\sigma_\mu^2 \left(1-\rho_\nu\right)^2 d_\alpha^2 + \sigma_e^2\right)\left(\mathbf{E_N} \otimes \mathbf{\bar{J}_T^\nu}\right) + \sigma_e^2 \left(\mathbf{E_N} \otimes \mathbf{E_T^\nu}\right). \tag{2.158}$$

In order to get the estimates of the other numerous parameters involved in the model, we first need an estimation of the correlation coefficient ρ_ν. The autocovariance function of the error term $\tilde{u}$ is written as

$$\tilde{\gamma}(h) = E\left(\tilde{u}_{it}\tilde{u}_{i,t-h}\right) = \frac{N-1}{N}\left[\sigma_\mu^2 + \gamma_\nu(h)\right] = \frac{N-1}{N}\left(\sigma_\mu^2 + \rho_\nu^h \sigma_\nu^2\right), \text{ for } h = 0,1,\ldots,t. \tag{2.159}$$

We deduce from it that

$$\rho_\nu = \frac{\tilde{\gamma}(1) - \tilde{\gamma}(2)}{\tilde{\gamma}(0) - \tilde{\gamma}(1)}, \tag{2.160}$$

and therefore

$$\hat{\rho}_\nu = \frac{\hat{\tilde{\gamma}}(1) - \hat{\tilde{\gamma}}(2)}{\hat{\tilde{\gamma}}(0) - \hat{\tilde{\gamma}}(1)} \tag{2.161}$$

where $\hat{\tilde{\gamma}}(h) = \dfrac{1}{N(T-h)} \sum_{i=1}^{N} \sum_{t=h+1}^{T} \hat{\tilde{u}}_{it}\hat{\tilde{u}}_{i,t-h}$ with $\hat{\tilde{u}}_{it}$ are the OLS residuals of the within equation (2.142). It then leads to a convergent estimator of ρ_ν (see Baltagi, 2005). We deduce

$$\hat{\alpha} = \sqrt{\frac{1+\hat{\rho}_\nu}{1-\hat{\rho}_\nu}}, \tag{2.162}$$

and

$$\hat{d}_\nu^2 = \left(1-\hat{\rho}_\nu\right)^2 \left[\hat{\alpha}^2 + (T-1)\right]. \tag{2.163}$$

Furthermore, the BQU estimates are

$$\begin{cases} \hat{\sigma}_e^2 = \dfrac{\hat{u}^{*\prime}\left(\mathbf{E_N} \otimes \mathbf{E_T^v}\right)\hat{u}^*}{\operatorname{trace}\left(\mathbf{E_N} \otimes \mathbf{E_T^v}\right)} = \dfrac{\hat{u}^{*\prime}\left(\mathbf{E_N} \otimes \mathbf{E_T^v}\right)\hat{u}^*}{(N-1)(T-1)} \\ \hat{\sigma}_\mu^2 \left(1-\hat{\rho}_\nu\right)^2 \hat{d}_\alpha^2 + \hat{\sigma}_e^2 = \dfrac{\hat{u}^{*\prime}\left(\mathbf{E_N} \otimes \mathbf{\bar{J}_T^v}\right)\hat{u}^*}{\operatorname{trace}\left(\mathbf{E_N} \otimes \mathbf{\bar{J}_T^v}\right)} = \dfrac{\hat{u}^{*\prime}\left(\mathbf{E_N} \otimes \mathbf{\bar{J}_T^v}\right)\hat{u}^*}{N-1}. \end{cases} \tag{2.164}$$

The estimate of σ_e^2 is then obtained:

$$\hat{\sigma}_e^2 = \frac{\hat{u}^{*\prime}\left(\mathbf{E_N} \otimes \mathbf{E_T^v}\right)\hat{u}^*}{(N-1)(T-1)}, \tag{2.165}$$

$\hat{u}^*$ being the vector of OLS residuals from the transformed equation (2.145). We can therefore determine

$$\hat{\sigma}_\nu^2 = \frac{\hat{\sigma}_e^2}{1-\hat{\rho}_\nu^2} \tag{2.166}$$

and then

$\hat{\sigma}_\mu^2 = \dfrac{N}{N-1}\hat{\tilde{\gamma}}(0) - \hat{\sigma}_\nu^2$ as in equation (2.151).

We now need to determine $\hat{\sigma}_\lambda^2$ and $\hat{\rho}_\lambda$. The autocovariance function of the error term u is given by

$$\gamma(h) = \rho_\nu^h \sigma_\nu^2 + \sigma_\mu^2 + \rho_\lambda^h \sigma_\lambda^2 = \frac{N}{N-1}\tilde{\gamma}(h) + \rho_\lambda^h \sigma_\lambda^2. \tag{2.167}$$

It comes, likewise in equation (2.154) that

$$\hat{\sigma}_{\lambda}^{2} = \hat{\gamma}(0) - \frac{N}{N-1}\hat{\hat{\gamma}}(0) = \hat{\gamma}(0) - \hat{\sigma}_{\mu}^{2} - \hat{\sigma}_{\nu}^{2}.$$

We hence find the correlation coefficient:

$$\rho_{\lambda} = \frac{\left[\gamma(1)-\gamma(2)\right] - \frac{N}{N-1}\left[\tilde{\gamma}(1)-\tilde{\gamma}(2)\right]}{\left[\gamma(0)-\gamma(1)\right] - \frac{N}{N-1}\left[\tilde{\gamma}(0)-\tilde{\gamma}(1)\right]}, \tag{2.168}$$

leading to

$$\hat{\rho}_{\lambda} = \frac{\left[\hat{\gamma}(1)-\hat{\gamma}(2)\right] - \frac{N}{N-1}\left[\hat{\hat{\gamma}}(1)-\hat{\hat{\gamma}}(2)\right]}{\left[\hat{\gamma}(0)-\hat{\gamma}(1)\right] - \frac{N}{N-1}\left[\hat{\hat{\gamma}}(0)-\hat{\hat{\gamma}}(1)\right]}. \tag{2.169}$$

As in the previous subsection, $\hat{\gamma}(h) = \frac{1}{N(T-h)}\sum_{i=1}^{N}\sum_{t=h+1}^{T}\hat{u}_{it}\hat{u}_{it-h}$ with $\hat{u}_{it}$ denoting the OLS residuals of $y = \mathbf{X}\eta + u.$ The variance σ_{ε}^{2} is estimated by

$$\hat{\sigma}_{\varepsilon}^{2} = \left(1-\hat{\rho}_{\lambda}^{2}\right)\hat{\sigma}_{\lambda}^{2}. \tag{2.170}$$

The within estimator $\eta_{W} = \left(\mathbf{X}^{*}\mathbf{X}^{*'}\right)^{-1}\mathbf{X}^{*'}y^{*}$ as well as the within-GLS-type one $\eta_{WGLS} = \left(\mathbf{X}^{**}\mathbf{X}^{**'}\right)^{-1}\mathbf{X}^{**'}y^{**}$ and all the variance components parameters are now known or obtainable. Moreover, the "true" GLS estimator η_{GLS}, see equation (2.57), can be determined because the AR(1) parameters ρ_{ν} and ρ_{λ} have been estimated and their knowledge entitles us to build the matrices involved in the determination of $\left(\mathbf{\Sigma}^{**}\right)^{-1}$, say matrices $\mathbf{C_1}$, $\mathbf{C_2}$, $\mathbf{P}$, $\mathbf{\Lambda}$, $\mathbf{K_T}$, $\mathbf{L_T}$, $\mathbf{S}$, and $\mathbf{S_T}$. This was not possible under the general model of the previous subsection.

2.3.3. Feasible Double MA(1) Model

We assume that $v_{it} = e_{it} - \rho_v e_{i,t-1}$ and $\lambda_t = \varepsilon_t - \rho_\lambda \varepsilon_{t-1}$ where

$$\begin{cases} e_{it} \sim IIN\left(0, \sigma_e^2\right) \\ \varepsilon_t \sim IIN\left(0, \sigma_\varepsilon^2\right). \end{cases}$$

Once again, deviations from individual means lead to the model $\tilde{y} = \tilde{\mathbf{X}}\eta + \tilde{u}$ with $\tilde{u}_{it} = \tilde{\mu}_i + \tilde{v}_{it}$. Here, $\tilde{v}_{it}$ is such that :

$$\tilde{v}_{it} = \left(v_{it} - \frac{1}{N}\sum_i^N v_{it} \right) = \tilde{e}_{it} - \rho_v \tilde{e}_{i,t-1} \tag{2.171}$$

where

$$\tilde{e}_{it} = e_{it} - \frac{1}{N}\sum_i^N e_{it} \text{ and } \tilde{e}_{it} \sim IID\left(0, \mathrm{Var}\left(\tilde{e}_{it}\right) = \frac{N-1}{N}\sigma_e^2 \right) \tag{2.172}$$

As a consequence, $\tilde{v}_{it}$ follows an MA(1) process with coefficient ρ_v since $\tilde{e}_{it}$ are white noises of variance $\frac{N-1}{N}\sigma_e^2$. Note that $\tilde{\mu}_i$ is defined according to equation (2.143b). We can write, according to the within principle,

$$\tilde{u} = \left(\mathbf{E_N} \otimes \mathbf{I_T}\right) u = \tilde{\mu} \otimes \iota_T + \tilde{v}$$

where $\tilde{\mu}$ and $\tilde{v}$ are such that

$$E\left(\tilde{\mu}\tilde{\mu}'\right) = E\left(\mathbf{E_N}\mu\mu'\mathbf{E_N}\right) = \sigma_\mu^2 \mathbf{E_N}$$

and

$E(\tilde{\nu}\tilde{\nu}') = E\left[(\mathbf{E_N} \otimes \mathbf{I_T})\nu\nu'(\mathbf{E_N} \otimes \mathbf{I_T})\right] = \sigma_\nu^2(\mathbf{E_N} \otimes \mathbf{\Gamma_v})$ since $E(\nu\nu') = \sigma_\nu^2(\mathbf{I_N} \otimes \mathbf{\Gamma_v})$ with a matrix $\mathbf{\Gamma_v}$ defined as a Toeptlitz one:

$$E(\nu_i\nu_i') = \sigma_\nu^2 \begin{bmatrix} 1 & \frac{-\rho_\nu}{1+\rho_\nu^2} & 0 & 0 & \cdots & 0 \\ \frac{-\rho_\nu}{1+\rho_\nu^2} & 1 & \frac{-\rho_\nu}{1+\rho_\nu^2} & 0 & \cdots & 0 \\ 0 & \frac{-\rho_\nu}{1+\rho_\nu^2} & 1 & \frac{-\rho_\nu}{1+\rho_\nu^2} & \ddots & \vdots \\ 0 & 0 & \frac{-\rho_\nu}{1+\rho_\nu^2} & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \frac{-\rho_\nu}{1+\rho_\nu^2} \\ 0 & 0 & \cdots & 0 & \frac{-\rho_\nu}{1+\rho_\nu^2} & 1 \end{bmatrix} = \sigma_\nu^2 \mathbf{\Gamma_v},$$

i.e., $\mathbf{\Gamma_v} = \text{Toeplitz}\left(1, r_\nu = \frac{-\rho_\nu}{1+\rho_\nu^2}, 0, \ldots, 0\right)$.

The variance-covariance matrix of $\tilde{u}$ is still given by equation (2.144):

$$\tilde{\mathbf{\Sigma}} = E(\tilde{u}\tilde{u}') = \sigma_\mu^2(\mathbf{E_N} \otimes \iota_T\iota_T') + \sigma_\nu^2(\mathbf{E_N} \otimes \mathbf{\Gamma_v}).$$

Here, we set $\mathbf{C_v} = \mathbf{C_T}$, $\mathbf{C_T}$ denoting the correlation correction matrix as defined by Baltagi and Li (1994) in their orthogonalizing algorithm presented in subsection 1.4.2. We then transformed the within model by matrix $(\mathbf{I_N} \otimes \mathbf{C_v})$. The new error term $u^* = (\mathbf{I_N} \otimes \mathbf{C_v})\tilde{u}$ yields the same covariance matrix as in subsection 2.3.1. Keeping the same notations, its spectral decomposition is given by equation (2.148):

$$\mathbf{\Sigma^*} = (\sigma_\mu^2 d_\nu^2 + \sigma_\nu^2)(\mathbf{E_N} \otimes \bar{\mathbf{J}}_\mathbf{T}^\mathbf{v}) + \sigma_\nu^2(\mathbf{E_N} \otimes \mathbf{E}_\mathbf{T}^\mathbf{v}).$$

Because of the moving-average nature of the process, linear estimation of the correlation parameter ρ_v is not easily obtainable. The BQU estimators are not going to be investigated since they do not prove useful in this case.

The autocovariance function of the within error term $\tilde{u}_{it} = \left(\mathbf{E_N} \otimes \mathbf{I_T}\right)u_{it} = \tilde{\mu}_i + \tilde{\nu}_{it}$ is given by equation (2.150):

$$\tilde{\gamma}(h) = E\left(\tilde{u}_{it}\tilde{u}_{i,t-h}\right) = \frac{N-1}{N}\left[\sigma_\mu^2 + \gamma_\nu(h)\right].$$

As a consequence,

$$\hat{\sigma}_\mu^2 = \frac{N}{N-1}\hat{\tilde{\gamma}}(j) \text{ for some } j \geq 2 \tag{2.173}$$

and

$$\hat{\sigma}_\nu^2 = \hat{\gamma}_\nu(0) = \frac{N}{N-1}\hat{\tilde{\gamma}}(0) - \hat{\sigma}_\mu^2 \tag{2.174}$$

where $\hat{\tilde{\gamma}}(s) = \dfrac{1}{N(T-s)}\displaystyle\sum_{i=1}^{N}\sum_{t=s+1}^{T}\hat{\tilde{u}}_{it}\hat{\tilde{u}}_{i,t-s}$, $s = 1,\ldots,t-1$ is the empirical autocovariance function and $\hat{\tilde{u}}_{it}$ s are the OLS residuals of the within equation (2.142). We also deduce for some $j \geq 2$,

$$\hat{r}_\nu = \frac{N}{(N-1)\hat{\sigma}_\nu^2}\left[\hat{\tilde{\gamma}}(1) - \hat{\tilde{\gamma}}(j)\right]. \tag{2.175}$$

Since one of our goals is to perform the within-type estimator η_{WGLS}, we apply the Baltagi and Li (1994) matrix $\mathbf{C_T}$ to the data (for instance to the within transformed dependent vector $\tilde{y}$). Moreover, $\mathbf{C_T}$ will be applied to the vector of constants to get estimates of the α_t s. We have, in the straight line of Baltagi and Li (1994), as in subsection 1.4.2, the following steps:

Step 1: Compute $y_{i1}^{*} = \frac{y_{i1}}{\sqrt{\hat{g}_{\nu,1}}}$ and $y_{it}^{*} = \frac{y_{it} - \frac{\hat{r}_{\nu} y_{i,t-1}^{*}}{\sqrt{\hat{g}_{\nu,t-1}}}}{\sqrt{\hat{g}_{\nu,t}}}$ for $t = 2,\ldots,T$ where $g_{\nu,1} = 1$ and $\hat{g}_{\nu,t} = 1 - \frac{\hat{r}_{\nu}^{2}}{\hat{g}_{\nu,t-1}}$ for $t = 2,\ldots,T$.

Step 2: Compute $y^{**} = \sigma_{\nu} \mathbf{\Sigma}^{*-1/2} y^{*}$ using the fact that $\iota_{T}^{\nu} = \mathbf{C_T} \iota_{T} = \left(\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_T\right)'$. The estimates of the α_t s are obtained as: $\hat{\alpha}_1 = 1$ and $\hat{\alpha}_t = \frac{1 - \frac{\hat{r}_{\nu}}{\sqrt{\hat{g}_{\nu,t-1}}}}{\sqrt{\hat{g}_{\nu,t}}}$ for $t = 2,\ldots,T$.

We then deduce

$$\hat{d}_{\nu}^{2} = \sum_{t=1}^{T} \hat{\alpha}_{t}^{2}. \tag{2.176}$$

The autocovariance function $\gamma(h) = E\left(u_{it} u_{it-h}\right) = \gamma_{\nu}(h) + \sigma_{\mu}^{2} + \gamma_{\lambda}(h)$ of the initial composite error term u and its empirical counterpart $\hat{\gamma}(h) = \frac{1}{N(T-h)} \sum_{i=1}^{N} \sum_{t=h+1}^{T} \hat{u}_{it} \hat{u}_{i,t-h}$, ($\hat{u}_{it}$ being the OLS residuals of the initial two-way model) permit the estimation of σ_{λ}^{2} and r_{λ}. The former is given by equation (2.154):

$$\hat{\sigma}_{\lambda}^{2} = \hat{\gamma}(0) - \hat{\sigma}_{\nu}^{2} - \hat{\sigma}_{\mu}^{2}.$$

The later is then deduced :

$$\hat{r}_{\lambda} = \frac{\hat{\gamma}(1) - \hat{r}_{\nu} \hat{\sigma}_{\nu}^{2} - \hat{\sigma}_{\mu}^{2}}{\hat{\sigma}_{\lambda}^{2}}. \tag{2.177}$$

The within estimator η_W and the within-type estimator η_{WGLS} are now obtainable. However, the "true" GLS estimator η_{GLS} can be estimated, providing the MA(1) parameters ρ_ν and ρ_λ are known, especially under the conditions $\Delta_\lambda = 1-4\hat{r}_\lambda^2 \geq 0$ and $\Delta_\nu = 1-4\hat{r}_\nu^2 \geq 0$. In other words, the estimates $\hat{r}_\lambda$ and $\hat{r}_\nu$ should both lie inside the open interval $\left(-\frac{1}{2},\frac{1}{2}\right)$ before one can get a direct estimator of $\eta_{GLS} = \left(\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}\mathbf{X}^{**}\right)^{-1}\mathbf{X}^{**\prime}\left(\mathbf{\Sigma}^{**}\right)^{-1}y^{**}$.

Finally, part 2 appears as a theoretical treatment of a complex but realistic correlation structure in the two-way error component model: the double autocorrelation situation. This part has dealt with some parsimonious models, especially the AR(1) and MA(1) ones, as well as with the general framework. Through a precise formula of the variance-covariance matrix of the errors, we have derived the GLS estimator and its asymptotic properties. Part 2 ends by suggesting some feasible counterparts to the models presented.

CONCLUDING REMARKS

This dissertation has investigated, through two separate parts, the issue of serial correlation in a two-way random effect model, when all time-varying components of the error term (i.e. both λ_t and v_{it}) were exhibiting either the same or different correlation patterns.

In part 1 we assumed that both λ_t and v_{it} were following the same process, with the same parameters. Under this somewhat strong assumption, some GLS transformations have been obtained after the spectral decomposition of the variance-covariance matrix of the composite error term. In the straight line of the Baltagi and Li (1992) treatment of the one-way correlated random effect model, we have derived a detailed two-stage GLS procedure that corrects for the autocorrelation in the disturbances. We have been able to show out the relations between the original data and their correlation-corrected counterparts. Furthermore, we have shown that the regression method could also be presented as a one-step GLS procedure, when the correlation correction matrix was "simple" enough. All these results have firstly been developed, as a theoretical application, in the context of some simple and well-known processes such as the AR(1) and MA(1) ones. Then, an identical autocorrelation model with an undefined pattern has been investigated, showing how general our GLS approach is. In order to definitely establish its applicability, a feasible version has been suggested in the last section of part 1. Whatever the process followed by both λ_t and v_{it} might be, the estimates of the involved parameters were obtainable. Again, the particular cases in which the correlated disturbances were distributed according to the AR(1) and MA(1) processes have been considered from a FGLS perspective. It came out that the moving-averages patterns were more involved and more complex than the autoregressive ones. Using the appropriate correction matrix provided by the literature has been recommended. Thus the ease of the implementation of the FGLS estimation procedure appeared.

Part 2 has tackled the most realistic structure: we allowed λ_t and v_{it} to follow different time series. Even when their processes were of the same type, the parameters were no longer identical. This double autocorrelation random effect model has been analyzed through three submodels. The first one considered that both λ_t and v_{it} were coming from an AR(1) process but of different parameters ρ_λ and ρ_v respectively. The second model assumed two different MA(1) processes. Lastly, an unrestricted double autocorrelation pattern has been investigated, allowing λ_t and v_{it} to follow any time series, independently. Solving a general framework often requires some invariant regularity in the diversity of the potential cases. This was achieved here through the variance-covariance matrix of the composite error. We have shown that, whatever the correlation structure might be, this matrix could be written under a precise formula. Thereafter, the inverse of our matrix of second order moments was derived, leading to a GLS estimator of the coefficient vector. Since the GLS transformations could not be detailed as in part 1 because of the more general framework retained here, general properties of our GLS estimator have instead been analyzed. Firstly, it appeared as an estimate obtained by pooling three uncorrelated estimators derived from the matrices involved in the inverse matrix formula proposed earlier. Actually, it is their weighted average. These underlying estimators have been labeled as covariance, between-time, and within-time estimators. Secondly, we proved that, under certain assumptions, the GLS and these three estimators of the coefficient vector are all asymptotically normally distributed. In particular, we established that the GLS estimator and the covariance one are asymptotically equivalent. Part 2 has ended its treatment of the general double autocorrelation issue by assessing the applicability of the GLS approach it had developed. We have suggested a within transformation matrix known as the "deviations from the individual mean" operator, resulting into a one-way error component model (it wipes out the time specificities). The remaining serial correlation was then carried by the error term v_{it}. Only one correlation-correction matrix was therefore needed. Combined to the use of the autocovariance functions of the true composite errors and of their transformed counterparts, these corrections entitled us to recover the variance estimates of the time-specific disturbances as well as of the estimates of other numerous parameters involved. As in part 1, these methods have been illustrated by the cases of AR(1) and

MA(1) double autocorrelation processes. The knowledge of the appropriate inversion matrix (provided by the literature) is the only requirement for the implementation of the FGLS methods.

This thesis finally appears as an essay on the autocorrelation issue in the context of the two-way random effect model. It has focused on deriving a GLS feasible method, in several cases, and from a theoretical perspective only. It opens the door to numerous extensions. Particularly, one is an empirical investigation of the results obtained in this study. Some Monte-Carlo studies will be welcome to assess the small-sample properties of our estimates, and even to compare them with estimators deduced from other methods. This will be the subject of future research that we would like to conduct. The implications on testing serial correlation, the case of unbalanced panel data, the double autocorrelation structure when spatial correlation is allowed, the heterogeneity in the time-series parameters across units, the presence of time-varying covariates and the use of other estimation procedures such as ML or Instrumental Variables (hereafter IV) methods, etc., all are potential investigation topics.

REFERENCES

Amemiya, T. (1971), "The Estimation of the Variances in a Variance-Components Model", *International Economic Review* 12, 1-13.

Balestra, P. and M. Nerlove (1966), "Pooling Cross-Section and Time-Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas", *Econometrica* 34, 585-612.

Balestra, P. (1973), "Best Quadratic Unbiased Estimators of the Variance-Covariance Matrix in Normal Regression", *Journal of Econometrics*, 2, 17-28.

Balestra, P. (1980), "A Note on the Exact Transformation Associated with the First Order Moving Average Process", *Journal of Econometrics* 14, 381-394.

Baltagi, B.H. (1981), "Pooling: An Experimental Study of Alternative Testing and Estimation Procedures in a Two-Way Error Components Model", *Journal of Econometrics* 17, 21-49.

Baltagi, B.H. (2005), *Econometric Analysis of Panel Data*, John Wiley and Sons, Third Edition.

Baltagi, B.H. (2006), *Panel Data Econometrics: Theoretical Contributions and Empirical Applications*, Elsevier Science, Amsterdam.

Baltagi, B.H., G. Bresson and A. Pirotte, (2005), "Adaptive Estimation of Heteroskedastic Error Component Models", *Econometric Reviews*, 24, pp.39-58.

Baltagi, B.H., C. Kao and L. Liu (2007), "Asymptotic Properties of Estimators for the Linear Panel Regression Model with Individual Effects and Serially Correlated Errors: The Case of Stationary and Nonstationary Regressors and Residuals", *Centre for Policy Research*, Working Paper 93, Syracuse, New-York.

Baltagi, B.H. and W. Kramer (1997), "A Simple Linear Trend Model with Error Components", *Econometric Theory*, 13, 463-463.

Baltagi, B.H. and Q. Li (1991), “A Transformation That Will Circumvent the Problem of Autocorrelation in Error Component Model”, *Journal of Econometrics*, 48, 385-393.

Baltagi B.H. and Q. Li (1992a) “Prediction in the One-way Error Component Model with Serial Correlation”, *Journal of Forecasting*, 11, 561-567.

Baltagi B.H. and Q. Li (1992b), “An Approximate Transformation for the Error Component Model with MA(q) Disturbances; Problem 92.4.3”, *Econometric Theory* 8, 582-583.

Baltagi, B.H. and Q. Li (1994), “Estimating Error Component Models with General MA(q) Disturbances”, *Econometric Theory*, 10, 396-408.

Baltagi, B.H. and Q. Li (1997), “Monte Carlo Results on Pure and Pretest Estimators of an Error Component Model with Autocorrelated Disturbances”, *Annales d’Economie et de Statistiques* 48, 69-82.

Bawn, K. and F. Rosenbluth (2006), “Short versus Long Coalitions: Electoral Accountability and the Size of the Public Sector”, *American Journal of Political Science* 50, 251-265.

Bharghava, A., L. Franzini and W. Narendranathan (1982), “Serial Correlation and Fixed Effects”, *Review of Economic Studies* 49, 533-549.

Cameron, A.C. and P.K. Trivedi (2005), *Microeconometrics: Methods and Applications*, First Edition, Cambridge University Press, New York.

Enders, W., A. Sachsida and T. Sandler (2006), “The Impact of Transnational Terrorism on U.S. Foreign Direct Investment” *Political Research Quarterly* 59, 517-531.

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

Galbraith, J.W. and V. Zinde-Walsh (1995), "Transforming the Error Component Model for Estimation with general ARMA Disturbances", *Journal of Econometrics* 66, 349-355.

Gordin, M. (1969), "The Central Limit Theorem for Stationary Processes", *Soviet Mathematical Dokl*, 10, 1174-1176.

Greene, W.H. (2008), *Econometric Analysis*, Prentice Hall, New Jersey.

Hadri, K. (2000), "Testing for Stationarity in Heterogeneous Panel Data", *Econometrics Journal* 3, 148-161.

Hadri, K. and R. Larsson, (2005), "Testing For Stationarity In Heterogeneous Panel Data Where The Time Dimension Is Finite," *Econometrics Journal* 8, 55-69.

Hamilton, J. (1994), *Time Series Analysis*, Princeton University Press, Princeton.

Hecock, R.D. (2006), "Electoral Competition, Globalization, and Subnational Education Spending in Mexico, 1999-2004", *American Journal of Political Science* 50, 950-961.

Hoeffler, A. (2002), "The Augmented Solow Model and the African Growth Debate", *Oxford Bulletin of Economics and Statistics* 64, 135-158.

Hsiao, C. (2003), *Analysis of Panel Data*, Cambridge University Press, Cambridge.

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lutkepohl and T.C. Lee (1985), *The Theory and Practice of Econometrics*, John Wiley and Sons, New York.

Kao, C. and J. Emerson, (2002a), "Testing for Structural Change of a Time Trend Regression in Panel Data: Part 1", *Journal of Propagations in Probability and Statistics*, 2, 57-75.

Kao, C. and J. Emerson, (2002a), "Testing for Structural Change of a Time Trend Regression in Panel Data: Part 2", *Journal of Propagations in Probability and Statistics*, 2, 207-250.

Karlsson, S. and J. Skoglund (2004), “Maximum-likelihood Based Inference in the Two-Way Random Effects Model with Serially Correlated Time Effects”, *Empirical Economics*, 29, 79-88.

Kelleher, C.A. and S.W. Yackee (2006), “Who’s Whispering in Your Ear? The Influence of Third Parties over State Agency Decisions”, *Political Research Quarterly* 59, 629-643.

Kiefer, N.M. (1980), “Estimation of Fixed Effects Models for Times Series of Cross Sections with Arbitrary Intertemporal Covariance”, *Journal of Econometrics* 14, 195-202.

Klevmarken, N.A. (1989), “Panel Studies: What Can We Learn From Them? Introduction”, *European Economic Review* 33, 523-529.

Lee, C.S. (2005), “Income Inequality, Democracy and Public Sector Size”, *American Sociological Review* 70, 158-181.

Lillard, L.A. and R.J. Willis (1978) ‘Dynamic Aspects of Earning Mobility’, *Econometrica*, 46, 985-1012.

Lillard, L.A. and Y. Weiss (1979), “Components of Variation in Panel Earnings Data: American Scientists 1960-1970”, *Econometrica* 47, 437-454.

MaCurdy, T.A. (1982), “The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis”, *Journal of Econometrics* 18, 83-114.

Maddala, G.S. (1971a), “Generalized Least Squares with an Estimated Variance Covariance Matrix”, *Econometrica* 39, 23-34.

Maddala, G.S. (1971b), “The Use of Variance Components Models in Pooling Cross-Section and Time Series Data”, *Econometrica* 39, 341-358.

Magnus, J.R. and A.D. Woodland (1988), "On the Maximum Likelihood Estimation of Multivariate Regression Models Containing Serially Correlated Error Components", *International Economic Review* 29, 707-725.

N'Gbo, A.G.M. (1994), "L'Efficacité Productive des Scop Françaises: Estimation et Simulation à partir d'une Frontière de Production Stochastique", *Revue Economique*, 45, 115-128.

Nerlove, M. (1971), "A Note on Error Components Models", *Econometrica* 39, 383-396.

Pesaran, M.H. (1974), "Exact Maximum Likelihood Estimation of a Regression Equation with a First Order Moving Average Errors", *The Review of Economic Studies*, Vol. XL, 529-538.

Phillips, P.C.B. and H. Moon (1999), "Linear Regression Limit Theory for Nonstationary Panel Data", *Econometrica* 67, 1057-1111.

Phillips, P.C.B. and H. Moon (2000), "Nonstationary Panel Data Analysis: An Overview of some recent developments, *Econometric Reviews* 19, 263-286.

Prais, S.J. and C.B. Winsten (1954), "Trend Estimators and Serial Correlation", Unpublished Cowles Commission Discussion Paper: Stat No. 383, Chicago.

Prucha, I.R. (1984), "On the Asymptotic Efficiency of Feasible Aitken Estimators for Seemingly Unrelated Regression Models with Error Components", *Econometrica* 52, 203-207.

Revankar, N.S. (1979), "Error Component Models with Serial Correlated Time Effects", *Journal of the Indian Statistical Association* 17, 137-160.

Schmidt, P. (1983), "A Note on a Fixed Effect Model with Arbitrary Interpersonal Covariance", *Journal of Econometrics* 22, 391-393.

South, S.J., K. Crowder and E. Chavez (2005), "Migration and Spatial Assimilation Among US Latinos: Classical Versus Segmented Trajectories", *Demography* 42, 497-521.

Stasavage, D. (2005), “Democracy and Education Spending in Africa”, *American Journal of Political Science* 49, 343-358.

Swamy P.A.V.B. and S.S. Arora (1972), “The Exact Finite Sample Properties of the Estimators of Coefficients in the Error Components Regression Models”, *Econometrica* 40, 261-275.

Theil, H. (1971), *Principles of Econometrics*, John Wiley and Sons, New York.

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

Wansbeek, T.J. and A. Kapteyn (1982), “A Simple Way to obtain the Spectral Decomposition of Variance Components Models for Balanced Data”, *Communications in Statistics* A11, 2105-2112.

Wansbeek, T.J. and A. Kapteyn (1983), “A Note on Spectral Decomposition and Maximum Likelihood Estimation of ANOVA Models with Balanced Data”, *Statistics and Probability Letters* 1, 213-15.

Ziliak, J.P. (1997), “Efficient Estimation with Panel Data when Instruments are Predetermined: An Empirical Comparison of Moment-Condition Estimators”, *Journal of Business and Economic Statistics*, 15, 419-431.

TABLE OF CONTENTS

ABSTRACT iii
RESUME iv
ACKNOWLEDGEMENTS v
BRIEF CONTENTS vii
INTRODUCTION 1
Part 1: A GLS ESTIMATION OF THE TWO-WAY RANDOM EFFECT MODEL WITH IDENTICAL TIME STRUCTURES 10
1.1. Identical AR(1) Error Structure 14
1.1.1. Specification of the Model 14
1.1.2. Variance-Covariance of the Transformed Errors and Its Spectral Decomposition ... 15
1.1.3. GLS Transformation 18
1.1.4. BQU Estimates 28
1.2. Identical MA(1) Error Structure 28
1.2.1. Specification of the Model 29
1.2.2. Variance-Covariance of the Transformed Errors and Its Spectral Decomposition ... 30
1.2.3. GLS Transformation 34
1.2.4. BQU Estimates 45
1.3. General Model with Identical Error Structure 46
1.3.1. Specification of the Model 46
1.3.2. Variance-Covariance of the Transformed Errors and Its Spectral Decomposition ... 48
1.3.3. GLS Transformation 50
1.3.4. BQU Estimates 58
1.4. FGLS Estimation of an Identical Time Structure Model 59
1.4.1. AR(1) Model 59
1.4.2. MA(1) Model 62
1.4.3. General Model 68
1.4.4. The Swamy and Arora Method and other FGLS Approaches 70

Part 2: A GLS ESTIMATION OF THE TWO-WAY RANDOM EFFECT MODEL WITH DIFFERENT TIME STRUCTURES 74
2.1. Autocorrelation and General Expression of the Variance-Covariance Matrix 75
2.1.1. Double AR(1) Error Structure 75
2.1.2. Double MA(1) Error Structure....80
2.1.3. General Double Error Structure 81
2.2. GLS Estimation and Properties of the Estimators....84
2.2.1. Inverse of the Variance- Covariance Matrix of the Overall Errors 84
2.2.2. Interpretation of the GLS Estimator 92
2.2.3. Asymptotic Properties of the GLS Estimator 104
2.3. FGLS Estimation of a Double Autocorrelation Model 118
2.3.1. Principle: The Within Approach....119
2.3.2. Feasible Double AR(1) Model....124
2.3.3. Feasible Double MA(1) Model....129
CONCLUDING REMARKS....135
REFERENCES 138
TABLE OF CONTENTS....144