TA session# 6

Jun Sakamoto

November 15,2018

Contents

1 **Review of Fixed Effect Model** **1**

2 **Review of Random Effect Model** **2**

3 **Hausman test** **4**

1 Review of Fixed Effect Model

Estimation formula by panel data is given by

$$y_{it} = X_{it}\beta + v_i + u_{it} \tag{1}$$

v_i is a individual effect. If the individual effects are all zero, it can be estimated by normal OLS. However, if the individual effect is nonzero and correlates with the explanatory variable, normal OLSE is biased.

$$\begin{aligned} \beta_{OLSE} &= (X'X)^{-1}X'y = (X'X)^{-1}X'(X'\beta + v_i + u_{it}) \\ &= \beta + (X'X)^{-1}X'(v_i + u_{it}) \neq \beta \end{aligned} \tag{2}$$

It is necessary to use a fixed effect model. To estimate the fixed effect model, there are LSDV estimator and Within estimator and these estimators coincide. LSDV is estimated from the following equation.

$$y_{it} = X_{it}\beta + D_i v_i + u_{it} \tag{3}$$

D_i is a dummy variable for each individual. LSDV is a normal OLS estimation with dummy variables added and can be estimated with a simple method. However, when the number of individuals is very large, the number of explanatory variables also increases, so there is a problem that it takes time to calculate the computer.

In order to avoid this problem, Within estimator is used. Within estimator is is obtained from the following equation.

$$\tilde{y}_{it} = \tilde{X}_{it}\beta + u_{it}$$
$$\tilde{y}_{it} = y_{it} - \bar{y}, \tilde{X} = X_{it} - \bar{X}_i$$

$\tilde{y}$ and $\tilde{X}$ are the deviation from the average for each individual. By using the variables thus converted, the same estimate as LSDV can be obtained.

2 Review of Random Effect Model

The fixed effect model allows the unobserved individual effects to be correlated with the included variables. If the individual effect are strictly uncorrelated with the regressors, then it might be appropriate to model the individual specific constant terms as randomly distributed across cross-sectional units. We consider the following regression model.

$$y_{it} = X_{it}\beta + v_i + \epsilon_{it}$$

In contrast to the fixed effect model, we assume the variable v_i is not correlated with the explanatory variables. For simplicity, we assume

$$E(v_i|X) = E(u_{it}|X) = 0 \text{ for all } i$$
$$E(v_i|X) = \sigma_v^2$$
$$V(u_{it}|X) = \sigma_u^2$$
$$Cov(v_i, v_j|X) = 0 \text{ for } i \neq j$$
$$Cov(u_{it}, u_{js}|X) = 0 \text{ for } i \neq j \text{ and } t \neq s$$
$$Cov(v_i, u_{jt}|X) = 0 \text{ for all } i, j \text{ and } t$$

Then,

$$\Omega = E[w_i w_i'] = \begin{pmatrix} \sigma_u^2 + \sigma_v^2 & \sigma_v^2 & \cdots & \sigma_v^2 \\ \sigma_v^2 & \sigma_u^2 + \sigma_v^2 & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ \sigma_v^2 & \cdots & \cdots & \sigma_u^2 + \sigma_v^2 \end{pmatrix}$$

where $w_i = v_i + u_i$. Thus, the variance-covariance matrix is $I_n \bigotimes \Omega$. We can get efficient estimator by the GLS, that is,

$$\hat{\beta} = (\textstyle\sum_{i=1}^{N} X_i \Omega^{-1} X_i)^{-1} (\sum_{i=1}^{N} X_i \Omega^{-1} y_i)$$

"A note on the proper econometric specification of the gravity equation" Peter Egger(2000) Economic letters 66,25-31
Matyas (1997) argued that the correct gravity specification is a three-way model. The corresponding reduced form equation to estimate the world volume of trade in such a model reads

$$X_{ijt} = \beta_0 + \beta_1 RLFAC_{ijt} + \beta_2 GDPT_{ijt} + \beta_3 SIMILAR + \beta_4 DIST + \alpha_i + \gamma_j + \delta_t + u_{ijt}$$

$RLFAC_{ijt} = |ln\frac{K_{jt}}{N_{jt}} - ln\frac{K_{it}}{N_{it}}|$:measures the distance between the two countries in terms of relative factor endowments. ;
$GDPT_{ijt} = ln(GDP_{it} + GDP_{jt})$:The two countries for given relative size and factor endowments. ;
$SIMILAR = ln[1 - (\frac{GDP_{it}}{GDP_{it}+GDP_{jt}})^2 - (\frac{GDP_{jt}}{GDP_{it}+GDP_{jt}})^2]$: The relative size of two countries in terms of GDP. ;
$DIST_{ij}$:The log of the distance variable which is a proxy for transportation costs ;

Result is as bellow.

Table 1
Estimation results[a]

	FEM		REM		OLS	
	β	t	β	t	β	t
RLFAC	0.03	0.9	0.06	2.1[h]	0.14	3.6[h]
GDPT	0.28	4.8[h]	1.01	26.0[h]	1.39	71.4[h]
SIMILAR	−0.02	−0.7	0.34	12.1[h]	0.55	21.1[h]
DIST	−1.08	−48.6[h]	−1.13	−49.2[h]	−1.23	−50.7[h]
CONST.	19.64	13.2[h]	5.41	1.2	−7.47	−12.9[h]
N	2184	–	2184	–	2184	–
$\bar{R}^2$	0.95	–	0.75	–	0.89	–
σ	0.40	–	0.43	–	0.59	–
$\hat{\sigma}_x$	–	–	0.47	–	–	–
$\hat{\sigma}_m$	–	–	0.64	–	–	–
LR-X[b] χ^2	–	1278.9[h]	–	–	–	–
	–	(14)[g]	–	–	–	–
LR-M[c] χ^2	–	893.9[h]	–	–	–	–
	–	(14)[g]	–	–	–	–
LR-T[d] χ^2	–	216.1[h]	–	–	–	80.8[h]
	–	(12)[g]	–	–	–	(12)[g]
Hausman[e] χ^2	–	286.1[h]	–	–	–	–
	–	(16)[g]	–	–	–	–
LM[f] χ^2	–	–	–	–	–	226.4[h]
	–	–	–	–	–	(28)[g]

[a] Note: country and time effects are not reported.
[b] Likelihood ratio test, Greene (1997, p. 161): fixed export effects.
[c] Likelihood ratio test: fixed import effects.
[d] Likelihood ratio test: fixed time effects.
[e] Hausman χ^2 statistic: $(\hat{\beta}_{lsdv} - b_{gls})'\{Var[\hat{\beta}_{lsdv}] - Var[b_{gls}]\}^{-1}(\hat{\beta}_{lsdv} - b_{gls})$, Greene (1997, p. 633).
[f] Breusch–Pagan Lagrange multiplier test, Baltagi (1995), p. 62: Testing for random effects. Note, that the test was computed for the average year:

$$LM_1 = 182/2(M-1)\left[\left(\sum_x \left(1/12 \sum_m u_{xm}\right)^2\right) \Big/ \left(1/12 \sum_x \sum_m u_{xm}^2\right)\right]$$

and

$$LM_2 = 182/2(X-1)\left[\left(\sum_m \left(1/12 \sum_x u_{xm}\right)^2\right) \Big/ \left(1/12 \sum_x \sum_m u_{xm}^2\right)\right]$$

with $LM = LM_1 + LM_2$. As we observe 12 years, the corresponding residuals and residual squares are divided by this number to obtain time averages. X and M are the group sizes for exporters and importers, each 14 in our case.
[g] Degrees of freedom in parenthesis.
[h] Significant at 1%.

This article saying,

A panel framework has many advantages vis-a-vis the cross-section approach. First of all it allows to disentangle country-specific and time-specific effects. The present paper demonstrates that the proper econometric specification of a gravity model in most applications would be one of fixed country and time effects. This was demonstrated by the Hausman χ 2-test and was motivated by the explanation of country effects as widely predetermined because of geographical, historical, or political contexts.

3 Hausman test

Consider the following model.

$$y_{it} = v_i + x_i'\beta + u_{it}$$

where v_i denote the individual effect. If $Cov(v_i, x_{it}) = 0, \forall i, t$ then we have to use random effect model and estimate using not OLS but GLS since GLS is more efficient than OLS. In contrast, if $Cov(v_i, x_{it}) \neq 0, \forall i, t$ then we need to use fixed effect model. Thus we need to decide which model to use. That is

$$H_0 : Cov(v_i, x_{it}) = 0, \forall i, t$$
$$H_1 : Cov(v_i, x_{it}) \neq 0, \forall i, t$$

where H_0 denote null hypothesis. You will remember that

	H_0	H_1
$\hat{\beta}_{FE}$	*Consistent, not efficient*	*consistent*
$\hat{\beta}_{RE}$	*Consistent, efficient*	*not Consistent*

The key idea of Hausman test is that under H_0 both $\hat{\beta}_{FE}$ and $\hat{\beta}_{RE}$ are consistent. Therefore, it is expected that $P(|\hat{\beta}_{FE} - \hat{\beta}_{RE}| > \varepsilon) \to 0$, that is $\hat{\beta}_{FE} - \hat{\beta}_{RE} \to_p 0$. In contrast, if H_1 is correct, then only $\hat{\beta}_{FE}$ is consistent and $\hat{\beta}_{FE} - \hat{\beta}_{RE} \to_p 0$ does not hold. These facts imply that if the difference between $\hat{\beta}_{FE}$ and $\hat{\beta}_{RE}$ is large, then we should use fixed effect model and vice versa. Hausman test is one of wald statistic and given by

$$H := (\hat{\beta}_{FE} - \hat{\beta}_{RE})'(Asy.V[\hat{\beta_{FE}}] - Asy.V[\hat{\beta}_{RE}])^{-1}(\hat{\beta}_{FE} - \hat{\beta}_{RE}) \to_d \chi(K)$$

where $Asy.V[]$ denotes asymptotic variance and $\chi(K)$ denotes a chi-squared distribution with d.f. K.

Empirical Example 1

"Has the crisis affected the behavior of the rating agencies? Panel evidence from the Eurozone" P.Boumparis, C. Milas and T.Panagioditis (2015) Economic letters 136 pp118-124.

This paper revisits the determinants of credit rating decisions for the Eurozone countries and confirm the role of crisis. The regression model is following.

$$CRA_{it} = \alpha_0 + \mu_i + \textstyle\sum_{i=1}^{9} \alpha_i x_{it} + \sum_{i=1}^{9} b_i \bar{x}_{it} + \sum_{j=1}^{3} c_j D_{crisis} x_{jt} + u_{it}$$

where x_{it} includes nine variables, namely GDP per capita, growth rate of GDP, government debt, inflation rate, unemployment rate, current account, external balance, log reserves, regulatory quality. D_{crisis} takes the value of 1 for the years 2009∼ 2013 and 0 otherwise. Three variables (government debt, current account and externalbalance) interact with the crisis dummy in line with Gros (2011) who argues that the external sector was of vital importance duringthe crisis.

Table 3
Credit rating models-Fitch.

Fitch	Pooled OLS				Fixed effects				Random effects			
	coef.	*p*-val	coef.	*p*-val	coef.	*p*-val	coef.	*p*-val	coef.	*p*-val	coef.	*p*-val
Log GDP per capita	10.970	0.000	11.086	0.000	4.039	0.365	8.825	0.017	9.195	0.000	9.370	0.000
Log GDP per capita cavg	−4.644	0.729			5.716	0.589			−1.586	0.874		
GDP growth rate	0.149	0.001	0.149	0.001	0.130	0.000	0.135	0.000	0.152	0.000	0.152	0.000
GDP growth rate cavg	−0.172	0.135	−0.222	0.002	−0.176	0.004	−0.123	0.004	−0.181	0.001	−0.208	0.000
Government debt	−0.032	0.000	−0.033	0.000	−0.043	0.014	−0.040	0.034	−0.024	0.013	−0.024	0.009
Government debt cavg	0.301	0.033	0.270	0.002	0.202	0.121	0.111	0.000	0.251	0.033	0.260	0.007
Inflation rate	−0.281	0.000	−0.287	0.000	−0.107	0.119	−0.154	0.008	−0.177	0.026	−0.179	0.010
Inflation rate cavg	−0.381	0.251	−0.247	0.127	−0.323	0.202	−0.349	0.007	−0.408	0.040	−0.372	0.037
Unemployment rate	−0.142	0.000	−0.143	0.000	−0.218	0.004	−0.200	0.009	−0.180	0.001	−0.174	0.001
Unemployment rate cavg	−0.723	0.297	−0.561	0.085	−0.278	0.649			−0.552	0.336	−0.612	0.118
Current account	−0.013	0.001	−0.013	0.001	0.000	0.985			−0.002	0.735		
Current account cavg	−0.007	0.864			−0.015	0.558			−0.014	0.566		
External balance	−0.062	0.001	−0.062	0.001	−0.008	0.669			−0.070	0.008	−0.079	0.011
External balance cavg	−0.667	0.065	−0.758	0.013	−0.642	0.005			−0.630	0.002	−0.675	0.002
Log reserves	1.439	0.000	1.440	0.000	−0.133	0.568			0.845	0.004	0.951	0.001
Log reserves cavg	−3.421	0.548			−0.955	0.623			−2.527	0.272		
Regulatory quality	0.165	0.659			1.716	0.003	1.953	0.002	1.022	0.059	0.867	0.115
Regulatory quality cavg	12.704	0.093	10.963	0.056	8.350	0.115	10.491	0.034	10.791	0.031	10.260	0.021
Government debt $*D_{crisis}$	−0.036	0.000	−0.036	0.000	−0.027	0.003	−0.024	0.000	−0.034	0.000	−0.034	0.000
Current account $*D_{crisis}$	0.021	0.000	0.022	0.000	0.009	0.248	0.014	0.013	0.012	0.140	0.010	0.043
External balance $*D_{crisis}$	−0.117	0.000	−0.118	0.000	−0.057	0.074	−0.075	0.009	−0.082	0.004	−0.079	0.001
Constant	−1.107	0.904	−63.166	0.000	−28.989	0.598	−37.046	0.053	−17.393	0.748	−50.247	0.000
* Robust standard erros												
R squared – Within					0.877		0.853		0.850		0.844	
Between					0.377		0.638		0.882		0.896	
Overall	0.901		0.901		0.511		0.677		0.872		0.880	
Pesaran cross sectional independence test					−1.19	Pr = 0.2353	1.82	Pr = 0.0691	−1.37	Pr = 0.1696	−1.33	Pr = 0.18
Hausman specification test									54.14	Pr = 0.00		

Note: The estimation is carried out in Stata and the robust standard errors are derived using the *vce(robust)* option in Stata.

In this case, H_0 must be rejected. Notice that since only $\hat{\beta}_{FE}$ is consistent estimator. This results suggest that government debt and cumulative current account exert a stronger positive impact on credit ratings post-2008 compared to the period before.

Empirical Example 2

Check the Hausman test of $\hat{\beta}_{FE}$ and $\hat{\beta}_{RE}$.

1.We generate samples $z \sim U[-2,2]$, $e \sim N(0,1)$ and $u \sim N(0,1)$.(n=4,T=25)
2.Individual effects be given by $\mu_i = 0,1,2,3$.(Fixed effect)
3.X is made by z, μ and e.(We assume true beta is 1)
4.Y is made by X, μ and u.(We assume true beta is 1)
5.We estimate $\hat{\beta}_{FE}$ and $\hat{\beta}_{RE}$.
So, we have to use $\hat{\beta}_{FE}$ in this example.

Estimates of $\hat{\beta}_{FE}$ and $\hat{\beta}_{RE}$ are

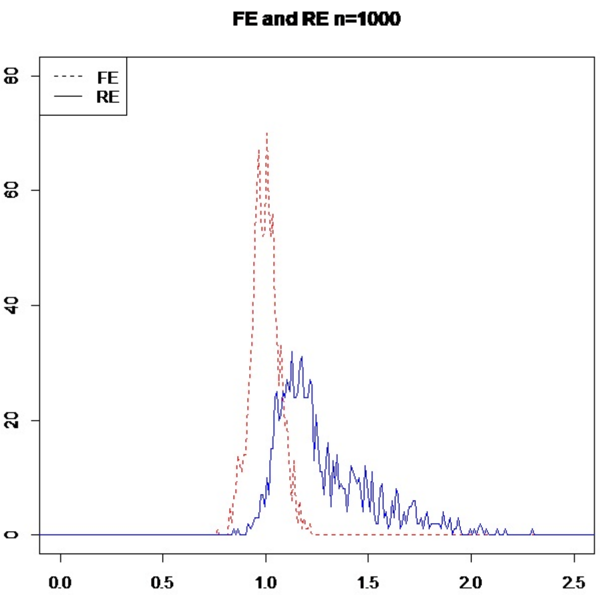


Figure 1: Histogram

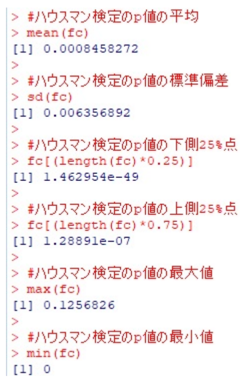

```
> #ハウスマン検定のp値の平均
> mean(fc)
[1] 0.0008458272
>
> #ハウスマン検定のp値の標準偏差
> sd(fc)
[1] 0.006356892
>
> #ハウスマン検定のp値の下側25%点
> fc[(length(fc)*0.25)]
[1] 1.462954e-49
>
> #ハウスマン検定のp値の上側25%点
> fc[(length(fc)*0.75)]
[1] 1.28891e-07
>
> #ハウスマン検定のp値の最大値
> max(fc)
[1] 0.1256826
>
> #ハウスマン検定のp値の最小値
> min(fc)
[1] 0
```

Figure 2: A result of Hausman test

We consider the second situation.

1.We generate samples $X \sim U(-2, 2)$ and $u \sim N(0, 1)$.(n=4,T=25)
2.Individual effects be given by $\mu_i \sim N(0, 1)$.(Rondom effect)
3.y is made by X, μ and u.(We assume true beta is 1)
4.We estimate $\hat{\beta}_{FE}$ and $\hat{\beta}_{RE}$.
So, we have to use $\hat{\beta}_{RE}$ in this example.

Estimates of $\hat{\beta}_{FE}$ and $\hat{\beta}_{RE}$ are

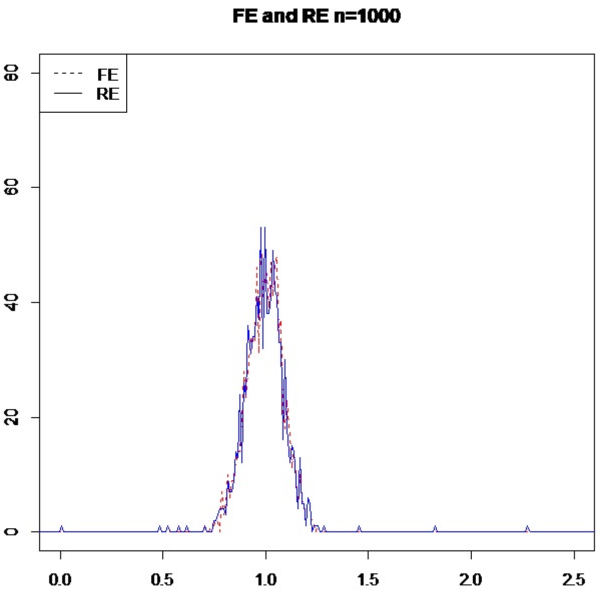
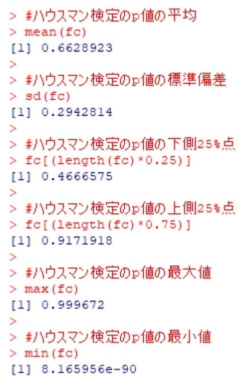


Figure 3: Histogram



```
> #ハウスマン検定のp値の平均
> mean(fc)
[1] 0.6628923
>
> #ハウスマン検定のp値の標準偏差
> sd(fc)
[1] 0.2942814
>
> #ハウスマン検定のp値の下側25%点
> fc[(length(fc)*0.25)]
[1] 0.4666575
>
> #ハウスマン検定のp値の上側25%点
> fc[(length(fc)*0.75)]
[1] 0.9171918
>
> #ハウスマン検定のp値の最大値
> max(fc)
[1] 0.999672
>
> #ハウスマン検定のp値の最小値
> min(fc)
[1] 8.165956e-90
```

Figure 4: A result of Hausman test