



UJIAN AKHIR SEMESTER GASAL 2019/2020

Mata Kuliah : **Ekonometrika 2 (CSPD) - C**
Program : S1 IEI
Tanggal Ujian : Selasa, 17 Desember 2019
Waktu : 150 Menit (2 ½ Jam)
Sifat Ujian : Closed Book/Closed Notes
Tim Pengajar : I Dewa Gede Karma Wisana
Aufa Doarest/Dhaniel Ilyas
Djoni Hartono

Petunjuk Umum:

1. Perhatikan, lembar soal ini terdiri dari **8 Halaman dan 4 Soal**. Bobot penilaian untuk setiap soal tersedia di akhir pertanyaan.
2. **Boleh** menggunakan kalkulator. **Tidak boleh** menggunakan telepon seluler/komputer/laptop sebagai kalkulator.
3. Sifat ujian adalah tutup buku (*closed book*) dan tutup catatan (*closed notes*).

Pertanyaan 1 (25 poin):

Seseorang bisa memilih 2 jenis asuransi atau tidak memiliki asuransi sama sekali. Preferensi mengenai asuransi dari 616 sample dapat dilihat pada tabel di bawah ini:

insure	Freq.	Percent	Cum.
Indemnity	294	47.73	47.73
Prepaid	277	44.97	92.69
Uninsure	45	7.31	100.00
Total	616	100.00	

Seorang peneliti kemudian berusaha untuk melihat korelasi antara karakteristik demografi dan jenis asuransi yang digunakan.

Anda diminta untuk membantu peneliti tersebut dengan menjawab beberapa pertanyaan di bawah ini.



- a. Apa metode ekonometrika yang paling tepat untuk menjawab pertanyaan di atas? Jelaskan
[5 poin]

Jika kemudian peneliti memilih untuk menggunakan multinomial logit, hasil estimasi dapat dilihat pada tabel di bawah ini.

```
. mlogit insure nonwhite
```

```
Iteration 0: log likelihood = -556.59502
Iteration 1: log likelihood = -551.78935
Iteration 2: log likelihood = -551.78348
Iteration 3: log likelihood = -551.78348
```

```
Multinomial logistic regression      Number of obs   =      616
                                      LR chi2(2)      =       9.62
                                      Prob > chi2     =     0.0081
Log likelihood = -551.78348          Pseudo R2      =     0.0086
```

insure	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Indemnity	(base outcome)					
Prepaid						
nonwhite	.6608212	.2157321	3.06	0.002	.2379942	1.083648
_cons	-.1879149	.0937644	-2.00	0.045	-.3716896	-.0041401
Uninsure						
nonwhite	.3779586	.407589	0.93	0.354	-.4209011	1.176818
_cons	-1.941934	.1782185	-10.90	0.000	-2.291236	-1.592632



. mlogit, rrr

```
Multinomial logistic regression      Number of obs   =      616
                                      LR chi2(2)      =      9.62
                                      Prob > chi2     =      0.0081
Log likelihood = -551.78348          Pseudo R2      =      0.0086
```

insure	RRR	Std. Err.	z	P> z	[95% Conf. Interval]	
Indemnity (base outcome)						
Prepaid						
nonwhite	1.936382	.4177397	3.06	0.002	1.268702	2.955442
_cons	.8286853	.0777011	-2.00	0.045	.6895682	.9958685
Uninsure						
nonwhite	1.459302	.5947956	0.93	0.354	.656455	3.244036
_cons	.1434263	.0255612	-10.90	0.000	.1011414	.2033896

Note: _cons estimates baseline relative risk for each outcome.

- Apa perbedaan antara estimasi dengan menggunakan *relative risk ratio* (rrr) dan tanpa rrr? Jelaskan. Apakah hasil akan berbeda kalo kita menggunakan base yang berbeda. [5 poin]
- Jelaskan intepretasi dari hasil regresi dari kedua tabel di atas. [5 poin]
- Hitung *predicted probability* untuk menggunakan prepaid insurance bagi white dan nonwhite [5 poin]
- Hitung *predicted relative risk* dari memilih asuransi indemnity dibandingkan prepaid (bagi whites) [5 poin]

Pertanyaan 2 (25 poin):

Seorang peneliti berusaha melihat faktor yang menentukan upah pekerja wanita. Dia memiliki 753 observasi, namun dia hanya bisa mengobservasi upah untuk 428 wanita yang bekerja, sementara 325 lainnya tidak bisa diobservasi karena merupakan sample wanita yang tidak bekerja. Dengan informasi yang tersedia ini, peneliti kemudian mengestimasi persamaan regresi dan beranggapan bahwa upah merupakan fungsi dari pendidikan, pengalaman (*exper*) dan kuadrat dari pengalaman (*expersq*). Hasil estimasi dari persamaan ini adalah:



```
. reg lwage educ exper expersq
```

Source	SS	df	MS	Number of obs =	428
Model	35.0222967	3	11.6740989	F(3, 424) =	26.29
Residual	188.305144	424	.444115906	Prob > F =	0.0000
Total	223.327441	427	.523015084	R-squared =	0.1568
				Adj R-squared =	0.1509
				Root MSE =	.66642

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
educ	.1074896	.0141465	7.60	0.000	.0796837 .1352956
exper	.0415665	.0131752	3.15	0.002	.0156697 .0674633
expersq	-.0008112	.0003932	-2.06	0.040	-.0015841 -.0000382
_cons	-.5220406	.1986321	-2.63	0.009	-.9124667 -.1316144

- a. Berdasarkan hasil estimasi di atas, tuliskan bentuk persamaan regresinya? Jelaskan, apakah terdapat potensi masalah yang timbul dari estimasi di atas? [8 poin]

Selanjutnya, mengingat jumlah wanita yang tidak bekerja sekitar 43% dari total sampel, sang peneliti mengestimasi model tersebut dengan Heckit. Dia beranggapan bahwa upah merupakan fungsi dari pendidikan, pengalaman (*exper*) dan kuadrat dari pengalaman (*expersq*). Sementara variable yang terdapat dalam *selection equation* adalah; *nwifeinc*, *educ*, *exper*, *expersq*, *age*, *kidslt6*, dan *kidsage6*. Hasil estimasi dengan Heckit diberikan di bawah ini:





```
. heckman lwage educ exper expersq, select( nwifeinc educ exper expersq age kidslt6 kidsge6) twostep
```

```
Heckman selection model -- two-step estimates      Number of obs      =      753
(regression model with sample selection)          Censored obs       =      325
                                                    Uncensored obs     =      428

                                                    Wald chi2(3)       =      51.53
                                                    Prob > chi2        =      0.0000
```

lwage	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
lwage						
educ	.1090655	.015523	7.03	0.000	.0786411	.13949
exper	.0438873	.0162611	2.70	0.007	.0120163	.0757584
expersq	-.0008591	.0004389	-1.96	0.050	-.0017194	1.15e-06
_cons	-.5781032	.3050062	-1.90	0.058	-1.175904	.019698
select						
nwifeinc	-.0120237	.0048398	-2.48	0.013	-.0215096	-.0025378
educ	.1309047	.0252542	5.18	0.000	.0814074	.180402
exper	.1233476	.0187164	6.59	0.000	.0866641	.1600311
expersq	-.0018871	.0006	-3.15	0.002	-.003063	-.0007111
age	-.0528527	.0084772	-6.23	0.000	-.0694678	-.0362376
kidslt6	-.8683285	.1185223	-7.33	0.000	-1.100628	-.636029
kidsge6	.036005	.0434768	0.83	0.408	-.049208	.1212179
_cons	.2700768	.508593	0.53	0.595	-.7267473	1.266901
mills						
lambda	.0322619	.1336246	0.24	0.809	-.2296376	.2941613
rho	0.04861					
sigma	.66362875					

Berdasarkan hasil estimasi di atas:

- b. Apakah keputusan untuk melakukan regresi dengan Heckit keputusan yang tepat? Berikan argumentasi anda (Hint: gunakan salah satu nilai koefisien regresi yang tersedia dalam tabel di atas) [8 poin]
- c. Interpretasikan hasil yang tertera pada tabel di atas, utamanya pada variabel yang menjelaskan upah [9 poin]





Pertanyaan 3 (25 poin):

Berdasarkan data, diperoleh informasi bahwa terdapat 753 wanita yang sudah menikah dan termasuk informasi mengenai data jam kerjanya. Dari 753 wanita, 428 adalah sudah bekerja dan tentunya memiliki jam kerja, sedangkan 325 wanita lainnya memiliki jam kerja 0 jam. Hasil estimasi dengan metode estimasi OLS dan MLE, diberikan pada tabel berikut ini:

Dependent Variable: <i>hours</i>		
Independent Variables	Linear (OLS)	Tobit (MLE)
<i>nwifeinc</i>	-3.45 (2.54)	-8.81 (4.46)
<i>educ</i>	28.76 (12.95)	80.65 (21.58)
<i>exper</i>	65.67 (9.96)	131.56 (17.28)
<i>exper</i> ²	-.700 (.325)	-1.86 (0.54)
<i>age</i>	-30.51 (4.36)	-54.41 (7.42)
<i>kidslt6</i>	-442.09 (58.85)	-894.02 (111.88)
<i>kidsge6</i>	-32.78 (23.18)	-16.22 (38.64)
<i>constant</i>	1,330.48 (270.78)	965.31 (446.44)
Log-likelihood value	—	-3,819.09
<i>R</i> -squared	.266	.274
$\hat{\sigma}$	750.18	1,122.02

Sumber: Wooldridge 2013, Halaman 601.

Berdasarkan tabel di atas:

- Apa yang anda ketahui mengenai Tobit regression? [7 poin]
- Apa interpretasi dari variabel pendidikan (*educ*)? [5 poin]



- c. Metode estimasi manakah yang lebih tepat digunakan untuk persoalan di atas? Jelaskan. [7 poin]
- d. Dalam hal variabel pendidikan (*educ*) dan *kidslt6* (jumlah anak yang umurnya kurang dari 6 tahun), apa yang bisa dikatakan mengenai nilai koefisien regresi hasil estimasi OLS dan MLE? [6 poin]

Pertanyaan 4 (25 poin):

Permasalahan yang akan dibahas dalam soal ini adalah analisis tentang faktor-faktor yang mempengaruhi berat bayi lahir rendah (*low birth weight*) dan fokus studi ini adalah dampak Program Bantuan Bagi Keluarga dengan Anak Dengan Kondisi Khusus. Berdasarkan data tahun 1987 dan 1990, sebuah model pada level negara disusun sebagai berikut:

$$\begin{aligned} lowbrth_{it} = & \beta_0 + \beta_1 d90_t + \beta_2 afdcprc_{it} + \beta_3 lphypc_{it} + \beta_4 lbedspc_{it} + \dots \\ & + \beta_5 lpcinc_{it} + \beta_6 lpopul_{it} + \alpha_i + \varepsilon_{it}, \end{aligned}$$

dimana *lowbrth* adalah persentase kelahiran yang diklasifikasi sebagai berat lahir rendah; dan variabel independen utama yang menjadi perhatian adalah *afdcprc* (persentase penduduk yang menerima program peningkatan kesejahteraan keluarga, *Aid to Families with Dependent Children* (AFDC)). Variabel lain yang digunakan dan bertindak sebagai kontrol terhadap kualitas kesehatan dan tingkat pendapatan adalah *lphypc* (logaritma dari jumlah tenaga medis *-physicians-* per kapita), *lbedspc* (logaritma dari jumlah tempat tidur di rumah sakit per kapita), *lpcinc* (logaritma dari pendapatan per kapita), *lpopul* (logaritma dari jumlah penduduk), dan *d90* adalah *dummy* tahun. Catatan mengenai program AFDC yang penting adalah bahwa keikutsertaan dalam program ini menyebabkan perempuan dari kelompok miskin berhak untuk memperoleh program nutrisi tambahan dan pelayanan kehamilan (*prenatal care*).

- [a] Apakah yang anda ketahui dengan data panel? [5 poin]
- [b] Apa keunggulan dan kelemahan penggunaan data panel? [5 poin]
- [c] Apakah tanda yang idealnya diharapkan dari parameter estimasi variabel *afdcprc* dan *pcinc*? [5 poin]
- [d] Dengan melihat semua hasil pengolahan data. Estimator untuk parameter β yang manakah yang Anda lebih pilih, di antara kedua metode tersebut (*random effect* dan *fixed effect*)? Jelaskan argumen Anda berdasarkan penilaian yang obyektif! [10 poin]



Hasil Pengolahan DATA

. xtreg lowbrth d90 afdcprc lphyc lbedspc lpcinc lpopul, re

Random-effects GLS regression
Group variable: state

Number of obs = 100
Number of groups = 50

R-sq: within = 0.2018
between = 0.1957
overall = 0.1955

Obs per group: min = 2
avg = 2.0
max = 2

Random effects u_i ~ Gaussian
corr(u_i, X) = 0 (assumed)

Wald chi2(6) = 22.67
Prob > chi2 = 0.0009

lowbrth	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
d90	.4936745	.1904448	2.59	0.010	.1204096	.8669393
afdcprc	-.0950182	.0795823	-1.19	0.232	-.2509967	.0609604
lphyc	.3186868	.8837395	0.36	0.718	-1.413411	2.050784
lbedspc	.4257289	.5284831	0.81	0.420	-.610079	1.461537
lpcinc	-1.591219	.9883224	-1.61	0.107	-3.528295	.3458577
lpopul	.76624	.8512532	0.90	0.368	-.9021856	2.434666
_cons	19.22492	8.181265	2.35	0.019	3.189935	35.2599
sigma_u	1.0494018					
sigma_e	.18464549					
rho	.96997022	(fraction of variance due to u_i)				

. xtreg lowbrth d90 afdcprc lphyc lbedspc lpcinc lpopul, fe

Fixed-effects (within) regression
Group variable: state

Number of obs = 100
Number of groups = 50

R-sq: within = 0.3839
between = 0.1741
overall = 0.1679

Obs per group: min = 2
avg = 2.0
max = 2

corr(u_i, Xb) = -0.9394

F(6, 44) = 4.57
Prob > F = 0.0011

lowbrth	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
d90	.1060055	.3090658	0.34	0.733	-.5168757	.7288868
afdcprc	-.176075	.0903732	-1.95	0.058	-.3582102	.0060601
lphyc	5.894553	2.816694	2.09	0.042	.2178803	11.57123
lbedspc	-1.576216	.8852147	-1.78	0.082	-3.360249	.207817
lpcinc	-.8454894	1.356771	-0.62	0.536	-3.579883	1.888904
lpopul	3.441163	2.872176	1.20	0.237	-2.347328	9.229655
_cons	-4.014527	22.97884	-0.17	0.862	-50.32534	42.29628
sigma_u	3.0975317					
sigma_e	.18464549					
rho	.99645917	(fraction of variance due to u_i)				

F test that all u_i=0: F(49, 44) = 59.46 Prob > F = 0.0000



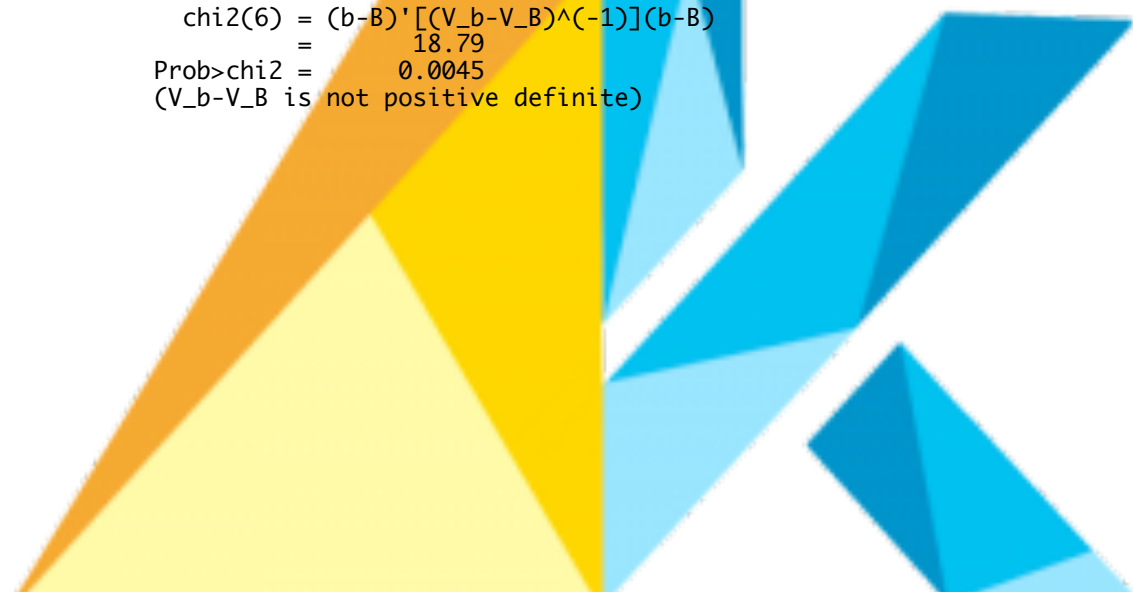
. hausman fixed

	---- Coefficients ----		(b-B)	sqrt(diag(V_b-V_B))
	(b)	(B)	Difference	S.E.
	fixed	.		
d90	.1060055	.4936745	-.3876689	.2434183
afdcprc	-.176075	-.0950182	-.0810569	.0428248
lphyc	5.894553	.3186868	5.575867	2.674466
lbedspc	-1.576216	.4257289	-2.001945	.7101483
lpcinc	-.8454894	-1.591219	.7457292	.9295416
lpopul	3.441163	.76624	2.674923	2.743131

b = consistent under Ho and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(6) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 18.79
Prob>chi2 = 0.0045
(V_b-V_B is not positive definite)



Kanopi FEBUI
Unity in Development