

## **Trabajo Práctico N° 5:** **Modelos de Variable Dependiente Discreta.**

### **Ejercicio 1.**

El archivo “wagepan.dta” contiene los datos utilizados por Vella y Verbeek (1998). Estos datos contienen información para 545 hombres que trabajaron cada año de 1980 a 1987. Utilizar los datos para analizar el impacto de la escolaridad ( $educ_{it}$ ) en la probabilidad de estar afiliado a un sindicato ( $union_{it}$ ). Las variables se describen en el conjunto de datos. Observar que la educación no cambia con el tiempo.

(a) Usar Pooled OLS para estimar el modelo:

$$P(union_{it} = 1 | educ_{it}) = \beta_0 + \beta_1 educ_{it} \quad (1)$$

De acuerdo a los resultados obtenidos, ¿tiene impacto un año más de escolaridad sobre la probabilidad de estar afiliado a un sindicato?

POLS:

Source	SS	df	MS	Number of obs	=	4,360
Model	.030271965	1	.030271965	F(1, 4358)	=	0.16
Residual	804.314682	4,358	.184560505	Prob > F	=	0.6855
Total	804.344954	4,359	.18452511	R-squared	=	0.0000
				Adj R-squared	=	-0.0002
				Root MSE	=	.42961

  

union	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
educ	-.0015092	.0037264	-0.40	0.686	-.0088148	.0057964
_cons	.261795	.0443282	5.91	0.000	.1748892	.3487009

Por lo tanto, se puede observar que un año más de escolaridad no tiene impacto sobre la probabilidad de estar afiliado a un sindicato, ya que la variable *educ* no es estadísticamente significativa.

(b) Usar Pooled Probit para estimar el modelo:

$$P(union_{it} = 1 | educ_{it}) = \Phi(\beta_0 + \beta_1 educ_{it}) \quad (2)$$

Comentar sobre el impacto de un año más de educación en la probabilidad de estar afiliado a un sindicato.

### Pooled Probit:

Probit regression  
Log likelihood = -2422.7142

Number of obs = 4,360  
LR chi2(1) = 0.17  
Prob > chi2 = 0.6758  
Pseudo R2 = 0.0000

	union	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
	educ	-.0051242	.0122467	-0.42	0.676	-.0291273	.0188789
	_cons	-.6331107	.1454783	-4.35	0.000	-.9182429	-.3479784

### Efectos marginales (promedio) en Pooled Probit:

Average marginal effects  
Model VCE: OIM

Number of obs = 4,360

Expression: Pr(union), predict()  
dy/dx wrt: educ

		Delta-method				[95% conf. interval]	
		dy/dx	std. err.	z	P> z		
	educ	-.0016074	.0038414	-0.42	0.676	-.0091364	.0059216

### Efectos marginales (condicionales) en Pooled Probit:

Conditional marginal effects  
Model VCE: OIM

Number of obs = 4,360

Expression: Pr(union), predict()  
dy/dx wrt: educ  
At: educ = 11.76697 (mean)

		Delta-method				[95% conf. interval]	
		dy/dx	std. err.	z	P> z		
	educ	-.0016074	.0038416	-0.42	0.676	-.0091369	.005922

Por lo tanto, se puede observar que un año más de escolaridad no tiene impacto sobre la probabilidad de estar afiliado a un sindicato, ya que la variable *educ* no es estadísticamente significativa.

### **(c) Usar Pooled Logit para estimar el modelo:**

$$P(\text{union}_{it} = 1 \mid \text{educ}_{it}) = \Lambda(\beta_0 + \beta_1 \text{educ}_{it}) = \frac{e^{\beta_0 + \beta_1 \text{educ}_{it}}}{1 + e^{\beta_0 + \beta_1 \text{educ}_{it}}} \quad (3)$$

Comentar sobre el impacto de un año más de educación en la probabilidad de estar afiliado a un sindicato. Compute el error estándar para esta estimación.

### Pooled Logit:

Logistic regression  
Log likelihood = -2422.7197  
Number of obs = 4,360  
LR chi2(1) = 0.16  
Prob > chi2 = 0.6857  
Pseudo R2 = 0.0000

	union	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
	educ	-.0081586	.0201411	-0.41	0.685	-.0476344	.0313172
	_cons	-1.034725	.2393549	-4.32	0.000	-1.503852	-.5655975

### Efectos marginales (promedio) en Pooled Logit:

Average marginal effects  
Model VCE: OIM  
Number of obs = 4,360

Expression: Pr(union), predict()  
dy/dx wrt: educ

		Delta-method				[95% conf. interval]	
		dy/dx	std. err.	z	P> z		
	educ	-.0015051	.0037154	-0.41	0.685	-.008787	.0057769

### Efectos marginales (condicionales) en Pooled Logit:

Conditional marginal effects  
Model VCE: OIM  
Number of obs = 4,360

Expression: Pr(union), predict()  
dy/dx wrt: educ  
At: educ = 11.76697 (mean)

		Delta-method				[95% conf. interval]	
		dy/dx	std. err.	z	P> z		
	educ	-.0015051	.0037155	-0.41	0.685	-.0087873	.0057771

Por lo tanto, se puede observar que un año más de escolaridad no tiene impacto sobre la probabilidad de estar afiliado a un sindicato, ya que la variable *educ* no es estadísticamente significativa.

(d) Estimar la siguiente extensión del modelo (2):

$$P(\text{union}_{it} = 1 \mid \text{educ}_{it}, c_i) = \Phi(\beta_0 + \beta_1 \text{educ}_{it} + c_i), \quad (4)$$

donde  $c_i$  son efectos no observables individuales. Usar el modelo Probit de efectos aleatorios. ¿Cuál es el problema que surge al momento de estimar el efecto parcial de interés?

### RE Probit:

```
Random-effects probit regression
Group variable: nr

Random effects u_i ~ Gaussian

Integration method: mvaghermite

Log likelihood = -1672.7504

Number of obs   = 4,360
Number of groups = 545

Obs per group:
    min = 8
    avg = 8.0
    max = 8

Integration pts. = 12

Wald chi2(1)    = 1.04
Prob > chi2     = 0.3080
```

	union	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
	educ	-.0502723	.0493148	-1.02	0.308	-.1469276	.0463829
	_cons	-.8013208	.5835448	-1.37	0.170	-1.945048	.342406
	/lnsig2u	1.099022	.1137102			.8761544	1.32189
	sigma_u	1.732406	.0984961			1.549725	1.936622
	rho	.7500768	.0213163			.7060247	.789496

LR test of rho=0: chibar2(01) = 1499.93      Prob >= chibar2 = 0.000

### Efectos marginales (promedio) en RE Probit:

```
Average marginal effects
Model VCE: OIM

Expression: Pr(union=1 | u_i=0), predict(pu0)
dy/dx wrt: educ

Number of obs = 4,360
```

		Delta-method				[95% conf. interval]	
		dy/dx	std. err.	z	P> z		
	educ	-.0076285	.0075519	-1.01	0.312	-.02243	.007173

### Efectos marginales (condicionales) en RE Probit:

```
Conditional marginal effects
Model VCE: OIM

Expression: Pr(union=1 | u_i=0), predict(pu0)
dy/dx wrt: educ
At: educ = 11.76697 (mean)

Number of obs = 4,360
```

		Delta-method				[95% conf. interval]	
		dy/dx	std. err.	z	P> z		
	educ	-.0076024	.0074781	-1.02	0.309	-.0222592	.0070543

El problema que surge al momento de estimar el efecto parcial de interés es que éste depende de  $c_i$ , el cual no es estimado, por lo que no es posible estimar la magnitud del efecto parcial, a menos que se imponga el valor de  $c_i$ . Este valor impuesto puede ser  $c_i =$

0, lo cual tiene sentido ya que se está asumiendo que la distribución de  $c_i$  es  $c_i | X_i \sim \text{Normal}(0, \sigma_c^2)$ .

(e) Estimar la siguiente extensión del modelo (3):

$$P(\text{union}_{it} = 1 | \text{educ}_{it}, c_i) = \Lambda(\beta_0 + \beta_1 \text{educ}_{it} + c_i), \quad (5)$$

donde  $c_i$  son efectos no observables individuales. Usar el modelo Logit de efectos aleatorios. ¿Surge el mismo problema que en el inciso anterior al momento de estimar el efecto parcial de interés?

### RE Logit:

```
Random-effects logistic regression          Number of obs   =   4,360
Group variable: nr                        Number of groups =    545

Random effects u_i ~ Gaussian              Obs per group:
                                         min =         8
                                         avg =        8.0
                                         max =         8

Integration method: mvaghermite            Integration pts. =    12

Wald chi2(1) =    0.98
Prob > chi2   =    0.3227

Log likelihood = -1670.7204
```

	union	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
	educ	-.0831898	.0841158	-0.99	0.323	-.2480538	.0816742
	_cons	-1.497913	.9974077	-1.50	0.133	-3.452797	.4569698
	/lnsig2u	2.248281	.1156237			2.021663	2.4749
	sigma_u	3.077571	.17792			2.747885	3.446812
	rho	.7421998	.0221233			.6965274	.7831387

LR test of rho=0: chibar2(01) = 1504.00      Prob >= chibar2 = 0.000

### Efectos marginales (promedio) en RE Logit:

```
Average marginal effects          Number of obs = 4,360
Model VCE: OIM

Expression: Pr(union=1 | u_i=0), predict(pu0)
dy/dx wrt:  educ
```

		Delta-method				[95% conf. interval]	
		dy/dx	std. err.	z	P> z		
	educ	-.0059833	.0061477	-0.97	0.330	-.0180325	.0060659

Efectos marginales (condicionales) en RE Logit:

Conditional marginal effects  
Model VCE: OIM

Number of obs = 4,360

Expression:  $\Pr(\text{union}=1 \mid u_i=0)$ , predict(pu0)  
dy/dx wrt: educ  
At: educ = 11.76697 (mean)

		Delta-method				
		dy/dx	std. err.	z	P> z	[95% conf. interval]
educ		-.0059476	.0060407	-0.98	0.325	-.0177872 .005892

Sí, surge el mismo problema que en el inciso anterior a la hora de estimar el efecto parcial de interés.

**(f)** *Computar el denominado estimador Logit de efectos fijos para el modelo (5). ¿Se puede computar el efecto de un año más de educación sobre la probabilidad de estar afiliado a un sindicato? Explicar.*

El efecto de un año más de educación sobre la probabilidad de estar afiliado a un sindicato no se puede computar, ya que, en este método no se identifican los coeficientes de los regresores que no varían en el tiempo, como es el caso de la variable *educ* en este ejercicio. Sin embargo, incluso si el coeficiente estuviera identificado (en el caso de que la variable varíe en el tiempo), sucedería lo de los incisos anteriores de que se debería imponer el valor de  $c_i$  sumado a que, en este caso, es difícil saber qué valor imponer, ya que no se está asumiendo que se conoce la distribución de  $c_i$ .

**(g)** *Considerar la siguiente extensión del modelo (4):*

$$P(\text{union}_{it} = 1 \mid \text{educ}_{it}, \text{black}_{it}, \text{married}_{it}, c_i) = \Phi(\beta_0 + \beta_1 \text{educ}_{it} + \beta_2 \text{black}_{it} + \beta_3 \text{married}_{it} + c_i), \quad (6)$$

donde  $\text{black}_{it}$  es una variable binaria que toma valor 1 si la persona es afroamericana y  $\text{married}_{it}$  es una variable binaria que toma valor 1 si la persona es casada. Asumir la siguiente versión de Mundlak (1978) del supuesto de Chamberlain (1980):

$$c_i \mid X_i \sim \text{Normal}(\psi + \xi \overline{\text{married}_i}, \sigma_a^2) \quad (7)$$

El modelo dado por (6) y (7) es un caso de lo que, en la literatura, se denomina modelo Probit de efectos aleatorios de Chamberlain. Al asumir sólo (6) y (7), se tiene que:

$$\begin{aligned} P(\text{union}_{it} = 1 \mid \text{educ}_{it}, \text{black}_{it}, \text{married}_{it}, c_i) &= \\ \Phi[(\beta_0 + \beta_1 \text{educ}_{it} + \beta_2 \text{black}_{it} + \beta_3 \text{married}_{it} + \psi + \xi \overline{\text{married}_i}) (1 + \sigma_a^2)^{-\frac{1}{2}}] &= \\ \Phi(\beta_{0,a} + \beta_{1,a} \text{educ}_{it} + \beta_{2,a} \text{black}_{it} + \beta_{3,a} \text{married}_{it} + \xi_a \overline{\text{married}_i}). \end{aligned}$$

*Usar Pooled Probit para estimar el modelo. Estimar el efecto de la escolaridad sobre la probabilidad de estar sindicalizado para una persona afroamericana casada.*

### Pooled Probit:

```
Probit regression                                Number of obs = 4,360
                                                Wald chi2(4) = 15.55
                                                Prob > chi2 = 0.0037
Log pseudolikelihood = -2389.9119              Pseudo R2 = 0.0136
```

(Std. err. adjusted for 545 clusters in nr)

			Robust			
	union	Coefficient	std. err.	z	P> z	[95% conf. interval]
	educ	-.0031662	.0219937	-0.14	0.886	-.0462731 .0399408
	black	.4707704	.1297179	3.63	0.000	.216528 .7250127
	married	.0425222	.0532229	0.80	0.424	-.0617927 .1468371
mean_married		.2052462	.1316841	1.56	0.119	-.0528499 .4633423
_cons		-.8279919	.2671794	-3.10	0.002	-1.351654 -.3043298

### Efectos marginales (condicionales) en Pooled Probit:

```
Conditional marginal effects                    Number of obs = 4,360
Model VCE: Robust
```

```
Expression: Pr(union), predict()
dy/dx wrt: educ
At: educ = 11.47692
      black = 1
      married = 1
      mean_married = .6923077
```

			Delta-method			
		dy/dx	std. err.	z	P> z	[95% conf. interval]
	educ	-.0012358	.0085842	-0.14	0.886	-.0180605 .0155888

## Ejercicio 2.

Considerar los datos del ejercicio previo para analizar la probabilidad de estar afiliado a un sindicato según la situación de afiliación sindical del año previo.

(a) Usar Pooled Probit para estimar el modelo:

$$P(\text{union}_{it} = 1 \mid \text{union}_{it-1}) = \Phi(\psi + \rho \text{union}_{it-1}) \quad (8)$$

A continuación, obtener una estimación para

$$P(\text{union}_{it} = 1 \mid \text{union}_{it-1} = 1)$$

y para

$$P(\text{union}_{it} = 1 \mid \text{union}_{it-1} = 0).$$

Comentar sobre el efecto marginal de estar afiliado a un sindicato en el año t-1 en la probabilidad de estar afiliado a un sindicato en el año t.

### Pooled Probit:

Probit regression	Number of obs = 3,815
	LR chi2(1) = 1416.84
	Prob > chi2 = 0.0000
Log likelihood = -1406.9979	Pseudo R2 = 0.3349

union	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
union						
L1.	1.953334	.055067	35.47	0.000	1.845404	2.061263
_cons	-1.348154	.0328877	-40.99	0.000	-1.412613	-1.283695

El efecto marginal de estar afiliado a un sindicato en el año t-1 en la probabilidad de estar afiliado a un sindicato en el año t es 0,639.

(b) Adicionar al modelo el conjunto completo de variables binarias temporales. Volver a estimar las probabilidades solicitadas para cada año de la muestra.



Pooled Probit (con variables binarias temporales):

Probit regression	Number of obs = 3,815
	LR chi2(7) = 1432.42
	Prob > chi2 = 0.0000
Log likelihood = -1399.2079	Pseudo R2 = 0.3386

	union	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
union							
l1.		1.967931	.0554571	35.49	0.000	1.859237	2.076625
d82		.0412915	.0960685	0.43	0.667	-.1469994	.2295824
d83		-.0513702	.0981709	-0.52	0.601	-.2437816	.1410412
d84		.0046824	.0980729	0.05	0.962	-.1875369	.1969016
d85		-.1584597	.1003127	-1.58	0.114	-.3550689	.0381495
d86		-.1384916	.1007772	-1.37	0.169	-.3360114	.0590282
d87		.1744346	.0957908	1.82	0.069	-.013312	.3621812
_cons		-1.339921	.0712489	-18.81	0.000	-1.479566	-1.200275

- El efecto marginal de estar afiliado a un sindicato en el año 1981 en la probabilidad de estar afiliado a un sindicato en el año 1982 es 0,651.
- El efecto marginal de estar afiliado a un sindicato en el año 1982 en la probabilidad de estar afiliado a un sindicato en el año 1983 es 0,636.
- El efecto marginal de estar afiliado a un sindicato en el año 1983 en la probabilidad de estar afiliado a un sindicato en el año 1984 es 0,646.
- El efecto marginal de estar afiliado a un sindicato en el año 1984 en la probabilidad de estar afiliado a un sindicato en el año 1985 es 0,614.
- El efecto marginal de estar afiliado a un sindicato en el año 1985 en la probabilidad de estar afiliado a un sindicato en el año 1986 es 0,618.
- El efecto marginal de estar afiliado a un sindicato en el año 1986 en la probabilidad de estar afiliado a un sindicato en el año 1987 es 0,667.

(c) *Estimar un modelo de efectos no observables dinámico. Usar el modelo Probit de efectos aleatorios incluyendo  $union_{i,80}$  como una variable explicativa adicional. Luego, promediar las probabilidades estimadas a lo largo de  $union_{i,80}$  para obtener la probabilidad promedio de estar afiliado a un sindicato en el año 1987 dado que estaba afiliado en el período anterior.*

### RE Probit:

```

Random-effects probit regression
Group variable: nr

Random effects u_i ~ Gaussian

Integration method: mvaghermite

Log likelihood = -1293.5235

Number of obs   = 3,815
Number of groups = 545

Obs per group:
    min = 7
    avg = 7.0
    max = 7

Integration pts. = 12

Wald chi2(8) = 335.28
Prob > chi2 = 0.0000
    
```

	union	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
union	L1.	.8886017	.0923747	9.62	0.000	.7075507	1.069653
	d82	.0398224	.1134088	0.35	0.725	-.1824547	.2620995
	d83	-.0636919	.1163085	-0.55	0.584	-.2916524	.1642686
	d84	-.0119455	.116519	-0.10	0.918	-.2403185	.2164276
	d85	-.224562	.1192335	-1.88	0.060	-.4582554	.0091315
	d86	-.2661777	.1199779	-2.22	0.027	-.5013301	-.0310253
	d87	.1277544	.1136075	1.12	0.261	-.0949122	.3504211
	union80	1.479094	.1659071	8.92	0.000	1.153922	1.804266
	_cons	-1.791908	.1172376	-15.28	0.000	-2.02169	-1.562127
/lnsig2u		.2114328	.1648497			-.1116665	.5345322
sigma_u		1.111507	.0916157			.9456968	1.306388
rho		.5526622	.0407552			.4721123	.6305396

LR test of rho=0: chibar2(01) = 160.37      Prob >= chibar2 = 0.000

La probabilidad promedio de estar afiliado a un sindicato en el año 1987 dado que estaba afiliado en el período anterior es 0,397.