

Week 11. Panel Models 2

Types of Panel Models

- FEM: Fixed Effects Model

$$y_{it} = \alpha + \sum \beta_k X_{ikt} + u_i + e_{it}$$

or,

$$y_{it} = \alpha_i + \sum \beta_k X_{ikt} + e_{it}$$

- FDM: First Difference Model

$$\Delta y_i = \alpha + \sum \beta_k \Delta X_{ik} + e_i$$

- REM: Random Effects Model

$$(y_{it} - \theta \bar{y}_i) = \alpha(1 - \theta) + \sum \beta_k (X_{ikt} - \theta \bar{X}_{ik}) + (e_{it} - \theta \bar{e}_i)$$

- CRE (The Correlated Random Effects Model) = Hybrid Model

$$y_{it} = \alpha + \sum \beta_k X_{ikt} + \sum \gamma_l \bar{X}_{ik} + \sum \delta_m Z_i + \epsilon_i + e_{it}$$

where Z_i is a vector of time-invariant covariates,

$$u_i = \sum \gamma_l \bar{X}_{ik} + \sum \delta_m Z_i + \epsilon_i$$

- RCM: Random Coefficients Model (or Mixed Model, or Linear Mixed Model or Multilevel Growth Curve Model)
- There are many other panel models which we will not discuss.

FEM: Interaction of Time-varying covariates \times Time-invariant covariates and Interaction of Time-varying covariates \times Time-varying covariates

- Interaction of Time-varying covariates(X_1) \times Time-invariant covariates(X_2): As X_1 changes by 1-unit, how much y will change differently by X_2 .
- Interaction of Time-varying covariates(X_1) \times Time-varying covariates(X_3): As X_1 increases by 1-unit, how much y will change differently when X_3 increases by 1-unit.
- In the below example,
 - **SSS**: Subjective social standing (1: lowest, 6: highest)
 - **highim2pop**: % high-status immigrants by region
 - **lowercl**: Low-class natives, time-invariant class status
 - **highercl**: High-class natives, time-invariant class status
 - **rank1**: Relative income rank by year, time-varying variable.
 - **marst**, **edu**: marital status, and education. Both are time-varying.

Fixed-effects (within) regression
Group variable: pid

Number of obs = 155,306
Number of groups = 28,241

R-sq:

within = 0.0313
between = 0.1736
overall = 0.1417

Obs per group:

min = 1
avg = 5.5
max = 10

(Std. Err. adjusted for 28,241 clusters in pid)

	SSS	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
highim2pop		-.0929533	.0467272	-1.99	0.047	-.1845409	-.0013656
1.lowercl		0	(omitted)				
lowercl#c.highim2pop							
1		.021822	.0987126	0.22	0.825	-.1716593	.2153034
1.uppercl		0	(omitted)				
uppercl#c.highim2pop							
1		.1930963	.0809319	2.39	0.017	.0344659	.3517267
rank1		.0040681	.0001519	26.79	0.000	.0037704	.0043658
c.highim2pop#c.rank1		.0007305	.0005505	1.33	0.184	-.0003484	.0018094
age		-.0183471	.0028889	-6.35	0.000	-.0240094	-.0126848
c.age#c.age		.0000653	.0000256	2.55	0.011	.0000151	.0001154
female		0	(omitted)				
marst							
2		.0652948	.0210354	3.10	0.002	.0240644	.1065252
3		-.1318094	.0515807	-2.56	0.011	-.23291	-.0307088
4		-.0898101	.0383994	-2.34	0.019	-.1650748	-.0145455
5		-.0009762	.0346178	-0.03	0.978	-.0688289	.0668764
famnum		-.011906	.004652	-2.56	0.010	-.0210241	-.0027879
edu							
2		-.0502888	.0191458	-2.63	0.009	-.0878155	-.0127621
3		-.09479	.019817	-4.78	0.000	-.1336322	-.0559478
4		-.1348723	.0317038	-4.25	0.000	-.1970132	-.0727313
5		-.0671691	.0547998	-1.23	0.220	-.1745793	.0402411

Random Coefficients Model: RCM

- In a Nutshell: In OLS, we have one intercept and one coefficient for each variable (e.g., time). In RCM, this assumption can be relaxed by allowing the intercept and slope to vary across individuals and be predicted by other covariates. That is, OLS models estimate means, while RCM models estimate both means and variance components.
- Random Coefficients Model = a multi-level model.
Here the 1st level is individual and the 2nd level is individual-time.

- Key Background Idea: Independence of the observations is a key assumption of many standard statistical methods including OLS. Common examples of data structures that do not fit into such a framework arise in longitudinal analysis, in which observations are made on subjects at subject-specific sequences of time points, and in studies that involve subjects (units) occurring naturally in clusters (or groups), such as individuals within families, school children within classrooms, employees within companies, and the like. The assumption of independence of the observational units is not tenable when observations within a cluster tend to be more similar than observations in general. REM is a relaxation of this assumption of independence of the observational units.

- In OLS,

$$y_{it} = \alpha + \sum \beta_k X_{ikt} + e_{it}$$

If there is only one independent variable, T in OLS,

$$y_{it} = \alpha + \beta T_{it} + e_{it}$$

We have usual assumptions such as normality, independence and equal variance (homoscedasticity) of the deviations e_{it} . In this case, $e_{it} \sim N(0, \sigma^2)$, i.i.d., implies that the regressions within the clusters/individuals/groups i have a common vector of coefficients β .

- We only have one intercept and one coefficient. Thus, they are fixed across observations.
- This restriction (= coefficients estimated are fixed) can be relaxed by allowing the regressions to differ in their intercepts and slopes. In REM, we model how systematically intercept and coefficients across clusters/groups/individuals.
-

$$y_{it} = (\alpha + b_{0i}) + (\beta + b_{1i})T_{it} + \varepsilon_{it}$$

Here, α and β are fixed effects and b_{0i} and b_{1i} are random effects.

(1)

$$b_{0i} = b_{00} + u_{0i}$$

$$b_{1i} = b_{10} + u_{1i}$$

- Now we have three error components: ε_{it} , u_{0i} , and u_{1i} . Recall that, in OLS, we only have one error component ε_{it} which is normally distributed with mean 0, $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$. Like OLS, we assume that ε_{it} is normally distributed with mean 0 and variance of σ_ε^2 .
- The other two error components are also assumed to be normally distributed with mean 0. That is, $u_{0i} \sim N(0, \sigma_{u_0}^2)$ and $u_{1i} \sim N(0, \sigma_{u_1}^2)$

- The main concern is how these three error components ($= \varepsilon_{it}$, u_{0i} , and u_{1i}) are correlated with each other.
- For all regression models, ε_{it} is assumed to be uncorrelated with other components. Otherwise, the coefficients estimated are biased. Thus, the error structure of the other two errors (= the variance-covariance matrix of u_{0i} and u_{1i}) is uncorrelated with ε_{it} . Let's say the variance-covariance matrix of u_{0i} and u_{1i} is G . The covariance between G and ε_{it} is assumed to be zero. Then,

$$\text{Var} \begin{bmatrix} u_{0i}, u_{1i} \\ \varepsilon_{it} \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & \sigma_{\varepsilon_{it}}^2 \end{bmatrix}$$

- For u_{0i} and u_{1i} , it is assumed that u_{0i} and u_{1i} are jointly normally distributed with mean 0, the variances of $\sigma_{u_0}^2$ and $\sigma_{u_1}^2$, and the covariance of $\rho(u_0, u_1)$. That is,

$$\text{Var} \begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix} = \begin{bmatrix} \sigma_{u_0}^2 & \rho_{u_0, u_1} \\ \rho_{u_0, u_1} & \sigma_{u_1}^2 \end{bmatrix} \sim N \begin{bmatrix} \sigma_{u_0}^2 & \rho_{u_0, u_1} \\ \rho_{u_0, u_1} & \sigma_{u_1}^2 \end{bmatrix}$$

Sometimes it is assumed that $\rho_{u_0, u_1} = 0$.

- Now we add independent variables to account for the random effects.

$$y_{it} = (\alpha + b_{0i}) + (\beta + b_{1i})T_{it} + \varepsilon_{it}$$

Here, α and β are fixed effects and b_{0i} and b_{1i} are random effects.

(2)

$$\begin{aligned} b_{0i} &= b_{00} + \sum b_{0k}X_{ikt} + u_{0i} \\ b_{1i} &= b_{10} + \sum b_{1k}X_{ikt} + u_{1i} \end{aligned}$$

- Combining the fixed and random components, equation (2) becomes equation (3):

$$\begin{aligned} y_{it} &= (\alpha + b_{00} + \sum b_{0k}X_{ikt} + u_{0i}) + (\beta + b_{10} + \sum b_{1k}X_{ikt} + u_{1i})T_{it} + \varepsilon_{it} \\ &= (\alpha + b_{00}) + (\beta + b_{10})T_{it} + \sum b_{0k}X_{ikt} + \sum b_{1k}(X_{ikt} \times T_{it}) \\ &\quad + u_{0i} + u_{1i}T_{it} + \varepsilon_{it} \\ &= \alpha' + \beta'T_{it} + \sum b_{0k}X_{ikt} + \sum b_{1k}(X_{ikt} \times T_{it}) + (u_{0i} + u_{1i}T_{it} + \varepsilon_{it}) \end{aligned} \tag{3}$$

- That is, T_{it} , X_{ikt} , and $(X_{ikt} \times T_{it})$ are a set of covariates you need to add in your model.

RCM: Random Intercept without Other Covariates and Fixed Slope

$$y_{it} = (\alpha + b_{0i}) + \beta T_{it} + \varepsilon_{it}$$

$$b_{0i} = b_{00} + u_{0i}$$

$$\text{Var} \begin{bmatrix} u_{0i} \\ \varepsilon_{it} \end{bmatrix} = \sim N \begin{bmatrix} \sigma_{u_0}^2 & 0 \\ 0 & \sigma_{\varepsilon_{it}}^2 \end{bmatrix}$$

```
. mixed lnwage time || ind:
```

```
Performing EM optimization:
```

```
Performing gradient-based optimization:
```

```
Computing standard errors:
```

```
Mixed-effects ML regression
```

```
Group variable: ind
```

```
Number of obs      =      840
```

```
Number of groups   =      42
```

```
Obs per group:
```

```
min =      20
```

```
avg =     20.0
```

```
max =      20
```

```
Wald chi2(1)      =     369.69
```

```
Prob > chi2       =     0.0000
```

```
Log likelihood = 1273.1363
```

```
-----
```

lnwage	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
time	.0051957	.0002702	19.23	0.000	.0046661 .0057254
_cons	2.523098	.0406276	62.10	0.000	2.44347 2.602727

```
-----
```

```
-----
```

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]
ind: Identity			
var(_cons)	.0689465	.0150676	.0449252 .1058118
var(Residual)	.0020395	.0001021	.0018489 .0022498

```
-----
```

```
LR test vs. linear model: chibar2(01) = 2708.06      Prob >= chibar2 = 0.0000
```

RCM: Random Intercept and Random Slope without Other Covariates

$$y_{it} = (\alpha + b_{0i}) + (\beta + b_{1i})T_{it} + \varepsilon_{it}$$

$$b_{0i} = b_{00} + u_{0i}$$

$$b_{1i} = b_{10} + u_{1i}$$

$$Var \begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix} = \sim N \begin{bmatrix} \sigma_{u_0}^2 & 0 \\ 0 & \sigma_{u_1}^2 \end{bmatrix}$$

```
. mixed lnwage time || ind: time
```

Performing EM optimization:

Performing gradient-based optimization:

Computing standard errors:

Mixed-effects ML regression

Group variable: ind

Number of obs = 840

Number of groups = 42

Obs per group:

min = 20

avg = 20.0

max = 20

Wald chi2(1) = 37.83

Prob > chi2 = 0.0000

Log likelihood = 1477.8734

lnwage	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
time	.0051957	.0008447	6.15	0.000	.00354	.0068514
_cons	2.523098	.0455192	55.43	0.000	2.433882	2.612314

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
ind: Independent				
var(time)	.0000284	6.54e-06	.0000181	.0000446
var(_cons)	.0868322	.0189902	.0565617	.1333027
var(Residual)	.0010321	.0000531	.0009331	.0011416

LR test vs. linear model: chi2(2) = 3117.53

Prob > chi2 = 0.0000

RCM: Random Intercept net of Other Covariates and Fixed Slope

$$y_{it} = (\alpha + b_{0i}) + (\beta + b_{1i})T_{it} + \varepsilon_{it}$$

$$b_{0i} = b_{00} + \sum b_{0k}X_{ikt} + u_{0i}$$

```
. mixed lnwage time union part || ind:
```

```
Performing EM optimization:
```

```
Performing gradient-based optimization:
```

```
Computing standard errors:
```

```
Mixed-effects ML regression
```

```
Group variable: ind
```

```
Number of obs      =      840
```

```
Number of groups   =      42
```

```
Obs per group:
```

```
min =      20
```

```
avg =     20.0
```

```
max =      20
```

```
Wald chi2(3)      =     830.51
```

```
Prob > chi2       =     0.0000
```

```
Log likelihood =  1409.9854
```

lnwage	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
time	.0066589	.0004251	15.66	0.000	.0058257	.0074922
union	.5420962	.0553771	9.79	0.000	.4335592	.6506333
part	-.7703482	.0766178	-10.05	0.000	-.9205164	-.6201801
_cons	2.486188	.0344645	72.14	0.000	2.418639	2.553738

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
ind: Identity				
var(_cons)	.0354159	.007816	.0229797	.0545823
var(Residual)	.0014989	.0000751	.0013588	.0016536

```
LR test vs. linear model: chibar2(01) = 2313.07
```

```
Prob >= chibar2 = 0.0000
```

RCM: Random Intercept net of Other Covariates and Random Slope

$$y_{it} = (\alpha + b_{0i}) + (\beta + b_{1i})T_{it} + \varepsilon_{it}$$

$$b_{0i} = b_{00} + \sum b_{0k}X_{ikt} + u_{0i}$$

$$b_{1i} = b_{10} + u_{1i}$$

$$\text{Var} \begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix} = \sim N \begin{bmatrix} \sigma_{u_0}^2 & 0 \\ 0 & \sigma_{u_1}^2 \end{bmatrix}$$

```
. mixed lnwage time union part || ind: time
```

Performing EM optimization:

Performing gradient-based optimization:

Computing standard errors:

Mixed-effects ML regression

Group variable: ind

Number of obs = 840

Number of groups = 42

Obs per group:

min = 20

avg = 20.0

max = 20

Wald chi2(3) = 156.00

Prob > chi2 = 0.0000

Log likelihood = 1511.0846

lnwage	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
time	.0056969	.0007439	7.66	0.000	.0042389	.0071549
union	.31924	.0676072	4.72	0.000	.1867322	.4517477
part	-.6651201	.0943489	-7.05	0.000	-.8500404	-.4801997
_cons	2.525942	.0390991	64.60	0.000	2.449309	2.602575

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
ind: Independent				
var(time)	.0000148	3.66e-06	9.08e-06	.000024
var(_cons)	.0459613	.010538	.0293244	.0720371
var(Residual)	.0010129	.0000526	.000915	.0011214

LR test vs. linear model: chi2(2) = 2515.26

Prob > chi2 = 0.0000

RCM: Random Intercept net of Other Covariates and Random Slope, Unstructured

$$y_{it} = (\alpha + b_{0i}) + (\beta + b_{1i})T_{it} + \varepsilon_{it}$$

$$b_{0i} = b_{00} + \sum b_{0k}X_{ikt} + u_{0i}$$

$$b_{1i} = b_{10} + u_{1i}$$

$$Var \begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix} = \sim N \begin{bmatrix} \sigma_{u_0}^2 & \rho(u_0, u_1) \\ \rho(u_0, u_1) & \sigma_{u_1}^2 \end{bmatrix}$$

```
. mixed lnwage time union part || ind: time, covariance(unstructured)
```

Performing EM optimization:

Performing gradient-based optimization:

Computing standard errors:

Mixed-effects ML regression

Number of obs = 840

Group variable: ind

Number of groups = 42

Obs per group:

min = 20

avg = 20.0

max = 20

Wald chi2(3) = 126.99

Log likelihood = 1516.0186

Prob > chi2 = 0.0000

lnwage	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
time	.0056362	.0007628	7.39	0.000	.0041412	.0071313
union	.2843386	.0689311	4.12	0.000	.1492361	.419441
part	-.5958343	.0982291	-6.07	0.000	-.7883599	-.4033088
_cons	2.52603	.040325	62.64	0.000	2.446994	2.605066

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
ind: Unstructured				
var(time)	.0000158	3.95e-06	9.72e-06	.0000258
var(_cons)	.0496985	.0115182	.0315549	.0782744
cov(time,_cons)	-.000438	.0001722	-.0007756	-.0001005
var(Residual)	.0010069	.000052	.0009099	.0011143

LR test vs. linear model: chi2(3) = 2525.13

Prob > chi2 = 0.0000

RCM: Random Intercept net of Other Covariates and Random Slope net of Other Covariates, Unstructured

$$y_{it} = (\alpha + b_{0i}) + (\beta + b_{1i})T_{it} + \varepsilon_{it}$$

$$b_{0i} = b_{00} + \sum b_{0k}X_{ikt} + u_{0i}$$

$$b_{1i} = b_{10} + \sum b_{1k}X_{ikt} + u_{1i}$$

$$Var \begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix} = \sim N \begin{bmatrix} \sigma_{u_0}^2 & \rho(u_0, u_1) \\ \rho(u_0, u_1) & \sigma_{u_1}^2 \end{bmatrix}$$

```
. mixed lnwage time union part c.time##c.union c.time##c.part || ind: time, covariance(unstructured)
```

Performing EM optimization:

Performing gradient-based optimization:

Computing standard errors:

Mixed-effects ML regression

Group variable: ind

Number of obs = 840

Number of groups = 42

Obs per group:

min = 20

avg = 20.0

max = 20

Wald chi2(5) = 248.34

Prob > chi2 = 0.0000

Log likelihood = 1581.4253

lnwage	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
time	.0176998	.0014519	12.19	0.000	.0148543	.0205454
union	.4372819	.0668816	6.54	0.000	.3061963	.5683675
part	-.1465969	.1114184	-1.32	0.188	-.364973	.0717793
time	0	(omitted)				
union	0	(omitted)				
c.time#c.union	-.0483779	.0041552	-11.64	0.000	-.056522	-.0402338
time	0	(omitted)				
part	0	(omitted)				
c.time#c.part	-.0451379	.0070325	-6.42	0.000	-.0589214	-.0313545
_cons	2.455194	.0433314	56.66	0.000	2.370266	2.540122

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]
---------------------------	----------	-----------	----------------------

-----+-----				
ind: Unstructured				
var(time)		.0000391	.0000131	.0000202 .0000754
var(_cons)		.0605239	.014725	.0375695 .097503
cov(time,_cons)		-.0006508	.0003655	-.0013672 .0000656
-----+-----				
var(Residual)		.0007963	.0000425	.0007173 .0008841

LR test vs. linear model: chi2(3) = 2654.55			Prob > chi2 = 0.0000	

RCM vs REM

- If we assume only intercepts differ across clusters/groups/individuals and the difference is random, then,

$$y_{it} = (\alpha + b_{0i}) + \beta T_{it} + \varepsilon_{it}$$

This is REM. Thus, REM is a special case of RCM. If you estimate a simple random-intercept RCM, you will get the identical results with REM with the option of MLE (maximum likelihood estimation).

```
. xtset ind time

* =====
* REM
* =====
. xtreg lnwage time union part, re mle

Random-effects ML regression              Number of obs      =       840
Group variable: ind                      Number of groups   =       42

Random effects u_i ~ Gaussian              Obs per group:
                                         min =           20
                                         avg  =          20.0
                                         max  =           20

                                         LR chi2(3)        =       577.48
Log likelihood = 1409.9854                Prob > chi2        =       0.0000

-----+-----
      lnwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      time |   .0066589   .0004252    15.66  0.000   .0058255   .0074924
     union |   .5420962   .0553895     9.79  0.000   .4335349   .6506576
      part |  -.7703483   .0771461    -9.99  0.000  -.9215517  -.6191448
     _cons |   2.486188   .0344729    72.12  0.000   2.418623   2.553754
-----+-----
    /sigma_u |   .1881911   .020766             .1515908   .2336282
    /sigma_e |   .0387161   .0009696             .0368617   .0406638
       rho   |   .9593946   .0088249             .9388778   .9739306
-----+-----

LR test of sigma_u=0: chibar2(01) = 2313.07      Prob >= chibar2 = 0.000

* =====
* RCM with two-level random intercept
* =====
. mixed lnwage time union part || ind:

Performing EM optimization:
Performing gradient-based optimization:
Computing standard errors:

Mixed-effects ML regression              Number of obs      =       840
Group variable: ind                      Number of groups   =       42
```

```

Obs per group:
      min =      20
      avg =     20.0
      max =      20

Wald chi2(3)      =     830.51
Prob > chi2       =     0.0000

Log likelihood = 1409.9854

-----+-----
      lnwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      time |   .0066589   .0004251    15.66   0.000    .0058257    .0074922
     union |   .5420962   .0553771     9.79   0.000    .4335592    .6506333
      part |  -.7703482   .0766178    -10.05   0.000   -.9205164   -.6201801
      _cons |   2.486188   .0344645    72.14   0.000    2.418639    2.553738
-----+-----

Random-effects Parameters |   Estimate  Std. Err.      [95% Conf. Interval]
-----+-----
ind: Identity |
      var(_cons) |   .0354159   .007816     .0229797    .0545823
-----+-----
      var(Residual) |   .0014989   .0000751     .0013588    .0016536
-----+-----

LR test vs. linear model: chibar2(01) = 2313.07      Prob >= chibar2 = 0.0000

```