

TRABAJO PRÁCTICO 3

- 01/11/2017

GRUPO**ANTÓN MARÍA PAZ****FIGLIOLIA JULIETA ALDANA****PONCIO FEDERICO****SABATER ANNA****WANG JIA QI****PUNTO 1****Ejercicio 1a**

Se estiman los siguientes modelos y se interpreta:

$$\text{Demanda: } \ln Q_t = \beta_0 + \beta_1 \ln P_t + \beta_2 LUN_t + \beta_3 MAR_t + \beta_4 MIER_t + \beta_5 JUE_t + \varepsilon_t$$

$$\text{Oferta: } \ln Q_t = \alpha_0 + \alpha_1 \ln P_t + \alpha_2 TOR_t + u_t$$

La elasticidad precio de la oferta de la pescadilla, α_1 , no es estadísticamente distinta de cero a un nivel del 5% de significancia porque la variable reporta un p-valor de 0.054 y su intervalo de confianza incluye al cero. El signo del coeficiente estimado es negativo y esto no es razonable con la teoría económica, de lo contrario, lo intuitivo sería que ante un aumento en el precio de la pescadilla se espera una suba en la cantidad ofertada por el productor.

La elasticidad precio de la demanda de pescadilla, β_1 , es estadísticamente significativa a un nivel del 5% porque computa un p-valor de 0.001. Es razonable que el signo del coeficiente estimado sea negativo dado que ante un aumento del 1% en el precio de la pescadilla se espera una baja de 0.56% en la cantidad demandada. Es decir, que la demanda de pescadilla es inelástica porque su cantidad reacciona de manera menos que proporcional.

Ejercicio 1b

La variable precio $\ln P_t$ es endógena porque es probable que estuviera correlacionada con el término de error debido a las interacciones entre la oferta y la demanda, lo cual significa que la variable está siendo determinada “dentro” del modelo, en vez de fuera de éste.

Existe causalidad simultánea cuando los precios determinan las cantidades y las cantidades también a los precios. En otras palabras, en la regresión de Y sobre X, además del vínculo causal de interés que va desde X hacia Y, existe un vínculo causal desde Y hacia X. Esta causalidad inversa provoca que X esté correlacionado con el término de error en la regresión poblacional de interés. En consecuencia, la estimación por MCO recogerá ambos efectos y el estimador MCO está sesgado e inconsistente.

A modo de ejemplo, se quiere estimar la elasticidad precio de la demanda de pescadilla, β_1 , del modelo del Punto 1a. Supongamos que ocurre un shock positivo en la demanda que no es observable por ninguna de las variables explicativas del modelo y dicho efecto es capturado por el término de error, por ejemplo, está de moda comer pescadilla. Ante este shock positivo en el error, se espera un aumento en la cantidad demandada de pescadilla y el nuevo precio de equilibrio será mayor. Es decir, que la variable explicativa precio y el término de error están correlacionados positivamente. Por esto, se viola uno de los supuestos de mínimos cuadrados, $E(u_i | X_{1i}, \dots, X_{ki}) \neq 0$ lo cual significa que el estimador MCO es sesgado e inconsistente.

Ejercicio 1c

La variable TOR_t puede ser usada como instrumento para $\ln P_t$ en la función de **demanda** si cumple con las condiciones, conocidas como condición de relevancia y condición de exogeneidad:

1. Relevancia del instrumento: $\text{corr}(TOR_t, \ln P_t) \neq 0$: una variación en los días de tormenta está relacionada con la variación en el precio. Se espera una relación positiva entre TOR_t y $\ln P_t$ porque en los días de tormenta aumenta el precio de la pescadilla ya que hay una restricción en la oferta de pescados dado que es más difícil pescar.
2. Exogeneidad del instrumento: $\text{corr}(TOR_t, u_i) = 0$: es decir que los días de tormenta no afectan de forma directa a la cantidad demandada de pescadilla pero sí lo hacen de manera indirecta a través del precio. De esta manera, el consumo de la pescadilla no correlaciona con los días de tormenta.

Las variables LUN_t , MAR_t , $MIER_t$, JUE_t pueden ser usadas como instrumentos de la variable $\ln P_t$ en la función de oferta si cumplen con las dos condiciones, de relevancia y exogeneidad:

1. Relevancia del instrumento: los días de la semana están correlacionados con el precio de la pescadilla, por ejemplo, se espera un signo negativo si los días miércoles no se consume pescadilla y cae el precio, o, si los días lunes está de moda comer pescadilla y sube el precio.
2. Exogeneidad del instrumento: la cantidad ofrecida de pescadilla no depende de los días de semana. (A lo sumo, la cantidad ofrecida podría depender de las cuestiones climáticas pero las variables instrumentales evaluadas no se refieren a esto).

Ejercicio 1d

Según la primer etapa de 2SLS de STATA, el instrumento TOR es individualmente significativo al 1% ya que reporta un p-valor 0.000. El signo positivo reportado es razonable y coincide con la argumentación planteada en el ejercicio anterior. Se espera una relación positiva entre los días de tormenta y el precio de la pescadilla porque en los días de tormenta aumenta el precio de la pescadilla ya que hay una restricción en la oferta de pescados dado que es más difícil pescar. Si ocurre el aumento de un día de tormenta en la semana se estima una suba de 34.64% en el precio del pescado.

Además, TORt superó el test F de fortaleza porque se computó un valor F de 22.92, el cual es mayor a 10 y, por lo tanto, hay evidencia estadística a favor de la fortaleza del instrumento.

En consecuencia, se concluye que la condición de relevancia de TOR se cumple estadísticamente mientras que la condición de exogeneidad se argumenta de manera teórica. El instrumento es válido.

La estimación 2SLS sugiere que ante una suba en el precio de un 1% disminuye la cantidad demandada de pescadilla en un 1.11%, manteniendo constante el resto de las variables. Esta elasticidad de demanda por 2SLS se incrementó en valores absolutos con respecto a MCO, lo cual indica que la demanda de pescadilla es elástica, en vez de inelástica. Esto tiene sentido con la realidad económica dado que el pescado suele ser el tipo de carne relativamente más elástico del mercado.

Ante estos resultados obtenidos, se concluye que la elasticidad de demanda por MCO estaba sesgada y era inconsistente.

Ejercicio 1e

De acuerdo a la primer etapa de 2SLS de STATA, los cuatro instrumentos *LUNt*, *MARt*, *MIERt*, *JUEt* no son estadísticamente significativos a un nivel del 10% porque se computaron los p-valores 0.575, 0.404, 0.487 y 0.809, respectivamente. A su vez, el test F de fortaleza de los instrumentos reportó un valor de 0.5384 que es menor a 10, entonces, hay evidencia a favor de la debilidad de los instrumentos para explicar el precio de la pescadilla. Por un lado, se observa que éstos no cumplen la condición de relevancia. Las consecuencias que puede tener esto sobre las estimaciones de elasticidad de oferta es que éstas sean aún más sesgadas que por MCO. Se sugieren las siguientes recomendaciones: no utilizar los cuatro días de la semana como instrumentos en forma conjunta y estimar el modelo solamente con un día, el más representativo para el mercado.

Por otro lado, se evaluó la condición de exogeneidad con el test de sobreidentificación. Con un p-valor alto de 0.8699 no se rechaza la hipótesis nula de exogeneidad de los instrumentos. Hay suficiente evidencia estadística a favor de que los instrumentos son exógenos.

En suma, se cumple la condición de exogeneidad pero no la de relevancia, entonces se concluye que los cuatro instrumentos no son válidos para explicar el precio de la pescadilla.

PUNTO 2

Ejercicio 2a

En la base de datos hay 753 mujeres en total. Hay menos observaciones para la variable *wage* que para *inlf* porque mientras que *inlf* refleja la participación (=1) o no (=0) de una mujer en el mercado laboral, *wage* mide el salario y este sólo se hace presente cuando la mujer trabaja (*inlf*=1).

La media de *inlf* es 0.5684. Es decir, el 56.84% de las mujeres de la muestra participan en el mercado laboral mientras que el 43.16% no lo hace.

Ejercicio 2b

Se estima el siguiente modelo de probabilidad lineal sobre la participación laboral de la mujer y se interpreta:

$$\text{inlf}_i = \beta_0 + \beta_1 \text{nwifeinc}_i + \beta_2 \text{educ}_i + \beta_3 \text{exper}_i + \beta_4 \text{expersq}_i + \beta_5 \text{age}_i + \beta_6 \text{kidslt6}_i + \beta_7 \text{kidsge6}_i + \varepsilon_i$$

- **Nwifeinc**: el coeficiente estimado es estadísticamente significativo al 5% (pero no al 1%) e indica que ante un aumento de mil dólares en el ingreso del marido disminuye en 0.34 puntos porcentuales (de 100) la probabilidad de participación en el mercado laboral de la mujer, manteniendo constante el resto de las variables. Esta predicción con signo negativo tiene sentido ya que si el sueldo del marido aumentara lo suficiente, no habría necesidad de que la mujer trabaje porque se cubre la cuota del hogar.
- **Educ**: el coeficiente de esta variable es significativo al 1% y predice que ante el aumento de un año en la educación de una mujer se espera una suba de 3.8 puntos porcentuales en su probabilidad de participar en el mercado laboral, sobre una base de 100, manteniendo las demás características constantes. Esta predicción con signo positivo es razonable y, aunque parezca pequeña, hay que considerar que, entre la escolaridad y la universidad, por ejemplo, se acumulan entre 15 y 20 años de educación, llevando a un aumento en 57 y 76 puntos porcentuales, respectivamente. Cabe mencionar que el modelo lineal falla en capturar los efectos que pueden tener los años de educación de distintos niveles en la probabilidad de que la mujer participe en el mercado laboral.

- **Age:** el coeficiente de esta variable es significativa al 1% y estima que frente a un aumento de un año en la edad de la mujer se predice una caída de 1.6 puntos porcentuales en la probabilidad de que participe en el mercado laboral, sobre una base de 100, y manteniendo todo lo demás constante.
 Este efecto es esperable ya que en el mercado laboral se tiende a aumentar las barreras de entrada a una persona de mayor edad y, sobre todo, si es femenina. (Realidad de “techo de cristal” para las mujeres.)
- **Exper:** el coeficiente estimado es significativo al 1% e indica que ante el aumento de un año en la experiencia laboral de una mujer se predice un incremento de 3.95 puntos porcentuales en la probabilidad de que participe en el mercado laboral, sobre una base de 100, y manteniendo todo lo demás constante.
 Este impacto positivo en la probabilidad tiene sentido ya que la experiencia es valorada en el mercado laboral como atributo.
- **Expersq:** el coeficiente es significativa al 1% y estima que ante el aumento de un año en el cuadrado de la experiencia de la mujer se predice una caída de 0.05 puntos porcentuales en la probabilidad de que trabaje, sobre una base de 100, y manteniendo todo lo demás constante.
 Esta variable tiene un efecto negativo dado que captura los rendimientos marginales decrecientes de la experiencia. Su magnitud es pequeña porque la tasa de decrecimiento es baja.
- **Kidslt6:** el coeficiente es estadísticamente significativo al 1% y predice que por cada hijo adicional menor a seis años que tiene una mujer disminuye en 26.18 puntos porcentuales la probabilidad de que participe en el mercado laboral, sobre una base de 100, y manteniendo todo lo demás constante.
 Esto es sumamente razonable ya que los niños menores a 6 años requieren mayores cuidados y, frente a la organización social contemporánea (y más aún en 1971), las mujeres del hogar son quienes se encargan de atenderlos, casi exclusivamente.
- **Kidsge6:** el coeficiente de esta variable no es estadísticamente significativo ya que reporta un pvalor de 0.337. De todas maneras, estimaría que ante un aumento en uno de la cantidad de hijos entre 6 y 18 años que tiene una mujer se incrementa en 1.3 puntos porcentuales la probabilidad de participación en el mercado laboral, sobre una base de 100, y manteniendo todo lo demás constante.
 Una posible explicación teórica para este efecto positivo podría ser que, a medida que crecen los hijos en una familia, éstos requieren menos cuidado intensivo pero mayor cantidad de comida y recursos para continuar con su crecimiento. Esto impulsaría a las mujeres a participar en el mercado laboral.
- **Constante:** ésta predeciría que una mujer sin educación, de cero años, sin hijos, y cuyo marido no percibe sueldo, tiene una probabilidad del 58.55% de participar en el mercado laboral. Esta constante no tiene interpretación en la vida real.

Con respecto al cambio de probabilidad entre estas dos mujeres hipotéticas, llamemos a la probabilidad estimada de la primera, \hat{p}_1 , y la segunda, \hat{p}_2 .

$$\Delta p = \hat{p}_1 - \hat{p}_2 = 0.0379953educ_1 - 0.0379953educ_2 = 0.0379953(educ_1 - educ_2) = 0.0379953 \cdot 10 = 0.379953$$

El diferencial en probabilidad entre estas dos mujeres es de 37.99 puntos porcentuales.

Ejercicio 2c

La ecuación que describe las características de este conjunto de mujeres es:

$$\hat{p}_i = 0.5855192 - 0.0034052 \cdot 50 + 0.0394924 \cdot 5 - 0.0005963 \cdot 25 - 0.0160908 \cdot 30 - 0.2618105 \cdot 1 + 0.0130122 \cdot 0 + 0.0379953 \cdot educ_i$$

$$\hat{p}_i = -0.14672080 + 0.0379953 \cdot educ_i$$

Si no tiene años de educación formal, la probabilidad estimada es -14.67%. Esto no tiene sentido porque las probabilidades no pueden ser negativas. En el modelo éstas dejan de serlo a partir de los 3.8615513 años o, aproximadamente, 4 años de educación en adelante. No hay en la muestra una mujer que tenga menos de 4 años de educación ya que el valor mínimo reportado en la variable *educ* es 5.

Esto también se puede ver en el histograma de *educ* y en el gráfico de \hat{p}_i .

Ejercicio 2d

Se estima nuevamente el modelo teórico propuesto en el punto 2b y se compara cada uno, probabilidad lineal (MPL), logístico (logit) y normal (probit).

| VARIABLES | MPL | Probit | Logit |
|-----------|-------------------------|-------------------------|------------------------|
| Nwifeinc | -0.00341** (0.00145) | -0.0120** (0.00484) | -0.0213** (0.00842) |
| educ | 0.0380*** (0.00738) | 0.131*** (0.0253) | 0.221*** (0.0434) |
| age | -0.0161*** (0.00248) | -0.0529*** (0.00848) | -0.0880*** (0.0146) |

| | | | |
|--------------|----------------------------|---------------------------|--------------------------|
| Exper | 0.0395*** (0.00567) | 0.123*** (0.0187) | 0.206*** (0.0321) |
| expersq | -0.000596*** (0.000185) | -0.00189*** (0.000600) | -0.00315*** (0.00102) |
| kidslt6 | -0.262*** (0.0335) | -0.868*** (0.119) | -1.443*** (0.204) |
| kidsge6 | 0.0130 (0.0132) | 0.0360 (0.0435) | 0.0601 (0.0748) |
| Constant | 0.586*** (0.154) | 0.270 (0.509) | 0.425 (0.860) |
| Observations | 753 | 753 | 753 |

Ejercicio 2e

Para los tres modelos se computa el cambio marginal de la probabilidad de estar trabajando para dos mujeres "promedio" con 20 y 25 años de experiencia.

Cuando la mujer tiene 20 años de experiencia, STATA reporta que la probabilidad lineal esperada es de 80 puntos porcentuales, la probabilidad logit y probit esperadas son ambas 82 puntos porcentuales. Aquí se evidencia que los modelos no lineales capturan efectos similares y mayores (con respecto al lineal) de la experiencia sobre la probabilidad de que la mujer "promedio" participe en el mercado laboral.

Cuando la mujer tiene 25 años de experiencia, STATA reporta que la probabilidad lineal esperada es de 86.9 puntos porcentuales mientras que la probabilidad esperada logit es 86.7 y la probit, 87.7 puntos porcentuales. Se observa que también los modelos no lineales computan efectos levemente similares de la experiencia y difieren poco con respecto del modelo lineal.

El cambio en la probabilidad cuando la mujer tiene 20 y 25 años de experiencia es de 6 puntos porcentuales en el modelo lineal y de 4 puntos porcentuales en ambos modelos no lineales. Por lo tanto, se concluye que los modelos no lineales capturan probabilidades esperadas similares y de mayor magnitud, a diferencia del modelo de probabilidad lineal.

Ejercicio 2f

En primer lugar, se observa que la experiencia de las mujeres de la muestra computó un máximo de 45 años y un mínimo de cero. A partir del gráfico de las tres probabilidades estimadas en función de exper, se puede argumentar que los efectos de los años de educación en una mujer promedio son crecientes hasta cierto punto donde alcanza lo que parece ser un punto de saturación. En el modelo lineal, este punto es claro, al tratarse de una cuadrática; sin embargo en los modelos probit y logit necesitaríamos tener datos de mujeres de mayor edad para ver la tendencia en éstas áreas.

Los tres modelos, lineal, logit y probit, estiman cálculos muy similares de estas probabilidades pero principalmente se nota una mayor convergencia entre la estimación logit y probit. Al principio, a medida que aumentan los años de experiencia de la mujer, la función de probabilidad se incrementa rápidamente pero este crecimiento se modera paulatinamente cuando se alcanzan los 20 años de experiencia, punto en el cual comienza a decaer la probabilidad de que la mujer trabaje, por la cercanía a la jubilación.

Por último, se observa que, con respecto al modelo lineal, las estimaciones logit y probit capturaron un leve mayor impacto de la experiencia sobre la probabilidad de la participación laboral femenina. En los extremos mínimos y máximos años de experiencia, el modelo lineal sobreestima la probabilidad de que la mujer trabaje mientras que en los valores intermedios, la subestima, a diferencia de los modelos no lineales. Por lo tanto, el modelo lineal podría no ser tan errado para los valores extremos pero sí podría serlo para determinar con precisión el punto de saturación en la experiencia.

Ejercicio 2g

En el modelo probit se evalúa la hipótesis nula de que la educación y la edad son conjuntamente iguales a cero. En particular, se testea lo siguiente:

$$H_0: \text{age} = \text{educ} = 0$$

$$H_1: \text{age} \neq 0 \text{ ó } \text{educ} \neq 0$$

Se computa un estadístico LR=80.19 y un p-valor de 0.0000. Ante estos resultados, se rechaza la hipótesis nula por lo que las variables educación y edad no son conjuntamente iguales a 0.