Econometrics 3: Assignment I

Laura B. Gomez Cely 2050216

April 2020

1 Question 1

Think about a hypothetical model for capital investment in one particular industry. Data are available for many countries (a country is the cross sectional unit) and a fixed number of T years:

$$log(invest_{it}) = \alpha_i + \gamma_t + z_{it}\beta + \delta_1 tax_{it} + \delta_2 disaster_{it} + \epsilon_{it}$$

The variable tax_{it} measures the marginal tax rate on capital (gains) in a country at time t and $disaster_{it}$ is a dummy variable indicating by value 1 if a significant natural disaster occurred in a country at time t (for example, a major flood, a hurricane, or an earthquake). Variables z_{it} contain other factors affecting capital investment and the t represents different time-specific intercepts.

a) Why is it important to allow for aggregate time effects?
It is important to allow aggregate time affects because we can control factors that are constant for countries but change over time. In this case for example, some periods of time where a global crisis or a pandemic affected all countries around that decrease the general investment.

b)What kinds of variables are captured in individual effects?
The individual effects capture the unobservable heterogeneity between countries such as cultural aspects, governments issues and other unobservable variables that can affect the amount of investment across the countries.

c)Interpreting the equation in a causal fashion, what sign does economic reasoning suggest for δ_1?
I will expect that the sign will be negative because an investor always decide the inversion according to the gains. In this case, if exist a tax rate on the gains of the capital, it will reduce the profit.

d)Explain how you would estimate this model; be specific about the assumptions you are making.
Since α_i could be unobservable characteristics about the government of each

country or cultural aspects, I think this variable could be related with the type of taxes a country could impose. So, following this idea, $E[\alpha_i|x_{i1},...,x_{iT}] \neq 0$. Therefore, in this case, we could use Fixed Effects. The way to eliminate α_i would be within.

$log(invest_{it}-\overline{invest_i}) = (\alpha_i-\alpha_i)+\gamma_t-\overline{\gamma}+(z_{it}-\overline{z_i})\beta+\delta_1(tax_{it}-\overline{tax_i})+ \delta_2(disaster_{it}-\overline{disaster_i})+(\epsilon_{it}-\overline{\epsilon_i})$

$log(inv\tilde{e}st_{it}) = \tilde{\gamma_t}+\tilde{z_{it}}+ta\tilde{x}_{it}+disa\tilde{s}ter_{it}+\tilde{\epsilon_{it}}$

We can estimate this regression through least squares because we suppose:
-Strict exogeneity assumption
$E(\tilde{\epsilon_{it}}|x_{i1},...,x_{iT},_i)=0$

-Full-rank condition has to be satisfied: $rank(E(\tilde{x_{it}}\tilde{x}_i^T))=p$ (Time-invariant variables cannot be included in $\tilde{x}_{it}=x_{it}x_{it}^T$, however as we are not interesting in invariant variables it is not a problem.

2 Question 2

$\Omega = \sigma_\epsilon^2 IT+\sigma_\alpha^2(\iota T\iota T^T)$
$\Omega^{-1} = (\frac{1}{\sigma_\epsilon^2 IT})*(IT-\frac{\sigma_\alpha^2}{\sigma_\epsilon^2+T\sigma_\alpha^2}(\iota T\iota T^T))$
$\Omega^{-1/2} = (\frac{1}{\sigma_\epsilon^2 IT})*(IT-(1-(\frac{\sigma_\epsilon^2}{\sqrt{\sigma_\epsilon^2+T\sigma_\alpha^2}}))(\iota T\iota T^T)$

a)Verify that the expressions for Ω^1 and $\Omega^{-1/2}$ on slide 39 are correct
There is two ways to show it

Verify Ω^{-1}
1) using the definition of an inverse of a matrix

$\Omega^{-1} = \sigma_\epsilon^{-2}IT-\frac{\sigma_\epsilon^{-2}\sigma_\alpha^2\sigma_\epsilon^{-2}}{1+T\sigma_\alpha^2\sigma_\epsilon^{-2}}(\iota T\iota T^T)$
if now we multiply and divide by σ_ϵ^4
$\Omega^{-1} = \sigma_\epsilon^{-2}IT-\frac{\sigma_\alpha^2}{\sigma_\epsilon^4+T\sigma_\alpha^2\sigma_\epsilon^2}(\iota T\iota T^T)$
$\Omega^{-1} = \frac{1}{\sigma_\epsilon^{-2}}(IT-\frac{\sigma_\alpha^2}{\sigma_\epsilon^2+T\sigma_\alpha^2})(\iota T\iota T^T)$

2)using the following expression $\Omega*\Omega^{-1}=IT$

$\Omega*\Omega^{-1} = \sigma_\epsilon^2 IT+\sigma_\alpha^2(\iota T\iota T^T)*(\frac{1}{\sigma_\epsilon^2 IT})*(IT-\frac{\sigma_\alpha^2}{\sigma_\epsilon^2+T\sigma_\alpha^2}(\iota T\iota T^T))$

$\Omega*\Omega^{-1} = \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2}ITIT+\frac{\sigma_\alpha^2}{\sigma_\epsilon^2}(\iota T\iota T^T)-\frac{\sigma_\alpha^2}{\sigma_\epsilon^2+T\sigma_\alpha^2}(\iota T\iota T^T)-\frac{\sigma_\alpha^4}{\sigma_\epsilon^2(\sigma_\epsilon^2+T\sigma_\alpha^2}(\iota T\iota T^T)(\iota T\iota T^T)$

$\Omega*\Omega^{-1} = IT+\frac{\sigma_\alpha^2}{\sigma_\epsilon^2(\sigma_\epsilon^2+T\sigma_\alpha^2}(\iota T\iota T^T)[(\sigma_\epsilon^2 IT+\sigma_\alpha^2)-\sigma_\epsilon^2-T\sigma_\alpha^2]$

Verify $\Omega^{-1/2}$
using the following expression $\Omega^{-1/2} * \Omega^{-1/2} = \Omega^{-1}$
$\Omega^{-1/2} * \Omega^{-1/2} = (\frac{1}{\sigma_\epsilon^2}) * (IT - (1 - (\frac{\sigma_\epsilon^2}{\sqrt{\sigma_\epsilon^2 + T\sigma_\alpha^2}}))(\iota T \iota T^T) * (\frac{1}{\sigma_\epsilon^2 IT}) * (IT - (1 - (\frac{\sigma_\epsilon^2}{\sqrt{\sigma_\epsilon^2 + T\sigma_\alpha^2}}))(\iota T \iota T^T)$

$\Omega^{-1/2} * \Omega^{-1/2} = (\frac{1}{\sigma_\epsilon^2}) * [IT - 2(\frac{\iota T \iota T^T}{T}) + 2(\frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + T\sigma_\alpha^2})(\frac{\iota T \iota T^T}{T}) + ((\frac{\iota T \iota T^T}{T}) - \frac{\sigma_\epsilon^2}{\sqrt{\sigma_\epsilon^2 + T\sigma_\alpha^2}}(\frac{\iota T \iota T^T}{T}))^2$

$\Omega^{-1/2} * \Omega^{-1/2} = (\frac{1}{\sigma_\epsilon^2}) * [IT - 2(\frac{\iota T \iota T^T}{T}) + 2(\frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + T\sigma_\alpha^2})(\frac{\iota T \iota T^T}{T}) + ((\frac{\iota T \iota T^T}{T}) - 2\frac{\sigma_\epsilon}{\sigma_\epsilon^2 + T\sigma_\alpha^2}(\frac{\iota T \iota T^T}{T})) + \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + T\sigma_\alpha^2}(\frac{\iota T \iota T^T}{T}))$

$\Omega^{-1/2} * \Omega^{-1/2} = (\frac{1}{\sigma_\epsilon^2}) * [IT - (\frac{\iota T \iota T^T}{T}) + \frac{\sigma_\epsilon}{\sigma_\epsilon^2 + T\sigma_\alpha^2}(\frac{\iota T \iota T^T}{T})]$

$\Omega^{-1/2} * \Omega^{-1/2} = (\frac{1}{\sigma_\epsilon^2}) * [IT - \frac{\sigma_\epsilon^2 + T\sigma_\alpha^2 + \sigma_\epsilon^2}{\sigma_\epsilon^2 + T\sigma_\alpha^2}(\frac{\iota T \iota T^T}{T})]$

$\Omega^{-1/2} * \Omega^{-1/2} = (\frac{1}{\sigma_\epsilon^2}) * [IT - \frac{\sigma_\alpha^2}{\sigma_\epsilon^2 + T\sigma_\alpha^2}(\frac{\iota T \iota T^T}{T})]$

b. How does the variance matrix $\Omega = var[(\alpha_i + \epsilon_{it})_{t=1}^T | x_{i1}, ..., x_{iT}]$ change if you assume $var(\epsilon_i) = var(\epsilon_i | x_{i1}, ..., x_{iT}) = \Sigma$ (instead of var(ϵ_i) = $\sigma_\epsilon^2 IT$) and $var(\alpha_i) = var(\alpha_i | x_{i1}, ..., x_{iT})$, where $\epsilon_i = (\epsilon_{i1}, ..., \epsilon_{iT})^T$? Describe its general structure.

Given that $var(\epsilon_i) = var(\epsilon_i | x_{i1}, ..., x_{iT}) = \Omega$ instead of $\sigma_\epsilon^2 IT$, the Ω matrix will change in the following aspects.

$(v_{it}) = (\epsilon_{it} + \alpha_i)$
$var(\epsilon_{it} + \alpha_i | x_{i1}, ..., x_{iT})_{t=1}^T = E[(\epsilon_{it} + \alpha_i) * (\epsilon_{it} + \alpha_i) | x_{i1}, ..., x_{iT}] =$
$E(\epsilon_{it}^2) + E(\epsilon_{it}\alpha_i) + E(\epsilon_{it}\alpha_i) + E(\alpha_i^2) =$
$\sigma_\epsilon^2 + \sigma_\alpha^2$

$E[v_{it}, v_{is} | x_{i1}, ..., x_{iT}] = E[(\epsilon_{it} + \alpha_i)(\epsilon_{is} + \alpha_i) | x_{i1}, ..., x_{iT})_{t=1}^T$
$E[v_{it}, v_{is} | x_{i1}, ..., x_{iT}] = E(\epsilon_{it}\epsilon_{is}) + E(\alpha_i^2)$
$E(\epsilon_{it}\epsilon_{is}) = 0$
$E[v_{it}, v_{is} | x_{i1}, ..., x_{iT}] = E(\alpha_i^2)$
Therefore, given the new assumption, the Ω matrix will have in the diagonals $\sigma_\epsilon^2 + \sigma_\alpha^2 = \sigma_\alpha^2(\iota T \iota T^T) + \Sigma$

Off the diagonals, the matrix will have σ_α^2

In conclusion, the matrix will be different under the new assumption because in this case we have autocorrelation because the off-diagonal elements are not equal to zero. Also, the diagonal elements of the matrix changed because the variance between the errors is now Σ different than σ_ϵ^2 and also, include the variance of the individual factor σ_i^2

c)Is the random effects estimator under the assumption of point (b) consistent? Is it efficient, and if not, can you construct a feasible efficient GLS estimator without further assumptions (and how)?

Consistent

It is consistent because the strict exogeneity assumption is still satisfied.
$\hat{\beta} = [X^T[I_n \bigotimes \Omega_{new}]^{-1}X]^{-1}X^T[I_n \bigotimes \Omega_{new}]^{-1}Y$

$\hat{\beta} = \beta + [X^T[I_n \bigotimes \Omega_{new}]^{-1}X]^{-1}X^T[I_n \bigotimes \Omega_{new}]^{-1}v_{it}$

$plim(T^{-1}X^T[I_n \bigotimes \Omega_{new}]^{-1}X]^{-1}) \longrightarrow A_{-1}$

$plim(T^{-1}X^T[I_n \bigotimes \Omega_{new}]^{-1}v_i) \longrightarrow E(X^TV_i) = 0)$

By slutsky theorem $plim(\hat{\beta} = \beta + A_{-1} * 0 = \beta$, so it is consistent

Efficient

However, this estimator is not efficient because in this case we have autocorrelation.To create a feasible estimator first we need to estimate $be\hat{t}a_n = (X^TX)^{-1}(X^TY)$. After that we have to compute the residuals $N^n = \{\hat{v}_i\}_{i=1}^n = Y - X\beta^n$

Then, we construct the matrix $\hat{\Omega}_n = \sum_{i=1}^n \hat{v}_n\hat{v}_n^T/n$. Thus, we can transform the variables X and Y like
$\tilde{Y} = [I_n \bigotimes \hat{\Omega}_n^{1/2}]Y$
$\tilde{X} = [I_n \bigotimes \hat{\Omega}_n^{1/2}]X$

Finally we get the generalized RE estimator
$\hat{\beta} = (\tilde{X}^T\tilde{X})^{-1}(\tilde{X}^T\tilde{Y})$

d)How does the variance matrix $\Omega = var[(\alpha_i + \epsilon_{it}]_{t=1}^T|x_{i1}, ..., x_{iT}]$ change if $var(\alpha_i) \neq var(\alpha_i|x_{i1}, ..., x_{iT})$, but on the other hand $var(\epsilon_i) = var(\epsilon_i|x_{i1}, ..., x_{iT}) = \sigma_\epsilon^2 IT$? Describe its general structure.

Given that $var(\alpha_i) \neq var(\alpha_i|x_{i1}, ..., x_{iT})$, I expect that the variance will change across individuals. So, in this sense, $\Omega = \sigma_\epsilon^2 + \sigma_{\alpha i}^2$. So, in this case the matrix will have heteroscedasticity due to the change of the variance across individuals.

e) Is the random effects estimator under the assumption of point (d) consistent? Is it efficient, and if not, can you construct an efficient GLS estimator without further assumptions (and how)?

This estimator is consistent because the strict assumption still being true $E(v_i, X) = 0$, therefore,

$\hat{\beta} = [X^T[I_n \bigotimes \Omega_{new}]^{-1}X]^{-1}X^T[I_n \bigotimes \Omega_{new}]^{-1}Y$

$\hat{\beta} = \beta + [X^T[I_n \bigotimes \Omega_{new}]^{-1}X]^{-1}X^T[I_n \bigotimes \Omega_{new}]^{-1}v_{it}$

$plim(T^{-1}X^T[I_n \bigotimes \Omega_{new}]^{-1}X]^{-1}) \longrightarrow A_{-1}$

$plim(T^{-1}X^T[I_n \bigotimes \Omega_{new}]^{-1}v_i) \longrightarrow E(X^T(\alpha_i + \epsilon_{it}) = 0)$

By slutsky theorem $plim(\hat{\beta} = \beta + A_{-1} * 0 = \beta$, so it is consistent

However, because of the heteroscedasticity (the diagonal elements of the matrix change across individuals) and autocorrelation, the matrix is not efficient. To create a feasible estimator first we need to estimate $be\hat{t}a_n = (X^TX)^{-1}(X^TY)$. After that we have to compute the residuals $N^n = \{\hat{v}_i\}_{i=1}^n = Y - X\beta^n$

Then, we construct the matrix $\hat{\Omega}_n = \sum_{i=1}^n \hat{v}_n\hat{v}_n^T/n$. Thus, we can transform the variables X and Y like
$\tilde{Y} = [I_n \bigotimes \hat{\Omega}_n^{1/2}]Y$
$\tilde{X} = [I_n \bigotimes \hat{\Omega}_n^{1/2}]X$

Finally we get the generalized RE estimator
$\hat{\beta} = (\tilde{X}^T\tilde{X})^{-1}(\tilde{X}^T\tilde{Y})$

f) How would you estimate the variance matrix of the standard random effect estimator if (b) and/or (d) are valid?
After estimating the $\hat{\beta}$ using the steps from part (c) and (e), this estimate will be asymtotically normal when $n \Longrightarrow \infty$. Therefore, $\hat{\beta}$ has the variance matrix as follows:
$\tilde{y} = \tilde{x}\beta + \tilde{v}$
$\hat{\beta} = \beta + (\tilde{x}^T\tilde{x})^{-1}(\tilde{x}^T\tilde{v})$
$Var\hat{\beta} = (\tilde{x}^T\tilde{x})^{-1}(\tilde{x}^T\tilde{v})(\tilde{v}^T\tilde{x})(\tilde{x}^T\tilde{x})^{-1}$
$Var\hat{\beta} = E[(x^T\hat{\Omega}_i^{-1}x)^{-1})(x^T\hat{\Omega}_i^{-1}v)(v^T\hat{\Omega}_i^{-1}x)(x^T\hat{\Omega}_i^{-1}x)^{-1})|x_{i1}, ..., x_{iT}]$
$= ((x^T\hat{\Omega}_i^{-1}x)^{-1})\sigma^2$

Finally, we would estimate $\hat{\Omega}_i$ like that
$\hat{\Omega}_i = \frac{\sum_{i=1}^n \hat{v}_i\hat{v}_i^T}{n}$

3 Exercise 3

a) Write down the formula of the within-group estimator of this model.

1) $y_{it} = \overline{x}_i\lambda + \alpha_i + \beta x_{it} + \epsilon_{it}$

The average over time
2) $\overline{y}_{it} = \overline{x}_i\lambda + \alpha_i + \beta\overline{x}_i + \overline{\epsilon}_i$

1)-2)

$$y_{it} - \overline{y}_{it} = (x_{it} - \overline{x}_i)\beta + \epsilon_{it} - \overline{\epsilon}_i$$

$\tilde{y}_{it} = \beta\tilde{x}_{it} + \tilde{\epsilon}_{it}$

b) Write down the formula of the pooled instrumental variable estimator of the above model:

$$\alpha_i$$

is then treated as an error term, and because of endogeneity, x_{it} is instrumented by$x_{it}xi$. Show it is identical to the within-group estimator.

$\hat{\beta}_{iv} = \frac{cov(z_{it},y_{it})}{cov(z_{it},x_{it})}$
z_{it} is an instrument for endogenous variable x_{it}

$z_{it} = x_{it} - \overline{xi}$

so,

$$\hat{\beta}_{iv} = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \overline{x}_{it}) * y_{it}}{\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \overline{x}_{it})^T * x_{it}}$$

$y_{it} = \alpha_i^* + x_{it}\beta + \epsilon_{it}$
$y_{it} = \overline{x_i}\lambda + \alpha_i + x_{it}\beta\epsilon_{it}$
$y_i = \beta x_i + IT\alpha^i + \epsilon_{it}$

Transformartion of the model using idempotent matrix $Q = IT - (\frac{\iota T^T \iota T}{T})$ where

$Q * Q^T = Q$ $x_i = Q * x_i = x_i^T - Tx_{it}$
$y_i = Q * y_i = y_i^T Ty_{it}$

$$\beta_{FE} = \frac{\sum_{i=1}^{n}\overline{x}_i\overline{y}_i}{\sum_{i=1}^{n}\overline{x}_i\overline{x}_i}$$

$$\beta_{FE} = \frac{\sum_{i=1}^{n}Qx_iQy_i}{\sum_{i=1}^{n}Qx_iQx_i}$$

$$\hat{\beta}_{iv} = \frac{\sum_{i=1}^{n} \sum_{t=1}^{T} (x_{it} - \overline{x}_{it}) * y_{it}}{\sum_{i=1}^{n} \sum_{t=1}^{T} (x_{it} - \overline{x}_{it})^T * x_{it}}$$

Therefore

$$\hat{\beta}_{iv} = \hat{\beta}_{FE}$$

c)Next, consider the between estimator, that is, the least squares estimator applied to the model $\overline{y}_i = intercept - slope\overline{x}_i + \overline{\epsilon}_i$. What parameters does this estimator identify?

If we rewrite the equation, we know that $\overline{y}_i = \alpha_i - slope\beta + \lambda\overline{x}_i + \overline{\epsilon}_i$ So, in this case, the between estimator can identify α_i, β and λ

d)Substituting for

$$\alpha_i^*$$

in the model equation, $y_{it} = \beta x_{it} + \overline{x}_i\lambda + \alpha_i +_{it}$ parameters β and α can be estimated by the random effect estimator. One can even show that this random effect estimator of β is exactly equal to the within-group estimator (you do not have to prove this). Does this mean there is no real difference between the random effects and fixed effects estimation strategy? Discuss.

The approach between random effect estimator of β and within group estimator is similar in the sense we do not estimate alpha. In the case of within group effects we delete this term including α_i and the invariant part of the x, $\overline{x}_i * \lambda$ and we estimate β. In the second case, we treat

$$\alpha_i$$

as a random variable, but imposing a distribution assumption on it, so in this case the error v_i becomes $\alpha i + \epsilon it$ and $\overline{x}_i$ is treated as a constant and in this way β is estimate.