Topics in Data Analysis (2012)
Midterm

Name : Han Myoung Park

1.1

	SS	df	MS
Model	51040 = 25520 * 2	2 (K-1)	25520 = 317 * 80.4263
Residual	56057 = 80 * 697	697 (n-K)	80.4263 = 8.9681^2
Total	107098 = 51040 + 56057	699 (n-1)	153 = $\frac{107098}{699}$

N = 700
F(2, 697) = 317
P = .0000
R^2 = .4766 = $\frac{51040}{107098}$
Adj. R^2 = .4751
SEE = 8.981

n = 700, k = 2, K = 3

Is given

1.2 1). H_0: $\beta_1 = \beta_2 = 0$
H_1: At least one of β_1 and β_2 is not zero
2) α = .05
3) F = 317 (2, 697) p-value = .0
4) reject H_0 since 0 < .05
5) At least one of β_1 and β_2 is not zero

1.3

	Parameter estimates	SE	t	95% CI	
Hexp	.0133	.0005 = $\sqrt{2.812e-07}$	25.15 = $\frac{.0133}{.0005}$	.0123	.0144
Index	-.0178	.0282 = $\sqrt{.0008}$	-0.63 = $\frac{-.0178}{.0282}$	-.0732	.0376

.0133 ± 1.96 · .0005
-.0178 ± 1.96 · .0282

1.4 1) H_0: $\beta_1 = 0$ H_a: $\beta_1 \neq 0$
2) α = .05
3) 95% CI = [.0123, .0144]
4) reject H_0 since the hypothesized value 0 does not exist within 95% CI
5) $\beta_1 \neq 0$

2.1 SEE = 8.9904 R^2 = .4770 from LSDV

2.2 69.0909 = 69.2550 + (-.1642)

2.3 .0005319 = .00053 × $\sqrt{\frac{698}{693}}$

2.4 1) H_0: All u_i included in LSDV1 is zero
H_a: At least one u_i in LSDV1 is not zero
2) $\alpha = .05$
3) $F_1 = .13 \quad (4, 693) \quad p: .9697$ (df)
4) Do not reject H_0 since $.9697 > .05$
5) All u_i included in LSDV1 is zero

3.1 1) H_0: All u_i included in LSDV1 is zero
H_a: At least one u_i in LSDV1 is not zero
2) $\alpha = .05$
3) $F_1 = 71.20 \quad (139, 558) \quad$ p-value: $.0000$ (df)
4) Reject H_0 since $.0000 < .05$
5) At least one u_i in LSDV1 is not zero.

3.2 1) $H_0: \sigma_u^2 = 0 \qquad H_a: \sigma_u^2 \neq 0$
2) $\alpha = .05$
3) $LM(\chi^2) = 1023.91 \quad (1) \quad$ p-value: $.0000$ (df)
4) Reject H_0 since $.0000 < .05$
5) $\sigma_u^2 \neq 0$ (significant random effect)

3.3 $\hat{\theta} = .8220 = 1 - \sqrt{\dfrac{\hat{\sigma}_v^2}{T\hat{\sigma}_u^2 + \hat{\sigma}_v^2}} = 1 - \sqrt{\dfrac{(3.3860)^2}{5 \cdot (8.3683)^2 + (3.3860)^2}}$

3.4 $\rho = .8595 = \dfrac{(8.3683)^2}{(8.3683)^2 + (3.3860)^2}$

σ_u^2 accounts for 86 percent of the composite error variance, and thus indicates a significant random effect in the data.

3.5 1) $H_0: E(\varepsilon_i|X) = 0$ Gauss-Markov assumption 2
$H_a: E(\varepsilon_i|X) \neq 0$ (individual effects are correlated with regressors)
2) $\alpha = .05$
3) $Hausman(\chi^2) = 8.64 \quad (2) \quad$ p-value: $.0133$ (df)
4) Reject H_0 since $.0133 < .05$
5) $E(\varepsilon_i|X) \neq 0$ Individual effects are correlated with at least one regressor. A fixed effect model is better

4.1 Chow (F1) = 18.5633

$$= \frac{(e'e - \sum e_i'e_i)/(n-1)(k+1)}{\sum e_i'e_i / n(T-k-1)}$$

$e'e : 56059, \sum e_i'e_i = 1959$

$n = 160 \quad T = 5 \quad k = 2$

$$= \frac{(56059 - 1959)/(160-1)(2+1)}{1959 / 160(5-2-1)}$$

4.2 1) $H_0 : \beta_{ik} = \beta_k$ H_a = At least one $\beta_{ik} \neq \beta_k$

2) $\alpha = .05$

df

3) Chow (F1) = 18.5633 (477, 280) p-value : .0000

4) Reject H_0 since .0000 < .05

5) $\beta_{ik} \neq \beta_k$ (Each group has different parameter estimates)

4.3
— 2.4 says no significant fixed time effect (Fixed time effect X)
— 3.1 says a significant fixed effect (pooled OLS X)
— 3.2 says a significant random effect (pooled OLS X)
— 3.5 says a fixed effect model is better than its random counterpart (Random effect model X)
— 4.2 says data are not poolable (Both fixed/random effect models are problematic because they assume constant slopes across groups.

Therefore, my model is a hierarchical linear regression model.

4.4 I need to fit a hierarchical linear model first to get parameter estimates.