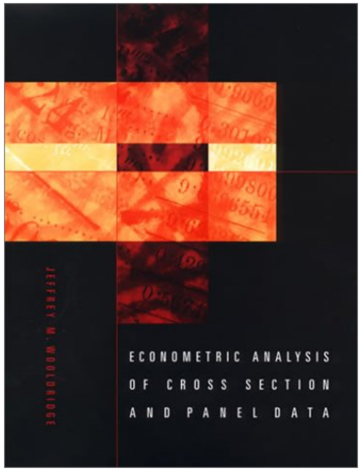
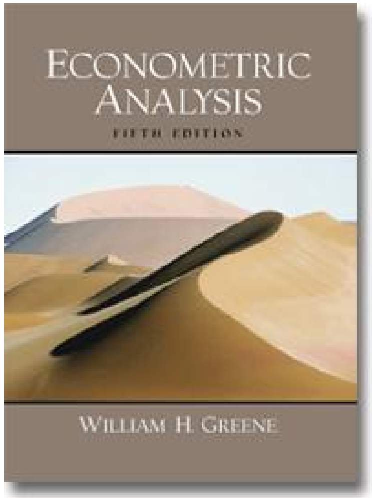
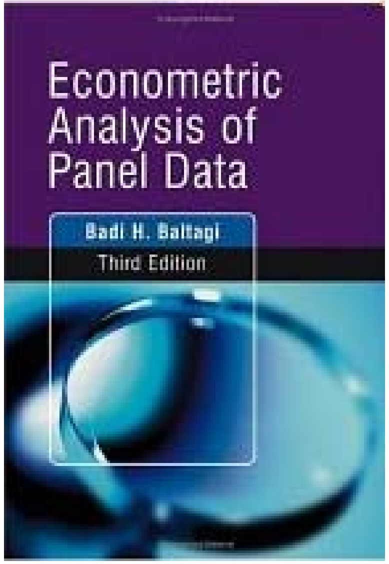
Econometric Analysis of Panel Data



William Greene
Department of Economics
Stern School of Business





Econometric Analysis of Panel Data

5. Random Effects Linear Model

The Random Effects Model

- The random effects model

  $y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + c_i + \varepsilon_{it}$, observation for person i at time t

  $\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + c_i\mathbf{i} + \boldsymbol{\varepsilon}_i$, T_i observations in group i

  $= \mathbf{X}_i\boldsymbol{\beta} + \mathbf{c}_i + \boldsymbol{\varepsilon}_i$, note $\mathbf{c}_i = (c_i, c_i, \ldots, c_i)'$

  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{c} + \boldsymbol{\varepsilon}$, $\Sigma_{i=1}^{N} T_i$ observations in the sample

  $c = (\mathbf{c}_1', \mathbf{c}_2', \ldots \mathbf{c}_N')'$, $\Sigma_{i=1}^{N} T_i$ by 1 vector

- c_i is uncorrelated with $\mathbf{x}_{it}$ for all t;
  - $E[c_i \mid \mathbf{X}_i] = 0$
  - $E[\varepsilon_{it} \mid \mathbf{X}_i, c_i] = 0$

Error Components Model

Generalized Regression Model

$$y_{it} = \mathbf{x}_{it}'\mathbf{b} + \varepsilon_{it} + u_i$$

$$E[\varepsilon_{it} \mid \mathbf{X}_i] = 0$$

$$E[\varepsilon_{it}^2 \mid \mathbf{X}_i] = \sigma_\varepsilon^2$$

$$E[u_i \mid \mathbf{X}_i] = 0$$

$$E[u_i^2 \mid \mathbf{X}_i] = \sigma_u^2$$

$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{\varepsilon}_i + u_i\mathbf{i}$ for T_i observations

$$\text{Var}[\boldsymbol{\varepsilon}_i + u_i\mathbf{i}] = \begin{bmatrix} \sigma_\varepsilon^2 + \sigma_u^2 & \sigma_u^2 & \cdots & \sigma_u^2 \\ \sigma_u^2 & \sigma_\varepsilon^2 + \sigma_u^2 & \cdots & \sigma_u^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_u^2 & \sigma_u^2 & \cdots & \sigma_\varepsilon^2 + \sigma_u^2 \end{bmatrix}$$

Notation

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_N \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_N \end{bmatrix} + \begin{bmatrix} u_1\mathbf{i}_1 \\ u_2\mathbf{i}_2 \\ \vdots \\ u_N\mathbf{i}_N \end{bmatrix} \quad \begin{matrix} T_1 \text{ observations} \\ T_2 \text{ observations} \\ \vdots \\ T_N \text{ observations} \end{matrix}$$

$$= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} + \mathbf{u} \quad \Sigma_{\mathbf{i=1}}^{\mathbf{N}} T_i \text{ observations}$$

$$= \mathbf{X}\boldsymbol{\beta} + \mathbf{w}$$

In all that follows, except where explicitly noted, X, $\mathbf{X}_i$ and $\mathbf{x}_{it}'$ contain a constant term as the first element.

To avoid notational clutter, in those cases, $\mathbf{x}_{it}'$ etc. will simply denote the counterpart without the constant term.

Use of the symbol K for the number of variables will thus be context specific but will usually include the constant term.

Notation

$$\text{Var}[\boldsymbol{\varepsilon}_i + u_i\mathbf{i}] = \begin{bmatrix} \sigma_\varepsilon^2 + \sigma_u^2 & \sigma_u^2 & \cdots & \sigma_u^2 \\ \sigma_u^2 & \sigma_\varepsilon^2 + \sigma_u^2 & \cdots & \sigma_u^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_u^2 & \sigma_u^2 & \cdots & \sigma_\varepsilon^2 + \sigma_u^2 \end{bmatrix}$$

$$= \sigma_\varepsilon^2 \mathbf{I}_{\mathbf{T_i}} + \sigma_u^2 \mathbf{ii}' \quad T_i \times T_i$$

$$= \sigma_\varepsilon^2 \mathbf{I}_{\mathbf{T_i}} + \sigma_u^2 \mathbf{ii}'$$

$$= \boldsymbol{\Omega}_i$$

$$\text{Var}[\mathbf{w} \mid \mathbf{X}] = \begin{bmatrix} \boldsymbol{\Omega}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Omega}_N \end{bmatrix}$$

(Note these differ only in the dimension T_i)

Regression Model-Orthogonality

$$\text{plim}\frac{1}{\#\,\text{observations}}\mathbf{X'w}=\mathbf{0}$$

$$\text{plim}\frac{1}{\Sigma_{i=1}^{N}T_i}\Sigma_{\mathbf{i=1}}^{\mathbf{N}}\mathbf{X}_{\mathbf{i}}'\mathbf{w}_{\mathbf{i}}=\text{plim}\frac{1}{\Sigma_{i=1}^{N}T_i}\Sigma_{\mathbf{i=1}}^{\mathbf{N}}\mathbf{X}_{\mathbf{i}}'(\boldsymbol{\varepsilon}_{\mathbf{i}}+u_i\mathbf{i})=\mathbf{0}$$

$$\text{plim}\frac{1}{\Sigma_{i=1}^{N}T_i}\left[\Sigma_{\mathbf{i=1}}^{\mathbf{N}}T_i\frac{\mathbf{X}_{\mathbf{i}}'\boldsymbol{\varepsilon}_{\mathbf{i}}}{T_i}+\Sigma_{\mathbf{i=1}}^{\mathbf{N}}T_iu_i\frac{\mathbf{X}_{\mathbf{i}}'\mathbf{i}_{\mathbf{i}}}{T_i}\right]$$

$$\text{plim}\left[\Sigma_{\mathbf{i=1}}^{\mathbf{N}}f_i\frac{\mathbf{X}_{\mathbf{i}}'\boldsymbol{\varepsilon}_{\mathbf{i}}}{T_i}+\Sigma_{\mathbf{i=1}}^{\mathbf{N}}f_i\frac{\mathbf{X}_{\mathbf{i}}'\mathbf{i}_{\mathbf{i}}}{T_i}u_i\right],\ 0<f_i=\frac{T_i}{\Sigma_{i=1}^{N}T_i}<1$$

$$\text{plim}\left[\Sigma_{\mathbf{i=1}}^{\mathbf{N}}f_i\frac{\mathbf{X}_{\mathbf{i}}'\boldsymbol{\varepsilon}_{\mathbf{i}}}{T_i}+\Sigma_{\mathbf{i=1}}^{\mathbf{N}}f_i\overline{\mathbf{x}}_iu_i\right]=\mathbf{0}$$

Convergence of Moments

$$\frac{\mathbf{X'X}}{\Sigma_{i=1}^{N} T} = \Sigma_{i=1}^{N} f_i \frac{\mathbf{X_i'X_i}}{T_i} = \text{a weighted sum of individual moment matrices}$$

$$\frac{\mathbf{X'\Omega X}}{\Sigma_{i=1}^{N} T} = \Sigma_{i=1}^{N} f_i \frac{\mathbf{X_i'\Omega_i X_i}}{T_i} = \text{a weighted sum of individual moment matrices}$$

$$= \sigma_\varepsilon^{\mathbf{2}} \Sigma_{i=1}^{N} f_i \frac{\mathbf{X_i'X_i}}{T_i} + \sigma_u^{\mathbf{2}} \Sigma_{i=1}^{N} f_i \bar{\mathbf{x}}_{\mathbf{i}} \bar{\mathbf{x}}_{\mathbf{i}}'$$

Note asymptotics are with respect to N. Each matrix $\frac{\mathbf{X_i'X_i}}{T_i}$ is the moments for the T_i observations. Should be 'well behaved' in micro level data. The average of N such matrices should be likewise.

T or T_i is assumed to be fixed (and small).

Random vs. Fixed Effects

- Random Effects
  - Small number of parameters
  - Efficient estimation
  - Objectionable orthogonality assumption ($c_i \perp \mathbf{X}_i$)
- Fixed Effects
  - Robust – generally consistent
  - Large number of parameters

Ordinary Least Squares

- Standard results for OLS in a GR model
  - Consistent
  - Unbiased
  - Inefficient
- True Variance

$$\text{Var}[\mathbf{b} \mid \mathbf{X}] = \frac{\mathbf{1}}{\sum_{i=1}^{N} T_i} \left[\frac{\mathbf{X}'\mathbf{X}}{\sum_{i=1}^{N} T_i} \right]^{-1} \frac{\mathbf{X}'\boldsymbol{\Omega}\mathbf{X}}{\sum_{i=1}^{N} T_i} \left[\frac{\mathbf{X}'\mathbf{X}}{\sum_{i=1}^{N} T_i} \right]^{-1}$$

$$\rightarrow \mathbf{0} \times \rightarrow \mathbf{Q}^{-1} \times \rightarrow \mathbf{Q}^{*} \times \rightarrow \mathbf{Q}^{-1}$$

$\rightarrow$ **0** as $N \rightarrow \infty$ with our convergence assumptions

Estimating the Variance for OLS

$$\text{Var}[\mathbf{b}\,|\,\mathbf{X}] = \frac{\mathbf{1}}{\sum_{i=1}^{N} T_i}\left[\frac{\mathbf{X}'\mathbf{X}}{\sum_{i=1}^{N} T_i}\right]^{-1}\left(\frac{\mathbf{X}'\mathbf{\Omega}\mathbf{X}}{\sum_{i=1}^{N} T_i}\right)\left[\frac{\mathbf{X}'\mathbf{X}}{\sum_{i=1}^{N} T_i}\right]^{-1}$$

$$\frac{\mathbf{X}'\mathbf{\Omega}\mathbf{X}}{\sum_{i=1}^{N} T} = \sum_{i=1}^{N} f_i \frac{\mathbf{X}_i'\mathbf{\Omega}_i\mathbf{X}_i}{T_i}, \text{ where } = \mathbf{\Omega}_i = E[\mathbf{w}_i\mathbf{w}_i'\,|\,\mathbf{X}_i]$$

In the spirit of the White estimator, use

$$\frac{\widehat{\mathbf{X}'\mathbf{\Omega}\mathbf{X}}}{\sum_{i=1}^{N} T} = \sum_{i=1}^{N} f_i \frac{\mathbf{X}_i'\hat{\mathbf{w}}_i\hat{\mathbf{w}}_i'\mathbf{X}_i}{T_i}, \quad \hat{\mathbf{w}}_i = \mathbf{y}_i - \mathbf{X}_i\mathbf{b}$$

Hypothesis tests are then based on Wald statistics.

THIS IS THE 'CLUSTER' ESTIMATOR

Mechanics

$$\text{Est.Var}[\mathbf{b} \mid \mathbf{X}] = \left[\mathbf{X}'\mathbf{X}\right]^{-1} \left(\Sigma_{i=1}^{N} \mathbf{X}_i' \hat{\mathbf{w}}_i \hat{\mathbf{w}}_i' \mathbf{X}_i\right) \left[\mathbf{X}'\mathbf{X}\right]^{-1}$$

$\hat{\mathbf{w}}_i'$ = set of T_i OLS residuals for individual i.

$\mathbf{X}_i$ = T_ixK data on exogenous variable for individual i.

$\mathbf{X}_i'\hat{\mathbf{w}}_i$ = K x 1 vector of products

$(\mathbf{X}_i'\hat{\mathbf{w}}_i)(\hat{\mathbf{w}}_i'\mathbf{X}_i) =$ KxK matrix (rank 1, outer product)

$\left(\Sigma_{i=1}^{N} \left(\mathbf{X}_i'\hat{\mathbf{w}}_i\right)\left(\hat{\mathbf{w}}_i'\mathbf{X}_i\right)\right)$ = sum of N rank 1 matrices. Rank $\leq$ K.

We could compute this as $\left(\Sigma_{i=1}^{N} \mathbf{X}_i' \left(\hat{\mathbf{w}}_i \hat{\mathbf{w}}_i'\right) \mathbf{X}_i\right) = \left(\Sigma_{i=1}^{N} \mathbf{X}_i' \left(\hat{\mathbf{\Omega}}_i\right) \mathbf{X}_i\right)$.

Why not do it that way?

Cornwell and Rupert Data

Cornwell and Rupert Returns to Schooling Data, 595 Individuals, 7 Years
Variables in the file are

EXP	= work experience, EXPSQ = EXP^2
WKS	= weeks worked
OCC	= occupation, 1 if blue collar,
IND	= 1 if manufacturing industry
SOUTH	= 1 if resides in south
SMSA	= 1 if resides in a city (SMSA)
MS	= 1 if married
FEM	= 1 if female
UNION	= 1 if wage set by unioin contract
ED	= years of education
BLK	= 1 if individual is black
LWAGE	= log of wage = dependent variable in regressions

These data were analyzed in Cornwell, C. and Rupert, P., "Efficient Estimation with Panel Data: An Empirical Comparison of Instrumental Variable Estimators," Journal of Applied Econometrics, 3, 1988, pp. 149-155. See Baltagi, page 122 for further analysis. The data were downloaded from the website for Baltagi's text.

OLS Results

Residuals	Sum of squares	=	522.2008
	Standard error of e	=	.3544712
Fit	R-squared	=	.4112099
	Adjusted R-squared	=	.4100766

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
Constant	5.40159723	.04838934	111.628	.0000	
EXP	.04084968	.00218534	18.693	.0000	19.8537815
EXPSQ	-.00068788	.480428D-04	-14.318	.0000	514.405042
OCC	-.13830480	.01480107	-9.344	.0000	.51116447
SMSA	.14856267	.01206772	12.311	.0000	.65378151
MS	.06798358	.02074599	3.277	.0010	.81440576
FEM	-.40020215	.02526118	-15.843	.0000	.11260504
UNION	.09409925	.01253203	7.509	.0000	.36398559
ED	.05812166	.00260039	22.351	.0000	12.8453782

Alternative Variance Estimators

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]
Constant	5.40159723	.04838934	111.628	.0000
EXP	.04084968	.00218534	18.693	.0000
EXPSQ	-.00068788	.480428D-04	-14.318	.0000
OCC	-.13830480	.01480107	-9.344	.0000
SMSA	.14856267	.01206772	12.311	.0000
MS	.06798358	.02074599	3.277	.0010
FEM	-.40020215	.02526118	-15.843	.0000
UNION	.09409925	.01253203	7.509	.0000
ED	.05812166	.00260039	22.351	.0000
Robust				
Constant	5.40159723	.10156038	53.186	.0000
EXP	.04084968	.00432272	9.450	.0000
EXPSQ	-.00068788	.983981D-04	-6.991	.0000
OCC	-.13830480	.02772631	-4.988	.0000
SMSA	.14856267	.02423668	6.130	.0000
MS	.06798358	.04382220	1.551	.1208
FEM	-.40020215	.04961926	-8.065	.0000
UNION	.09409925	.02422669	3.884	.0001
ED	.05812166	.00555697	10.459	.0000

Generalized Least Squares

$$\hat{\boldsymbol{\beta}} = [\mathbf{X}'\boldsymbol{\Omega}^{\mathbf{-1}}\mathbf{X}]^{-1}[\mathbf{X}'\boldsymbol{\Omega}^{\mathbf{-1}}\mathbf{y}]$$

$$= [\Sigma_{i=1}^{N}\mathbf{X}_{\mathbf{i}}'\boldsymbol{\Omega}_{\mathbf{i}}^{\mathbf{-1}}\mathbf{X}_{\mathbf{i}}]^{-1}[\Sigma_{i=1}^{N}\mathbf{X}_{\mathbf{i}}'\boldsymbol{\Omega}_{\mathbf{i}}^{\mathbf{-1}}\mathbf{y}_{\mathbf{i}}]$$

$$\boldsymbol{\Omega}_{\mathbf{i}}^{\mathbf{-1}} = \frac{1}{\sigma_{\varepsilon}^{2}}\left[I_{T_i} - \frac{\sigma_{\varepsilon}^{2}}{\sigma_{\varepsilon}^{2} + T_i\sigma_{u}^{2}}\mathbf{ii}' \right]$$

(note, depends on i only through T_i)

Panel Data Algebra (1)

$\boldsymbol{\Omega}_{\mathbf{i}} = \sigma_\varepsilon^2 \mathbf{I} + \sigma_u^2 \mathbf{ii}'$, depends on 'i' because it is $T_i \times T_i$

$\boldsymbol{\Omega}_{\mathbf{i}} = \sigma_\varepsilon^2 [\mathbf{I} + \rho^{\mathbf{2}} \mathbf{ii}'],\ \rho^2 = \sigma_u^2 / \sigma_\varepsilon^2$

$\boldsymbol{\Omega}_{\mathbf{i}} = \sigma_\varepsilon^2 [\mathbf{I} + \rho^{\mathbf{2}} \mathbf{ii}'] = \sigma_\varepsilon^2 [\mathbf{A} + \mathbf{bb}'],\ \mathbf{A} = \mathbf{I},\ \mathbf{b} = \rho \mathbf{i}.$

Using (A-66) in Greene (p. 822)

$$\boldsymbol{\Omega}_{\mathbf{i}}^{\mathbf{-1}} = \frac{1}{\sigma_\varepsilon^2}\left[\mathbf{A}^{\mathbf{-1}} - \frac{1}{1 + \mathbf{b}'\mathbf{A}^{\mathbf{-1}}\mathbf{b}} \mathbf{A}^{\mathbf{-1}}\mathbf{bb}'\mathbf{A}^{\mathbf{-1}}\right]$$

$$= \frac{1}{\sigma_\varepsilon^2}\left[\mathbf{I} - \frac{1}{1 + T_i\rho^2}\rho^2 \mathbf{ii}'\right] = \frac{1}{\sigma_\varepsilon^2}\left[\mathbf{I} - \frac{\sigma_u^2}{\sigma_\varepsilon^2 + T_i\sigma_u^2}\mathbf{ii}'\right]$$

Panel Data Algebra (2)

(Based on Wooldridge p. 286)

$$\mathbf{\Omega_i} = \sigma_\varepsilon^2\mathbf{I}+\sigma_u^2\mathbf{ii}' = \sigma_\varepsilon^2\mathbf{I}+T_i\sigma_u^2\mathbf{i(i'i)^{-1}i}' = \sigma_\varepsilon^2\mathbf{I}+T_i\sigma_u^2\mathbf{P_D^i}$$

$$= \sigma_\varepsilon^2\mathbf{I}+T_i\sigma_u^2(\mathbf{I}-\mathbf{M_D^i})$$

$$= (\sigma_\varepsilon^2+T_i\sigma_u^2)[\mathbf{P_D^i}+\eta(\mathbf{I}-\mathbf{P_D^i})],\ \eta_i = \sigma_\varepsilon^2/(\sigma_\varepsilon^2+T_i\sigma_u^2)$$

$$= (\sigma_\varepsilon^2+T_i\sigma_u^2)[\mathbf{P_D^i}+\eta\mathbf{M_D^i}]$$

$$= (\sigma_\varepsilon^2+T_i\sigma_u^2)\mathbf{S_i}$$

$\mathbf{S_i}^{-1} = \mathbf{P_D^i}+(1/\eta_i)\mathbf{M_D^i}$ (Prove by multiplying. $\mathbf{P_D^i M_D^i} = \mathbf{0}$.)

$$\mathbf{S_i}^{-1/2} = \mathbf{P_D^i}+(1/\sqrt{\eta_i})\mathbf{M_D^i} = \frac{1}{1-\theta_i}\left[\mathbf{I}-\theta_i\mathbf{P_D^i}\right],\ \theta_i=1-\sqrt{\eta_i}$$

(Note $\mathbf{S_i^a} = \mathbf{P_D^i}+\eta_i^a\mathbf{M_D^i}$)

Panel Data Algebra (2, cont.)

$$\boldsymbol{\Omega}_{\mathbf{i}}^{-1/2} = \frac{1}{\sqrt{\sigma_\varepsilon^2 + T_i\sigma_u^2}}(\mathbf{P}_{\mathbf{D}}^{\mathbf{i}} + (1/\sqrt{\eta_i})\mathbf{M}_{\mathbf{D}}^{\mathbf{i}})$$

$$= \frac{1}{\sqrt{\sigma_\varepsilon^2 + T_i\sigma_u^2}}\frac{1}{1-\theta_i}\left[\mathbf{I} - \theta_i\mathbf{P}_{\mathbf{D}}^{\mathbf{i}}\right],$$

$$\theta_i = 1 - \sqrt{\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T_i\sigma_u^2}} = 1 - \frac{\sigma_\varepsilon}{\sqrt{\sigma_\varepsilon^2 + T_i\sigma_u^2}}$$

$$\boldsymbol{\Omega}_{\mathbf{i}}^{-1/2} = \frac{1}{\sigma_\varepsilon}[\mathbf{I} - \theta_i\mathbf{P}_{\mathbf{D}}^{\mathbf{i}}]$$

$$\text{Var}[\boldsymbol{\Omega}_{\mathbf{i}}^{-1/2}(\boldsymbol{\varepsilon}_i + u_i\mathbf{i})] = \sigma_\varepsilon^2\mathbf{I}$$

$\boldsymbol{\Omega}_{\mathbf{i}}^{-1/2}\mathbf{y}_i = (1/\sigma_\varepsilon)(\mathbf{y}_i - \theta_i\bar{y}_i\mathbf{i})$ for the GLS transformation.

GLS (cont.)

GLS is equivalent to OLS regression of

$$y_{it}* = y_{it} - \theta_i \bar{y}_{i}\text{.} \text{ on } \mathbf{x}_{it}* = \mathbf{x}_{it} - \theta_i \bar{\mathbf{x}}_{i}\text{.},$$

$$\text{where } \theta_i = 1 - \frac{\sigma_\varepsilon}{\sqrt{\sigma_\varepsilon^2 + T_i \sigma_u^2}}$$

$$\text{Asy.Var}[\hat{\boldsymbol{\beta}}] = [\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}]^{-1} = \sigma_\varepsilon^2[\mathbf{X}' * \mathbf{X}*]^{-1}$$

Estimators for the Variances

$$y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \varepsilon_{it} + u_i$$

With a consistent estimator of $\boldsymbol{\beta}$, say $\mathbf{b}_{OLS}$,

$\Sigma_{i=1}^{N}\Sigma_{t=1}^{T_i}(y_{it} - \mathbf{x}'_{it}\mathbf{b})^2$ estimates $\Sigma_{i=1}^{N}\Sigma_{t=1}^{T_i}(\sigma_\varepsilon^2 + \sigma_U^2)$

Divide by something to estimate $\sigma^2 = \sigma_\varepsilon^2 + \sigma_U^2$

With the LSDV estimates, a_i and $\mathbf{b}_{LSDV}$,

$\Sigma_{i=1}^{N}\Sigma_{t=1}^{T_i}(y_{it} - a_{\mathbf{i}} - \mathbf{x}'_{it}\mathbf{b})^2$ estimates $\Sigma_{i=1}^{N}\Sigma_{t=1}^{T_i}\sigma_\varepsilon^2$

Divide by something to estimate σ_ε^2

Estimate σ_U^2 with $\widehat{(\sigma_\varepsilon^2 + \sigma_U^2)} - \hat{\sigma}_\varepsilon^2$.

Feasible GLS

Feasible GLS requires (only) consistent estimators of σ_ε^2 and σ_u^2.

Candidates:

From the robust LSDV estimator: $\hat{\sigma}_\varepsilon^2 = \dfrac{\sum_{i=1}^{N}\sum_{t=1}^{T_i}(y_{it} - a_i - \mathbf{x}_{it}'\mathbf{b}_{LSDV})^2}{\sum_{i=1}^{N} T_i - K - N}$

From the pooled OLS estimator: $\widehat{\sigma_\varepsilon^2 + \sigma_u^2} = \dfrac{\sum_{i=1}^{N}\sum_{t=1}^{T_i}(y_{it} - a_{OLS} - \mathbf{x}_{it}'\mathbf{b}_{OLS})^2}{\sum_{i=1}^{N} T_i - K - 1}$

From the group means regression: $\widehat{\sigma_\varepsilon^2 / \bar{T} + \sigma_u^2} = \dfrac{\sum_{i=1}^{N}(\bar{y}_{it} - \tilde{a} - \bar{\mathbf{x}}_i'\tilde{\mathbf{b}}_{MEANS})^2}{N - K - 1}$

(Wooldridge) Based on $E[w_{it}w_{is} \mid \mathbf{X}_i] = \sigma_u^2$ if $t \neq s$, $\hat{\sigma}_u^2 = \dfrac{\sum_{i=1}^{N}\sum_{t=1}^{T_i-1}\sum_{s=t+1}^{T_i}\hat{w}_{it}\hat{w}_{is}}{\sum_{i=1}^{N} T_i - K - N}$

There are many others.

x´ does not contain a constant term in the preceding.

Practical Problems with FGLS

All of the preceding regularly produce negative estimates of σ_u^2.

Estimation is made very complicated in unbalanced panels.

A bulletproof solution (originally used in TSP, now LIMDEP and others).

From the robust LSDV estimator: $\hat{\sigma}_\varepsilon^2 = \dfrac{\sum_{i=1}^{N}\sum_{t=1}^{T_i}(y_{it} - a_i - \mathbf{x}_{it}'\mathbf{b}_{LSDV})^2}{\sum_{i=1}^{N} T_i}$

From the pooled OLS estimator: $\widehat{\sigma_\varepsilon^2 + \sigma_u^2} = \dfrac{\sum_{i=1}^{N}\sum_{t=1}^{T_i}(y_{it} - a_{OLS} - \mathbf{x}_{it}'\mathbf{b}_{OLS})^2}{\sum_{i=1}^{N} T_i} \geq \hat{\sigma}_\varepsilon^2$

$$\hat{\sigma}_u^2 = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T_i}(y_{it} - a_{OLS} - \mathbf{x}_{it}'\mathbf{b}_{OLS})^2 - \sum_{i=1}^{N}\sum_{t=1}^{T_i}(y_{it} - a_i - \mathbf{x}_{it}'\mathbf{b}_{LSDV})^2}{\sum_{i=1}^{N} T_i} \geq 0$$

x´ does not contain a constant term in the preceding.

Computing Variance Estimators

Using full list of variables (FEM and ED are time invariant)

OLS sum of squares = 522.2008.

$\widehat{\sigma_\varepsilon^2+\sigma_u^2}$ = 522.2008 / (4165 - 9) = 0.12565.

Using full list of variables and a generalized inverse (same as dropping FEM and ED), LSDV sum of squares = 82.34912.

$\widehat{\sigma_\varepsilon^2}$ = 82.34912 / (4165 - 8-595) = 0.023119.

$\widehat{\sigma_u^2}$ = 0.12565 - 0.023119 = 0.10253

Both estimators are positive. We stop here. If $\widehat{\sigma_u^2}$ were negative, we would use estimators without DF corrections.

Application

```
+----------------------------------------------------+
| Random Effects Model: v(i,t) = e(i,t) + u(i)        |
| Estimates:  Var[e]              =   .231188D-01     |
|             Var[u]              =   .102531D+00     |
|             Corr[v(i,t),v(i,s)] =   .816006         |
| (High (low) values of H favor FEM (REM).)           |
|             Sum of Squares          .141124D+04     |
|             R-squared              -.591198D+00     |
+----------------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
EXP	.08819204	.00224823	39.227	.0000	19.8537815
EXPSQ	-.00076604	.496074D-04	-15.442	.0000	514.405042
OCC	-.04243576	.01298466	-3.268	.0011	.51116447
SMSA	-.03404260	.01620508	-2.101	.0357	.65378151
MS	-.06708159	.01794516	-3.738	.0002	.81440576
FEM	-.34346104	.04536453	-7.571	.0000	.11260504
UNION	.05752770	.01350031	4.261	.0000	.36398559
ED	.11028379	.00510008	21.624	.0000	12.8453782
Constant	4.01913257	.07724830	52.029	.0000	

Testing for Effects: LM Test

Breusch and Pagan Lagrange Multiplier statistic

Assuming normality (and for convenience now, a balanced panel)

$$LM = \frac{NT}{2(T\text{-}1)}\left[\frac{\Sigma_{i=1}^{N}(T\bar{e}_i^2)}{\Sigma_{i=1}^{N}\Sigma_{i=1}^{N}e_{it}^2} - 1\right]^2 = \frac{NT}{2(T\text{-}1)}\left[\frac{\Sigma_{i=1}^{N}[(T\bar{e}_i^2) - \mathbf{e}_i'\mathbf{e}_i]}{\Sigma_{i=1}^{N}\mathbf{e}_i'\mathbf{e}_i}\right]^2$$

Converges to chi-squared[1] under the null hypothesis of no common effects. (For unbalanced panels, the scale in front becomes $(\Sigma_{i=1}^{N}T_i)^2 / [2\Sigma_{i=1}^{N}T_i(T_i - 1)]$.)

Testing for Effects: Moments

Wooldridge (page 265) suggests based on the off diagonal elements

$$Z=\frac{\sum_{i=1}^{N}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}e_{it}e_{is}}{\sqrt{\sum_{i=1}^{N}\left(\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}e_{it}e_{is}\right)^2}}$$

which converges to standard normal. ("We are not assuming any particular distribution for the ε_{it}. Instead, we derive a similar test that has the advantage of being valid for any distribution...") It's convenient to examine Z^2 which, by the Slutsky theorem converges (also) to chi-squared with one degree of freedom.

Testing (2)

$\Sigma_{t=1}^{T-1}\Sigma_{s=t+1}^{T} e_{it}e_{is}$ = 1/2 of the sum of all off diagonal elements of

$\mathbf{e_i e_i'}$ = 1/2 the sum of all the elements minus the diagonal elements.

$\Sigma_{t=1}^{T-1}\Sigma_{s=t+1}^{T} e_{it}e_{is} = 1/2[\mathbf{i}'(\mathbf{e_i e_i'})\mathbf{i} - \mathbf{e_i' e_i}]$. But, $\mathbf{i}'\mathbf{e_i} = T\bar{e}_i$. So,

$\Sigma_{t=1}^{T-1}\Sigma_{s=t+1}^{T} e_{it}e_{is} = (1/2)[(T\bar{e}_i)^2 - \mathbf{e_i' e_i}]$

$$Z^2 = \frac{\left\{\Sigma_{i=1}^{N}[(T\bar{e}_i)^2 - \mathbf{e_i' e_i}]\right\}^2}{\Sigma_{i=1}^{N}[(T\bar{e}_i)^2 - \mathbf{e_i' e_i}]^2} = LM \times \frac{2(T-1)}{NT}\frac{[\Sigma_{i=1}^{N}\mathbf{e_i' e_i}]^2}{\Sigma_{i=1}^{N}[(T\bar{e}_i)^2 - \mathbf{e_i' e_i}]^2}$$

Note, also

$$Z = \frac{\Sigma_{i=1}^{N} r_i}{\sqrt{\Sigma_{i=1}^{N} r_i^2}} = \frac{\sqrt{N}\ \bar{r}}{s_r}, \text{where } r_i = (T\bar{e}_i)^2 - \mathbf{e_i' e_i}.$$

The claim that one function of $[\bar{e}_i, \mathbf{e_i' e_i}]_{i=1,\ldots,N}$ is more valid than the other seems a little dubious.

Application: Cornwell-Rupert

```
+----------------------------------------------------+
| Ordinary    least squares regression                |
| LHS=LWAGE    Mean                 =   6.676346      |
|              Standard deviation   =   .4615122      |
| Model size   Parameters           =          7      |
|              Degrees of freedom   =       4158      |
| Residuals    Sum of squares       =   556.3030      |
|              Standard error of e  =   .3657745      |
| Fit          R-squared            =   .3727592      |
|              Adjusted R-squared   =   .3718541      |
+----------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant      5.66098218       .04685914   120.808   .0000
 FEM           -.39478212       .02603413   -15.164   .0000    .11260504
 ED             .05688005       .00267743    21.244   .0000   12.8453782
 OCC           -.11220205       .01464317    -7.662   .0000    .51116447
 SMSA           .15504405       .01233744    12.567   .0000    .65378151
 MS             .09569050       .02133490     4.485   .0000    .81440576
 EXP            .01043785       .00054206    19.256   .0000   19.8537815
+--------------------------------------------------+
| Random Effects Model: v(i,t) = e(i,t) + u(i)      |
| Estimates:  Var[e]              =   .235368D-01   |
|             Var[u]              =   .110254D+00   |
|             Corr[v(i,t),v(i,s)] =   .824078       |
| Lagrange Multiplier Test vs. Model (3) = 3797.07  |
| ( 1 df, prob value =   .000000)                   |
| (High values of LM favor FEM/REM over CR model.)  |
+--------------------------------------------------+
 Constant      4.24669585       .07763394    54.702   .0000
 FEM           -.34715010       .04681514    -7.415   .0000    .11260504
 ED             .11120152       .00525209    21.173   .0000   12.8453782
 OCC           -.03908144       .01298962    -3.009   .0026    .51116447
 SMSA          -.03881553       .01645862    -2.358   .0184    .65378151
 MS            -.06557030       .01815465    -3.612   .0003    .81440576
 EXP            .05737298       .00088467    64.852   .0000   19.8537815
```

Testing for Effects

```
Regress; lhs=lwage;rhs=fixedx,varyingx;res=e$
Matrix  ; tebar=7*gxbr(e,person)$
Calc    ; list;lm=595*7/(2*(7-1))*
              (tebar'tebar/sumsqdev - 1)^2$
Create  ; e2=e*e$
Matrix  ; e2i=7*gxbr(e2,person)$
Matrix  ; ri=dirp(tebar,tebar)-e2i$
Matrix  ; sumri=ri'1$
Calc    ; list;z2=(sumri)^2/ri'ri$
```

```
LM    =  .37970675705025540D+04
Z2    =  .16533465085356830D+03
```

Hausman Test for FE vs. RE

Estimator	Random Effects $E[c_i|\mathbf{X}_i] = 0$	Fixed Effects $E[c_i|\mathbf{X}_i] \neq 0$
FGLS (Random Effects)	Consistent and Efficient	Inconsistent
LSDV (Fixed Effects)	Consistent Inefficient	Consistent Possibly Efficient

Hausman Test for Effects

Basis for the test, $\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE}$

Wald Criterion: $\hat{\mathbf{q}} = \hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE}$; $W = \hat{\mathbf{q}}'[Var(\hat{\mathbf{q}})]^{-1}\hat{\mathbf{q}}$

A lemma (Hausman (1978)): Under the null hypothesis (RE)

$$\sqrt{nT}[\hat{\boldsymbol{\beta}}_{RE} - \boldsymbol{\beta}] \xrightarrow{d} N[\mathbf{0},\mathbf{V}_{RE}] \text{ (efficient)}$$

$$\sqrt{nT}[\hat{\boldsymbol{\beta}}_{FE} - \boldsymbol{\beta}] \xrightarrow{d} N[\mathbf{0},\mathbf{V}_{FE}] \text{ (inefficient)}$$

Note: $\hat{\mathbf{q}} = (\hat{\boldsymbol{\beta}}_{FE} - \boldsymbol{\beta})-(\hat{\boldsymbol{\beta}}_{RE} - \boldsymbol{\beta})$. The lemma states that in the joint limiting distribution of $\sqrt{nT}[\hat{\boldsymbol{\beta}}_{RE} - \boldsymbol{\beta}]$ and $\sqrt{nT}\ \hat{\mathbf{q}}$, the limiting covariance, $\mathbf{C}_{Q,RE}$ is $\mathbf{0}$. But, $\mathbf{C}_{Q,RE} = \mathbf{C}_{FE,RE} - \mathbf{V}_{RE}$. Then, $Var[\mathbf{q}] = \mathbf{V}_{FE} + \mathbf{V}_{RE} - \mathbf{C}_{FE,RE} - \mathbf{C}'_{FE,RE}$. Using the lemma, $\mathbf{C}_{FE,RE} = \mathbf{V}_{RE}$. It follows that $Var[\mathbf{q}]=\mathbf{V}_{FE} - \mathbf{V}_{RE}$. Based on the preceding

$$H=(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE})' [Est.Var(\hat{\boldsymbol{\beta}}_{FE}) - Est.Var(\hat{\boldsymbol{\beta}}_{RE})]^{-1}(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE})$$

β does not contain the constant term in the preceding.

Computing the Hausman Statistic

$$\text{Est.Var}[\hat{\boldsymbol{\beta}}_{\textbf{FE}}] = \hat{\sigma}_\varepsilon^2 \left[\Sigma_{i=1}^N \mathbf{X}_i' \left(\mathrm{I} - \frac{1}{T_i} \mathbf{ii}' \right) \mathbf{X}_i \right]^{-1}$$

$$\text{Est.Var}[\hat{\boldsymbol{\beta}}_{\textbf{RE}}] = \hat{\sigma}_\varepsilon^2 \left[\Sigma_{i=1}^N \mathbf{X}_i' \left(\mathrm{I} - \frac{\hat{\gamma}_i}{T_i} \mathbf{ii}' \right) \mathbf{X}_i \right]^{-1}, 0 \leq \hat{\gamma}_i = \frac{T_i \hat{\sigma}_u^2}{\hat{\sigma}_\varepsilon^2 + T_i \hat{\sigma}_u^2} \leq 1$$

As long as $\hat{\sigma}_\varepsilon^2$ and $\hat{\sigma}_u^2$ are consistent, as N $\rightarrow \infty$, $\text{Est.Var}[\hat{\boldsymbol{\beta}}_{\textbf{FE}}] - \text{Est.Var}[\hat{\boldsymbol{\beta}}_{\textbf{RE}}]$ will be nonnegative definite. In a finite sample, to ensure this, both must be computed using the same estimate of $\hat{\sigma}_\varepsilon^2$. The one based on LSDV will generally be the better choice.

Note that columns of zeros will appear in $\text{Est.Var}[\hat{\boldsymbol{\beta}}_{\textbf{FE}}]$ if there are time invariant variables in **X**.

β does not contain the constant term in the preceding.

Hausman Test

```
+--------------------------------------------------+
| Random Effects Model: v(i,t) = e(i,t) + u(i)     |
| Estimates:  Var[e]              =   .235236D-01  |
|             Var[u]              =   .133156D+00  |
|             Corr[v(i,t),v(i,s)] =   .849862      |
| Lagrange Multiplier Test vs. Model (3) = 4061.11 |
| ( 1 df, prob value =  .000000)                   |
| (High values of LM favor FEM/REM over CR model.) |
| Fixed vs. Random Effects (Hausman)     = 2632.34 |
| ( 4 df, prob value =  .000000)                   |
| (High (low) values of H favor FEM (REM).)        |
+--------------------------------------------------+
```

Variable Addition Test

- Asymptotic equivalent to Hausman
- Also equivalent to Mundlak formulation
- In the random effects model, using FGLS
  - Only applies to time varying variables
  - Add expanded group means to the regression (i.e., observation i,t gets same group means for all t.
  - Use standard F or Wald test to test for coefficients on means equal to 0. Large F or chi-squared weighs against random effects specification.

Fixed vs. Random Effects

$$\hat{\boldsymbol{\beta}}_{\text{Model}} = \left[\Sigma_{i=1}^{N} \mathbf{X}_i' \left(I - \frac{\hat{\gamma}_{i,\text{Model}}}{T_i} \mathbf{ii}' \right) \mathbf{X_i} \right]^{-1} \left[\Sigma_{i=1}^{N} \mathbf{X}_i' \left(I - \frac{\hat{\gamma}_{i,\text{Model}}}{T_i} \mathbf{ii}' \right) \mathbf{y_i} \right]$$

$\hat{\gamma}_{\text{Model}} = 1$ for fixed effects.

$\hat{\gamma}_{i,\text{Model}} = \sqrt{\dfrac{T_i \hat{\sigma}_u^2}{\hat{\sigma}_\varepsilon^2 + T_i \hat{\sigma}_u^2}}$ for random effects.

As $T_i \to \infty, \hat{\gamma}_{i,RE} \to 1,$ random effects becomes fixed effects

As $\hat{\sigma}_u^2 \to 0, \hat{\gamma}_{i,RE} \to 0,$ random effects becomes OLS (of course)

As $\hat{\sigma}_u^2 \to \infty, \hat{\gamma}_{i,RE} \to 1,$ random effects becomes fixed effects

For the C&R application, $\hat{\sigma}_u^2$=.133156, $\hat{\sigma}_\varepsilon^2$=.0235231, $\hat{\gamma} = .975384$.

Looks like a fixed effects model. Note the Hausman statistic agrees.

β does not contain the constant term in the preceding.