

Taller #10

Respuestas Sugeridas

Econometría 06216

Profesor: Julio César Alonso C.

Monitor: Manuel Serna Cortés.

Notas:

- Recuerde que tres preguntas, seleccionadas al azar, serán calificadas.
- Este taller puede subirse en la plataforma Moodle hasta la 7:10 del 27 de abril de 2009. **Sólo se recibirán talleres en formato pdf.** Cualquier otro formato no será calificado.

INSTRUCCIONES:

- Este taller debe ser escrito en computador.
- Cuando sea posible, debe mostrar el procedimiento efectuado para llegar a sus resultados.
- Este taller es un trabajo en pareja. Por tanto el taller debe reflejar **únicamente** el trabajo de la pareja.
- Si bien no es necesario reportar todos los números decimales, sí lo es hacer los cálculos con **todos** ellos.

En la Isla de Laurasia, un joven economista ha sido contratado por el ministerio de Hacienda para realizar un estudio sobre la sostenibilidad del endeudamiento del Estado mediante la adquisición de crédito internacional. Para esto cuenta con la siguiente base de datos: índice de producción industrial (IPI_i), tasa de desempleo (TD_i), índice de Riesgo país (IRP_i), Gasto Público (G_i , medido en millones de britanios) e Impuestos (T_i , también medido en millones de britanios). Las observaciones con las que cuenta el economista corresponden a las 100 provincias de Laurasia y están consignadas en el archivo T10-01-09.xlsx.

En principio, el estudio se centrará en estudiar los factores que afectan la probabilidad de obtener crédito internacional, de acuerdo con esto, el economista recolectó datos sobre la aprobación del crédito para cada una de las provincias. De esta manera, se construyó una variable dicotómica ($Credito_i$) que asigna 1 a la aprobación del crédito y 0 en caso contrario.

1. De acuerdo con la información anterior:
 - a. El economista desea que las pendientes de su modelo estimado representen los cambios marginales en la probabilidad de obtener crédito internacional debido a cambios en las variables explicativas. Escriba el modelo que se ajusta a este requerimiento.

Respuesta sugerida:

$$Credito_i = \beta_0 + \beta_1 IPI_i + \beta_2 TD + \beta_3 IRP + \beta_4 T_i + \beta_5 G_i + \varepsilon_i \tag{1}$$

Por tanto, al realizar las derivadas parciales, las pendientes corresponden a los cambios marginales en la probabilidad de obtener crédito internacional.

b. Estime el modelo y repórtelo en una tabla.

Respuesta sugerida:

El modelo que debieron estimar en este punto corresponde a la corrección de MCP dividiendo por la variable $w_i = \sqrt{\hat{y}_i(1 - \hat{y}_i)}$ donde $\hat{y}_i = (\hat{\beta}_{OLS})^T \mathbf{x}_i^T$, pues como sabemos por la construcción del modelo de probabilidad lineal (MPL), éste tendrá un error heteroscedástico, su varianza es de la forma $\sigma_i^2 = E[y_i|x_i](1 - E[y_i|x_i])$.

Tabla 1. Estimación modelo 1 MPL

Tabla 1. Resultados de la estimación del modelo (1)			Tabla 2. Resultados de la estimación del modelo (2)		
Variable dependiente: Creditoi			Variable dependiente: Creditoi		
Estadísticos t entre parentesis			Estadísticos t entre parentesis		
	MCO	Sig	MCP	Sig	
Constante	-36,672535 (-7,16)	***	-0,3273947 (-1,520)		
IPIi	0,2267634 (6,74)	***	0,1567854 (5,28)	***	
TDi	0,0083692 (0,231)		-0,1551817 (-3,183)	***	
IRPi	0,2537492 (6,825)	***	0,1377503 (5,066)	***	
Ti	0,0278719 (0,742)		-0,1686983 (-3,998)	***	
Gi	-0,0463367 (-1,337)		-0,1077282 (-2,338)	**	
R²	0,556		0,9031		
R̄²	0,5323		0,8971		
F	23,54	***	149,17	***	
No. de obs.	100		86		

(***) Nivel de significancia: 1%

(**) Nivel de significancia: 5%

(*) Nivel de significancia: 10%

MCO: Mínimos Cuadrados Ordinarios

MCP: Mínimos Cuadrados Ponderados

c. Comente la significancia global del modelo estimado e intérprete la bondad de ajuste.

Respuesta sugerida:

La prueba de significancia muestra que al menos una de las variables explicativas es significativa con un nivel de confianza del 99%. Respecto a la bondad de ajuste, el R cuadrado no se puede interpretar por la particularidad del modelo, es decir no es una buena medida para determinar el ajuste del modelo.

2. Interprete los coeficientes estimados en el punto anterior de acuerdo con su significancia.

Respuesta sugerida:

Para empezar, $\hat{\beta}_1$ no es significativo, dado que no se rechaza ni con un 90% de confianza la hipótesis nula de que el coeficiente sea estadísticamente igual a cero. De aquí que se pueda decir que la probabilidad de la provincia i de obtener crédito internacional que no depende del índice de riesgo país, índice de producción industrial, los impuestos, el gasto público y la tasa de desempleo es igual a 0. Ceteris Paribus.

Por otro lado, $\hat{\beta}_2$ un nivel de confianza del 99%, nos dice que un aumento de una unidad en el índice de producción industrial de la provincia i , en promedio, implicará un aumento de la probabilidad de que la provincia i obtenga un crédito internacional en 15.68 puntos porcentuales. Ceteris Paribus.

Por su parte, $\hat{\beta}_3$ un nivel de confianza del 99%, nos dice que un aumento de un punto porcentual en la tasa de desempleo de la provincia i , en promedio, implicará una reducción de la probabilidad de que la provincia i obtenga un crédito internacional en 15.52 puntos porcentuales. Ceteris Paribus.

De otro lado, $\hat{\beta}_4$ con un nivel de confianza del 99%, nos dice que un aumento de una unidad en el índice de riesgo país de la provincia i , en promedio, implicará un aumento de la probabilidad de que la provincia i obtenga un crédito internacional en 13.78 puntos porcentuales. Ceteris Paribus.

En este orden de ideas, $\hat{\beta}_5$ un nivel de confianza del 99%, nos dice que un aumento de un millón de británicos en los impuestos de la provincia i , en promedio, implicará una reducción de la probabilidad de que la provincia i obtenga un crédito internacional en 16.87 puntos porcentuales. Ceteris Paribus.

Finalmente, $\hat{\beta}_6$ un nivel de confianza del 95%, nos dice que un aumento de un millón de británicos en el gasto público de la provincia i , en promedio, implicará una reducción de la probabilidad de que la provincia i obtenga un crédito internacional en 10.77 puntos porcentuales. Ceteris Paribus.

3. El economista presenta sus resultados preliminares a uno de sus compañeros investigadores. Dicho compañero no luce muy satisfecho con la elección del modelo, pues considera que tanto en la teoría como en la práctica, este modelo contiene muchos problemas.

- a. ¿Usted está de acuerdo con la afirmación del compañero investigador? Justifique su respuesta con el mayor número de argumentos que considere pertinentes.

Respuesta sugerida:

Argumentos teóricos:

- Suponer que los efectos marginales en la probabilidad para todas las observaciones son **constantes**. Un claro ejemplo son los años de educación y la probabilidad de obtener empleo, pues no es lo mismo un año más de educación, si se pasa de noveno a décimo, que si se pasa de undécimo a la universidad, claramente, el obtener el título de bachiller debe incrementar la probabilidad de obtener trabajo en una mayor proporción que pasar de noveno a décimo.
- $\sigma_i^2 = E[y_i|x_i](1 - E[y_i|x_i])$. Heteroscedasticidad
- Suponer una distribución de probabilidad lineal no permite restringir los valores de predicción entre cero y 1.
 - Implica probabilidades sin sentido
 - Varianzas negativas.
- El R cuadrado no tiene sentido.

Argumentos prácticos:

- Implica doble estimación, aumentando el nivel de error.
- Se deben eliminar los datos menores a cero y mayores a uno para poder realizar MCP. En este caso, se disminuyó el número de observaciones en 14% por ciento.
- La eliminación de datos implica una menor muestra.

- b. Comente que otras opciones existen para la estimación de una variable dependiente discreta. Justifique por qué estas opciones son más adecuadas que estimar por el primer modelo.

Respuesta sugerida:

Como se vio en clase, la estimación de modelos Logit o Probit resulta mucho más adecuada que por MPL, la razón es que restringe los valores de predicción entre cero y 1 por la forma de sus funciones de densidad de probabilidad acumulada asintóticas para dichos valores, no contiene problemas teóricos de heteroscedasticidad, ni tampoco de eliminación de datos. Además, lo más importante es que los incrementos en la probabilidad dado cambios en las variables explicativas, ya no son constantes para estos dos modelos. Esto se evidencia en que en el rango medio, tanto la función de densidad normal y logística, tienen pendientes mucho más inclinadas que la de MPL, mientras que en los extremos, las pendientes son mucho más planas.

4. Continuando con la pregunta anterior:

- a) Estime las alternativas que planteó. Repórtelas en una tabla (incluya la razón de máxima verosimilitud y el logaritmo de la función de máxima verosimilitud)

Respuesta sugerida: Ver tabla 2

- b) Su compañero de investigación nuevamente analiza sus resultados e interpreta los coeficientes estimados de los modelos de acuerdo con su significancia. Justifique por

qué es correcto o no interpretar los coeficientes de los modelos. En caso de estar en desacuerdo, encuentre los resultados de los modelos que si se pueden interpretar para cada una de las variables y repórtelos en la Tabla que realizó en el literal anterior.

Respuesta sugerida: Ver tabla 2. No se pueden interpretar los coeficientes por si solos, pues recuerden que estos corresponden a la solución por máxima verosimilitud del siguiente problema:

$$Max_{\beta} L(\beta) = Max_{\beta} [P(y)] = Max_{\beta} \left[\prod_{\forall \text{ital que } y_i=1} P(y=1) \prod_{\forall \text{ital que } y_i=0} P(y=0) \right]$$

$$Max_{\beta} L(\beta) = Max_{\beta} \left[\prod_{\forall \text{ital que } y_i=1} 1 - F(-z_i) \prod_{\forall \text{ital que } y_i=0} F(-z_i) \right]$$

Donde $F(-z_i)$ corresponde a la función de densidad normal para Probit o la función logística para Logit. De ahí que lo que realmente se interprete sea los efectos marginales que están dados por:

$$\frac{\partial [y_i | \mathbf{x}_i]}{\partial X_j} = \phi(z_i) \beta_j \quad \text{y,} \quad \frac{\partial [y_i | \mathbf{x}_i]}{\partial X_j} = \Lambda(z_i) [1 - \Lambda(z_i)] \beta_j$$

Respectivamente. No obstante estos dependen de cada observación, por tanto, se promedian.

Tabla 2. Estimación de Logit y Probit.

VARIABLE DEPENDIENTE: <i>Credito_i</i>				
Estadísticos t entre paréntesis				
	Probit EMV		Logit EMV	
	Coeficientes	Efecto Marginal	Coeficientes	Efecto Marginal
constante	-734.90 (-3.64) ***	--	-1246.51 (-3.64) ***	--
<i>IPI_i</i>	4.28 (3.58) ***	0.352	7.29 (3.58) ***	0.339
<i>TD_i</i>	0.96 (2.45) **	0.079	1.56 (2.45) **	0.073
<i>IRP_i</i>	4.55 (3.58) ***	0.375	7.71 (3.58) ***	0.359
<i>T_i</i>	0.76 (2.28) **	0.063	1.34 (2.28) **	0.062
<i>G_i</i>	-0.41 (-1.28)	-0.033	-0.73 (-1.28)	-0.034
LRI	0.78		0.78	
Wald	14.80 **		13.31 **	
ln(L)	-15.01		-15.14	
# de Obs.	100.00		100.00	

(*) nivel de significancia: 10%

(**) nivel de significancia: 5%

(***) nivel de significancia: 1%

Wald: corresponde al estadístico de Wald que comprueba la significancia conjunta de todas las pendientes.

EMV: Estimadores de Máxima Verosimilitud

5. De acuerdo con los resultados obtenidos en punto anterior:

- a. Para los modelos estimados en el punto 4, se puede afirmar que el efecto marginal del IPI es igual al efecto marginal del IRP en la probabilidad de obtener crédito internacional.

Respuesta sugerida:

Es importante aclarar que la afirmación se hace sobre los efectos marginales promedio y no sobre los coeficientes, como normalmente se trabaja, no obstante, la prueba de hipótesis es equivalente pues tenemos que para probit:

El efecto marginal es:

$$\frac{\partial [y_i | \mathbf{x}_i]}{\partial X_j} = \phi(z_i) \beta_j$$

Y en promedio es:

$$\frac{\sum_{i=1}^n \phi(z_i) \beta_j}{n}$$

Por tanto debemos comparar el efecto marginal promedio de la variable IPI con la variable IRP que corresponden a los coeficientes β_1 y β_3 respectivamente.

$$\frac{\sum_{i=1}^n \phi(z_i) \beta_1}{n} = \frac{\sum_{i=1}^n \phi(z_i) \beta_3}{n}$$

$$\sum_{i=1}^n \phi(z_i) \beta_1 = \sum_{i=1}^n \phi(z_i) \beta_3$$

$$\beta_1 \sum_{i=1}^n \phi(z_i) = \beta_3 \sum_{i=1}^n \phi(z_i)$$

De esta manera nuestra prueba de hipótesis corresponde a:

Ho: $\beta_1 = \beta_3$

Ha: No Ho.

El resultado de esta prueba con un test de Wald, muestra que el estadístico es 0.33 con un p-valor de 0.56546, concluyendo así que no existe suficiente evidencia para afirmar que los efectos marginales promedio de las variables IPI y IRP son estadísticamente diferentes.

Para Logit el procedimiento es análogo:

Ho: $\beta_1 = \beta_3$

Ha: No Ho.

El resultado de esta prueba con un test de Wald, muestra que el estadístico es 0.23 con un p-valor de 0.62968, concluyendo así que no existe suficiente evidencia para afirmar que los efectos marginales promedio de las variables IPI y IRP son estadísticamente diferentes.

- b. ¿Qué relación encuentra entre los coeficientes estimados?

Respuesta sugerida:

Se cumple la hipótesis de Amemiya. Es decir que se cumple que $\hat{\beta}_{Probit} \approx 0.625 \hat{\beta}_{Logit}$ pues la diferencia porcentual promedio es del 7% entre los coeficientes estimados por probit y los estimados por la relación de Amemiya, lo cual es bastante bajo.

- c. El economista realiza la siguiente afirmación: con un nivel de significancia del 1%, la probabilidad media en ausencia de las demás variables de obtener crédito internacional, para los modelos estimados en el punto 4, es estadísticamente igual cero. ¿Está en capacidad de realizar dicha afirmación con los resultados obtenidos hasta el momento?

Respuesta sugerida:

No está en capacidad de realizar la afirmación, pues a diferencia de MPL, el intercepto no corresponde a la media de la variable dependiente debido a la transformación que se realiza. Usted no cuenta con la ecuación matemática que relacione el intercepto con la probabilidad promedio en ausencia de los regresores, por tanto no cuenta con los elementos necesarios para realizar dicha afirmación.

6. Teniendo en cuenta los resultados de las estimaciones, interprete los efectos marginales de los modelos del punto 4 de acuerdo con su significancia. ¿Coinciden estos efectos marginales entre los modelos? ¿Coinciden con lo esperado?

Respuesta sugerida:

- La probabilidad promedio de obtener crédito internacional aumenta en 35.2 puntos porcentuales (en el modelo probit) y en 33.9 puntos porcentuales (en el modelo logit), ante un incremento en un punto del IPI, con un nivel de significancia del 1%. Coincide con el signo esperado.
- La probabilidad promedio de obtener crédito internacional aumenta en 7.9 puntos porcentuales (en el modelo probit) y en 7.3 puntos porcentuales en el (modelo logit), ante un incremento en un punto porcentual de la TD, con un nivel de significancia del 5%. El signo esperado no es claro y puede ser objeto de discusión.
- La probabilidad promedio de obtener crédito internacional aumenta en 37.5 puntos porcentuales (en el modelo probit) y en 35.9 puntos porcentuales (en el modelo logit), ante un incremento en un punto del IRP, con un nivel de significancia del 1%. No coincide con el signo esperado.
- La probabilidad promedio de obtener crédito internacional aumenta en 6.3 puntos porcentuales (en el modelo probit) y en 6.2 puntos porcentuales (en el modelo logit), ante un incremento en un millón de británicos de T, con un nivel de significancia del 5%. Coincide con el signo esperado.
- La probabilidad promedio de obtener crédito internacional no se afecta, ante un incremento en un millón de británicos de G.
- La probabilidad media que no depende de las demás variables, no es un efecto marginal, además no se cuenta con la herramienta para encontrarla.