Time Invariant Variables and Panel Data Models : A Generalised Frisch-Waugh Theorem and its Implications

Jaya Krishnakumar

No 2004.01

Cahiers du département d'économétrie
Faculté des sciences économiques et sociales
Université de Genève

Décembre 2003

Département d'économétrie
Université de Genève, 40 Boulevard du Pont-d'Arve, CH -1211 Genève 4
http://www.unige.ch/ses/metri/

Time Invariant Variables and Panel Data Models: A Generalised Frisch-Waugh Theorem and its Implications

by

Jaya Krishnakumar[1]
University of Geneva

August 2004

Abstract

Mundlak (1978) showed that when individual effects are correlated with the explanatory variables in an error component (EC) model, the GLS estimator is given by the *within*. In this paper we bring out some additional interesting properties of the *within* estimator in Mundlak's model and go on to show that the *within* estimator remains valid in an extended EC model *with time invariant variables* and correlated specific effects. Adding an auxiliary regression to take account of possible correlation between the explanatory variables and the individual effects, we find that not only the elegant results obtained by Mundlak but also the above mentioned special features carry over to the extended case with interesting interpretations. We obtain these results using a generalised version of the Frisch-Waugh theorem, stated and proved in the paper. Finally, for both the EC models with and without time invariant variables we have shown that the estimates of the coefficients of the auxiliary variables can also be arrived at by following a two-step procedure.

Keywords: panel data, error components, correlated effects, *within* estimator.

JEL Classification codes : C23

[1]Corresponding author: Department of Econometrics, University of Geneva, UNI-MAIL, 40, Bd. du Pont d'Arve, CH-1211, GENEVA 4, Switzerland. Email: jaya.krishnakumar@metri.unige.ch. The author would like to thank an anonymous referee, Pietro Balestra, Badi Baltagi, Geert Dhaene and the participants at the 2004 panel data conference for useful comments and suggestions.

1 Introduction

This paper is concerned with the issue of time invariant variables in panel data models. We try to look into an 'old' problem from a new angle or rather in an extended framework. It is well-known that when time invariant variables are present, the *within* transformation wipes them out and hence does not yield estimates for their coefficients. However they can be retrieved by regressing the means of the *within* residuals on these variables (see Hsiao (1986) e.g.). Hausman and Taylor (1981) provide an efficient instrumental variable estimation of the model when the individual effects are correlated with some of the time invariant variables and some of the X's. Valid instruments are given by the other time invariant and time varying variables in the equation.

Suppose we consider the case in which the individual effects are correlated with all the explanatory variables. The earliest article dealing with this issue in panel data literature is that of Mundlak (1978) where the author looked at the error component model with individual effects and possible correlation of these individual effects with the explanatory variables (or rather their means). He showed that upon taking this correlation into account the resulting GLS estimator is the *within*. Thus the question of choice between the *within* and the random effects estimators was both "arbitrary and unnecessary" according to Mundlak.

Note that the question of correlation arises only in the random effects framework as the fixed effects are by definition non-stochastic and hence cannot be linked to the explanatory variables. We point this out because Mundlak's conclusion may often be interpreted wrongly that the fixed effects *model* is the correct specification. What Mundlak's study shows is that the estimator is the same (the *within*) whether the effects are considered fixed or random.

Now what happens to Mundlak's results when time invariant variables are present in the model? Do they still carry over? Or do they have to be modified? If so in what way? Are there any neat interpretations as in Mundlak's case? This paper is an attempt to answer these questions and go beyond them interpreting the results in a way that they keep the same elegance as in Mundlak's model.

The answers to the above questions follow smoothly if we go through a theorem extending the Frisch-Waugh result from the classical regression to the generalised regression. Thus we start in Section 2 by stating a generalised version of Frisch-Waugh theorem and giving its proof. In this section we also explain the important characteristic of this new theorem which makes it more than just a straightforward extension of the classical Frisch-Waugh theorem and point out in what way it is different from a similar theorem derived by Fiebig, Bartels and Krämer (1996). The next section briefly recalls Mundlak's case and puts the notation in place. Section 4 brings out some interesting features of Mundlak's model which enable the known results. Section 5 presents the model with time invariant variables and discusses it from the point of view of correlated effects. Relationships between the different estimators are established and compared with the previous case. Finally we conclude with a summary of our main results.

2 The Generalised Frisch-Waugh Theorem

Theorem 1

In the generalised regression model:

$$y = X_1\beta_1 + X_2\beta_2 + u \tag{1}$$

with $E(u) = 0$ and $V(u) = V$, positive definite, non-scalar, the GLS estimator of a subvector of the coefficients, say β_2, can be written as

$$\hat{\beta}_{2,gls} = (R_2'V^{-1}R_2)^{-1}R_2'V^{-1}R_1 \tag{2}$$

where

$$R_1 = y - X_1(X_1'V^{-1}X_1)^{-1}X_1'V^{-1}y$$
$$R_2 = X_2 - X_1(X_1'V^{-1}X_1)^{-1}X_1'V^{-1}X_2$$

The proof of this theorem is given in Appendix 1.

Let us note an important property in the above formula for $\hat{\beta}_{2,gls}$ in that it represents a generalised regression of the residuals of GLS of y on X_1 on the GLS residuals of X_2 on X_1 with the *same initial* V as the variance covariance matrix in all the three regressions. An additional feature is that one can even replace R_1 by y in (2) and our result still holds (as in the classical case).

Fiebig, Bartels and Krämer (1996) arrive at the GLS estimator $\hat{\beta}_2$ through a different route (applying M_1 to (1) and then (true) GLS on the transformed model). They also show that using a (mistaken) original V for their transformed model leads to a different estimator (which they call the pseudo GLS) and derive conditions under which pseudo GLS is equal to true GLS. Baltagi (2000) refers to Fiebig, Bartels and Krämer (1996) in the context of his Example 3 while mentioning a special case examined by Baltagi and Krämer (1995) in which pseudo GLS equals true GLS.

Our $\hat{\beta}_2$ is the same as their true GLS on the initial model but obtained through different transformations and has an interesting interpretation in terms of (GLS) residuals of auxiliary regressions as in the classical Frisch-Waugh result.

Corollary 1:

If in model (1) above we further have orthogonality between X_1 and X_2 in the metric V^{-1} i.e. if

$$X_1'V^{-1}X_2 = 0$$

then

$$\begin{aligned}
\hat{\beta}_{1,gls} &= (X_1'V^{-1}X_1)^{-1}X_1'V^{-1}y \\
\hat{\beta}_{2,gls} &= (X_2'V^{-1}X_2)^{-1}X_2'V^{-1}y
\end{aligned}$$

3 The Known Case: Mundlak's Model

Let us briefly recall Mundlak's result for a panel data model with only individual effects. The model is:

$$y = X\,\beta + (I_N \otimes \iota_T)u + w \qquad (3)$$

We have the usual assumptions $E(u) = 0$, $V(u) = \sigma_u^2 I_N$, $E(w) = 0$, $V(w) = \sigma_w^2 I_{NT}$ and independence between u and w. Thus denoting $\varepsilon = (I_N \otimes \iota_T)u + w$

we have $V(\varepsilon) \equiv \Sigma = \lambda_1 P + \lambda_2 Q$ with $\lambda_1 = \sigma_w^2 + T\sigma_u^2$, $\lambda_2 = \sigma_w^2$, $P = \frac{1}{T}(I_N \otimes \iota_T \iota_T')$ and $Q = I_{NT} - P$. Q is the well-known *within* transformation matrix.

When there is correlation between the individual effects u and the explanatory variables X, it is postulated using:

$$u = \bar{X}\gamma + v \tag{4}$$

where $\bar{X} = \frac{1}{T}(I_N \otimes \iota_T')X$ and $v \sim (0, \sigma_v^2 I_N)$. Here one should leave out the previous assumption $E(u) = 0$. Substituting (4) into (3) we get

$$y = X\,\beta + (I_N \otimes \iota_T)\bar{X}\gamma + (I_N \otimes \iota_T)\; v + w \tag{5}$$

Applying GLS to (5) Mundlak showed that

$$\begin{aligned} \hat{\beta}_{gls} &= \hat{\beta}_w \\ \hat{\gamma}_{gls} &= \hat{\beta}_b - \hat{\beta}_w \end{aligned} \tag{6}$$

where $\hat{\beta}_w$ and $\hat{\beta}_b$ are the *within* and the *between* estimators respectively.

Hence Mundlak concluded that the *within* estimator should be the preferred option in all circumstances.

4 Some interesting features

In this section we highlight some additional results for the above model which have interesting interpretations and lead us to the more general case of a model with time invariant variables.

Why *within* is GLS for β

Let us first look at the GLS estimation of the full model (5). Note that the additional term $(I_N \otimes \iota_T)\bar{X}$ can be written as PX.

Thus the augmented model becomes

$$y = X\beta + PX\gamma + \tilde{\varepsilon} \tag{7}$$

with $\tilde{\varepsilon} = (I_N \otimes \iota_T)v + w$ and $V(\tilde{\varepsilon}) \equiv \tilde{\Sigma} = \tilde{\lambda_1} P + \tilde{\lambda_2} Q$ with $\tilde{\lambda_1} = \sigma_w^2 + T\sigma_v^2$, $\tilde{\lambda_2} = \sigma_w^2$.

Splitting X into its two orthogonal components QX and PX let us rewrite the above equation as

$$y = QX\beta + PX(\beta + \gamma) + \tilde{\varepsilon} \tag{8}$$

Noticing that QX and PX are such that $X'Q\tilde{\Sigma}^{-1}PX = 0$ we can apply Corollary 1 to obtain

$$\begin{aligned} \hat{\beta}_{gls} &= (X'Q\tilde{\Sigma}^{-1}QX)^{-1}X'Q\tilde{\Sigma}^{-1}y \\ &= (X'QX)^{-1}X'Qy = \hat{\beta}_w \end{aligned}$$

and

$$\begin{aligned} \widehat{(\beta + \gamma)}_{gls} &= (X'P\tilde{\Sigma}^{-1}PX)^{-1}X'P\tilde{\Sigma}^{-1}y \\ &= (X'PX)^{-1}X'Py = \hat{\beta}_b \end{aligned}$$

Thus we get back Mundlak's result (6):

$$\hat{\gamma}_{gls} = \hat{\beta}_b - \hat{\beta}_w$$

This result can be further explained intuitively. Looking at model (7) we have X and PX as explanatory variables. Thus the coefficient of X i.e. β measures the effect of X on y holding that of PX constant. Holding the effect of PX constant means that we are only actually measuring the effect of QX on y with β . Hence it is not surprising that we get $\hat{\beta}_w$ as the GLS estimator on the full model (7). However in the case of γ, it is the effect of PX holding X constant. Since X contains PX and QX as its components, we are only holding the QX component constant letting the PX component vary along with the PX which is explicitly in the equation whose combined effect is β and γ. Now the effect of PX on y is estimated by none other than the *between* estimator. So we have $\widehat{(\beta + \gamma)}_{gls} = \hat{\beta}_b$ i.e. result (6) once again.

Within also equals an IV for β

As the X's are correlated with the error term $\varepsilon = (I_N \otimes \iota_T)u + w$, the GLS estimator will be biased but one could use instrumental variables. Various IV sets have been proposed in the literature (cf. Hausman and Taylor (1981), Amemiya and McCurdy (1986) and Breusch *et al.* (1989)) and relative efficiency discussed at length. We will not go into that discussion here. Instead we point out that choosing the simple valid instrument QX also leads to the *within* estimator. Indeed, premultiplying equation (3) by $X'Q$ we have

$$X'Qy = X'QX\beta + X'Q\varepsilon \tag{9}$$

and applying GLS we get the *within* estimator

$$\hat{\beta}_{IV} = (X'QX)^{-1}X'Qy = \hat{\beta}_w \tag{10}$$

GLS for γ is equivalent to a two-step procedure

As far as γ is concerned, we observe that GLS on the full model is equivalent to the following two step procedure:

Step 1: *Within* regression on model (3)

Step 2: Regression of *within* estimates of individual effects on $\bar{X}$ which gives $\hat{\gamma}$.

The individual effects estimates can be written as

$$\begin{aligned} u^* &= \frac{1}{T}(I_N \otimes \iota'_T)\left[I_{NT} - X(X'QX)^{-1}X'Q\right]y \\ &= u + \frac{1}{T}(I_N \otimes \iota'_T)\left[I_{NT} - X(X'QX)^{-1}X'Q\right]\varepsilon \end{aligned}$$

substituting (3) for y. Thus we have

$$u^* = \bar{X}\gamma + v + \frac{1}{T}(I_N \otimes \iota'_T)\left[I_{NT} - X(X'QX)^{-1}X'Q\right]\varepsilon$$

or

$$u^* = \bar{X}\gamma + w^* \tag{11}$$

denoting $w^* = v + \frac{1}{T}(I_N \otimes \iota'_T)\left[I_{NT} - X(X'QX)^{-1}X'Q\right]\varepsilon$.

It is interesting to verify that

$$V(w^*)\bar{X} = \bar{X}A$$

with A non-singular and hence we can apply OLS on (11). Thus we obtain

$$\begin{aligned} \hat{\gamma} &= (\bar{X}'\bar{X})^{-1}\bar{X}'u^* \\ &= (\bar{X}'\bar{X})^{-1}\bar{X}'(\bar{y} - \bar{X}\hat{\beta}_w) \\ &= \hat{\beta}_b - \hat{\beta}_w \end{aligned} \tag{12}$$

which is the same result as (6).

The above simple results not only show that we are able to arrive at the same estimator by various ways but also provide useful insight into the interesting connections working within the same model due to the special decomposition of the variance covariance structure of EC models.

5 Extension to the case with time invariant variables

Now let us see what happens when time invariant variables come in. The new model is

$$y = X\beta + (I_N \otimes \iota_T)Z\delta + (I_N \otimes \iota_T)u + w = X\beta + CZ\delta + \varepsilon \tag{13}$$

where Z is a $N \times p$ matrix of observations on p time-invariant variables relating to the N individuals and $C \equiv I_N \otimes \iota_T$.

5.1 Without correlated effects

Applying Theorem 1 on (13) and simplifying (see Appendix 2) one can obtain that $\hat{\beta}_{gls}$ is a weighted combination of the '*within*' and '*between*' (in fact an '*extended between*', see below) estimators i.e.

$$\hat{\beta}_{gls} = W_1\hat{\beta}_{eb} + W_2\hat{\beta}_w \tag{14}$$

where $\hat{\beta}_w$ is the same as before,

$$\hat{\beta}_{eb} = \left[X'\left(\frac{1}{T\lambda_1}CM_ZC'\right)X\right]^{-1}X'\left(\frac{1}{T\lambda_1}CM_ZC'\right)y \tag{15}$$

and W_1, W_2 are weight matrices defined in Appendix 2.

The estimator given in (15) is in fact the *between* estimator of β for an EC model with time invariant variables (as the *between* transformation changes the X's into their means but keeps the Z's as such; hence we have the transformation M_Z in between to eliminate the $Z's$). We call it the '*extended between*' estimator and abbreviate it as '*eb*'.

Turning to $\hat{\delta}_{gls}$, Theorem 1 implies

$$\hat{\delta}_{gls} = (F_2'\Sigma^{-1}F_2)^{-1}F_2'\Sigma^{-1}F_1 \tag{16}$$

where F_2 are residuals of CZ on X and F_1 are residuals of y on X. However for the former we should in fact be talking of residuals of Z on $\bar{X}$ as X is time varying and Z is time invariant. This means that in order to obtain $\hat{\delta}$ we should be regressing the individual *means* of residuals of y on X on those of Z on $\bar{X}$. Redefining F_1 and F_2 in this way and simplifying the expressions, we get

$$\begin{aligned} \hat{\delta}_{gls} &= (Z'M_{\bar{X}}Z)^{-1}Z'M_{\bar{X}}\frac{1}{T}(I_N \otimes \iota_T')(I_{NT} - X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1})y \\ &= (Z'M_{\bar{X}}Z)^{-1}Z'M_{\bar{X}}\bar{y} \end{aligned} \tag{17}$$

5.2 With correlated effects

Now suppose that the individual effects are correlated with the X's and the Z's. The above estimators become inconsistent. Writing the auxiliary regression as

$$u = \bar{X}\gamma + Z\phi + v \tag{18}$$

and substituting u in (13) we get

$$\begin{aligned} y &= X\beta + CZ\delta + (I_N \otimes \iota_T)\bar{X}\gamma + (I_N \otimes \iota_T)Z\delta + (I_N \otimes \iota_T)v + w \\ &= X\beta + CZ(\delta + \phi) + PX\gamma + (I_N \otimes \iota_T)v + w \end{aligned} \tag{19}$$

Within is still GLS for β

If we apply Theorem 1 to our model (19) above then we have the result that $\hat{\beta}_{gls}$ on (19) is the same as $\hat{\beta}_{gls}$ on the following model:

$$R_1 = R_2\beta + \varepsilon$$

where

$$R_1 = y - \tilde{Z}(\tilde{Z}'\Sigma^{-1}\tilde{Z})^{-1}\tilde{Z}'\Sigma^{-1}y$$

and

$$R_2 = X - \tilde{Z}(\tilde{Z}'\Sigma^{-1}\tilde{Z})^{-1}\tilde{Z}'\Sigma^{-1}X$$

with

$$\tilde{Z} = [(I_N \otimes \iota_T)Z \quad PX] = (I_N \otimes \iota_T)[Z \quad \bar{X}] = C\bar{Z}$$

In other words,

$$\hat{\beta}_{gls} = (R_2'\Sigma^{-1}R_2)^{-1}R_2'\Sigma^{-1}R_1 \tag{20}$$

Once again making use of some special matrix results, one can show (see Appendix 3) that $\hat{\beta}_{gls} = \hat{\beta}_w$ for the augmented EC model with time invariant variables and correlated effects.

How can we intuitively explain this? Again it is straightforward if we write the model as

$$y = QX\beta + PX(\beta + \gamma) + CZ(\delta + \phi) + \varepsilon$$

and notice that QX is orthogonal to both PX and CZ in the metric Σ^{-1}. Corollary 1 above tells us that $\hat{\beta}_{gls}$ is given by

$$\hat{\beta}_{gls} = (X'Q\Sigma^{-1}QX)^{-1}X'Q\Sigma^{-1}Qy = (X'QX)^{-1}X'Qy = \hat{\beta}_w$$

Within also equals an IV for β

Now it is easy to see that instrumenting X by QX in the new model (13) also leads to the *within* estimator for β coinciding with the GLS in the extended model. Of course transforming the model by the instrument matrix eliminates the time invariant variables just like the *within* transformation does. The coefficient estimates of the latter can always be retrieved in a second step by regressing the residual means on these same variables (see below).

GLS for γ is an *'extended' between - within*

From the above intuitive reasoning we can also deduce that the parameters γ, δ and ϕ should be estimated together whereas we could leave out β as QX is orthogonal to both PX and Z in the metric Σ^{-1}.

Writing

$$\theta = \begin{bmatrix} (\delta + \phi) \\ (\beta + \gamma) \end{bmatrix}$$

we have by Theorem 1

$$\hat{\theta} = \begin{bmatrix} \widehat{(\delta+\phi)} \\ \widehat{(\beta+\gamma)} \end{bmatrix} = (\tilde{Z}'\Sigma^{-1}\tilde{Z})^{-1}\tilde{Z}'\Sigma^{-1}y$$

Separate solutions for the two components of $\hat{\theta}$ can be obtained as yet another application of the same theorem:

$$\widehat{(\delta+\phi)} = (Z'M_{\bar{X}}Z)^{-1}Z'M_{\bar{X}}\bar{y}$$

$$\widehat{(\beta+\gamma)} = (\bar{X}'M_Z\bar{X})^{-1}\bar{X}'M_Z\bar{y}$$

where $\widehat{(\beta+\gamma)}$ can be recognised as the '*extended between*' estimator[2]. Once again the estimator of γ in the extended model is derived as the difference between the '*extended between*' and the *within* estimators:

$$\hat{\gamma}_{gls} = \widehat{(\beta+\gamma)} - \hat{\beta} = \hat{\beta}_{eb} - \hat{\beta}_w \tag{21}$$

<u>GLS for γ is again a two-step procedure</u>

The above result on $\hat{\gamma}_{gls}$ leads to another interpretation similar to that of result (12) obtained in the model without time invariant variables. We have

$$\begin{aligned}
\hat{\gamma}_{gls} &= (\bar{X}'M_Z\bar{X})^{-1}\bar{X}'M_Z\bar{y} - (X'QX)^{-1}X'Qy \\
&= (X'C'M_ZC'X)^{-1}X'CM_ZC'y - (X'QX)^{-1}X'Qy \\
&= (X'C'M_ZC'X)^{-1}X'CM_ZC'y - (X'C'M_ZC'X)^{-1}X'CM_ZC'X(X'QX)^{-1}X'Qy \\
&= (X'C'M_ZC'X)^{-1}X'CM_ZC'(I_{NT} - X(X'QX)^{-1}X'Q)y \\
&= (X'C'M_ZC'X)^{-1}X'CM_ZC'\hat{u}^*
\end{aligned}$$

which implies that $\hat{\gamma}_{gls}$ can be obtained by a two step procedure as follows:

Step 1: *Within* regression of model (13)

Step 2: Regressing the *within* residual means on the residuals of the means of the X's on Z.

Now a few additional remarks. Note the formula for $\widehat{(\delta+\phi)}$ is exactly the same as the one for $\hat{\delta}$ in the 'old' model (17) and this can be understood

[2] Here the *'between'* model is $\bar{y} = \bar{X}(\beta+\gamma) + Z(\delta+\phi) + \bar{\varepsilon}$.

if we look into the effect captured by this coefficient. In model (13) δ is the effect of Z on y holding that of X constant i.e. holding constant the effect of both the components QX and PX and the combined coefficient $(\delta + \phi)$ retains the same interpretation in the augmented model (19) too. However a major difference here is that one can only estimate the sum $(\delta + \phi)$ and cannot identify δ and ϕ separately. This is logical as both the coefficients are in a way trying to measure the same effect. Thus the inclusion of $Z\phi$ in the auxiliary regression (18) is redundant. The expression for $(\delta + \phi)$ can in fact be obtained by regressing $\hat{u}$ on $\bar{X}$ and Z. Thus, practically speaking δ and γ can be retrieved by regressing *within* residual means on $\bar{X}$ and Z.

Let us also mention that the Hausman specification tests are carried out in the same manner whether time invariant variables are present or not and the absence of correlation can be tested using any one of the differences $\hat{\beta}_b - \hat{\beta}_w$, $\hat{\beta}_{gls} - \hat{\beta}_w$, $\hat{\beta}_{gls} - \hat{\beta}_b$ or $\hat{\beta}_{gls} - \hat{\beta}_{ols}$ as shown in Hausman and Taylor (1981).

If we assume non-zero correlation between explanatory variables and the *combined* disturbance term (the individual effects *and* the genuine disturbance terms), for instance in the context of a simultaneity problem, then the whole framework changes, *within* estimator is no longer consistent and only instrumental variables procedures such as the generalised 2SLS (G2SLS) or the error component 2SLS (EC2SLS) are valid (see e.g. Krishnakumar (1988), Baltagi (1981)).

6 Concluding remarks

In this paper we have shown that Mundlak's approach and the *within* estimator remain perfectly valid even in an extended EC model with time invariant variables. Adding an auxiliary regression to take account of possible correlation between the explanatory variables and the individual effects one finds that the elegant results obtained by Mundlak (1978) as well as some additional interesting ones can be derived in the extended case too. These results are established by the application of a generalised version of the Frisch-Waugh theorem also presented in the paper. Further, it is shown that for both the models with and without time invariant variables, the estimates of the coefficients of the auxiliary variables can also be obtained by a two-step estimation procedure.

References

Amemiya, T. and T.E. McCurdy (1986), “Instrumental Variables Estimation of an Error Componenets Model”, *Econometrica,* 54, 869-881.

Baltagi, B.H. (1981), “Simultaneous Equations with Error Components”, *Journal of Econometrics,* 17, 189–200.

Baltagi, B.H. (2000), “Further Evidence on the Efficiency of Least Squares in Regression Models”, pp. 279-291 in Krishnakumar, J. and E. Ronchetti (eds.) (2000), *Panel Data Econometrics: Future Directions*, North Holland Elsevier, Amsterdam.

Baltagi, B.H. and W. Krämer (1995), “A mixed error component model”, Problem 95.1.4, *Econometric Theory,* 11, 192–193.

Breusch, T.S., G.E. Mizon and P. Schmidt (1989), “Efficient Estimation Using Panel Data, *Econometrica* 57, 695-700.

Cornwell, C., P. Schmidt and D. Wyhowski (1992), “Simultaneous Equations and Panel Data”, *Journal of Econometrics*, 51, 151–181.

Fiebig, D.G., R. Bartels and W. Krämer (1996), “The Frisch-Waugh Theorem and Generalised Least Squares Estimators”, *Econometric Reviews*, 15, 431–444.

Hausman, J.A. and W.E. Taylor (1981), “Identification in Linear Simultaneous Equations Models with Covariance Restrictions: An Instrumental Variables Interpretation”, *Econometrica* 51, 5, 1527–1549.

Hsiao, C. (1986), *Analysis of Panel Data*, Econometric Society Monographs, Cambridge University Press.

Krishnakumar, J. (1988): *Estimation of Simultaneous Equation Models with Error Components Structure.* Springer–Verlag, Berlin–Heidelberg.

Mundlak, Y. (1978), "On the Pooling of Time Series and Cross-Section Data", *Econometrica*, 46, 69-85.

Appendix 1

Proof of Theorem 1 :

Let us transform the original model (1) by $V^{-1/2}$ to get

$$y^* = X_1^*\beta_1 + X_2^*\beta_2 + u^*$$

where $y^* = V^{-1/2}y$, $X_1^* = V^{-1/2}X_1$, $X_2^* = V^{-1/2}X_2$ and $u^* = V^{-1/2}u$.

Now $V(u^*) = I_{NT}$ and hence we can apply the classical Frisch-Waugh theorem to obtain

$$\hat{\beta}_2 = (R_2^{*'}R_2^*)^{-1}R_2^{*'}R_1^*$$

where

$$R_1^* = y^* - X_1^*(X_1^{*'}X_1^*)^{-1}X_1^{*'}y^*$$
$$R_2^* = X_2^* - X_1^*(X_1^{*'}X_1^*)^{-1}X_1^{*'}X_2^*$$

Substituting the starred variables in terms of the non-starred ones and rearranging we get

$$\begin{aligned}
\hat{\beta}_2 &= [X_2'(V^{-1} - V^{-1}X_1(X_1'V^{-1}X_1)^{-1}X_1'V^{-1})X_2]^{-1} \\
&\quad X_2'(V^{-1} - V^{-1}X_1(X_1'V^{-1}X_1)^{-1}X_1'V^{-1})y \\
&= [X_2'V^{-1}(I_{NT} - X_1(X_1'V^{-1}X_1)^{-1}X_1'V^{-1})X_2]^{-1} \\
&\quad X_2'V^{-1}(I_{NT} - X_1(X_1'V^{-1}X_1)^{-1}X_1'V^{-1})y \\
&= [X_2'(I_{NT} - V^{-1}X_1(X_1'V^{-1}X_1)^{-1}X_1')V^{-1}(I_{NT} - X_1(X_1'V^{-1}X_1)^{-1}X_1'V^{-1})X_2]^{-1} \\
&\quad X_2'(I_{NT} - V^{-1}X_1(X_1'V^{-1}X_1)^{-1}X_1')V^{-1}(I_{NT} - X_1(X_1'V^{-1}X_1)^{-1}X_1'V^{-1})y \\
&= (R_2'V^{-1}R_2)^{-1}R_2'V^{-1}R_1 \qquad \text{q.e.d}
\end{aligned}$$

Appendix 2

Applying Theorem 1 on (13) yields:

$$\hat{\beta}_{gls} = (E_2'\Sigma^{-1}E_2)^{-1}E_2'\Sigma^{-1}E_1 \tag{22}$$

where

$$E_1 = y - CZ(Z'C'\Sigma^{-1}CZ)^{-1}Z'C'\Sigma^{-1}y = (I_{NT} - \frac{1}{T}CZ(Z'Z)^{-1}Z'C')y$$

and

$$E_2 = X - CZ(Z'C'\Sigma^{-1}CZ)^{-1}Z'C'\Sigma^{-1}X = (I_{NT} - \frac{1}{T}CZ(Z'Z)^{-1}Z'C')X$$

using $C'\Sigma^{-1}C = \frac{1}{\lambda_1}TI_N$ and writing $\bar{X} = \frac{1}{T}C'X$.

Since $PC = C$, $QC = 0$, $CC' = TP$ and $C'C = TI_N$ one can see that

$$\begin{aligned}
\hat{\beta}_{gls} &= (E_2'\Sigma^{-1}E_2)^{-1}E_2'\Sigma^{-1}E_1 \\
&= \left[X'\left(\frac{\lambda_2}{T\lambda_1}CM_ZC' + Q\right)X\right]^{-1}X'\left(\frac{\lambda_2}{T\lambda_1}CM_ZC' + Q\right)y \\
&= W_1\hat{\beta}_{eb} + W_2\hat{\beta}_w
\end{aligned}$$

where

$$\begin{aligned}
M_Z &= I_N - Z(Z'Z)^{-1}Z' \\
W_1 &= \left[X'\left(\frac{\lambda_2}{T\lambda_1}CM_ZC' + Q\right)X\right]^{-1}X'\left(\frac{\lambda_2}{T\lambda_1}CM_ZC'\right)X \\
W_2 &= \left[X'\left(\frac{\lambda_2}{T\lambda_1}CM_ZC' + Q\right)X\right]^{-1}X'QX
\end{aligned}$$

and

$$\hat{\beta}_{eb} = \left[X'\left(\frac{1}{T\lambda_1}CM_ZC'\right)X\right]^{-1}X'\left(\frac{1}{T\lambda_1}CM_ZC'\right)y$$

Appendix 3

We have from (20)

$$\hat{\beta}_{gls} = (R_2'\Sigma^{-1}R_2)^{-1}R_2'\Sigma^{-1}R_1 \tag{23}$$

Let us examine R_1 and R_2. We can write them as $R_1 = \tilde{M}y$ and $R_2 = \tilde{M}X$ where $\tilde{M} = I_N - \tilde{Z}(\tilde{Z}'\Sigma^{-1}\tilde{Z})^{-1}\tilde{Z}'\Sigma^{-1}$.

Noting once again that $PC = C$, $QC = 0$, $CC' = TP$, $C'C = TI_N$, $C'\Sigma^{-1}C = \frac{1}{\lambda_1}TI_N, \tilde{Z}'\Sigma^{-1} = \frac{1}{\lambda_1}TZ'C'$ and $\tilde{Z}'\Sigma^{-1}\tilde{Z} = \frac{T}{\lambda_1}Z'Z$, one can show that $\tilde{M} = I_{NT} - \frac{1}{T}\,C\bar{Z}(\bar{Z}'\bar{Z})^{-1}\bar{Z}'C' = I_{NT} - \frac{1}{T}\,CP_{\bar{Z}}C'$.

Further due to the partitioned nature of $\bar{Z}$ we also know that

$$P_{\bar{Z}} = P_{\bar{X}} + M_{\bar{X}}Z(Z'M_{\bar{X}}Z)^{-1}Z'M_{\bar{X}}$$

Hence

$$\tilde{M} = I_{NT} - \frac{1}{T}C\left(P_{\bar{X}} + M_{\bar{X}}Z(Z'M_{\bar{X}}Z)^{-1}Z'M_{\bar{X}}\right)C'$$

and

$$\tilde{M}X = (I_{NT} - \frac{1}{T}C\bar{X}) = (I_{NT} - P)X = QX$$

as $P_{\bar{X}}C'X = TP_{\bar{X}}\bar{X} = T\bar{X} = C'X$ and $M_{\bar{X}}C'X = 0$. Therefore

$$R_2'\Sigma^{-1}R_2 = X'\tilde{M}\Sigma^{-1}\tilde{M}X = \frac{1}{\lambda_2}X'QX$$

Similarly one can verify that

$$R_2'\Sigma^{-1}R_1 = X'\tilde{M}\Sigma^{-1}\tilde{M}y = \frac{1}{\lambda_2}X'Qy$$

Thus

$$\hat{\beta}_{gls} = (X'QX)^{-1}X'Qy = \hat{\beta}_w$$