Revista de Análisis Económico, Vol. 36, Nº 1, pp. 3-22 (Abril 2021)

ECUACIONES DE MINCER DE PAREJAS BAJO UN ESQUEMA DE SELECCION MUESTRAL BIVARIADA. UNA APLICACION AL CASO ARGENTINO

MINCER EQUATIONS FOR COUPLES WITH BIVARIATE SAMPLE SELECTIVITY. AN APPLICATION FOR ARGENTINA

JAVIER ALEJO*
Instituto de Economía (Iecon), Universidad de la República (UdelaR)

VICTOR FUNES LEAL**
Department of Agricultural and Consumer Economics,
University of Illinois at Urbana-Champaign

Resumen

Este documento explora el efecto de las decisiones laborales conjuntas sobre la estimación de ecuaciones de Mincer en parejas. El sesgo de selección muestral se genera por la relación entre la decisión de trabajar y el salario. La mayoría de la literatura considera modelos de participación laboral individuales como mecanismo de selección. Sin embargo, evidencia reciente muestran una mayor participación laboral de las mujeres y una mayor relevancia en la decisión conjunta de los cónyuges creando un mecanismo de selección conjunto. Una versión bivariada del método de Heckman da una solución a este problema. Los resultados indican que la decisión conjunta de la pareja es un factor relevante en el sesgo por selección.

Palabras clave: *Salarios, decisiones laborales conjuntas, sesgo por selección, método de Heckman.*

Clasificación JEL: *C51, D13, E24, J31.*

* Instituto de Economía (IECON). Universidad de la República. Gonzalo Ramírez 1926. Montevideo. Uruguay. E-mail: javier.alejo@ccee.edu.uy

** Department of Agricultural, Applied and Consumer Economics. University of Illinois at Urbana-Champaign. 1302 W. Gregory Dr. Champaign, Illinois. Estados Unidos. E-mail: victorf2@illinois.edu

Abstract

This working paper explores the effect of joint labor decisions on the study of wage regression models. The estimation of Mincer equations suffers from numerous sources of bias, including the sample selection problem generated by the fact that agents' decision to work is not independent of their wage levels. Most of the papers correct for this bias using a model of individual labor participation. However, recent trends in the labor market show greater participation of women in the labor force and seem to indicate that the joint decision of the spouses is increasingly relevant in determining the selection mechanism. A bivariate version of Heckman's method appears as an interesting alternative to solve this problem. Although the estimates are in line with the previous literature, the results indicate that the joint decision of the couple is a relevant factor in the selection bias.

Keywords: *Wages, joint labor decisions, selection bias, Heckman's method.*

JEL Classification: *C51, D13, E24, J31.*

1. INTRODUCCION

Existe numerosa evidencia acerca del aumento de la participación laboral femenina en el mercado de trabajo pero con una desaceleración significativa en dicho proceso que se inicia a partir de los 2000 (Chioda, 2016; Gasparini, Leonardo; Marchionni, Mariana; Badaracco, Nicolás; Serrano, 2015). Para el caso de América Latina, los cambios en la estructura educativa, matrimonial y la fecundidad favorecieron una mayor participación laboral de las mujeres y, por tanto, se esperaría que en ausencia de dichos eventos la desaceleración en la tasa de participación no debiera haber ocurrido.

Este empoderamiento de las mujeres en el plano laboral tiene su correlato en el proceso de toma de decisiones dentro del hogar. En particular, es posible que su poder de negociación en la pareja mejore y esta razón las acciones conjuntas sean más consensuadas e interdependientes. En consecuencia, esto implica un cambio de patrón en el comportamiento laboral tanto de mujeres y hombres en parejas heterosexuales.

Existe numerosa literatura teórica que analiza el proceso de acciones económicas conjuntas dentro de un hogar. Sin embargo, la literatura empírica es menos abundante y solo se limita validar o refutar modelos teóricos. Este trabajo se basa en esos modelos pero con un objetivo diferente, como es el de corregir un problema de estimación econométrica. En particular, se utiliza una extensión del conocido método en dos etapas (Heckman, 1979) para estimar ecuaciones de salarios en parejas donde prevalece un modelo de selección bivariada.

Ciertos ejercicios de simulación junto con la aplicación del método al caso de Argentina muestran que incorporar el análisis del comportamiento intrahogar en el estudio de los determinantes salariales puede ser un aspecto relevante si el objetivo es conocer los canales por los que opera el sesgo por selección.

El trabajo se organiza de la siguiente manera: en la Sección 2 se discuten los antecedentes de toma de decisiones intrahogar en el contexto del mercado laboral, mientras que en la Sección 3 se expone un marco teórico simple que resume las principales ideas de la literatura. La Sección 4 presenta los distintos aspectos metodológicos y en la Sección 5 se muestra los resultados de la aplicación para el caso argentino. Finalmente, la Sección 6 presenta las conclusiones principales del trabajo.

2. ANTECEDENTES

El estudio de las decisiones laborales intrahogar tiene sus orígenes en el campo de la teoría. Becker (1991) desarrolla la idea que la división sexual del trabajo se origina en la inversión diferencial entre hombres y mujeres en diferentes tipos de capital humano, la que genera diferentes ventajas comparativas, estas a su vez llevan a la especialización en diferentes tareas. Esto implica que el poder de negociación relativo del hombre y la mujer está determinado por sus productividades relativas, de manera tal que el cónyuge con mayor productividad generará más ingresos para el hogar y tendrá más poder de negociación.

Posteriormente se desarrolló numerosa literatura que modela la forma en la que una pareja decide conjuntamente su oferta laboral en función de los intereses dentro del hogar (algunos ejemplos son: Blundell *et al.*, 1998, 2007; Pierre-André Chiappori, 1988; Pierre-Andre Chiappori, 1992; van Klaveren *et al.*, 2008, entre otros). Por simplicidad, la mayoría de los análisis de decisiones laborales intrahogar se hacen modelando solo al jefe y su cónyuge. Sin desconocer que pueden existir otros miembros del hogar que aporten ingresos laborales al hogar, este trabajo seguirá la línea de mantener el análisis a nivel de parejas.

Si bien existe una gran diversidad en estos modelos, en general se analiza la forma en que se distribuyen los recursos dentro del hogar, analizando quién y cuánto trabaja cada miembro del hogar. Algunas variantes consideran decisiones centralizadas mientras que otras suponen un proceso de negociación entre partes. El resultado de equilibrio dependerá de las relaciones de poder dentro de la pareja y sus determinantes, así como los factores culturales, los salarios de reserva, etcétera.

Respecto del estudio empírico de las decisiones laborales intrahogar Blundell y Macurdy (1999) sostienen que el modelo tradicional de oferta laboral individual no resulta apto para estudiar la división de trabajo intrafamiliar ni tampoco permite estudiar el impacto distributivo de políticas públicas que afecten a, por ejemplo, mujeres casadas con hijos. En contraposición a estos, los autores denominan “modelos colectivos” a los modelos que tienen en cuenta la distribución intrafamiliar.

Ransom (1987) presenta el primer trabajo que utiliza una especificación similar en espíritu a este, ya que estima la oferta de trabajo conjunta de maridos y mujeres teniendo en cuenta la interdependencia de las decisiones de oferta laboral. Las ofertas de trabajo se estiman por Máxima Verosimilitud partiendo de la distribución de probabilidades conjunta de horas de trabajo de varones y mujeres. La novedad en este modelo es que permite a las mujeres decidir si trabajan o no por medio de la incorporación de dos "regímenes" en la función de verosimilitud. Otro detalle importante es que este modelo permite diferencias en preferencias que, si bien no son observables, son funciones de variables que sí lo son, capturando entonces su heterogeneidad. La estimación de la ecuación de horas trabajadas se estima utilizando una forma reducida que permite recuperar los coeficientes estructurales y estimar las elasticidades de las horas trabajadas respecto de las variables independientes.

El trabajo de Pierre-André Chiappori *et al.*, 2002, estima un modelo de oferta laboral conjunta por medio del Método Generalizado de Momentos (GMM) donde las horas de trabajo ofrecidas en el mercado dependen en simultáneo del salario horario del marido y la mujer, el ingreso no laboral y la razón de hombres con relación a mujeres en el distrito, utilizando datos del Panel Study of Income Dynamics para el año 1988. Los autores encuentran que los coeficientes de salarios, ingreso no laboral y razón de género son significativos. Respecto de este último, un incremento de la cantidad de hombres relativa a la de mujeres reduce la oferta laboral de las mujeres, mientras que incrementa la de los hombres, resultado que se mantiene en todas las especificaciones. Otro factor exógeno estudiado son las leyes de divorcio en cada estado de la muestra utilizando un conjunto de variables que representan diferentes causales de este (consentimiento mutuo, división de propiedad, obligación de mantener al cónyuge y si la ley considera a los títulos universitarios como bien ganancial). En los tres primeros casos los coeficientes resultan negativos para los hombres y positivos para las mujeres, mientras que el cuarto es positivo para ambos, lo que se interpreta como que las leyes de divorcio otorgan mayor poder de negociación a las esposas.

Un aspecto menos explorado en la literatura es la repercusión de la decisión laboral en parejas en la estimación de ecuaciones de salarios. Es usual que en los modelos salariales a la Mincer se corrija por sesgo de selección muestral provocado por el hecho de que la decisión de participar del mercado laboral no es una decisión independiente del salario. Incorporar el proceso de decisión intrahogar requiere un modelo más rico que contemple no solo aspectos observables (algunos de ellos, comunes a ambos cónyuges), sino también la relación con las fuentes inobservables contempladas por la pareja en la toma de sus decisiones laborales.

3. MARCO TEORICO: UN MODELO SIMPLE

En esta sección se presenta un marco teórico simple que sirve como motivación para la estrategia empírica. Siguiendo a la literatura, se analiza la toma de decisiones de una pareja bajo un modelo de decisión conjunta donde cada miembro j tiene una función de utilidad $U_j(c_j, l_j)$, siendo c_j el consumo individual y l_j las horas de ocio para $j = h, m$.

Suponiendo un modelo estático, el total de recursos del hogar viene dada por la suma de los ingresos laborales y no laborales de la pareja y por tanto el consumo de ambos debe asignarse a base de dicho presupuesto:

$$c_h + c_m = w_h \cdot (T - l_h) + w_m \cdot (T - l_m) + a$$

donde w_h y w_m son los salarios por hora de cada miembro de la pareja, T es su tiempo disponible y a representa el ingreso no laboral del hogar. [1] Siguiendo a Chiappori (1992), en este modelo simple las parejas ubicarán su decisión conjunta en alguna de las asignaciones Pareto-óptimas $c_j^*(w, U_m^0, a)$ y $l_j^*(w, U_m^0, a)$, dadas por:

$$\max_{(c_j, l_j)} \mathcal{L} = U_h(c_h, l_h) + \kappa\left[U_m^0 - U_m(c_m, l_m)\right] + \xi[w_h(T - l_h) + \ldots$$

$$\ldots + w_m(T - l_m) + y_p - c_h - c_m]$$

donde U_m^0 es la utilidad alcanzada por el miembro m pero acordada entre ambos. La elección de un valor U_m^0 particular tiene implícito un juego entre ambos miembros que define una regla para compartir los recursos del hogar y que depende del poder de negociación de cada uno de ellos. A modo de simplificación, supongamos que ese poder para negociar depende de un conjunto de variables exógenas $q = (q_h, q_m)$ observables en una encuesta y otro $\eta = (\eta_h, \eta_m)$ del que no tenemos información. En otras palabras, supongamos que $U_m^0 = B(q_h, q_m, \eta_h, \eta_m) = B(q, \eta)$. Luego, las elecciones de la pareja son $c_j^*(w, q, \eta)$ y $l_j^*(w, q, \eta)$, para $j = h, m$.[2]

Asimismo, supongamos que el mercado laboral opera en un equilibrio donde asigna salarios en función de las características del trabajador. Es decir, sea $w_j = w_j(x_j, u_j)$ una función de salarios para $j = h, m$, donde $x = (x_h, x_m)$ es el conjunto de características observables en una encuesta y $u = (u_h, u_m)$ las no observables. Para simplificar la notación, sea $z = (x, q)$ el conjunto de información observable en una encuesta.

Claramente los atributos u y η están correlacionados, ya que existen algunos factores inobservables que si bien determinan el poder de negociación dentro de la pareja también son retribuidos en el mercado laboral (inteligencia, talento, perseverancia,

---

1 Una interpretación más laxa incluiría a los ingresos de otros miembros ocupados del hogar en término a, monto que se considera exógeno en este problema.

2 Es fácil pensar que los ingresos no laborales de cada miembro también es una fuente de poder de negociación y es por eso que ha quedado englobada dentro del componente q.

etc.). Sin embargo, a diferencia de los atributos observables, no es posible distinguir entre ambos inobservables. Para ello definimos una función resumen $\psi = \psi(u,\eta)$, quedando así que las asignaciones de consumo y ocio de la pareja son $c_j^*(z,\psi)$ y $l_j^*(z,\psi)$.

4. ESTRATEGIA EMPIRICA: UN MODELO ESTIMABLE

El problema fundamental del sesgo por selección en la estimación de ecuaciones de Mincer es que solo se observan los salarios en aquellos individuos que trabajan. Como se discute en la sección previa, la decisión de trabajar no es independiente del salario. Aun controlando por las diferencias generadas por las características observables z, es claro que el mecanismo de selección induce una endogeneidad, porque aun pueden existir otros factores que correlacionan la decisión de entrar al mercado laboral con los salarios. Además, la toma de decisiones laborales conjunta hace interdependiente el mecanismo de selección dentro de una misma pareja generando una fuente adicional de endogeneidad en el proceso. [3] El mecanismo de selección bivariado está dado por el conjunto de ecuaciones:

$$S_h = 1[T - l_h^*(z,\psi) > 0]$$

$$S_m = 1[T - l_m^*(z,\psi) > 0]$$

Si consideramos que los inobservables ψ son una variable aleatoria, esto define un modelo de probabilidad conjunta para (S_h, S_m) condicional en los factores z, donde *a priori* parecería difícil sostener que sean independientes debido a que ambos son funciones de ψ. Además, porque el vector $z=(z_h, z_m)$, el modelo de probabilidad conjunta debe de ser función de factores observables de ambos miembros de la pareja. Una parametrización posible para este modelo es:

$$S_h = 1[z'\gamma_h + \varepsilon_h > 0]$$

$$S_m = 1[z'\gamma_m + \varepsilon_m > 0]$$

donde el vector $(\varepsilon_h, \varepsilon_m)$ tiene distribución normal bivariada con $corr(\varepsilon_h, \varepsilon_m) = \rho$, $E(\varepsilon_j) = 0$ y $V(\varepsilon_j) = 1$ para $j = h,m$. Por tanto, el parámetro ρ permite la posibilidad de incorporar la interdependencia en las decisiones laborales de la pareja vinculada a factores inobservables. Este modelo es conocido como *Probit bivariado* y es posible

---

[3] En este caso, estaríamos hablando de correlaciones intrapareja del tipo $COV(\psi_h, \psi_m) \neq 0$.

estimarlo consistentemente por el método de máxima verosimilitud (de Luca, 2008; Poirier, 1980).

Un supuesto usual en la estimación de ecuaciones de salarios $w_j(x_j,u_j)$ es que la especificación es lineal en parámetros y con error aditivo, es decir:

$$w_h = x_h^{'}\beta_h + u_h$$

$$w_m = x_m^{'}\beta_m + u_m$$

donde $E\left(u_j | x_j\right) = 0$ para $j = h,m$.

Solo es posible observar los salarios para los individuos que trabajan y debido a que dicha decisión no es independiente de (u_j, u_m), existe un problema de selección latente. En este caso, intervienen las decisiones de ambos miembros de la pareja y por tanto es necesario modelar ambas interrelaciones. Supongamos $\varepsilon = (\varepsilon_h, \varepsilon_m)$ es independiente de z y que los errores se relacionan de la siguiente manera:

$$E\left(u_h | \varepsilon\right) = \theta_1\varepsilon_h + \theta_2\varepsilon_m$$

$$E\left(u_m | \varepsilon\right) = \tau_1\varepsilon_h + \tau_2\varepsilon_m$$

Luego, bajo estos supuestos se puede mostrar que es posible expresar el valor esperado de los salarios de la siguiente manera:

$$E\left(w_h | S_h = 1, S_m = 1\right) = x_h^{'}\beta_h + \delta_{h1}\lambda_1^d\left(z'\gamma_h, z'\gamma_m\right) + \delta_{h2}\lambda_2^d\left(z'\gamma_h, z'\gamma_m\right)$$

$$E\left(w_m | S_m = 1, S_h = 1\right) = x_m^{'}\beta_m + \delta_{m1}\lambda_1^d(z'\gamma_h, z'\gamma_m) + \delta_{m2}\lambda_2^d(z'\gamma_h, z'\gamma_m)$$

donde,

$$\delta_{h1} = \left(\theta_1 + \theta_2\rho\right)$$

$$\delta_{h2} = \left(\theta_1\rho + \theta_2\right)$$

$$\delta_{m1} = \left(\tau_1 + \tau_2\rho\right)$$

$$\delta_{m2} = \left(\tau_1\rho + \tau_2\right)$$

Los términos $\lambda_1^d(.)$ y $\lambda_2^d(.)$ son funciones auxiliares similares a la inversa de la razón de Mills (ver Anexo A.1), cuya forma funcional depende de cada situación específica en donde se observan salarios, indicado mediante la siguiente variable:

$$d = \begin{cases} 1 & si\ S_h = 1 \wedge S_m = 1 \\ 2 & si\ S_h = 1 \wedge S_m = 0 \\ 3 & si\ S_h = 0 \wedge S_m = 1 \end{cases}$$

Por tanto, es posible estimar construir una estrategia de estimación en dos etapas análoga a la de Heckman, pero para el caso de selección bivariada es (Tauchmann, 2010).

Paso 1: obtener una estimación $(\hat{\gamma}_h,, \hat{\gamma}_m, \hat{\rho})$ con un modelo *Probit bivariado* y computar las variables $\hat{\lambda}_1 = \lambda_1^d \left(z'\hat{\gamma}_h, z'\hat{\gamma}_m \right)$ y $\hat{\lambda}_2 = \lambda_2^d \left(z'\hat{\gamma}_h, z'\hat{\gamma}_m \right)$.
Paso 2: estimar una regresión de MCO de w_j en función de x_j junto a $\hat{\lambda}_1$ y $\hat{\lambda}_2$ como regresores adicionales, para $j = h,m$.

Debido a que en cada paso se utilizan métodos consistentes, este procedimiento da como resultado estimaciones consistentes de $(\beta_j, \delta_{j1}, \delta_{j2})$, para $j = h,m$. Como paso adicional, se puede obtener estimaciones para la correlación de los inobservables entre las ecuaciones de selección y las de regresión:

$$\hat{\theta}_1 = \frac{\hat{\delta}_{11} - \hat{\delta}_{12}\hat{\rho}}{(1 - \hat{\rho}^2)}$$

$$\hat{\theta}_2 = \frac{\hat{\delta}_{12} - \hat{\delta}_{11}\rho}{(1 - \hat{\rho}^2)}$$

$$\hat{\tau}_1 = \frac{\hat{\delta}_{21} - \hat{\delta}_{22}\rho}{(1 - \hat{\rho}^2)}$$

$$\hat{\tau}_2 = \frac{\hat{\delta}_{22} - \hat{\delta}_{21}\hat{\rho}}{(1 - \hat{\rho}^2)}$$

El cómputo de los errores estándar de estos estimadores no es trivial, ya que involucra complicaciones similares al caso univariado (Heckman, 1979). Para este trabajo se opta por utilizar el método de *bootstrap* para realizar los ejercicios de inferencia, en particular:

i) *Sesgo de selección en ambas ecuaciones de salarios*: se puede implementar utilizando la siguiente hipótesis general $H_0 : \delta_{h1} = \delta_{h2} = \delta_{m1} = \delta_{m1} = 0$.

ii) *Sesgo de selección en la ecuación de salarios del miembro j*: es un caso particular de la hipótesis anterior, es decir $H_0 : \delta_{j1} = \delta_{j2} = 0$.

iii) *Sesgo de selección inducida por los inobservables del compañero*: para ello es útil la hipótesis general $H_0 : \theta_2 = \tau_1 = 0$, o bien alguna de sus versiones individuales.

Una vez computados las estimaciones y los errores estándar, las distintas pruebas de hipótesis se realizan mediante los procedimientos asintóticos usuales.

5. RESULTADOS

Como paso previo a la aplicación empírica a continuación se muestran los resultados de varios experimentos de Monte Carlo de la metodología expuesta previamente ya que el argumento utilizado para justificar el procedimiento de estimación es la propiedad de consistencia. El objetivo de estos ejercicios es estudiar el comportamiento del sesgo con tamaños de muestras finitas.

El diseño de los experimentos considera *n* datos generados por un modelo lineal muy simple (ver Anexo A.2) que difiere en los grados de correlación tanto entre los errores de las ecuaciones de selección como en su relación con las ecuaciones de salarios. Por un lado, para la correlación en la decisión de participar en el mercado laboral se consideraron tres valores alternativos para $\rho = \{0, 0.50,\ 0.75\}$. Por el otro, se alternaron los valores 0 y 1 para los parámetros $(\theta_1, \theta_2, \tau_1, \tau_2)$, a modo de explorar la identificación del origen activo en el proceso de selección de la muestra utilizada para estimar las ecuaciones de regresión. Para la simulación se consideraron 1000 réplicas del experimento con $n = \{5000, 100000\}$. [4] En particular, se analiza el sesgo de estimar β_1 y β_2 con tres métodos alternativos: MCO y Heckman en dos etapas en sus versiones univariada y bivariada.

En la Tabla 1 se muestran los resultados para el caso con muestras de 5000 observaciones y en la Tabla 2 para el caso con 100 mil observaciones. En ambas se observa que cuando está activa alguna fuente de selección muestral (mediante la correlación entre los inobservables) el método en dos etapas reduce notablemente el sesgo por selección que comete MCO. Sin embargo, en la Tabla 1 puede apreciarse que dicha corrección es similar en ambas variantes del método de Heckman.

---

4 El objetivo del Monte Carlo es poner a prueba el método con tamaños de muestra similares a los utilizados en las encuestas de hogares.

TABLA 1

SESGOS DE ESTIMACION CON UN TAMAÑO DE MUESTRA DE N = 5000

				MCO			Heckman			Bi–Heckman		
				ρ			ρ			ρ		
θ_1	θ_2	τ_1	τ_2	0,00	0,50	0,75	0,00	0,50	0,75	0,00	0,50	0,75
				β_1								
0	0	0	0	0,00	0,01	0,36	0,00	–0,05	0,31	0,00	–0,06	0,31
0	0	0	1	–0,19	0,02	–0,23	–0,18	–0,11	–0,14	–0,18	–0,12	–0,15
0	0	1	0	0,00	–0,18	0,07	0,09	–0,17	0,08	0,10	–0,18	0,06
0	1	0	0	7,18	–153,6	–239,9	13,81	9,00	–0,38	7,20	5,90	–2,12
1	0	0	0	–52,5	–52,8	–53,1	0,81	1,10	0,74	0,80	1,15	0,79
1	1	1	1	–55,5	–218,2	–297,7	–2,23	–5,02	–3,08	–2,66	–6,65	2,12
				β_2								
0	0	0	0	–0,18	0,44	–0,12	–0,24	0,46	–0,15	–0,24	0,44	–0,14
0	0	0	1	–86,66	–80,37	–83,73	–0,34	8,31	4,94	–0,33	5,71	1,70
0	0	1	0	0,72	–7,52	–11,27	0,37	–0,10	–0,59	0,07	–0,14	–0,15
0	1	0	0	0,26	–0,02	0,05	0,23	–0,03	0,06	0,23	–0,03	0,06
1	0	0	0	0,3	–0,01	0,08	0,41	–0,01	0,08	0,41	–0,01	0,06
1	1	1	1	–87,0	–96,5	–102,3	1,75	0,85	–3,55	1,88	2,27	–4,03

Nota: Cómputos realizados con datos simulados a base de 1000 experimentos.

Solo con tamaños grandes de la muestra (Tabla 2) la diferencia entre ambos métodos se hace palpable, siendo la opción bivariada la mejor en términos de insesgadez. Estos resultados muestran una baja tasa de convergencia en las propiedades del método en dos etapas. La razón principal es la baja potencia generada por la inclusión como regresores de las variables auxiliares $\hat{\lambda}_1$ y $\hat{\lambda}_2$, ya que están altamente correlacionadas con las x_j. Si bien este aspecto no es un problema grave, su presencia pone en tela de juicio los resultados estimados con tamaños de muestra moderados.

Asimismo, un aspecto favorable no menor es que aun en los casos en donde las dos variantes del método de Heckman son similares en términos de sesgo, el método bivariado muestra un mejor desempeño medido por el Error Cuadrático Medio (Tabla 3), indicando que es una ganancia notable en cuanto a la eficiencia relativa de este método.

TABLA 2

SESGOS DE ESTIMACION CON UN TAMAÑO DE MUESTRA DE N = 100000

				MCO			Heckman			Bi–Heckman		
				ρ			ρ			ρ		
θ_1	θ_2	τ_1	τ_2	0,00	0,50	0,75	0,00	0,50	0,75	0,00	0,50	0,75
				β_1								
0	0	0	0	–0,01	–0,04	0,02	0,01	–0,07	–0,01	0,01	–0,07	0,00
0	0	0	1	–0,03	–0,06	–0,01	–0,02	–0,03	–0,01	–0,02	–0,03	–0,01
0	0	1	0	–0,01	–0,05	0,00	–0,02	–0,05	0,01	–0,02	–0,05	0,01
0	1	0	0	1,09	–159,3	–240,6	1,20	1,72	–0,50	1,22	1,07	–0,37
1	0	0	0	–53,4	–53,5	–53,6	0,11	–0,05	–0,02	0,11	–0,08	–0,01
1	1	1	1	–54,8	–214,7	–295,4	–1,27	0,03	–0,43	0,01	–0,18	–0,38
				β_2								
0	0	0	0	0,02	0,02	0,10	0,01	0,03	0,07	0,01	0,03	0,08
0	0	0	1	–87,61	–87,75	–87,88	0,94	0,58	0,02	0,94	0,46	0,07
0	0	1	0	–0,24	–7,17	–11,20	–0,26	0,11	–0,21	–0,08	0,11	–0,04
0	1	0	0	–0,05	0,02	0,08	–0,04	0,02	0,09	–0,04	0,02	0,09
1	0	0	0	0,0	0,04	0,07	–0,02	0,03	0,11	–0,02	0,03	0,09
1	1	1	1	–88,1	–94,7	–99,0	–0,13	0,95	0,35	0,10	1,30	0,85

Nota: Cómputos realizados con datos simulados a base de 1000 experimentos.

TABLA 3

ERROR CUADRATICO MEDIO CON UN TAMAÑO DE MUESTRA DE N = 100000

				MCO			Heckman			Bi-Heckman		
				ρ			ρ			ρ		
θ_1	θ_2	τ_1	τ_2	0,00	0,50	0,75	0,00	0,50	0,75	0,00	0,50	0,75
				β_1								
0	0	0	0	0,03	0,03	0,03	0,03	0,04	0,03	0,03	0,04	0,03
0	0	0	1	0,03	0,03	0,03	0,03	0,03	0,04	0,03	0,03	0,03
0	0	1	0	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03
0	1	0	0	16,81	267,94	590,07	18,89	15,79	12,63	8,59	8,47	7,58
1	0	0	0	28,73	28,82	28,90	0,22	0,20	0,21	0,22	0,20	0,20
1	1	1	1	45,91	476,32	884,93	17,83	17,13	13,25	8,67	8,55	8,78
				β_2								
0	0	0	0	0,02	0,02	0,02	0,02	0,02	0,03	0,02	0,02	0,03
0	0	0	1	82,16	81,90	82,24	6,05	5,40	5,67	6,05	4,98	4,63
0	0	1	0	0,43	0,87	1,51	0,46	0,40	0,29	0,24	0,21	0,16
0	1	0	0	0,02	0,02	0,02	0,03	0,03	0,03	0,03	0,03	0,03
1	0	0	0	0,02	0,02	0,03	0,03	0,03	0,03	0,03	0,03	0,03
1	1	1	1	82,97	95,82	104,87	6,01	6,80	7,49	5,87	5,89	5,63

Nota: Cómputos realizados con datos simulados a base de 1000 experimentos.

Para la aplicación empírica se consideran datos de la Encuesta Permanente de Hogares elaborada por el Instituto Nacional de Estadística y Censos (INDEC) de Argentina. La muestra utilizada son parejas de hombres y mujeres mayores de 25 años de edad y que no hayan superado la edad jubilatoria. A la luz de los resultados de los ejercicios de simulación, se utiliza un *pool* de datos de dos años contiguos como forma de morigerar parcialmente la convergencia lenta del método bivariado.[5] Los periodos considerados en el análisis son 1997-1998, 2005-2006 y 2013-2014, quedando una muestra de 25774, 29394 y 30364 observaciones de parejas con datos completos, respectivamente. Las variables incluidas en los modelos de regresión de cada miembro de la pareja son las usuales en la literatura: logaritmo del salario horario en la ocupación principal como variable dependiente y la edad, los años de educación (y sus cuadrados) y variables regionales como regresores. En la ecuación de selección se incluyeron como regresores variables que representan las características tanto individuales como comunes a la pareja: edad, años de educación, asistencia al sistema educativo, estatus de jefe de hogar, número de hijos menores a 12 años, región en la que reside y una variable binaria que indica si en el hogar hay otros miembros ocupados.

La Tabla 4 muestra las estimaciones de las ecuaciones de salario para el periodo 1997-1998, la Tabla 5 para 2005-2006 y finalmente la Tabla 6 para el *pool* 2013-2014. De manera consistente con la evidencia empírica previa, el sesgo por selección estimado por el método univariado para el caso de las mujeres pierde relevancia en los años más recientes, posiblemente como consecuencia de su mayor participación en el mercado laboral. Además, la correlación entre los factores inobservables en las decisiones individuales de participar dentro de la pareja (medido por el parámetro ρ) muestra evidencia de una leve asociación positiva entre ambas decisiones en el primer y el último periodo, mientras que en la muestra de 2005-2006 dicha relación no es estadísticamente significativa. Este último caso indica que la corrección tradicional de Heckman no debería diferir demasiado de su versión bivariada.

En cuanto al efecto de la selección en la estimación de los coeficientes de la ecuación de Mincer, si bien hay una gran diferencia en las estimaciones basadas en MCO respecto del método en etapas univariado, en general no se observan grandes diferencias en los coeficientes estimados por la versión bivariada, particularmente en el periodo 2005-2009 en donde la correlación de los mecanismos de selección es baja.

---

5 En la especificación de cada modelo se incluye una *dummy* para controlar por el efecto año.

TABLA 4

ECUACIONES DE MINCER EN PAREJAS, PERIODO 1997-1998

Variables	Hombres			Mujeres		
	MCO	Heckman	BiHeckman	MCO	Heckman	BiHeckman
Edad	0,0585***	0,0256***	0,0308***	0,0409***	0,0475***	0,0479***
	(0,00692)	(0,00867)	(0,0102)	(0,00781)	(0,00868)	(0,00906)
Edad^2	–0,000601***	–0,000110	–0,000189	–0,000420***	–0,000495***	–0,000501***
	(8,00e–05)	(0,000122)	(0,000148)	(9,61e–05)	(0,000107)	(0,000113)
Educación	–0,00502	–0,00466	–0,00390	–0,0157*	–0,0264***	–0,0271***
	(0,00837)	(0,00521)	(0,00534)	(0,00846)	(0,00818)	(0,00895)
Educación^2	0,00470***	0,00468***	0,00467***	0,00518***	0,00623***	0,00626***
	(0,000384)	(0,000250)	(0,000255)	(0,000387)	(0,000483)	(0,000508)
Pampeana	–0,219***	–0,197***	–0,200***	–0,195***	–0,191***	–0,190***
	(0,0242)	(0,0163)	(0,0158)	(0,0241)	(0,0218)	(0,0218)
Cuyo	–0,338***	–0,353***	–0,346***	–0,344***	–0,343***	–0,342***
	(0,0283)	(0,0210)	(0,0218)	(0,0283)	(0,0260)	(0,0259)
NOA	–0,348***	–0,368***	–0,364***	–0,331***	–0,324***	–0,323***
	(0,0267)	(0,0185)	(0,0188)	(0,0266)	(0,0242)	(0,0256)
Patagonia	0,224***	0,212***	0,212***	0,226***	0,230***	0,231***
	(0,0273)	(0,0186)	(0,0178)	(0,0272)	(0,0255)	(0,0241)
NEA	–0,381***	–0,415***	–0,414***	–0,422***	–0,436***	–0,435***
	(0,0284)	(0,0188)	(0,0187)	(0,0283)	(0,0258)	(0,0269)
Dummy 1998	0,0649***	0,0530***	0,0560***	0,0694***	0,0774***	0,0773***
	(0,0145)	(0,0106)	(0,0106)	(0,0145)	(0,0132)	(0,0135)
λ		–0,488***			0,102*	
		(0,158)			(0,0613)	
λ_1			–0,374*			0,0291***
			(0,197)			(0,0103)
λ_2			–0,0651***			0,105
			(0,00776)			(0,0644)
τ_1 (θ_1)			–0,372*			0,0233**
			(0,197)			(0,00998)
τ_2 (θ_2)			–0,0441***			0,104
			(0,00897)			(0,0644)
ρ			0,0562***			0,0562***
			(0,0119)			(0,0119)
Observaciones	7,724	25,774	25,774	7,724	25,774	25,774

Fuente: Estimaciones propias a base de la EPH de INDEC.
Nota: Errores estándar entre paréntesis; estimación robusta para MCO; corrección por selección muestral en Heckman univariado y *bootstrap* de 500 réplicas en el Heckman bivariado. Los asteriscos indican la significatividad estadística: 1% (***), 5% (**) y 10% (*).

Analizando la ecuación salarial de las mujeres, el sesgo de selección solo está presente en el periodo 1997-1998 y los resultados muestran que en dicho sesgo no es relevante su decisión individual, debido a que se ve afectada fundamentalmente por la decisión conjunta con su pareja. Además, en el caso de los hombres los coeficientes asociados a las variables a auxiliares del método bivariado son individualmente significativas en todos los periodos, indicando que tanto las decisiones del hombre como las de su pareja son relevantes en el proceso de selección muestral.

TABLA 5

ECUACIONES DE MINCER EN PAREJAS, PERIODO 2005-2006

Variables	Hombres			Mujeres		
	MCO	Heckman	BiHeckman	MCO	Heckman	BiHeckman
Edad	0,0332***	–0,00377	–0,00316	0,0439***	0,0457***	0,0457***
	(0,00546)	(0,00799)	(0,00718)	(0,00626)	(0,00660)	(0,00657)
Edad^2	–0,000315***	0,000220**	0,000211**	–0,000444***	–0,000457***	–0,000457***
	(6,17e–05)	(0,000102)	(9,39e–05)	(7,49e–05)	(7,81e–05)	(7,86e–05)
Educación	–0,00307	–0,00488	–0,00506	–0,0168**	–0,00370	–0,00368
	(0,00739)	(0,00712)	(0,00608)	(0,00821)	(0,00788)	(0,00873)
Educación^2	0,00398***	0,00354***	0,00357***	0,00549***	0,00507***	0,00508***
	(0,000329)	(0,000332)	(0,000273)	(0,000359)	(0,000412)	(0,000460)
Pampeana	–0,0985***	–0,0979***	–0,0976***	–0,0691***	–0,0633***	–0,0631***
	(0,0177)	(0,0180)	(0,0147)	(0,0184)	(0,0173)	(0,0168)
Cuyo	–0,269***	–0,274***	–0,273***	–0,280***	–0,269***	–0,269***
	(0,0243)	(0,0247)	(0,0193)	(0,0252)	(0,0238)	(0,0251)
NOA	–0,350***	–0,344***	–0,344***	–0,300***	–0,302***	–0,302***
	(0,0207)	(0,0208)	(0,0175)	(0,0215)	(0,0201)	(0,0201)
Patagonia	0,333***	0,378***	0,379***	0,335***	0,345***	0,345***
	(0,0223)	(0,0230)	(0,0190)	(0,0231)	(0,0220)	(0,0207)
NEA	–0,433***	–0,408***	–0,409***	–0,426***	–0,420***	–0,420***
	(0,0243)	(0,0241)	(0,0204)	(0,0253)	(0,0234)	(0,0258)
Dummy 2005	–0,311***	–0,285***	–0,286***	–0,304***	–0,298***	–0,298***
	(0,0125)	(0,0135)	(0,0112)	(0,0129)	(0,0122)	(0,0119)
λ		–1,019***			–0,0313	
		(0,151)			(0,0520)	
λ_1			–1,001***			–0,00493
			(0,152)			(0,0105)
λ_2			–0,0323***			–0,0315
			(0,00649)			(0,0595)
τ_1 (θ_1)			–1,001***			–0,00496
			(0,152)			(0,0106)
τ_2 (θ_2)			–0,0331***			–0,0315
			(0,00967)			(0,0595)
ρ			–0,000741			–0,000741
			(0,0118)			(0,0118)
Observaciones	11,722	29,394	29,394	11,722	29,394	29,394

Fuente: Estimaciones propias a base de la EPH de INDEC.
Nota: Errores estándar entre paréntesis; estimación robusta para MCO; corrección por selección muestral en Heckman univariado y *bootstrap* de 500 réplicas en el Heckman bivariado. Los asteriscos indican la significatividad estadística: 1% (***), 5% (**) y 10% (*).

TABLA 6

ECUACIONES DE MINCER EN PAREJAS, PERIODO 2013-2014

Variables	Hombres			Mujeres		
	MCO	Heckman	BiHeckman	MCO	Heckman	BiHeckman
Edad	0,0239***	–0,00730	–0,00421	0,0316***	0,0338***	0,0336***
	(0,00467)	(0,00770)	(0,00951)	(0,00554)	(0,00599)	(0,00607)
Edad^2	–0,000225***	0,000184*	0,000143	–0,000309***	–0,000333***	–0,000331***
	(5,23e–05)	(9,77e–05)	(0,000123)	(6,57e–05)	(7,07e–05)	(7,24e–05)
Educación	0,00480	0,00373	0,00315	–0,00618	–0,00710	–0,00727
	(0,00680)	(0,00523)	(0,00573)	(0,00801)	(0,00775)	(0,00925)
Educación^2	0,00282***	0,00265***	0,00271***	0,00430***	0,00456***	0,00456***
	(0,000295)	(0,000236)	(0,000258)	(0,000339)	(0,000373)	(0,000420)
Pampeana	–0,0559***	–0,0321***	–0,0315***	–0,0372**	–0,0345**	–0,0347**
	(0,0156)	(0,0123)	(0,0117)	(0,0166)	(0,0158)	(0,0161)
Cuyo	–0,197***	–0,184***	–0,182***	–0,200***	–0,197***	–0,197***
	(0,0213)	(0,0171)	(0,0170)	(0,0226)	(0,0216)	(0,0224)
NOA	–0,299***	–0,276***	–0,275***	–0,273***	–0,277***	–0,277***
	(0,0175)	(0,0139)	(0,0137)	(0,0186)	(0,0178)	(0,0187)
Patagonia	0,261***	0,335***	0,334***	0,309***	0,315***	0,315***
	(0,0183)	(0,0147)	(0,0153)	(0,0194)	(0,0185)	(0,0186)
NEA	–0,322***	–0,311***	–0,310***	–0,331***	–0,323***	–0,324***
	(0,0216)	(0,0163)	(0,0163)	(0,0229)	(0,0221)	(0,0236)
Dummy 2013	–0,600***	–0,600***	–0,600***	–0,615***	–0,618***	–0,619***
	(0,0103)	(0,00813)	(0,00818)	(0,0110)	(0,0105)	(0,0108)
λ		–0,570***			–0,0141	
		(0,135)			(0,0448)	
λ_1			–0,504***			0,0166*
			(0,181)			(0,00988)
λ_2			–0,0456***			–0,0154
			(0,00576)			(0,0486)
τ_1 (θ_1)			–0,503***			0,0174*
			(0,181)			(0,00973)
τ_2 (θ_2)			–0,0208**			–0,0163
			(0,00908)			(0,0487)
ρ			0,0493***			0,0493***
			(0,0125)			(0,0125)
Observaciones	14,007	30,364	30,364	14,007	30,364	30,364

Fuente: Estimaciones propias a base de la EPH de INDEC.

Nota: Errores estándar entre paréntesis; estimación robusta para MCO; corrección por selección muestral en Heckman univariado y *bootstrap* de 500 réplicas en el Heckman bivariado. Los asteriscos indican la significatividad estadística: 1% (***), 5% (**) y 10% (*).

Aun en el periodo 2005-2009, donde pareciera no existir dependencia en las decisiones laborales de la pareja, el método bivariado revela que en la ecuación salarial de los hombres la decisión de su pareja juega un rol significativo en el sesgo por selección. El canal por el que opera la decisión de la mujer es mediante la correlación entre los inobservables de la selección conjunta y los salarios que paga el mercado, medidos por los parámetros θ_1 y θ_2. Este ejemplo muestra que si el objetivo es estimar consistentemente β_1 da lo mismo utilizar la versión univariada o bivariada cuando $\rho \approx 0$. Sin embargo, el método bivariado es más informativo si lo que se busca es entender el origen del sesgo por selección.

A modo de resumen formal de todo este análisis, la Tabla 7 muestra los resultados de varias pruebas de hipótesis conjuntas realizadas sobre los parámetros del Heckman bivariado.

TABLA 7

PRUEBAS DE HIPOTESIS A BASE DEL MODELO DE SELECCION BIVARIADO

HIPOTESIS NULA: AUSENCIA DE SESGO DE SELECCION

	χ^2	p-valor
Periodo 1998-1999		
En ambas ecuaciones de salarios	96,48	0,000
En la ecuación de salarios del hombre	82,96	0,000
En la ecuación de salarios de la mujer	8,77	0,013
Inducido por los inobservables del compañero	31,52	0,000
Periodo 2005-2006		
En ambas ecuaciones de salarios	69,74	0,000
En la ecuación de salarios del hombre	69,71	0,000
En la ecuación de salarios de la mujer	0,48	0,787
Inducido por los inobservables del compañero	11,74	0,003
Periodo 2013-2014		
En ambas ecuaciones de salarios	68,73	0,000
En la ecuación de salarios del hombre	62,66	0,000
En la ecuación de salarios de la mujer	3,26	0,196
Inducido por los inobservables del compañero	8,613	0,014

Fuente: Estimaciones propias a base de la EPH de INDEC.
Nota: Inferencia basada en un ejercicio de *bootstrap* de 500 réplicas del Heckman bivariado.

Los resultados de las pruebas están en línea con los comentarios realizados previamente y muestran que el uso del método bivariado brinda un análisis más rico revelando información interesante respecto del origen del sesgo por selección bajo esquemas de decisiones laborales interdependientes al interior del hogar.

6. CONCLUSIONES

En este trabajo se intenta dar un paso más en la estimación de ecuaciones de salarios sujeto a procesos de selección muestral basados en decisiones de un conjunto de individuos. En particular, se consideró como unidad de análisis la decisión de participar del mercado laboral de la pareja. Si bien este aspecto ha sido ampliamente estudiado tanto en la literatura teórica como empírica, sus consecuencias respecto del sesgo por selección en las ecuaciones de Mincer han sido muy poco exploradas.

Las estimaciones realizadas representan un aporte a la literatura empírica incorporando una versión bivariada del método de Heckman para estimar ecuaciones de salarios. Al igual que en el caso univariado, el método también consta de dos etapas, pero modelando una selección interdependiente entre los miembros de las parejas y sus posibles fuentes de endogeneidad en sus respectivas ecuaciones salariales. Los resultados de los experimentos de Monte Carlo muestran una convergencia asintótica relativamente lenta del método bivariado, lo que sugiere la necesidad de utilizar una cantidad de datos considerables en su implementación empírica. Sin embargo, el método bivariado parece tener un buen desempeño en cuanto a la eficiencia relativa respecto del método univariado.

Si bien algunos de los resultados de la aplicación empírica están en línea con la literatura previa, el método bivariado muestra evidencia cuantitativa de que la interdependencia en las decisiones laborales de las parejas es un factor importante en la determinación del sesgo por selección muestral. Por tanto, la incorporación de un mecanismo endógeno bivariado al estudio de las ecuaciones de Mincer en las aplicaciones empíricas con datos que contienen parejas aparece como una opción relevante para reducir y entender los sesgos de estimación.

REFERENCIAS

BECKER, G. S. (1991). *A Treatise on the Family*. Harvard University Press.

BLUNDELL, R.; P. A. CHIAPPORI; T. MAGNAC, y C. MEGHIR (2007). "Collective labour supply: Heterogeneity and non-participation". *Review of Economic Studies*, *74*(2), 417-445.

BLUNDELL, R.; A. DUNCAN, y C. MEGHIR (1998). "Estimating Labor Supply Responses Using Tax Reforms". *Econometrica*, *66*(4), 827-861.

BLUNDELL, R. y T. MACURDY (1999). "*Labor Supply: A Review of Alternative Approaches",* en O. C. Ashenfelter y D. E. Card, *Handbook of Labor Economics*, Vol. 3, Elsevier, pp. 1559-1695.

CHIAPPORI, P. A. (1988). "Rational Household Labor Supply". *Econometrica*, *56*(1), 63-90.

CHIAPPORI, P. A. (1992)."Collective Labor Supply and Welfare". *Journal of Political Economy*, *100*(3), 437-467.

CHIAPPORI, P. A.; B. FORTIN y G. LACROIX (2002). "Marriage Market, Divorce Legislation, and Household Labor Supply". *Journal of Political Economy*, *110*(1), 37-72.

CHIODA, L. (2016). *Work and Family: Latin American and Caribbean women in search of a new balance,* Latin American Development Forum Series. Washington, D.C. : World Bank Group.

DE LUCA, G. (2008). "SNP and SML Estimation of Univariate and Bivariate Binary-Choice Models". *The Stata Journal*, *8*(2), 190-220.

GASPARINI, L.; M. MARCHIONNI; N. BADARACCO; J. Serrano. (2015). "*Female Labor Force Participation in Latin America: Evidence of Deceleration*", Documento de Trabajo 181, Centro de Estudios Distributivos, Laborales y Sociales.

HECKMAN, J. J. (1979). "Sample Selection Bias as a Specification Error". *Econometrica*, *47*(1), 153-161.

POIRIER, D. J. (1980). "Partial observability in bivariate probit models". *Journal of Econometrics*, *12*(2), 209-217.

RANSOM, M. R. (1987). "An Empirical Model of Discrete and Continuous Choice in Family Labor Supply". *The Review of Economics and Statistics*, *69*(3), 465-472.

TAUCHMANN, H. (2010). "Consistency of Heckman-type two-step estimators for the multivariate sample-selection model". *Applied Economics*, *42*(30), 3895-3902.

VAN KLAVEREN, C.; B. VAN PRAAG; y H. MAASSEN VAN DER BRINK (2008). "A public good version of the collective household model: an empirical approach with an application to British household data". *Review of Economics of the Household*, *6*(2), 169-191.

ANEXO

A.1. Variables auxiliares para la segunda etapa

Sean (v_1, v_2) dos variables aleatorias con distribución normal estándar bivariada con parámetro de correlación ρ. Entonces, siguiendo a Tauchmann (2010), las formas funcionales de las variables auxiliares son:

$$\lambda_1^d(k_1,k_2)=\begin{cases} \dfrac{\phi(k_1)\Pr(v_2>k_2 \mid v_1=k_1)}{\Pr(d=1)} & si\ \ d=1 \\ \dfrac{\phi(k_1)\Pr(v_2<k_2 \mid v_1=k_1)}{\Pr(d=2)} & si\ \ d=2 \\ -\dfrac{\phi(k_1)\Pr(v_2>k_2 \mid v_1=k_1)}{\Pr(d=3)} & si\ \ d=3 \end{cases}$$

$$\lambda_2^d(k_1,k_2)=\begin{cases} \dfrac{\phi(k_2)\Pr(v_1>k_1 \mid v_2=k_2)}{\Pr(d=1)} & si\ \ d=1 \\ -\dfrac{\phi(k_2)\Pr(v_2>k_2 \mid v_2=k_2)}{\Pr(d=2)} & si\ \ d=2 \\ \dfrac{\phi(k_2)\Pr(v_1<k_1 \mid v_2=k_2)}{\Pr(d=3)} & si\ \ d=3 \end{cases}$$

donde

$$d=\begin{cases} 1 & si\ S_h=1 \wedge S_m=1 \\ 2 & si\ S_h=1 \wedge S_m=0 \\ 3 & si\ S_h=0 \wedge S_m=1 \end{cases}$$

y

$$\Pr(A)=\iint_{(v_1,v_2)\in A} \phi_2(v_1,v_2,\rho)\partial v_1 \partial v_2$$

A.2. Diseño del experimento de Monte Carlo

El modelo utilizado para realizar los ejercicios de simulación consta de las siguientes ecuaciones y funciones de distribución:

–Ecuaciones de selección

$$S_h = 1[-3 + 4z_h + \varepsilon_h > 0]$$

$$S_m = 1[-3 + 7z_m + \varepsilon_m > 0]$$

donde los regresores son $z_h = 0.2x_h + 0.4x_m + 0.4\zeta_h$ y $z_m = 0.4x_h + 0.2x_m + 0.4\zeta_m$, con $\zeta_j \sim U(0,1)$ para $j = \{h,m\}$.

–Ecuaciones de regresión

$$w_h = 1 + x_h + u_h$$

$$w_m = 1 + x_m + u_m$$

donde w_j es observable solo si $S_j = 1$ y los regresores $x_j \sim U(0,1)$ para $j = \{h,m\}$.

–Relación entre los inobservables

$$(\varepsilon_h, \varepsilon_m) \sim N_2(0, 0, \sigma_h = 4, \sigma_m = 24, \rho)$$

con $\rho = \{0, 0.50, 0.75\}$ y

$$u_h = \theta_1\varepsilon_h + \theta_2\varepsilon_m + \vartheta_h$$

$$u_m = \tau_1\varepsilon_h + \tau_2\varepsilon_m + \vartheta_m$$

con $\vartheta_j \sim N(0,1)$, $\theta_k = \{0,1\}$ y $\tau_k = \{0,1\}$ para $j = \{h,m\}$ y $k = \{1,2\}$.