REGRESSION SYSTEMS WITH RANDOM COEFFICIENTS: ESTIMATION PROCEDURES FOR THE UNBALANCED PANEL DATA CASE

ERIK BIØRN,

UNIVERSITY OF OSLO

16th International Conference on Panel Data, University of Amsterdam, July 3, 2010

Background & Motivation – 1

- How treat heterogeneity w.r.t. form of relationships across units in a panel data (PD) set?
  Frequently assumed: Common slope coefficients, unit specific intercepts (FE or RE models). If heterogeneity is more complex, inefficient (and inconsistent?) estimation may result.
- More appealing: Allow for heterogeneity also in slopes. Using *Random coefficients (RC)* to structure coefficient distribution. Expectation vector representing average coefficients, Covariance matrix indicating heterogeneity.
- A RC setup can be considered a way of representing certain kinds of disturbance *heteroskedasticity*.

Background & Motivation – 2

OVERVIEW:

- Relation to literature.
- Model. Rearranging observations
- Full Maximum Likelihood (ML) problem.
  Solving full system of first-order conditions complicated.
  Reason: Unbalance 'interact' with RC.
  GLS subproblem not complicated to solve.
- Simplified stepwise algorithm.
  Simplification relates to estimation of covariance matrices for disturbances and coefficient slacks.
- An application. Three-equation model for factor costs in manufacturing.

Relation to literature

ISSUES:
1. Single equation vs. regression system.
2. Balanced PD vs. unbalanced PD.
3. Intercept heterogeneity vs. slope heterogeneity.

EXISTING LITERATURE:

(a) RC in *single* regression equation with *balanced* PD: Swamy (1970), Hsiao (1975), Swamy & Mehta (1977).

(b) Regression *system* with random *intercepts* for *balanced* PD: Avery (1977), Baltagi (1980).

(c) *Single* regression equation with random *intercepts* for *unbalanced* PD: Biørn (1981), Baltagi (1985), Wansbeek & Kapteyn (1989).

(d) Regression *system* with random *intercepts* for *unbalanced* PD: Biørn (2004).

Random-Coefficient-SUR Model – 1

PRESENT PAPER:

Extends (a)–(d), allowing for heterogeneity across units only.

ELEMENTS:

- (i) System of linear, static regressions equations.
- (ii) Random heterogeneity in intercepts and slopes.
- (iii) Unbalanced panel data.
- (iv) Sample selection rules assumed *ignorable*.

Random-Coefficient-SUR Model – 2

ESSENTIALS:

- G **equations**, equation g $(g = 1, \dots, G)$ has K_g regressors.
- **Unbalanced panel**, units observed in from 1 to P periods.
- Arrange units in **blocks**: N_p = number of units observed in p periods (not necessarily the same or contiguous), (ip) indexes unit i in block p $(i = 1, \dots, N_p;\ p = 1, \dots, P)$.
- For each (i, p): $t\,(= 1, \dots, p)$ indexes no. of observation. In general, **observation no.** $t \neq$ **calendar period**.
- Formally: A **collection of** $P-1$ **balanced panels & one cross-section**. The data set in block p $(p = 2, \dots, P)$ constitutes a balanced PD set with p observations of N_p units. Block 1 defines a cross section – with observations originating from different calendar periods.

Random-Coefficient-SUR Model – 3

Parameter restrictions. The following model setup admits both:

- *Situation [A]: Each equation has a distinct coefficient vector.*
  Total number of coefficients $= \sum_{g=1}^{G} K_g$.
- *Situation [B]: Certain coefficients are common to at least two equations – e.g.*, cross-equational (symmetry) constraints following from agents' optimizing behaviour.
  Total number of coefficients $< \sum_{g=1}^{G} K_g$.

Random-Coefficient-SUR Model – 4

Eq. g for unit (ip) with its p observations stacked:

$$\mathbf{y}_{g(ip)} = \mathbf{X}_{g(ip)}\boldsymbol{\beta}_{g(ip)} + \mathbf{u}_{g(ip)}, \qquad \boldsymbol{\beta}_{g(ip)} = \boldsymbol{\beta}_g + \boldsymbol{\delta}_{g(ip)} \implies$$

$$\mathbf{y}_{g(ip)} = \mathbf{X}_{g(ip)}\boldsymbol{\beta}_g + \boldsymbol{\eta}_{g(ip)}, \qquad \boldsymbol{\eta}_{g(ip)} = \mathbf{X}_{g(ip)}\boldsymbol{\delta}_{g(ip)} + \mathbf{u}_{g(ip)},$$

$$g = 1, \ldots, G, \;\; i = 1, \ldots, N_p, \quad p = 1, \ldots, P,$$

$\boldsymbol{\beta}_g$ = fixed **mean vector**, $\boldsymbol{\delta}_{g(ip)}$ = random **shift/slack vector**.
Stacking by equations we get

$$\mathbf{y}_{(ip)} = \mathbf{X}_{(ip)}\boldsymbol{\beta} + \boldsymbol{\eta}_{(ip)}, \qquad \boldsymbol{\eta}_{(ip)} = \mathbf{X}_{(ip)}\boldsymbol{\delta}_{(ip)} + \mathbf{u}_{(ip)},$$

$$\boldsymbol{\eta}_{(ip)} = \text{'gross disturbance vector' for unit } (ip),$$

$$\mathsf{E}[\boldsymbol{\delta}_{(ip)}\boldsymbol{\delta}'_{(ip)}] = \boldsymbol{\Sigma}^{\delta}, \qquad \mathsf{E}[\mathbf{u}_{(ip)}\mathbf{u}'_{(ip)}] = \boldsymbol{\Sigma}^{u},$$

$$\mathsf{E}[\boldsymbol{\eta}_{(ip)}\boldsymbol{\eta}'_{(ip)}] = \mathbf{X}_{(ip)}\boldsymbol{\Sigma}^{\delta}\mathbf{X}'_{(ip)} + \mathbf{I}_p \otimes \boldsymbol{\Sigma}^{u} = \boldsymbol{\Omega}_{(ip)}.$$

Maximum Likelihood Problem. General setup – 1

Assume normality of disturbances & coefficient slacks.

Then the Log-likelihood function of all $(\mathbf{y}|\mathbf{X})$s for block p and the overall Log-likelihood function are, respectively:

$$\mathcal{L}_{(p)} = -\frac{GN_p p}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{N_p}[\ln|\boldsymbol{\Omega}_{(ip)}| + Q_{(ip)}],$$

$$\mathcal{L} = -\frac{G\sum_{p=1}^{P} N_p p}{2}\ln(2\pi) - \frac{1}{2}\sum_{p=1}^{P}\sum_{i=1}^{N_p}[\ln|\boldsymbol{\Omega}_{(ip)}| + Q_{(ip)}].$$

where

$$Q_{(ip)} = \boldsymbol{\eta}'_{(ip)}\boldsymbol{\Omega}_{(ip)}^{-1}\boldsymbol{\eta}_{(ip)} = \boldsymbol{\eta}'_{(ip)}[\mathbf{X}_{(ip)}\boldsymbol{\Sigma}^{\delta}\mathbf{X}'_{(ip)} + \mathbf{I}_p \otimes \boldsymbol{\Sigma}^{u}]^{-1}\boldsymbol{\eta}_{(ip)}.$$

Maximum Likelihood Problem. General setup – 2

First-order conditions, block p problem:

$$\sum_{i=1}^{N_p}\left(\frac{\partial Q_{(ip)}}{\partial \boldsymbol{\beta}}\right)=\mathbf{0} \Longrightarrow \textbf{GLS est. for block } p \textbf{ [cond. on } (\boldsymbol{\Sigma}^u,\boldsymbol{\Sigma}^\delta)\textbf{]},$$

$$\sum_{i=1}^{N_p}\left[\frac{\partial \ln|\boldsymbol{\Omega}_{(ip)}|}{\partial \boldsymbol{\Sigma}^u}+\frac{\partial Q_{(ip)}}{\partial \boldsymbol{\Sigma}^u}\right]=\mathbf{0},$$

$$\sum_{i=1}^{N_p}\left[\frac{\partial \ln|\boldsymbol{\Omega}_{(ip)}|}{\partial \boldsymbol{\Sigma}^\delta}+\frac{\partial Q_{(ip)}}{\partial \boldsymbol{\Sigma}^\delta}\right]=\mathbf{0}, \quad p=1,\ldots,P.$$

First-order conditions, overall problem:

'compromising' across blocks:

$$\sum_{p=1}^{P}\sum_{i=1}^{N_p}\left[\frac{\partial Q_{(ip)}}{\partial \boldsymbol{\beta}}\right]=\mathbf{0} \Longrightarrow \textbf{overall GLS est. [cond. on } (\boldsymbol{\Sigma}^u,\boldsymbol{\Sigma}^\delta)\textbf{]},$$

$$\sum_{p=1}^{P}\sum_{i=1}^{N_p}\left[\frac{\partial \ln|\boldsymbol{\Omega}_{(ip)}|}{\partial \boldsymbol{\Sigma}^u}+\frac{\partial Q_{(ip)}}{\partial \boldsymbol{\Sigma}^u}\right]=\mathbf{0},$$

$$\sum_{p=1}^{P}\sum_{i=1}^{N_p}\left[\frac{\partial \ln|\boldsymbol{\Omega}_{(ip)}|}{\partial \boldsymbol{\Sigma}^\delta}+\frac{\partial Q_{(ip)}}{\partial \boldsymbol{\Sigma}^\delta}\right]=\mathbf{0}.$$

Estimation. Stepwise setup – 1

Essentials of algorithm:

A. Initial OLS estimation of $\boldsymbol{\beta}_{(ip)}$ and $\boldsymbol{\beta}$.
Extract residuals & coefficient slacks.

B. Initial estimation of $\boldsymbol{\Sigma}^{u}$ and $\boldsymbol{\Sigma}^{\delta}$.

C. Revised FGLS estimation of $\boldsymbol{\beta}_{(ip)}$ and $\boldsymbol{\beta}$.
Extract residuals & coefficient slacks.

D. Revised estimation of $\boldsymbol{\Sigma}^{u}$ and $\boldsymbol{\Sigma}^{\delta}$.

Iterating between steps C & D?

Estimation. Stepwise setup – 2

Structure of algorithm for all blocks combined:

Similar algorithm can be applied block-wise – see paper.

1: OLS estimation of unit-specific coefficients for all units in blocks $p \in [q, P]$, where $q =$ lowest value of p permitting such estimation of all equations. Extract $\mathbf{u}$ residuals.

2: Compute estimator of $\boldsymbol{\beta}$ for blocks $p \in [q, P]$ jointly.

3: Estimate, from $\mathbf{u}$ residuals and coefficient slacks, $(\boldsymbol{\Sigma}^u, \boldsymbol{\Sigma}^\delta)$.

4: Compute estimators for $\boldsymbol{\Omega}_{(ip)}$ for $i = 1, \ldots, N_p;\ p \in [1, P]$.

5: Compute overall FGLS estimator of $\boldsymbol{\beta}$.

6: Extract revised residuals and coefficient slacks, and recompute estimators of $(\boldsymbol{\Sigma}^u, \boldsymbol{\Sigma}^\delta)$.

7: Recompute estimators of $\boldsymbol{\Omega}_{(ip)}$ for blocks $p \in [1, P]$ jointly.

8: Recompute overall FGLS estimator of $\boldsymbol{\beta}$.

Steps 6–8 can be iterated until convergence, by some criterion?

Application – 1

Basics:

- Pulp and Paper Industry in Norway, 1972–1993: $(T = P = 22)$.
- About two thirds of full data set selected for this example:
  $\sum N_p = 111$ firms, $\sum N_p p = 1891$ observations,
  including those observed $p = 22, 21, 20, 10, 7, 5$ times
  $\Longrightarrow$ 6 blocks among the full 22 – are included.

Panel design:

p	N_p	$N_p p$
22	61	1342
21	8	168
20	6	120
10	11	110
7	13	91
5	12	60
$\sum$	111	1891

Application – 2

- Gaps occur in time series for all blocks with $p = 2, \ldots, 21$.
  Only $p = 22$ block has contiguous observations.
- 'Cross-section block' mixes observations from different years.
- Average no. of observations per firm: 17.
- Why 'curtail' the data set in this way?
  One reason: N_p is, for several p, quite low $\rightarrow$ potentially 'volatile'/'unstable' estimates in unit- & block-specific regressions.
- $G = 3$ equations. Explain:

  log cost per unit output & cost shares materials & labour

  by means of

  log output & log (material price/labour price)
- Computations implemented by code programmed in Gauss.

Application – 3

Overall OLS Estimates, for all blocks.

Standard errors, from OLS formula *neglecting coefficient heterogeneity*, in parenthesis

LHS VAR	*logx*	*logpml*	OLS SER: $\widehat{\sigma}_u$
logcx	-0.3008 (0.0094)	0.6786 (0.0661)	0.2802
csm	-0.0439 (0.0013)	0.0244 (0.0093)	0.0340
csl	0.0283 (0.0015)	-0.0400 (0.0106)	0.0395

Means ($\widehat{\beta}$) and Standard deviations of OLS estimates for all firms

	$\widehat{\beta}$ = Mean of $\widehat{\beta}_{(ip)}$		Emp.st.dev. in $\widehat{\beta}_{(ip)}$ distribution	
LHS VAR	*logx*	*logpml*	*logx*	*logpml*
logcx	-0.2461	0.8074	0.8424	0.9871
csm	-0.0377	0.0823	0.0722	0.1360
csl	0.0369	-0.1138	0.0785	0.1413

Application – 4

Estimate of coefficient slack covariance matrix, $\widehat{\boldsymbol{\Sigma}}^{\delta}$

	LHS VAR: *logcx* con	*logx*	*logpml*	LHS VAR: *csm* con	*logx*	*logpml*	LHS VAR: *csl* con	*logx*	*logpml*
con	82.6957								
logx	-6.7030	0.7096							
logpml	-4.0870	0.0010	0.9744						
con	-2.3112	0.1591	0.1955	0.7651					
logx	0.1653	-0.0144	-0.0075	-0.0429	0.0052				
logpml	0.2056	-0.0082	-0.0316	-0.0813	-0.0004	0.0185			
con	0.2429	0.0143	-0.1378	-0.6685	0.0376	0.0713	0.8655		
logx	0.0079	-0.0018	0.0074	0.0358	-0.0047	0.0009	-0.0512	0.0062	
logpml	-0.1156	0.0056	0.0155	0.0757	0.0005	-0.0174	-0.0840	-0.0008	0.0200

Standard error of coefficient slacks $= (\text{diag}[\widehat{\boldsymbol{\Sigma}}^{\delta}])^{\frac{1}{2}}$

LHS VAR:	RHS VAR: *logx*	*logpml*
logcx	0.8424	0.9871
csm	0.0722	0.1360
csl	0.0785	0.1413

Application – 5

Overall FGLS Coefficient Estimates, versus Overall OLS estimates

	FGLS, RHS VAR:		OLS, RHS VAR:	
LHS VAR:	*logx*	*logpml*	*logx*	*logpml*
logcx	-0.2158 (0.0852)	0.9230 (0.1073)	-0.3008	0.6786
csm	-0.0367 (0.0076)	0.0742 (0.0144)	-0.0439	0.0244
csl	0.0327 (0.0083)	-0.1112 (0.0152)	0.0283	-0.0400

Application – 6

SOME CONCLUSIONS:

- The FGLS standard errors substantially exceed the simple OLS ones. The latter relate to a model which disregards coefficient heterogeneity. The former fully exploits it, 'weighting' firm-to-firm 'coefficient slacks' and 'disturbance noise'.
- Evidence of increasing returns to scale: elasticity of cost w.r.t. output $\in (0, 1)$.
- Increasing production reduces materials cost share and increases labour cost share. Non-homotheticity?
- The factor price ratio significantly affects cost shares in the FGLS estimates.
- Overall, with respect to sign, the FGLS results 'robustify' the OLS results. But the numerical values of the estimates depart substantially.

Thank you!