

**EJEMPLO EMPIRICO II SOBRE
ESTIMACION DE SISTEMAS DE ECUACIONES
UTILIZANDO EL SOFTWARE LIBRE GRETL**

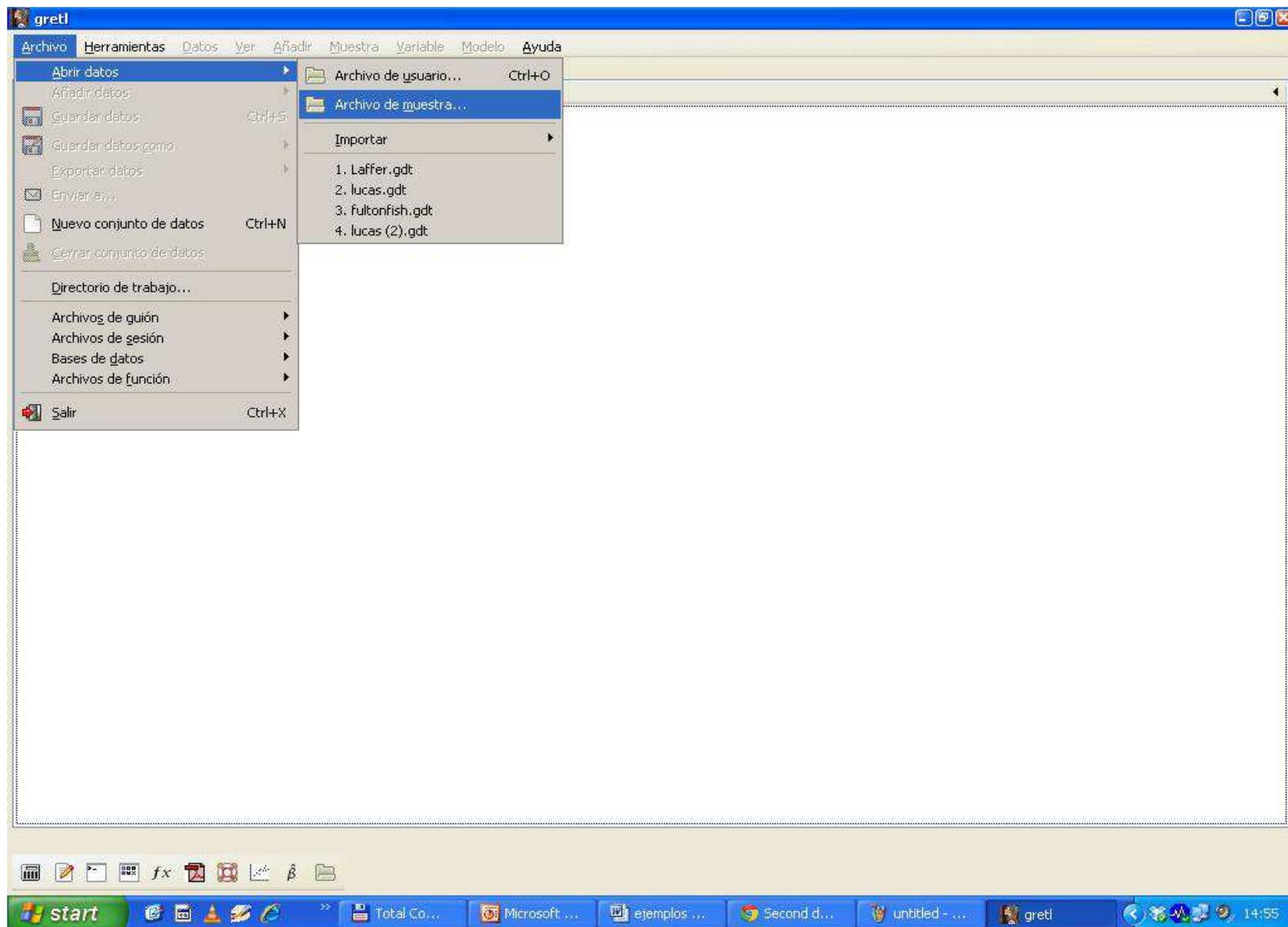
**OFERTA DE TRABAJO Y OFERTA SALARIAL (DEMANDA
DE TRABAJO)**

Ejemplo de estimación por MC2E, MC3E, MVIL y MVIC

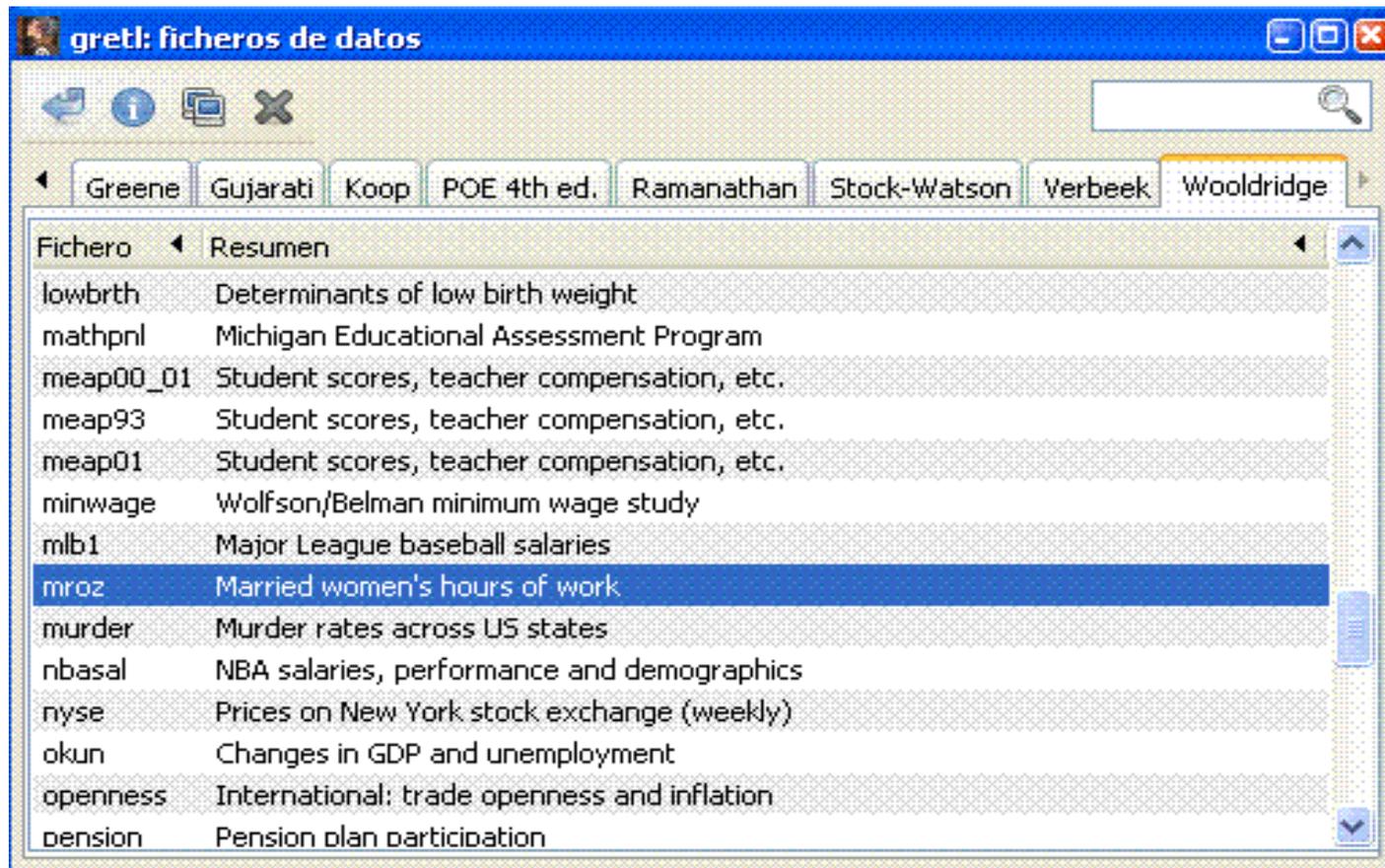
Marta Regúlez (UPV-EHU)

Fichero de datos y variables

- Usaremos el fichero de muestra en gretl `Mroz.gdt` asociado al libro de Wooldridge. Son datos utilizados por T.A. Mroz en su artículo de 1987, "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions" publicado en *Econometrica* 55, pp.765-799.
- Algunas de las variables que se muestran en el fichero son:
 - hours = horas trabajadas, 1975
 - kidslt6 = número de niños en la familia menores de 6 años
 - kidsge6 = número de niños en la familia entre 6 y 18 años
 - age = edad de la mujer en años
 - educ = años de educación
 - lwage = logaritmo del salario por hora
 - exper = experiencia laboral previa
 - nwifeinc = ingreso no salarial de la esposa
 - expersq = experiencia al cuadrado



Fichero de datos de muestra



gretl		
Archivo Herramientas Datos Ver Añadir Muestra Variable Modelo Ayuda		
mroz.gdt		
ID #	Nombre de variable	Etiqueta descriptiva
0	const	constante generada automáticamente
1	inlf	=1 if in labor force, 1975
2	hours	hours worked, 1975
3	kidslt6	# kids < 6 years
4	kidsge6	# kids 6-18
5	age	woman's age in yrs
6	educ	years of schooling
7	wage	estimated wage from earns., hours
8	repwage	reported wage at interview in 1976
9	hushrs	hours worked by husband, 1975
10	husage	husband's age
11	huseduc	husband's years of schooling
12	huswage	husband's hourly wage, 1975
13	faminc	family income, 1975
14	mtr	fed. marginal tax rate facing woman
15	motheduc	mother's years of schooling
16	fatheduc	father's years of schooling
17	unem	unem. rate in county of resid.
18	city	=1 if live in SMSA
19	exper	actual labor mkt exper
20	nwifeinc	(faminc - wage*hours)/1000
21	lwage	log(wage)
22	expersq	exper^2

Especificación 1. Forma Estructural

Son datos de sección cruzada y aunque el fichero tiene un total de 753 individuos, todos ellos mujeres, solamente consideramos las que se encontraban trabajando en 1975 que son un total de 428. Las variables consideradas endógenas son *hours* y *lwage*, siendo el resto variables exógenas.

Partimos de la siguiente forma estructural del modelo donde la primera ecuación es la función de oferta de mujeres casadas trabajadoras y la segunda ecuación es una función de oferta salarial ya que hemos normalizado con respecto a *lwage*. Esto implica que *lwage* es la variable dependiente de esa ecuación.

$$hours = \gamma_{11} + \beta_{12}lwage + \gamma_{12}educ + \gamma_{13}age + \gamma_{14}kidslt6 + \gamma_{15}kidsge6 + \gamma_{16}nwifeinc + u_1$$

$$lwage = \gamma_{21} + \beta_{21}hours + \gamma_{22}educ + \gamma_{23}exper + \gamma_{24}expersq + u_2$$

Identificación Forma Estructural

La primera ecuación satisface la condición de orden, necesaria para la identificación, con exceso. El número de variables exógenas excluidas de esa ecuación, $K_1^* = 2$, es mayor que el número de endógenas incluidas en la ecuación como explicativas, $G_1 = 1$. Por lo tanto, la ecuación estará sobreidentificada si se satisface la condición de rango. Esta condición, necesaria y suficiente para la identificación, se satisface si $\gamma_{23} \neq 0$ y \ o $\gamma_{24} \neq 0$.

La segunda ecuación satisface la condición de orden también con exceso. El número de variables exógenas excluidas de esa ecuación, $K_2^* = 4$, es mayor que el número de endógenas incluidas en la ecuación como explicativas, $G_2 = 1$. Por lo tanto, la ecuación estará sobreidentificada si se satisface la condición de rango. Esta condición, necesaria y suficiente para la identificación, se satisface si $\gamma_{13} \neq 0$ y, o $\gamma_{14} \neq 0$ y, o $\gamma_{15} \neq 0$ y, o $\gamma_{16} \neq 0$.

Estimación MCO Forma Estructural

Modelo → Mínimos Cuadrados Ordinarios

Estimaciones MCO de la forma Estructural

Función de oferta de mujeres casadas trabajadoras (N=428)

Variable dependiente: hours

<i>Variable</i>	<i>Coefficiente</i>	<i>Desv. típica</i>	<i>Estadístico t</i>	<i>valor p</i>	
const	2114,7	340,131	6,2173	<0,00001	***
lwage	-17,4078	54,2154	-0,3211	0,74830	
educ	-14,4449	17,9679	-0,8039	0,42189	
age	-7,72998	5,52945	-1,3980	0,16286	
kidslt6	-342,505	100,006	-3,4248	0,00068	***
kidsge6	-115,021	30,8293	-3,7309	0,00022	***
nwifeinc	-4,24581	3,65581	-1,1614	0,24614	

La estimación MCO del coeficiente que afecta a lwage indica una función de oferta con pendiente negativa, signo no esperado. Aunque no es significativamente distinto de cero.

Estimación MCO Forma Estructural

Función de demanda de trabajo (o mejor Función de oferta salarial dado que normalizamos con respecto del salario (lwage))

Variable dependiente: lwage

<i>Variable</i>	<i>Coefficiente</i>	<i>Desv. típica</i>	<i>Estadístico t</i>	<i>valor p</i>	
const	-0,461995	0,203848	-2,2664	0,02393	**
hours	-5,65486e-05	4,37823e-05	-1,2916	0,19721	
educ	0,106214	0,0141698	7,4958	<0,00001	***
exper	0,0447035	0,013387	3,3393	0,00091	***
expersq	-0,000858544	0,000394639	-2,1755	0,03014	**

El coeficiente estimado que acompaña a hours tampoco tiene el signo esperado, aunque tampoco es estadísticamente significativo, por lo que no hay evidencia de que la oferta salarial varíe con las horas trabajadas.

Puede que estos efectos no esperados sean consecuencia de los sesgos en la estimación por MCO si hours y lwage son ambas variables endógenas. Consideremos estimar por MC2E y MC3E, métodos de variables instrumentales que tienen en cuenta esa posible endogeneidad.

Estimación de la forma estructural por métodos alternativos a MCO

Para estimar un sistema de ecuaciones en Gretl por todos estos métodos de estimación elegimos:

Modelo → Ecuaciones simultáneas

En la ventana que aparece se han de escribir las ecuaciones del modelo de la siguiente forma:

```
equation  hours  const  lwage  educ  age  kidslt6  kidsge6  nwifeinc
equation  lwage  const  hours  educ  exper  expersq
endog     hours  lwage
```

Dado que en la tercera línea se han identificado cuales son las variables endógenas del sistema, Gretl asigna al resto la categoría de variables exógenas.

En la parte inferior de la ventana elegimos el estimador que se quiere utilizar, dentro de un listado que nos muestra el programa. Los resultados de estimar cada una de estas ecuaciones por los distintos métodos se muestran en el siguiente cuadro resumen:

Estimación MC2E de la forma estructural

gretl: sistema de ecuaciones simultáneas

Especificar ecuaciones simultáneas:
(Por favor, consulte la Ayuda para más información)

```
equation hours 0 lwage educ age kidslt6 kidsge6 nwifeinc  
equation lwage 0 hours educ exper expersq  
endog hours lwage
```

Estimación iterativa

Mostrar los detalles de las iteraciones

Estimador: Mínimos cuadrados en dos etapas (tsls)

Ayuda Cancelar Aceptar

Sistema de ecuaciones, Mínimos cuadrados en dos etapas

Ecuación 1: estimaciones MC2E

utilizando las 428 observaciones 1-428

Variable dependiente: hours

Instrumentos: const educ age kidslt6 kidsge6 nwifeinc exper expersq

VARIABLE	COEFICIENTE	DESV.TÍP.	ESTAD T	VALOR P	
const	2432,20	594,172	4,093	0,00004	***
lwage	1544,82	480,739	3,213	0,00131	***
educ	-177,449	58,1426	-3,052	0,00227	***
age	-10,7841	9,57735	-1,126	0,26017	
kidslt6	-210,834	176,934	-1,192	0,23342	
kidsge6	-47,5571	56,9179	-0,836	0,40341	
nwifeinc	-9,24912	6,48112	-1,427	0,15355	

Media de la var. dependiente = 1302,93

Desviación típica de la var. dependiente. = 776,274

Suma de cuadrados de los residuos = 7,13583e+008

Desviación típica de los residuos = 1301,91

Ecuación 2: estimaciones MC2E

utilizando las 428 observaciones 1-428

Variable dependiente: lwage

Instrumentos: const educ age kidslt6 kidsge6 nwifeinc exper expersq

VARIABLE	COEFICIENTE	DESV.TÍP.	ESTAD T	VALOR P	
const	-0,692790	0,306600	-2,260	0,02385	**
hours	0,000160806	0,000215408	0,747	0,45535	
educ	0,111118	0,0153319	7,247	<0,00001	***
exper	0,0326460	0,0180610	1,808	0,07068	*
expersq	-0,000676540	0,000442636	-1,528	0,12640	

Media de la var. dependiente = 1,19017

Desviación típica de la var. dependiente. = 0,723198

Suma de cuadrados de los residuos = 198,494

Desviación típica de los residuos = 0,68502

Estimación MC3E de la forma estructural

The screenshot shows a dialog box titled "gretl: sistema de ecuaciones simultáneas". The main area contains the following text:

```
equation hours 0 lwage educ age kidslt6 kidsge6 nwifeinc
equation lwage 0 hours educ exper expersq
endog hours lwage
```

Below the text area, there are two checkboxes:

- Estimación iterativa
- Mostrar los detalles de las iteraciones

The "Estimador" dropdown menu is set to "Mínimos cuadrados en tres etapas (3sls)".

At the bottom, there are four buttons: "Ayuda", "Limpiar", "Cancelar", and "Aceptar".

SISTEMA DE ECUACIONES, MÍNIMOS CUADRADOS EN TRES ETAPAS

Ecuación 1: estimaciones MC3E

utilizando las 428 observaciones 1-428

Variable dependiente: hours

Instrumentos: const educ age kidslt6 kidsge6 nwifeinc exper expersq

VARIABLE	COEFICIENTE	DESV.TÍP.	ESTAD T	VALOR P	
const	2504,80	535,892	4,674	<0,00001	***
lwage	1676,93	431,169	3,889	0,00010	***
educ	-205,027	51,8473	-3,954	0,00008	***
age	-12,2812	8,26153	-1,487	0,13713	
kidslt6	-200,567	134,268	-1,494	0,13523	
kidsge6	-48,6399	35,9514	-1,353	0,17608	
nwifeinc	0,367895	3,45152	0,107	0,91511	

Media de la var. dependiente = 1302,93

Desviación típica de la var. dependiente. = 776,274

Suma de cuadrados de los residuos = 8,01393e+008

Desviación típica de los residuos = 1368,36

Ecuación 2: estimaciones MC3E

utilizando las 428 observaciones 1-428

Variable dependiente: lwage

Instrumentos: const educ age kidslt6 kidsge6 nwifeinc exper expersq

VARIABLE	COEFICIENTE	DESV.TÍP.	ESTAD T	VALOR P	
const	-0,705110	0,304590	-2,315	0,02062	**
hours	0,000201031	0,000210881	0,953	0,34044	
educ	0,112970	0,0151452	7,459	<0,00001	***
exper	0,0208906	0,0142782	1,463	0,14344	
expersq	-0,000294293	0,000261380	-1,126	0,26020	

Media de la var. dependiente = 1,19017

Desviación típica de la var. dependiente. = 0,723198

Suma de cuadrados de los residuos = 203,333

Desviación típica de los residuos = 0,689258

Estimación MVIL de la forma estructural

gretl: sistema de ecuaciones simultáneas

Especificar ecuaciones simultáneas:
(Por favor, consulte la Ayuda para más información)

```
equation hours const lwage educ age kidslt6 kidsge6 nwifeinc  
equation lwage const hours educ exper expersq  
endog hours lwage
```

Estimación iterativa

Mostrar los detalles de las iteraciones

Estimador: Máxima Verosimilitud con Información Limitada (liml)

ESTIMACION POR MVIL

Endogenous variables: hours lwage

Exogenous variables: const educ age kidslt6 kidsge6 nwifeinc exper expersq

Sistema de ecuaciones, Máxima Verosimilitud con Información Limitada

Ecuación 1: estimaciones MVIL

utilizando las 428 observaciones 1-428

Variable dependiente: hours

VARIABLE	COEFICIENTE	DESV.TÍP.	ESTAD T	VALOR P
const	2449,33	616,070	3,976	0,00007 ***
lwage	1629,13	510,876	3,189	0,00143 ***
educ	-186,247	61,3963	-3,034	0,00242 ***
age	-10,9489	9,92583	-1,103	0,26999
kidslt6	-203,727	183,576	-1,110	0,26710
kidsge6	-43,9160	59,1775	-0,742	0,45802
nwifeinc	-9,51916	6,72509	-1,415	0,15693

Media de la var. dependiente = 1302,93

Desviación típica de la var. dependiente. = 776,274

Suma de cuadrados de los residuos = 7,66074e+008

Desviación típica de los residuos = 1348,95

Log-verosimilitud = -6016,92

Smallest eigenvalue = 1,00194

Contraste de sobreidentificación LR:

Chi-cuadrado(1) = 0,829301 con valor p 0,3625

Ecuación 2: estimaciones MVIL
 utilizando las 428 observaciones 1-428
 Variable dependiente: lwage

VARIABLE	COEFICIENTE	DESV.TÍP.	ESTAD T	VALOR P
const	-0,735315	0,324821	-2,264	0,02359 **
hours	0,000200855	0,000236189	0,850	0,39510
educ	0,112021	0,0156374	7,164	<0,00001 ***
exper	0,0304243	0,0189511	1,605	0,10840
expersq	-0,000643005	0,000454012	-1,416	0,15670

Media de la var. dependiente = 1,19017
 Desviación típica de la var. dependiente. = 0,723198
 Suma de cuadrados de los residuos = 202,892
 Desviación típica de los residuos = 0,692568
 Log-verosimilitud = -6017,97
 (Log-verosimilitud para wage = -940,696)
 Smallest eigenvalue = 1,00685
 Contraste de sobreidentificación LR:
 Chi-cuadrado(3) = 2,92124 con valor p 0,4039

Matriz de covarianzas cruzada residual
 (correlaciones por encima de la diagonal principal)

1,7899e+006	(-0,934)
-859,91	0,47405

logaritmo del determinante = 11,5996

Estimación MVIC de la forma estructural



gretl: sistema de ecuaciones simultáneas

Especificar ecuaciones simultáneas:
(Por favor, consulte la Ayuda para más información)

```
equation hours const lwage educ age kidslt6 kidsge6 nwifeinc  
equation lwage const hours educ exper expensq  
endog hours lwage
```

Estimación iterativa

Mostrar los detalles de las iteraciones

Estimador: Máxima Verosimilitud con Información Completa (fiml)

Ayuda Cancelar Aceptar

Sistema de ecuaciones, Máxima Verosimilitud con Información Completa

Se alcanzó la convergencia después de 15 iteraciones

Log-verosimilitud = -3853,14

Ecuación 1: estimaciones MVIC

utilizando las 428 observaciones 1-428

Variable dependiente: hours

VARIABLE	COEFICIENTE	DESV.TÍP.	ESTAD T	VALOR P	
const	2435,10	579,001	4,206	0,00003	***
lwage	1773,93	497,304	3,567	0,00036	***
educ	-216,729	61,8412	-3,505	0,00046	***
age	-10,5961	8,84614	-1,198	0,23099	
kidslt6	-167,984	143,024	-1,175	0,24019	
kidsge6	-40,8436	36,5103	-1,119	0,26327	
nwifeinc	1,24342	2,13017	0,584	0,55941	

Media de la var. dependiente = 1302,93

Desviación típica de la var. dependiente. = 776,274

Suma de cuadrados de los residuos = 8,68212e+008

Desviación típica de los residuos = 1424,27

Ecuación 2: estimaciones MVIC
 utilizando las 428 observaciones 1-428
 Variable dependiente: lwage

VARIABLE	COEFICIENTE	DESV.TÍP.	ESTAD T	VALOR P
const	-0,740600	0,314122	-2,358	0,01839 **
hours	0,000245573	0,000223153	1,100	0,27113
educ	0,113986	0,0156199	7,298	<0,00001 ***
exper	0,0171624	0,0142774	1,202	0,22934
expersq	-0,000238085	0,000226117	-1,053	0,29237

Media de la var. dependiente = 1,19017

Desviación típica de la var. dependiente. = 0,723198

Suma de cuadrados de los residuos = 209,093

Desviación típica de los residuos = 0,698952

Matriz de covarianzas cruzada residual
 (correlaciones por encima de la diagonal principal)

2,0285e+006	(-0,963)
-958,61	0,48853

logaritmo del determinante = 11,1855

- Primera ecuación estructural:

Función de oferta de mujeres casadas trabajadoras

Variable dependiente: hours

Explicativas	MCO	MC2E	MVIL	MC3E	MVIC
const	2114, 7* (340, 131)	2432, 20* (594, 172)	2449, 33* (616, 070)	2504, 80* (535, 892)	2435, 10* (579, 001)
lwage	-17, 4078 (54, 2154)	1544, 82* (480, 739)	1629, 13* (510, 876)	1676, 93* (431, 169)	1773, 93* (497, 304)
educ	-14, 4449 (17, 9679)	-177, 449* (58, 1426)	-186, 247* (61, 3963)	-205, 027* (51, 8473)	-216, 729* (61, 8412)
age	-7, 72998 (5, 52945)	-10, 7841 (9, 57735)	-10, 9489 (9, 92583)	-12, 2812 (8, 26153)	-10, 5961 (8, 84614)
kidslt6	-342, 505* (100, 006)	-210, 834 (176, 934)	-203, 727 (183, 576)	-200, 567 (134, 268)	-167, 984 (143, 024)
kidsge6	-115, 021* (30, 8293)	-47, 5571 (56, 9179)	-43, 9160 (59, 1775)	-48, 6399 (35, 9514)	-40, 8436 (36, 5103)
nwifeinc	-4, 24581 (3, 65581)	-9, 24912 (6, 48112)	-9, 51916 (6, 72509)	0, 367895 (3, 45152)	1, 24342 (2, 13017)

* significativo al 5% $N = 428$

Menor valor propio estimado = 1,00194

Contraste LR: Chi-cuadrado(1) = 0,829301 v.p. = 0,3625

Contraste de Sargan : TR2 = 0,858169 valor p = 0,354252

Contraste de Hausman - Chi-cuadrado(1) = 35,9481

Resultados obtenidos de la estimación de la primera ecuación estructural:

- a) La primera ecuación estructural o curva de oferta estimada tanto por MC2E, MVIL, MC3E y MVIC tiene una pendiente positiva y significativa². A efectos de comparación cuando esta ecuación de oferta de trabajo se estima por MCO, el coeficiente estimado asociado a *lwage* es negativo y no significativo, lo que implica la ausencia de efectos de variaciones en el salario sobre las horas trabajadas.
- b) El coeficiente estimado para *lwage* se interpreta como la variación estimada en el número de horas ofrecidas por una mujer trabajadora a una variación unitaria en el salario, manteniendo constantes los demás factores. Por ejemplo, utilizando la estimación MC2E, sería de un 15,4% de la variación en el salario. Utilizando MVIL, MC3E y MVIC es algo mayor.
- c) La elasticidad de la oferta de trabajo respecto del salario no es constante en este modelo, porque es *hours* y no $\log(hours)$ la variable dependiente en la ecuación de oferta, primera ecuación del modelo. Por ejemplo, utilizando el método MC2E, al nivel medio de horas trabajadas, que es de 1303, la elasticidad estimada es de 1540/1303 que es algo menor que 1,18. Esto implica que en respuesta a un incremento salarial del 1%, se da un incremento de más de un 1% en las horas trabajadas. La elasticidad será menor para niveles más altos de horas ofrecidas. A niveles más bajos, como por ejemplo $hours = 700$, la elasticidad es superior a 2.

Resultados obtenidos de la estimación de la primera ecuación estructural:

- d) También hay diferencias importantes para los efectos de la variable *educ* entre MCO y el resto de métodos que tienen en cuenta la posible endogeneidad de *lwage*. El coeficiente estimado que acompaña a esta variable es negativo en todos los casos pero no así su significatividad. Utilizando MCO no es significativo pasando a serlo con el resto de métodos, así como aumentando considerablemente su magnitud en valor absoluto.
- e) En cuanto al resto de variables, también cambia la significatividad de los coeficientes asociados a la existencia de hijos menores de 6 años, *kidslt6*, e hijos entre 6 y 18 años, *kidsge6*, pasando de ser ambas variables significativas con MCO a no serlo con el resto de métodos, aunque el signo negativo se mantiene y es el esperado. Por otro lado, las variables *age* y *nwifeinc* no son significativas en todos los casos.

- Segunda ecuación estructural:

Función de oferta salarial
Variable dependiente: lwage

Explicativas	MCO	MC2E	MVIL	MC3E	MVIC
const	-0,461995* (0,203848)	-0,69279* (0,3066)	-0,735315* (0,324821)	-0,705110* (0,304590)	-0,740600* (0,314122)
hours	-0,000056 (0,0000438)	0,000161 (0,000215)	0,000201 (0,0002362)	0,000201 (0,0002108)	0,0002456 (0,0002232)
educ	0,106214* (0,0141698)	0,111118* (0,0153319)	0,112021* (0,0156374)	0,112970* (0,0151452)	0,113986* (0,01562)
exper	0,0447035* (0,013387)	0,032646 (0,018061)	0,0304243 (0,0189511)	0,0208906 (0,0142782)	0,0171624 (0,0142774)
expersq	-0,000858* (0,000395)	-0,0006765 (0,0004426)	-0,000643 (0,000454)	0,00029433 (0,0002614)	-0,0002381 (0,0002261)

* significativo al 5% $N = 428$

Menor valor propio estimado = 1,00685

Contraste LR: Chi-cuadrado(3) = 2,92124 con valor p = 0,4039

Contraste de Sargan : TR2 = 2,94083 con valor p = 0,4008

Contraste de Hausman - Chi-cuadrado(1) = 1,14311

Contraste de sobreidentificación en todo el sistema de Hansen-Sargan:

Chi-cuadrado(4) = 4,10677 con valor p = 0,3917

Resultados obtenidos de la estimación de la segunda ecuación estructural:

- a) El coeficiente que acompaña a la variable *hours* no es significativo tanto utilizando MCO como el resto de métodos, aunque el signo cambia siendo negativo en el primer caso y positivo en el resto.
- b) En cuanto al resto de variables, el coeficiente que acompaña a *educ* se mantiene con signo positivo y significativo en todos los casos, no así para las variables que recogen el efecto sobre la oferta salarial del nivel de experiencia. Los signos de estos coeficientes se mantienen pero pasan de ser significativos con MCO a no serlo en el resto de métodos.

Contrastes de especificación y exogeneidad

a) Contrastes de Sargan y Hausman.

Al estimar por MC2E en Gretl utilizando

Modelo → Otros modelos lineales → Mínimos Cuadrados en 2 etapas

además de mostrar los resultados de la estimación por MC2E también nos muestra los resultados del Contraste de sobreidentificación de Sargan y el contraste de Hausman comparando MC2E y MCO. En cambio, si se utiliza la opción

Modelo → Ecuaciones simultáneas → Mínimos Cuadrados en 2 etapas

solamente se muestra el resultado del contraste de Sargan.

Contrastes de especificación y exogeneidad

Los resultados de estos contrastes en cada una de las ecuaciones indica lo siguiente:

- Tanto en el caso de la primera ecuación como en la segunda, el contraste de sobreidentificación de Sargan no rechaza la hipótesis nula de que todos los instrumentos son válidos.
- En cuanto a los resultados del contraste de Hausman, para la primera ecuación estructural, se rechaza la hipótesis nula de que la variable *lwage* no esté correlacionado con el término de perturbación de esa ecuación. Esto indica evidencia de endogeneidad de *lwage*. Considerando en cambio la segunda ecuación estructural, el contraste de Hausman no rechaza la hipótesis nula de incorrelación entre la variable *hours* y el término de perturbación de esa ecuación.

Contrastes de especificación y exogeneidad

- b) Los contrastes de sobreidentificación LR en ambas ecuaciones no rechazan la hipótesis nula, lo que indica que se aceptan como buenas las restricciones de identificación en cada ecuación.
- c) Por último el contraste de sobreidentificación de Hansen-Sargan sobre la bondad de las restricciones de sobreidentificación en el sistema no rechaza la hipótesis nula. Los instrumentos parecen adecuados.

Estimación de la Forma Reducida

La forma reducida del modelo es el siguiente sistema de dos ecuaciones donde en cada ecuación una de las variables endógenas es función de todo el conjunto de variables exógenas.

$$\begin{aligned} \text{hours} &= \pi_{11} + \pi_{12}\text{educ} + \pi_{13}\text{age} + \pi_{14}\text{kidslt6} + \pi_{15}\text{kidsge6} + \pi_{16}\text{nwifeinc} + \pi_{17}\text{exper} + \pi_{18}\text{expersq} + v_1 \\ \text{lwage} &= \pi_{21} + \pi_{22}\text{educ} + \pi_{23}\text{age} + \pi_{24}\text{kidslt6} + \pi_{25}\text{kidsge6} + \pi_{26}\text{nwifeinc} + \pi_{27}\text{exper} + \pi_{28}\text{expersq} + v_2 \end{aligned}$$

La estimación de la forma reducida nos da la posibilidad de analizar la bondad de los instrumentos para cada ecuación. Podemos contrastar el grado de correlación de los instrumentos, o variables exógenas, con las variables endógenas *hours* y *lwage*.

Estimación de la Forma Reducida: 1ª ecuación (para hours)

Forma Reducida: Ecuación para hours.

Estimaciones MCO utilizando las 428 observaciones 1–428
Variable dependiente: hours

Variable	Coefficiente	Desv. típica	Estadístico <i>t</i>	valor p
const	2056,64	346,484	5,9357	0,0000
educ	–22,788	16,4345	–1,3866	0,1663
age	–19,663	5,89403	–3,3362	0,0009
kidslt6	–305,72	96,4501	–3,1697	0,0016
kidsge6	–72,366	30,3610	–2,3835	0,0176
nwifeinc	0,443852	3,61350	0,1228	0,9023
exper	47,0051	14,5565	3,2291	0,0013
expersq	–0,513644	0,437358	–1,1744	0,2409

Bondad de los instrumentos para la variable endógena *hours* en la segunda ecuación estructural

Contraste de omisión de variables -

Hipótesis nula: los parámetros son cero para las variables

age

kidslt6

kidsge6

nwifeinc

Estadístico de contraste: $F(4, 420) = 4,80035$

con valor $p = P(F(4, 420) > 4,80035) = 0,00085263$

Se rechaza la hipótesis nula, por lo que alguno o algunos de los instrumentos *age*, *kidslt6*, *kidsge6*, *nwifeinc* disponibles para instrumentalizar *hours* en la segunda ecuación (ecuación de oferta salarial) están suficientemente correlacionados con la variable a instrumentalizar que es la variable *hours*.

Estimación de la Forma Reducida: 2ª ecuación (para lwage)

Forma Reducida: Ecuación para lwage

Estimaciones MCO utilizando las 428 observaciones 1–428
Variable dependiente: lwage

Variable	Coefficiente	Desv. típica	Estadístico <i>t</i>	valor p
const	−0,357997	0,318296	−1,1247	0,2613
educ	0,0998844	0,0150975	6,6160	0,0000
age	−0,00352039	0,00541452	−0,6502	0,5159
kidslt6	−0,0558725	0,0886034	−0,6306	0,5287
kidsge6	−0,0176485	0,0278910	−0,6328	0,5272
nwifeinc	0,00569422	0,00331952	1,7154	0,0870
exper	0,0407097	0,0133723	3,0443	0,0025
expersq	−0,000747326	0,000401777	−1,8601	0,0636

Bondad de los instrumentos para la variable endógena *lwage* en la primera ecuación estructural

Contraste de omisión de variables -

Hipótesis nula: los parámetros son cero para las variables
exper
expersq

Estadístico de contraste: $F(2, 420) = 8,25023$

con valor $p = P(F(2, 420) > 8,25023) = 0,000305887$

Se rechaza la hipótesis nula, por lo que para instrumentalizar a *lwage* en la primera ecuación estructural o función de oferta de trabajo, los instrumentos *exper* y *expersq* parecen estar correlacionadas con esta variable.

Cambio de normalización de la segunda ecuación estructural

En la nueva forma estructural del modelo la primera ecuación sigue siendo la misma, la función de oferta de mujeres casadas trabajadoras y la segunda ecuación es distinta ya que hemos normalizado con respecto a *hours* en lugar de *lwage*. Esto implica que *hours* es ahora la variable dependiente de esa ecuación.

$$hours = \gamma_{11} + \beta_{12}lwage + \gamma_{12}educ + \gamma_{13}age + \gamma_{14}kidslt6 + \gamma_{15}kidsge6 + \gamma_{16}nwifeinc + u_1$$

$$hours = \gamma_{21} + \beta_{21}lwage + \gamma_{22}educ + \gamma_{23}exper + \gamma_{24}expersq + u_2$$

¿Qué ocurre si en la segunda ecuación de la Forma Estructural normalizamos con respecto a *hours* y tenemos a *lwage* como endógena explicativa? En este caso en la segunda ecuación tenemos que utilizar como instrumentos para *lwage* a las variables exógenas que no aparecen en esa ecuación, esto es, *kidslt6*, *kidsge6*, *age* y *nwifeinc*.

Bondad de los instrumentos para la variable endógena lwage en la segunda ecuación

Forma Reducida: Ecuación para lwage

Estimaciones MCO utilizando las 428 observaciones 1–428
Variable dependiente: lwage

Variable	Coefficiente	Desv. típica	Estadístico <i>t</i>	valor p
const	−0,357997	0,318296	−1,1247	0,2613
educ	0,0998844	0,0150975	6,6160	0,0000
age	−0,00352039	0,00541452	−0,6502	0,5159
kidslt6	−0,0558725	0,0886034	−0,6306	0,5287
kidsge6	−0,0176485	0,0278910	−0,6328	0,5272
nwifeinc	0,00569422	0,00331952	1,7154	0,0870
exper	0,0407097	0,0133723	3,0443	0,0025
expersq	−0,000747326	0,000401777	−1,8601	0,0636

Contraste de omisión de variables -

Hipótesis nula: los parámetros son cero para las variables

kidslt6

kidsge6

age

nwifeinc

Estadístico de contraste: $F(4, 420) = 0,914214$ con valor $p = 0,455487$

Bondad de los instrumentos para la variable endógena $lwage$ en la segunda ecuación

Estos instrumentos no parecen estar correlacionadas con $lwage$, dado que no se rechaza la hipótesis nula. ¿Qué consecuencias tiene este hecho en la estimación de la forma estructural? Esto lo tratamos en la siguiente sección.

Estimación de esta nueva especificación de la Forma Estructural

Modelo → Ecuaciones simultáneas

En la ventana que aparece se han de escribir las ecuaciones del modelo de la siguiente forma:

```
equation  hours  const  lwage  educ  age  kidslt6  kidsge6  nwifeinc
equation  hours  const  lwage  educ  exper  expersq
endog     hours  lwage
```

¿Qué ocurre si en la segunda ecuación de la Forma Estructural normalizamos con respecto a *hours* y tenemos a *lwage* como endógena explicativa? En este caso en la segunda ecuación tenemos que utilizar como instrumentos para *lwage* a las variables exógenas que no aparecen en esa ecuación, esto es, *kidslt6*, *kidsge6*, *age* y *nwifeinc*.

Estimación de esta nueva especificación de la Forma Estructural

CAMBIO DE NORMALIZACION EN LA SEGUNDA ECUACION

Función de oferta de mujeres casadas trabajadoras

Explicativas	Variable dependiente: hours				
	MCO	MC2E	MVIL	MC3E	MVIC
const	2114,7 * (340,131)	2432,20 * (594,172)	2449,33 * (616,070)	1575,53 * (406,597)	2435,10 * (579,001)
lwage	-17,4078 (54,2154)	1544,82 * (480,739)	1629,13 * (510,876)	1649,56 * (425,531)	1773,93 * (497,304)
educ	-14,4449 (17,9679)	-177,449 * (58,1426)	-186,247 * (61,3963)	-197,057 * (52,2255)	-216,729 * (61,8412)
age	-7,72998 (5,52945)	-10,7841 (9,57735)	-10,9489 (9,92583)	6,36431 (4,29410)	-10,5961 (8,84614)
kidslt6	-342,505 * (100,006)	-210,834 (176,934)	-203,727 (183,576)	97,7946 (70,6825)	-167,984 (143,024)
kidsge6	-115,021 * (30,8293)	-47,5571 (56,9179)	-43,9160 (59,1775)	25,1756 (22,5259)	-40,8436 (36,5103)
nwifeinc	-4,24581 (3,65581)	-9,24912 (6,48112)	-9,51916 (6,72509)	-2,96603 (5,98872)	1,24342 (2,13017)

* significativo al 5% N=428

Menor valor propio estimado = 1,00194

Contraste de sobreidentificación LR: Chi-cuadrado(1) = 0,829301 v. p. 0,3625

Como es de esperar en los métodos de información limitada las estimaciones de la primera ecuación no cambian. No así en los de información completa ya que influye la especificación de la segunda ecuación en la estimación de la primera.

Estimación de esta nueva especificación de la Forma Estructural

Función de demanda de trabajo (normalización sobre hours)

Variable dependiente: hours

Explicativas	MCO	MC2E	MVIL	MC3E	MVIC
const	1025,57 * (221,765)	1584,15 * (520,055)	3660,92 (3224,49)	1525,46 (506,819)	3015,81 (1951,40)
lwage	-69,4665 (53,7838)	1000,53 (805,418)	4978,71 (5854,57)	938,526 (793,394)	4072,11 (3700,34)
educ	-15,0935 (16,6995)	-130,108 (89,2759)	-557,721 (633,548)	-124,260 (88,1676)	-464,165 (405,994)
exper	58,3612 * (14,7615)	13,8850 (39,1525)	-151,474 (252,721)	24,7100 (34,0601)	-69,8873 (116,278)
expersq	-0,893710 * (0,437687)	-0,0257319 (0,891065)	3,20134 (5,16668)	-0,377879 (0,639631)	0,969511 (1,68077)

* significativo al 5%

N=428

Menor valor propio estimado = 1,00685

Contraste de sobreidentificación LR: Chi-cuadrado(3) = 2,92124 v.p. 0,4039

(Es invariante a la normalización) No así el de Sargan basado en MC2E

Contraste de sobreidentificación de Sargan - $TR^2 = 8,14343$ valor p = 0,0431385

Contraste de Hausman - Chi-cuadrado(1) = 3,51544 con valor p = 0,0607996.

Estimación de esta nueva especificación de la Forma Estructural

Contraste de sobreidentificación de Hansen-Sargan en todo el sistema:

Chi-cuadrado(4) = 8,4736 con valor p 0,0757 (Este contraste tampoco es invariante)

El valor del estadístico de Hansen-Sargan es un valor bastante mayor que el que se obtenía con la normalización para *lwage*. Se rechaza la hipótesis nula al 1%. En este caso el contraste de sobre identificación señala que los instrumentos no son adecuados.

Los resultados para la primera ecuación no cambian porque es un método de información limitada y no influye el cambio en la normalización de la segunda ecuación. Pero en el caso de los coeficientes de la segunda ecuación ninguno es significativo

¿Y si estimamos por un método de Información Completa, por ejemplo MC3E o MVIC?

Observamos lo mismo que ocurría con los métodos de información limitada. La especificación de la segunda ecuación normalizando por *hours* en lugar de por *lwage* no es la más adecuada ya que no se pueden identificar bien los coeficientes con los instrumentos disponibles.

En el caso del método de estimación MC3E, a diferencia de lo que ocurría en los métodos de información limitada, la normalización de la segunda ecuación afecta tanto a la estimación de la primera ecuación como de la segunda. En cambio, en el caso de MVIC, la estimación de la primera ecuación no se ve afectada por la normalización elegida para la segunda ecuación.

Contraste de Sargan

Aunque Gretl muestra el valor de este estadístico cuando se estima por MC2E, en esta sección vamos a explicar como se calcula. Consideremos primero la primera ecuación estructural (Función de oferta de trabajo). El procedimiento es el siguiente:

- Estimar por MC2E y guardar los residuos que denotamos por *uhat1*.
- Realizamos la regresión de los residuos MC2E, *uhat1*, sobre todas las variables exógenas. Obtenemos los siguientes resultados:

Estimaciones MCO utilizando las 428 observaciones 1–428
Variable dependiente: *uhat1*

Variable	Coefficiente	Desv. típica	Estadístico <i>t</i>	valor p
const	177,486	621,684	0,2855	0,7754
educ	0,357340	29,4878	0,0121	0,9903
age	-3,4410	10,5754	-0,3254	0,7451
kidslt6	-8,5740	173,057	-0,0495	0,9605
kidsge6	2,45400	54,4756	0,0450	0,9641
nwifeinc	0,896436	6,48356	0,1383	0,8901
exper	-15,884	26,1182	-0,6082	0,5434
expersq	0,640840	0,784734	0,8166	0,4146

R^2

0,00200507

Contraste de Sargan

El valor del estadístico de Sargan se calcula con el tamaño muestral $T = 428$ y el coeficiente de determinación $R^2 = 0,00200507$ de esta última regresión. El valor muestral obtenido $TR^2 = 2,94$ es menor que el valor crítico $\chi^2(3)_{0,05} = 7,81473$. Por lo tanto, no se rechaza las restricciones de sobreidentificación, que son 3 en este caso. Los instrumentos y la especificación de la segunda ecuación parece adecuada dada esta normalización.

¿Qué nos dirá este contraste si utilizamos la normalización sobre *hours* en la segunda ecuación y consideramos estimar la demanda de trabajo en lugar de la oferta salarial?

En este caso procedemos del siguiente modo:

1. Estimamos por MC2E y guardamos los residuos en *uhat2*.
2. Realizamos la regresión de estos residuos MC2E, *uhat2*, sobre todas las variables exógenas. El resultado es el siguiente:

Contraste de Sargan

Estimaciones MCO utilizando las 428 observaciones 1–428
Variable dependiente: uhat2

Variable	Coefficiente	Desv. típica	Estadístico t	valor p
const	0,00407139	0,327084	0,0124	0,9901
kidslt6	-0,00671065	0,0910496	-0,0737	0,9413
kidsge6	-0,00601142	0,0286610	-0,2097	0,8340
age	-0,000358372	0,00556401	-0,0644	0,9487
educ	-0,00756862	0,0155143	-0,4878	0,6259
exper	0,000505048	0,0137414	0,0368	0,9707
nwifeinc	0,00562285	0,00341117	1,6484	0,1000
expersq	1,18111e-05	0,000412869	0,0286	0,9772

R^2

0,00687110

El valor del estadístico de Sargan en este caso es $TR^2 = 8,14343$ que es mayor que el valor crítico $\chi^2(3)_{0,05} = 7,81473$. En este caso se rechaza la hipótesis nula, confirmando lo que indicaba el contraste de razón de verosimilitudes LR de información limitada también. Dada esta normalización, los instrumentos no son adecuados para $lwage$. No se pueden identificar adecuadamente los parámetros de la ecuación de demanda de trabajo.