these two are both individual effect

individual can mean widely,person,firm,etc. now, we consider it as person

$$y = X\beta + \epsilon$$

if y = earning, x = IQ, we, noramlly, say people with higher IQ, earn more. But in this case, we can say, just because it is you, so you earn more. the independent variable is the person itself.
even though it hard for use to estimate how the person itself affect his earning, we can know person does affect the earning. and we can use in other models

Fixed Effects vs Random Effects

notice, fixed effect is not "Fixed",so do random effect

Chungsang Tom Lam[1]

[1]Department of Economics
Clemson University

March 26, 2020

The Model

Consider the model: *X1,,Xk are change over time t, but alpha will not change over time.*

only i here, because each individual i have their own effect, and will be along the time t

$$y_{it} = \beta_0 + x_{1it}\beta_1 + x_{2it}\beta_2 \ldots + x_{kit}\beta_k + \alpha_i + \epsilon_{it} \qquad (1)$$

Strict exogeneity assumption:

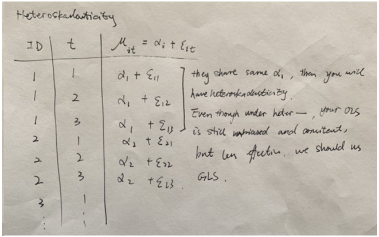
if this assumption doesn't hold, then you have to use IV to solve the problem.

$$E[\epsilon_{it}|x_i, \alpha_i] = 0 \qquad (2)$$

without alpha, if you regress y on x1 to xk,it works.
But if there's alpha, you still regress y on x1 to xk, with assumption below,

$E[\epsilon_{it}|x_i, \alpha_i] = 0$ $E[\alpha_i|x_i] = 0$ *alpha and epsilon are both not related to X, then OLS works*

you consider the error term as a combination of $\alpha_i + \epsilon_{it}$


Heteroskadasticity

ID	t	$\mu_{it} = \alpha_i + \varepsilon_{it}$
1	1	$\alpha_1 + \varepsilon_{11}$
1	2	$\alpha_1 + \varepsilon_{12}$
1	3	$\alpha_1 + \varepsilon_{13}$
2	1	$\alpha_2 + \varepsilon_{21}$
2	2	$\alpha_2 + \varepsilon_{22}$
2	3	$\alpha_2 + \varepsilon_{23}$.
3	1	
⋮	⋮	

they share same α_1, then you will have heteroskadasticity.
Even though under heter—, your OLS is still unbiased and consitent, but less effective. we should us GLS.

in real world, it's hard for us to find this individual effect alpha, because it is too subjective, cannot have data support. so the only thing we need to do is try to convince other this is the individual effect.

Random Effects

- In the random effects model, we have to make the "unrelated effects" assumption: alpha should not be related with X,

you should know, most of the time this assumption doesn't hold

$$E[\alpha_i|x_i] = 0 \quad (3)$$

so, you need to argue it.

with this assumption, we are ok to omit alpha, as it doesn't affect x

- Under this assumption, we do know that the Pooled OLS is unbiased and consistent. We just treat the whole $\alpha_i + \epsilon_{it} = \mu_{it}$ as the error term and run an OLS regression.
- However,since we know about the error μ_{it}, we can run a GLS estimation instead:

random effect is unbiased and consistent you can also use cluster standard error to solve this question

Random Effects

each one of them
have this matrix

$$\Omega_\mu = \begin{bmatrix} \Omega_{\mu,1} & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \dots & \dots & \vdots \\ 0 & \dots & \Omega_{\mu,i} & \dots & 0 \\ \vdots & \dots & \dots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & \Omega_{\mu,N} \end{bmatrix} \quad (4)$$

and

σ^2_α this term at off
diagonal position
because this is one
person matrix, he have
same alpha.

this is person i's matrix

row and column represent
different time periods.

$$\Omega_{\mu,i} = \begin{bmatrix} \sigma^2_\mu & \sigma^2_\alpha & \dots & \dots & \sigma^2_\alpha \\ \sigma^2_\alpha & \sigma^2_\mu & \sigma^2_\alpha & \dots & \sigma^2_\alpha \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \dots & \ddots & \sigma^2_\alpha \\ \sigma^2_\alpha & \sigma^2_\alpha & \dots & \sigma^2_\alpha & \sigma^2_\mu \end{bmatrix} \quad (5)$$

individual one period 1
and individual on period 2

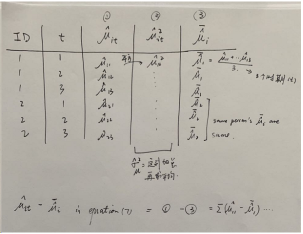
where $\sigma^2_\mu = \sigma^2_\alpha + \sigma^2_\epsilon$

Random Effects

the estimator we get, you can call it FGLS estimator or random effect estimator

- And we can estimate the Ω_μ by $\hat{\Omega}_\mu$
- First, we run the Pooled OLS and obtain the residual
- Then we use the residual $\hat{\sigma}_\mu$ to calculate the required elements

$$\underset{\text{mean}}{\hat{\sigma}_\mu^2} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\hat{\mu}_{it}^2 \tag{6}$$



ID	t	① $\hat{\mu}_{it}$	② $\hat{\mu}_{it}^2$	③ $\bar{\hat{\mu}}_i$
1	1	$\hat{\mu}_{11}$ 平方 →	$\hat{\mu}_{11}^2$	$\bar{\hat{\mu}}_1 = \frac{\hat{\mu}_{11} + \cdots \hat{\mu}_{13}}{3.}$ → 3个时期 (t)
1	2	$\hat{\mu}_{12}$	⋮	$\bar{\hat{\mu}}_1$
1	3	$\hat{\mu}_{13}$		$\bar{\hat{\mu}}_1$
2	1	$\hat{\mu}_{21}$	⋮	$\bar{\hat{\mu}}_2$
2	2	$\hat{\mu}_{22}$		$\bar{\hat{\mu}}_2$
2	3	$\hat{\mu}_{23}$		$\bar{\hat{\mu}}_2$

same person's $\bar{\hat{\mu}}_i$ are same.

$\hat{\sigma}_\mu^2$ = 这列加总 再求平均.

$\hat{\mu}_{it} - \bar{\hat{\mu}}_i$ in equation (7) = ① - ③ = $\Sigma(\hat{\mu}_{11} - \bar{\hat{\mu}}_1) \cdots$

$$\hat{\sigma}_\epsilon^2 = \frac{1}{NT-N}\sum_{t=1}^{T}\sum_{i=1}^{N}(\hat{\mu}_{it} - \bar{\hat{\mu}}_i)^2 \tag{7}$$

$$\hat{\sigma}_\alpha^2 = \hat{\sigma}_\mu^2 - \hat{\sigma}_\epsilon^2 \tag{8}$$

Notice, random effect require more assumptions especially this one: $E[\alpha_i|x_i] = 0$
fixed effect does not require this one. it mean, alpha and x can be related.
BUT both of them need this:

$E[\epsilon_{it}|x_i, \alpha_i] = 0$

Fixed Effects

we can use fixed effect and IV together if X is endogenous

- In the fixed effects model we do not need the unrelated effect assumption.
- However, we need to ensure that there are no time-invariant regressors in x.
- There are three ways to compute the fixed effects estimator

all of them are unbiased and consistent

  - Demeaned regression this means take away mean
  - Individual Dummies usually use this
  - First Differencing

Demeaned regression always gives you the same result (beta) as individual dummies

normally, when people say individual effect, they mean fixed effect, except they mention it is random effect

random effect: solve error term and run GLS,
fiex effect: try to put the omit variable alpha back to the regression

without this assumption, $E[\alpha_i|x_i] = 0$, in fixed effects, we cannot omit alpha

fixed effect is unbiased and consistent, but not blue.

Fixed Effects

For the demeaning regression, we can just take the time-average of the model

can we do both time fixed effects and dummy variable together? of course yes!
someone may say you will lose degree of freedom. it is really a tricky argument

this is for each individual over time, not a sum of all individuals

$$\bar{y}_i = \frac{1}{T}\sum_t y_{it}$$

$$\bar{x}_{1i} = \frac{1}{T}\sum_t x_{1it}$$

$$\vdots$$

$$\bar{x}_{ki} = \frac{1}{T}\sum_t x_{kit}$$

ID	Time
1	1
1	2
1	3
2	1
2	2
2	3
3	1
3	2
3	3

individual - time effect means each combination has an effect on y. if you have 9 obs, then you need 8 dummy, so you'll never have enough degree of freedom to do. that's why people say cannot do individual and time fixed effect together.

however, if we consider each ID ~~as a~~ and time period separately we have 3 IDs and 3 time period (1, 2 and 3), then we can do it.

$$y_{it} = \beta_0 + \beta_1 x_{1it} \ldots + I(i=2)\alpha_2 + I(i=3)\alpha_3 + I(t=2)\gamma_2 + I(t=3)\gamma_3 + \varepsilon_{it}$$

alpha bar = alpha as for each person alphas are same ie: (5+5+5)/3 = 5

$$\bar{\alpha}_i = \frac{1}{T}\sum_t \alpha_i = \alpha_i \tag{13}$$

$$\tag{14}$$

Fixed Effects

Therefore,

recall, $y_{it} = \beta_0 + x_{1it}\beta_1 + x_{2it}\beta_2 \ldots + x_{kit}\beta_k + \alpha_i + \epsilon_{it}$

$$y_{it} - \bar{y}_i$$
$$= (x_{1it} - \bar{x}_{1i})\beta_1 + (x_{2it} - \bar{x}_{2i})\beta_2 \ldots + (x_{kit} - \bar{x}_{ki})\beta_k$$
$$+ (\alpha_i - \alpha_i) + \epsilon_{it} - \bar{\epsilon}_i$$

now you are regressing $y_{it} - \bar{y}_i$ on $(x_{1it} - \bar{x}_{1i})$, $(x_{2it} - \bar{x}_{2i})$ ....rather y on x
this is a totally different regression, but the math tell us the betas will not change.True beta doesn't change, but the beta hat you get can be a little bit different.

$$y_{it} - \bar{y}_i$$
$$= (x_{1it} - \bar{x}_{1i})\beta_1 + (x_{2it} - \bar{x}_{2i})\beta_2 \ldots + (x_{kit} - \bar{x}_{ki})\beta_k$$
$$+ \epsilon_{it} - \bar{\epsilon}_i$$

Notice: through this method, different with what we recalled, we do not have individual effect(alpha) anymore.

Fixed Effects

Or we can have dummy variables for each individuals:

I is indicator Fn
means, for this I
when i=2,I = 1
otherwise I = 0

$$y_{it} = \beta_0 + x_{1it}\beta_1 + x_{2it}\beta_2 ... + x_{kit}\beta_k$$
$$+ I(i = 2)\alpha_2 + ... + I(i = N)\alpha_N$$
$$+ \epsilon_{it}$$

I = 1 when i = N
I = 0 when O.W

we know how many individual in data, then we add n-1 alphas

Which one is better?

for dummy, we can actually estimate the exact each person's individual effect(alpha),
but demeaning cannot. so if you care more in alpha, use dummy.
But,
if you have 1 million obs, then you have to estimate (1 million - 1) more variables.
demeaning will run faster(but lose 1 degree of freedom for each individual),dummy
give you too much computational burden.obs>100,000, your regression will run slow.
so demeaning is better than dummy variable

by controlling all other X and alpha, for X1,if ob1 and ob2 have different X and Y,
then we can attribute this difference to individual effect.

first differencing is tricky. sometimes biased, sometimes not, because it has a lag term, if the lag period is correlated with the error term, then biased.

Fixed Effects

Finally, we have the first differencing:

$$y_{it} = \beta_0 + x_{1it}\beta_1 + x_{2it}\beta_2 ... + x_{kit}\beta_k + \alpha_i + \epsilon_{it} \tag{15}$$

$$y_{it+1} = \beta_0 + x_{1it+1}\beta_1 + x_{2it+1}\beta_2 ... + x_{kit+1}\beta_k + \alpha_i + \epsilon_{it+1} \tag{16}$$

this eq estimate next period.

The difference between two equations gives:

$$\begin{aligned} & y_{it+1} - y_{it} \\ & = (x_{1it+1} - x_{1it})\beta_1 + (x_{2it+1} - x_{2it})\beta_2 ... + (x_{kit+1} - x_{kit})\beta_k \\ & + \epsilon_{it+1} - \epsilon_{it} \end{aligned}$$

by this method, we lose 1 time period data(ob) for each individual
first differencing is same as demeaning only when we have 2 period obs.
so if you only have 2 period, use first differencing!
but for this method, if you use (Yit+2 - Yit) or (Yit+3 - Yit) the model gives you same betas (slightly different). we don't know which one is better.