

Práctica de Ordenador PO-RE.1

Para la realización de este ejercicio utilizamos el fichero de datos *smoke* del libro de Wooldridge que tenéis como archivo de muestra en gretl. Son datos para 807 individuos varones residentes en distintos estados americanos en el año 1979. Las variables que están en este fichero son:

- *income*: renta familiar anual en miles de dólares.
- *lincome*: logaritmo de la renta familiar anual en miles de dólares.
- *cigs*: el número medio de cigarrillos fumados por día.
- *educ*: el número de años escolarizado.
- *age*: edad del individuo en años.
- *agesq*: edad del individuo en años elevado al cuadrado.
- *cigpric*: el precio de un paquete de cigarrillos en centavos.
- *lcigpric*: logaritmo del precio de un paquete de cigarrillos en centavos.
- *restaurn*: variable ficticia que es igual a la unidad si una persona reside en un estado donde hay restricciones al tabaquismo en los restaurantes, cero en otro caso.
- *white*: variable ficticia que es igual a la unidad si el individuo es blanco, cero en otro caso.

Puedes acceder a estos datos ejecutando gretl → En Archivo → Abrir datos → Archivo de muestra → Elige Wooldridge, fichero **smoke**¹

Considera la especificación del Modelo (1):

$$\text{lincome}_i = \beta_1 + \beta_2 \text{cigs}_i + \beta_3 \text{educ}_i + \beta_4 \text{age}_i + \beta_5 \text{agesq}_i + u_i \quad (1)$$

1. Muestra los resultados de la estimación por MCO del Modelo (1).
2. Comenta los resultados obtenidos sobre la bondad de ajuste, los coeficientes estimados y su significatividad. (“A realizar en casa”).
3. ¿Hay evidencia de que la relación entre la variable *lincome* y *age* sea cuadrática, manteniendo constante el resto de las variables explicativas? Muestra los resultados del contraste utilizado para tus conclusiones. (“A realizar en casa”).
4. Se cree que el consumo de cigarrillos puede estar determinado conjuntamente con la renta, tal que la variable **cigs** es un regresor estocástico correlacionado con el término de perturbación del Modelo (1). Explica las consecuencias que esto tiene sobre los resultados obtenidos en los apartados anteriores. (“A realizar en casa”).

¹Fichero *smoke*, disponible en gretl pestaña Wooldridge. Fuente: Wooldridge, J. M. (2003): *Introductory Econometrics*, fichero *smoke*.

5. Muestra los resultados de estimar el Modelo (1) por el método de Variables Instrumentales utilizando la variable **restaurn** como instrumento para **cigs**. ¿Son los resultados muy diferentes a los obtenidos por MCO? Comenta estos resultados.
6. Escribe la matriz de instrumentos Z y la matriz de variables explicativas del modelo, X . No pongas los valores muestrales, simplemente utiliza los nombres de las variables en las columnas. Escribe su dimensión. (“A realizar en casa”).
7. Escribe la expresión del estimador de VI utilizado. Escribe los elementos de $Z'X$ y $Z'Y$ utilizando sumatorios y su dimensión. ¿Qué características tiene que tener Z para que el estimador se pueda obtener? ¿Qué condiciones tiene que satisfacer Z para que el estimador tenga propiedades deseables y los contrastes sean válidos? (“A realizar en casa”).
8. Se dispone además de otro instrumento adicional para **cigs**, la variable **lcigpric**. Estima el Modelo (1) por Mínimos Cuadrados en 2 Etapas utilizando todos los instrumentos. Muestra el resultado obtenido en gretl. Compara estos resultados con los obtenidos en el apartado 5).
9. Calcula las correlaciones entre los instrumentos y la variable **cigs**. ¿Qué indican estas correlaciones sobre la bondad de estos instrumentos?
10. Realiza la regresión de la variable **cigs** sobre **todos los posibles instrumentos** incluida la constante:

$$cigs = \alpha_1 + \alpha_2 educ + \alpha_3 age + \alpha_4 age^2 + \alpha_5 lcigpric + \alpha_6 restaurn + u_i \quad (2)$$

Guarda la serie ajustada de la regresión $\widehat{cigs}_i$ $i = 1, \dots, 879$ y utiliza esta variable como instrumento para **cigs**. ¿Obtienes los mismos resultados que en el apartado 8)? ¿Por qué obtienes esos resultados? ¿Son las variables *lcigpric* y *restaurn* significativas?

11. Realiza el contraste de Hausman para los resultados del apartado 5). A la vista del resultado ¿Cómo estimarías los coeficientes del Modelo (1)?
12. Contrasta si la variable edad es significativa. ¿En cuánto se estima el cambio en la renta cuando el individuo tiene un año adicional y el resto de las características se mantienen constantes? ¿Es esta variación la misma para todos los individuos en la muestra?