

Nama : Yudhapratama Nugraha
NIM : 24021302

Homework II

Soal :

1. Lakukan estimasi model probabilitas linear dimana dependen variable adalah *chjob* dengan independent variable meliputi *hi*, *log(hwage)*, *hi * othhi*, *othhi*, *educ*, *exper*, *union*, *nfam*, dan *winc*.

- a. Apakah anda menemukan bukti bahwa terdapat fenomena job-lock pada spesifikasi ini? Tunjukan dengan perintah STATA dan outputnya

Jawaban :

- Penggunaan Linear Probability Model

Pertama, dengan melakukan perintah ; regress chjob loghwage othhi educ exper union nfam winc. Maka akan muncul table regresi sebagai berikut :

```
regress chjob loghwage hixothhi othhi educ exper union nfam winc
```

Source	SS	df	MS	Number of obs	=	2,978
				F(8, 2969)	=	15.30
Model	12.5274956	8	1.56593695	Prob > F	=	0.0000
Residual	303.953364	2,969	.10237567	R-squared	=	0.0396
				Adj R-squared	=	0.0370
Total	316.48086	2,977	.106308653	Root MSE	=	.31996

chjob	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
loghwage	-.0372481	.0116856	-3.19	0.001	-.0601609 -.0143353
hixothhi	-.1147447	.0201235	-5.70	0.000	-.154202 -.0752873
othhi	.0964776	.0187241	5.15	0.000	.0597641 .1331911
educ	-.0062467	.0023419	-2.67	0.008	-.0108386 -.0016548
exper	-.0039465	.0007195	-5.48	0.000	-.0053573 -.0025357
union	-.0671297	.0139016	-4.83	0.000	-.0943874 -.039872
nfam	-.0093936	.0051441	-1.83	0.068	-.0194799 .0006926
winc	3.50e-07	4.62e-07	0.76	0.449	-5.57e-07 1.26e-06
_cons	.3996513	.043673	9.15	0.000	.3140189 .4852836

Kedua, output regresi diproses dengan menggunakan excel. Probabilitas dihitung dengan mengkalkulasi kemungkinan pindah kerja apabila nilai variable predictor dalam nilai rata-ratanya.

Var	Beta	Mean (X)	Beta*X
loghwage	-0.0372481	2,2954	-0,0855

hixothhi	-.1147447	0,2817	-0,0323
othhi	.0964776	0,4060	0,0392
educ	-.0062467	12,8761	-0,0804
exper	-.0039465	19,1756	-0,0757
union	-.0671297	0,2482	-0,0167
nfam	-.0093936	3,3559	-0,0315
winc	0,0000	11969,4300	
_cons	.3996513	1,0000	0,3997
Probabilitas			0,1167

-Hasil dari regresi tidak bisa langsung serta merta diinterpretasikan, berdasarkan perhitungan melalui excel, apabila kondisi variable pada nilai rata ratanya maka peluang pindah kerja terjadi sebesar 10 %

- b. Tunjukkan bahwa model probabilitas linear tidak cukup baik untuk digunakan sebagai model estimasi (tunjukkan kelemahan model ini ketika menggunakan variabel dependen biner)

Jawaban :

Model linear tidak cukup baik sebagai model estimasi dengan kasus variable dependen bersifat biner. Hal ini dapat dilihat melalui pencarian probabilitas Ketika subjeknya memiliki variable ekstrim. Sebagai contoh :

Var	BETA	extreme(X)	ExtremexBeta
loghwage	-0,0372	5,2591	-0,1959
hixothhi	-0,1147	1,0000	-0,1147
othhi	0,0965	0,0000	0,0000
educ	-0,0062	0,0000	0,0000
exper	-0,0039	0,0000	0,0000
union	-0,0671	0,0000	0,0000
nfam	-0,0094	12,0000	-0,1127
winc	0,0000	1,0000	0,0000
_cons	0,3997	1,0000	0,3997
Probabilitas			-0,0237

Hasil perhitungan menunjukan nilai yang **tidak dimungkinkan**. Harusnya nilai probabilitas **antara (0-1)**

Selain itu, bisa dilakukan hettest

```
. hettest  
  
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity  
Ho: Constant variance  
Variables: fitted values of chjob  
  
chi2(1)      =    295.97  
Prob > chi2  =    0.0000
```

2. Gunakan model regresi probit untuk menjawab pertanyaan-pertanyaan berikut:
- Lakukan estimasi model probit dimana dependen variable adalah *chjob* dengan independent variable meliputi *hi*, *log(hwage)*, *othhi*, *educ*, *exper*, *union*, *nfam*, dan *winc*. Apakah anda menemukan bukti adanya fenomena job-lock? Jelaskan!
 - Tambahkan variable *othhi*hi* pada daftar variable independent dari pertanyaan (2a), dan lakukan re-estimasi model probit anda. Apakah estimasi anda tentang efek dari asuransi kesehatan berbasis pekerjaan terhadap perubahan pekerjaan berubah? Bagaimana perubahannya dan bagaimana perubahan ini berhubungan dengan hasil yang anda harapkan (jelaskan secara spesifik pada apa arti koefisien dari variable *othhi*hi* dalam model probit ini?
 - Lakukan estimasi (2b) diatas dengan menggunakan model logit. Apakah anda menemukan bukti adanya job-lock secara significant?

Jawaban :

- Penggunaan model Probit untuk estimasi. (tanpa variable *otthi*hi*)

```

. probit chjob hi loghwage othhi educ exper union nfam winc

Iteration 0:   log likelihood = -1097.9516
Iteration 1:   log likelihood = -1011.8755
Iteration 2:   log likelihood = -1010.3856
Iteration 3:   log likelihood = -1010.3836
Iteration 4:   log likelihood = -1010.3836

Probit regression               Number of obs   =       2,978
                               LR chi2(8)         =       175.14
                               Prob > chi2         =       0.0000
Log likelihood = -1010.3836     Pseudo R2       =       0.0798

```

chjob	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
hi	-.6263661	.0673109	-9.31	0.000	-.7582931	-.4944391
loghwage	-.1106355	.0617509	-1.79	0.073	-.231665	.0103941
othhi	.0543906	.067523	0.81	0.421	-.0779521	.1867333
educ	-.0160879	.0125098	-1.29	0.198	-.0406067	.0084308
exper	-.0171767	.0037389	-4.59	0.000	-.0245049	-.0098485
union	-.3493616	.0834693	-4.19	0.000	-.5129584	-.1857649
nfam	-.0418398	.027206	-1.54	0.124	-.0951625	.0114829
winc	1.32e-06	2.35e-06	0.56	0.573	-3.29e-06	5.93e-06
_cons	.1907169	.2234559	0.85	0.393	-.2472486	.6286824

Untuk mendapatkan hasil yang dapat diinterpretasikan, perlu pengolahan data lebih lanjut. Karena hasil regresinya hanya bisa mengindikasikan bahwa variable variable tersebut berkorelasi positif/negative dan signifikan atau tidak

Tanpa Var			
Othhi*hi			
Var	Beta	Mean (x)	Beta*X
hi	-0,62637	0,7518469	-0,47093
loghwage	-0,11064	2,29539	-0,25395
othhi	0,054391	0,4059772	0,022081
educ	-0,01609	12,876090	-0,20715
exper	-0,01718	19,175620	-0,32937
union	-0,34936	0,2481531	-0,0867
nfam	-0,04184	3,35594	-0,14041
winc	1,32E-06	11969,43	0,0158
_cons	0,190717	1	0,190717
Nilai z			-1,25992

Probability			18%
-------------	--	--	-----

Setelah perhitungan melalui excel dengan menghitung semua variable pada keadaan rata ratanya dalam mendapatkan nilai z. Nilai z lalu dihitung dengan fungsi :

=NORM.DIST (nilai z; Mean; standard_dev; cumulative)

Sampai sini bisa diinterpretasikan, bahwa probabilitas pindah kerja seseorang tanpa adanya variable otthi*hi adalah sebesar 18 %

b.) Penggunaan model Probit untuk estimasi. (dengan variable otthi*hi)

Probit regression		Number of obs		=	2,978	
		LR chi2(9)		=	177.16	
		Prob > chi2		=	0.0000	
Log likelihood = -1009.3715		Pseudo R2		=	0.0807	
chjob	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
hi	-.7168023	.0925523	-7.74	0.000	-.8982015	-.535403
loghwage	-.1028264	.06212	-1.66	0.098	-.2245795	.0189267
othhi	-.0648858	.1077053	-0.60	0.547	-.2759843	.1462126
educ	-.0139419	.0126142	-1.11	0.269	-.0386652	.0107814
exper	-.0167485	.0037536	-4.46	0.000	-.0241055	-.0093916
union	-.3325823	.0842249	-3.95	0.000	-.4976601	-.1675045
nfam	-.0406022	.0272035	-1.49	0.136	-.0939201	.0127156
winc	1.04e-06	2.37e-06	0.44	0.659	-3.60e-06	5.69e-06
hixothhi	.1897743	.1333403	1.42	0.155	-.071568	.4511166
_cons	.1943641	.2233577	0.87	0.384	-.2434089	.632137

Sama seperti sebelumnya, Untuk mendapatkan hasil yang dapat diinterpretasikan, perlu pengolahan data lebih lanjut. Karena hasil regresinya hanya bisa mengindikasikan bahwa variable variable tersebut berkorelasi positif/negative dan signifikan atau tidak. Maka ;

Var	Beta	Mean (x)	Beta*X
hi	-0,7168	0,751847	-0,53893
loghwage	-0,10283	2,29539	-0,23603
othhi	-0,06489	0,405977	-0,02634
educ	-0,01394	12,876090	-0,17952
exper	-0,01675	19,175620	-0,32116
union	-0,33258	0,248153	-0,08253
nfam	-0,0406	3,35594	-0,13626
winc	1,04E-06	11969,43	0,012448
hixothhi	0,189774	0,281733	0,053466
_cons	0,194364	1	0,194364
Nilai z			-1,26049
Probability			10%

Setelah perhitungan dengan menambahkan variable hi*otthi, terjadi perubahan. probabilitas menunjukan angka **10%**. maka yang terjadi adalah **pengurangan persentase probabilitas sebanyak 8%** .

c.) Penggunaan model Logit

```
. logit chjob hi loghwage othhi educ exper union nfam winc hixothhi
```

Iteration 0: log likelihood = -1097.9516
Iteration 1: log likelihood = -1021.6777
Iteration 2: log likelihood = -1009.6779
Iteration 3: log likelihood = -1009.6476
Iteration 4: log likelihood = -1009.6476

Logistic regression

Number of obs	=	2,978
LR chi2(9)	=	176.61
Prob > chi2	=	0.0000
Pseudo R2	=	0.0804

Log likelihood = -1009.6476

chjob	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
hi	-1.298916	.1683449	-7.72	0.000	-1.628866 - .9689665
loghwage	-.1880759	.1173028	-1.60	0.109	-.4179852 .0418334
othhi	-.0843305	.1858944	-0.45	0.650	-.4486767 .2800158
educ	-.0259127	.0237662	-1.09	0.276	-.0724936 .0206681
exper	-.032056	.0071388	-4.49	0.000	-.0460478 -.0180642
union	-.642309	.1685657	-3.81	0.000	-.9726916 -.3119263
nfam	-.069735	.0507982	-1.37	0.170	-.1692978 .0298277
winc	1.77e-06	4.38e-06	0.40	0.686	-6.81e-06 .0000104
hixothhi	.3207745	.2432066	1.32	0.187	-.1559018 .7974507
_cons	.4856604	.4120704	1.18	0.239	-.3219828 1.293304

Kemudian, output regresi diproses pada excel. Perbedaananya, perhitungan melalui excel dengan menghitung semua variable pada keadaan rata ratanya dalam mendapatkan nilai z. Nilai z lalu dihitung dengan fungsi :

$$=1/(1+EXP(\text{nilai } z))$$

Var	Beta	Mean X	Beta*Mean
hi	-1,2989	0,751847	-0,97659
loghwage	-0,1880759	2,29539	-0,43171
othhi	-0,0843305	0,405977	-0,03424
educ	-0,0259127	12,876090	-0,33365
exper	-0,032056	19,17562	-0,61469
union	-0,642309	0,248153	-0,15939
nfam	-0,069735	3,35594	-0,23403
winc	1,77E-06	11969,43	0,021186
hixothhi	0,3207745	0,281733	0,090373

_cons	0,4856604	1	0,48566
Nilai z			-2,18708
Probability			90%

Ternyata, setelah menggunakan model logit, terjadi perubahan yang signifikan. Probabilitas pindah kerja menunjukkan angka 90%. Hal ini kemungkinan terjadi kesalahan pada proses perhitungan yang dilakukan oleh penulis. Karena, setelah menggunakan Adjusted prediction dengan melakukan command : Margins, atmeans post pada STATA. Maka hasil yang diberikan tidak jauh berbeda dengan hasil pada nomor 2a. yaitu sekitar **10%**

```
. margins, atmeans post
```

Adjusted predictions		Number of obs	=	2,978
Model VCE : OIM				
Expression : Pr(chjob), predict()				
at	hi	=	.7518469	(mean)
	loghwage	=	2.295388	(mean)
	othhi	=	.4059772	(mean)
	educ	=	12.87609	(mean)
	exper	=	19.17562	(mean)
	union	=	.2481531	(mean)
	nfam	=	3.355944	(mean)
	winc	=	11969.43	(mean)
	hixothhi	=	.2817327	(mean)

	Delta-method					
	Margin	Std. Err.	z	P> z	[95% Conf. Interval]	
_cons	.1009181	.0059952	16.83	0.000	.0891677	.1126686