



Población y Salud en Mesoamérica

Revista electrónica publicada por el
Centro Centroamericano de Población
Universidad de Costa Rica, 2060 San José, Costa Rica
<http://ccp.ucr.ac.cr>

Población y Salud en Mesoamérica

Revista electrónica

Volumen 4, número 1, informe técnico 2

Julio - diciembre, 2006

Publicado 1 de julio, 2006

<http://ccp.ucr.ac.cr/revista/>

Cómo diagnosticar y corregir el problema de la endogeneidad: el número de hijos tenidos en la predicción de las preferencias de fecundidad en Costa Rica

María Isabel González

Cómo diagnosticar y corregir el problema de la endogeneidad: el número de hijos tenidos en la predicción de las preferencias de fecundidad en Costa Rica

María Isabel González¹

Resumen

Se ilustra un método sencillo para diagnosticar una posible endogeneidad del número de hijos tenidos en la predicción del número ideal de hijos, del deseo por un nuevo hijo y de la fecundidad no deseada. Se concluyó que el número de hijos tenidos es una variable exógena y se utilizó para estudiar la asociación entre las características de la madre y las preferencias de fecundidad. Se concluyó que el número ideal de hijos aumenta con el número de hijos tenidos y es más alto para las mujeres católicas que para las que no lo son, que la posibilidad de que una mujer desee un nuevo hijo es menor si se tiene mayor edad, si no se es católica y si se tiene más hijos y la posibilidad de que una mujer haya deseado su último hijo es mayor entre las de mayor edad y las que trabajan y menor entre las católicas y entre las que han tenido menos hijos.

Introducción.

El estudio de las preferencias de fecundidad suele abordarse desde el punto de vista de la predicción, a la luz de modelos que muestran los determinantes de la fecundidad. Bongaarts (1978), por ejemplo, establece un marco para analizar los determinantes próximos de la fecundidad, en el cual ubica a los factores socioeconómicos (en los cuales podrían ubicarse las preferencias), culturales y ambientales como determinantes indirectos y a las variables como control directo de la fecundidad, proporción de matrimonios y factores naturales inherentes a las parejas como determinantes directos.

En Costa Rica, Rosero et.al. (1982) plantearon un modelo para la explicación de la fecundidad (Gráfico 1) que identifica determinantes básicos, condicionantes del control y componentes intermedios. En este modelo las preferencias de fecundidad se ubican dentro de los componentes del control y se suponen influenciados o determinadas por factores individuales como edad, duración de la unión e hijos tenidos y factores estructurales como educación, clase social, y urbanización.

Un análisis para la predicción del número de hijos deseados (González y Chen-Mok, 2003) realizado con base en la Encuesta de Fecundidad de Costa Rica, 1999 mostró, sin embargo que la única variable que aportaba a la predicción del tamaño deseado de familia era el número de hijos tenidos vivos. Esto sugirió el planteamiento de un modelo alternativo que se muestra en el Gráfico 2. En este modelo se plantean como hipótesis:

¹ Máster en Estadística con Énfasis en Población, Universidad de Costa Rica. Catedrática Escuela de Estadística, Universidad de Costa Rica.

A.- Las características de la mujer determinan tanto el número de hijos tenidos como las preferencias de fecundidad.

B.- Las características del hogar y las características del compañero solamente afectan a las preferencias de fecundidad indirectamente, por medio del número tenido de hijos.

Sin embargo, el modelo descrito sugiere una posible endogeneidad de la variable “número de hijos tenidos”. Cuando una variable independiente es endógena en un modelo de regresión y esto no se corrige, los estimadores son sesgados e inconsistentes y pueden llevar a conclusiones erróneas e incluso aberrantes. Una variable es endógena cuando sus valores están determinados dentro del modelo y es predeterminada o exógena cuando sus valores se determinan fuera del modelo. Las variables endógenas se consideran estocásticas, y por ende a ellas se asocia un error, mientras que las exógenas son medidas sin error (Gujarati, 2003, Goldberger, 1964).

La importancia de la endogeneidad ha sido reconocida fundamentalmente por los econométricos, sin embargo es cada vez más frecuente encontrar estudios de otras disciplinas en las que se estudia y se toma en consideración para llegar a buenos resultados. Zohoori (1997) afirma que “muchos estudios epidemiológicos pasan por alto una fuente importante de sesgo en potencia, a saber, la endogeneidad potencial de las variables de comportamiento”, y llevó a cabo una investigación empírica con base en los resultados de la Encuesta Cebu de Salud Longitudinal y Nutrición de los efectos de la endogeneidad y la heterogeneidad no observadas en el análisis de los resultados en salud, demostrando que los valores estimados para los parámetros son diferentes si se corrige por endogeneidad.

Connelly et.al.(2002) estudiaron endogeneidad potencial de la estructura de la familia en un proyecto de investigación relativo al empleo de la mujer y estrategias familiares para la provisión de cuidados a los niños en Brasil.

Bollen et.al. abordaron el problema para la predicción de variables respuesta binarias y variables explicativas endógenas y utilizaron varios métodos en una aplicación que modelaba la diferencia entre el tamaño ideal de familia y el número real de hijos tenidos como un determinante del uso de métodos anticonceptivos, con base en la Encuesta de Demografía y Salud de Tunisia en 1988, encontrando que un modelo probit de una sola ecuación resultaba adecuado para la predicción.

Stern (1995) mostró que controlar por endogeneidad utilizando variables instrumentales hacía más pequeños los efectos de variables relativas a los niños sobre las decisiones para el cuidado de ellos a largo plazo en USA.

Sander examinó el efecto de la escolaridad de la mujer sobre la fecundidad, prestando especial atención al hecho de que la escolaridad fuera un determinante exógeno de la fecundidad, con datos de la US National Opinion Research Center's General Social Survey para los años 1985-91, utilizando la prueba de Hausman, con el cual demostró que la variable era exógena.

El objetivo de este trabajo es probar las hipótesis A y B planteadas anteriormente en tres de las variables que conforman las preferencias de fecundidad: el tamaño ideal de familia, la fecundidad no deseada y el deseo por un nuevo hijo, haciendo las correcciones necesarias en el caso de que la variable “número de hijos tenidos” resulte ser endógena en el modelo, con base en la información de la Encuesta Nacional de Salud Reproductiva realizada en Costa Rica en 1999. Se busca, también, ilustrar un método sencillo para diagnosticar, y, si es necesario, corregir el problema de la endogeneidad.

Metodología.

La encuesta

La Encuesta Nacional de Salud Reproductiva de 1999 fue realizada por el Programa Centroamericano de Población de la Escuela de Estadística de la Universidad de Costa Rica. El propósito de ésta fue recopilar información relativa a varios ámbitos entre los que se cuentan información general de los hogares, antecedentes y características de la entrevistada (incluyendo datos sobre la condición migratoria), historia de los embarazos, comportamiento post-parto, vacunas y salud infantil, planificación familiar, estado conyugal y preferencias de fecundidad. El presente trabajo se centra en el último de los temas mencionados: preferencias de fecundidad.

La muestra incluyó 50 conglomerados y un total de 1612 entrevistadas.

Población en estudio.

La población de interés de esta encuesta la constituyen todas las mujeres residentes en viviendas particulares, en Costa Rica, en 1999, con edades comprendidas entre 18 y 44 años.

VARIABLES A ANALIZAR.

Las variables que se seleccionaron para ser analizadas en este trabajo fueron las respuestas a las siguientes preguntas:

- Para la medición de la fecundidad no deseada:
¿Cuándo usted quedó embarazada, quería quedar embarazada en ese momento, quería esperarse un poco, o del todo no quería embarazarse?
- Para la medición del deseo de un nuevo hijo:
¿Le gustaría tener un/otro hijo o preferiría no tener más hijos?
- Para la medición del tamaño deseado de familia:
Si pudiera escoger el número de hijos, ¿cuántos le gustaría tener en toda su vida?

Además, las siguientes, con el fin de utilizarlas para medir asociaciones con las variables de interés o para predicción:

De la entrevistada:

- Edad en años cumplidos.
- Número de hijos tenidos vivos.
- Nivel educativo, medido como número de años aprobados.
- Si la entrevistada trabaja o no.
- Si la entrevistada es de la religión católica o no.

Del jefe o del principal sostén económico del hogar:

- Escolaridad (número de años de estudio aprobados)
- Ocupación, a un dígito, con las siguientes categorías:
 - 1.- Directores y gerentes
 - 2.- Profesionales
 - 3.- Técnicos
 - 4.- Oficinistas
 - 5.- Trabajadores de servicios personales, vendedores
 - 6.- Agricultores, trabajadores agropecuarios
 - 7.- Oficiales y operarios de industria
 - 8.- Operadores de máquinas, chóferes
 - 9.- Trabajadores no calificados de ventas y servicios
 10. Peones

Del hogar

- Índice de nivel económico
- Ubicación (rural, urbana)
- Número de miembros
- Duración de la unión

Todas las variables mencionadas corresponden a una pregunta en la encuesta, con excepción del índice económico que agrupa las siguientes variables: (codificadas 0, 1):

- tenencia de electrodomésticos (cocina, televisión, televisión con cable, refrigeradora, licuadora, teléfono)
- disponibilidad de electricidad
- disponibilidad de agua de cañería
- posesión de vehículo de transporte (bicicleta, motocicleta, automóvil o similar)
- servicio sanitario (cloaca)

El índice se calcula como :

$$I = \frac{\sum k_i * 1}{f_i} \quad \text{donde } k_i = 1 \text{ si tiene} \\ = 0 \text{ si no tiene}$$

f_i = frecuencia de personas que dijeron contar con el rubro respectivo, así, el índice es la suma del puntaje de cada uno de los rubros mencionados, ponderado por el inverso de la frecuencia del rubro respectivo. (Madrigal, 1986)

Análisis.

La endogeneidad es un problema que se puede corregir. El primer paso es la identificación de la o las variables independientes potencialmente endógenas en el modelo. En este caso la variable que se identificó como potencialmente endógena en el modelaje tanto del número ideal de hijos como del deseo de un nuevo hijo o la fecundidad no deseada fue el número tenido de hijos. Las características de la mujer se plantean como asociadas tanto con el número tenido de hijos como con las tres variables dependientes mencionadas. Por otro lado, el marco teórico supone una asociación de las características del hogar y de las características del compañero con el número tenido de hijos, pero no identifica una asociación directa con las variables dependientes que se desea modelar (ver Grafico 2). Esto permite que se utilicen como variables instrumentales o variables de identificación en el análisis de endogeneidad.

El procedimiento que se utilizó para evaluar la posible endogeneidad del número de hijos tenido fue la prueba de Hausman, de la siguiente manera (Maddala, 1996, Gujarati, 2003)

1.- Se planteó un sistema de ecuaciones simultáneas:

$$Y = \beta_0 + \beta_1 H + \beta_i X_i + \mu \quad (1)$$

$$H = \gamma_0 + \gamma_i Z_i + \gamma_j Q_j + v \quad (2)$$

Donde: Y = número ideal de hijos ó deseo por un nuevo hijo ó deseo por el último hijo tenido

H = número de hijos tenidos

X_i = características de la mujer: edad en años cumplidos, nivel educativo, condición de trabajo, religión

Z_i = características del hogar: índice de nivel económico, ubicación, número de miembros, duración de la unión

Q_j = características del compañero: escolaridad, ocupación

Las variables X_i, Z_i, y Q_j son exógenas, H es potencialmente endógena. Z_i y Q_j son instrumentales.

2.- Se probó si Z_i, y Q_j son independientes de Y agregándolas a la ecuación (1) y evaluando la significancia de sus coeficientes de regresión.

3.- Se evaluó la ecuación (2) y se obtuvieron los residuos estimados

4.- Los residuos estimados de la ecuación (2) se agregaron a la ecuación (1) y se evaluó la significancia de su coeficiente de regresión. Si el coeficiente es significativo la variable número de hijos tenidos es endógena.

Para la predicción del número ideal de hijos se utilizó regresión de mínimos cuadrados y para la predicción del número deseado de hijos y del deseo por el último hijo tenido, ambas dicotómicas, se utilizó regresión logística.

Los datos se procesaron con el paquete estadístico STATA8, utilizando los comandos svyreg svyprobit y svylogit, que corrigen por una posible multicolinealidad provocada por el diseño de la muestra.

Resultados

Diagnóstico de endogeneidad.

Para la predicción del número ideal de hijos, la variable número de hijos tenidos vivos, como el deseo por un nuevo hijo y la fecundidad no deseada resultaron ser variables exógenas. Este resultado puede apreciarse en el Cuadro 1, que presenta las probabilidades asociadas a cada uno de los coeficientes de regresión para cada una de las variables dependientes modeladas. La **ECUACIÓN 0** estima cada variable dependiente con todas las variables independientes en consideración. Su objetivo es mostrar que en los tres modelos todas las variables correspondientes al hogar y al compañero resultaron ser no significativas, lo cual apoya lo planteado en la hipótesis B y llena el primer requisito para que sean usadas como variables instrumentales o de identificación.

La **ECUACIÓN 1** y la **ECUACIÓN 2** muestran los resultados obtenidos, para las tres variables dependientes, en el sistema de ecuaciones simultáneas planteado en la metodología para la prueba de Hausman.. Como se puede observar en la ecuación 2, la duración de la unión, el índice de artefactos, el número de miembros del hogar y la escolaridad del jefe resultaron altamente significativos en la predicción del número de hijos tenidos, lo que comprueba la hipótesis B y las hace útiles como variables de identificación, porque cumplen, además, con el requisito de no estar asociadas con el número ideal de hijos, ni con el deseo por un nuevo hijo ni con el deseo por el último hijo tenido.

Por último, los residuos de esta ecuación resultaron no significativos al ser agregados a la ecuación 1, (esto se puede observar en la **ECUACIÓN 3**) con lo que se concluye que la variable número de hijos tenidos es exógena en la predicción del número ideal de hijos, del deseo por un nuevo hijo y del deseo por el último hijo tenido.

Si el número de hijos tenidos hubiera resultado endógena en alguno de los modelos evaluados, esta endogeneidad se habría controlado con sólo utilizar los valores predichos por la ecuación 2 en lugar de usar los datos reales.

La predicción de las preferencias de fecundidad.

Para la predicción del número ideal de hijos el modelo muestra dos variables significativas a un nivel de significancia del 5%: el número de hijos tenidos vivos y la religión. Por cada hijo adicional tenido vivo el número ideal de hijos aumenta en $=.26$ y no pertenecer a la religión católica lo disminuye en $.22$. La educación y la condición de trabajo no aportan de manera significativa a la explicación del número ideal de hijos, como tampoco lo hace la edad.

En el caso del deseo por un nuevo hijo resultaron significativas la religión, los hijos tenidos vivos y la edad de la entrevistada. Resultaron no significativas la educación y la condición de trabajo.

Las razones de ventaja obtenidas para las variables que resultaron significativas fueron:

Religión	0.56
Hijos tenidos vivos	0.50
Edad de la entrevistada	0.95

las cuales indican que por cada mujer de religión católica que desea un hijo adicional hay la posibilidad de que 0.56 mujeres no católicas lo deseen, por cada mujer que desea un hijo adicional hay una posibilidad de que haya 0.50 mujeres con un hijo tenido menos que lo deseen y una posibilidad de que haya 0.95 mujeres que lo deseen entre las que tienen un año menos de edad.

La fecundidad no deseada, medida por medio del deseo por el último hijo tenido, resultó estar asociada con los hijos tenidos vivos, la condición de trabajo, la religión y la edad de la entrevistada. Solamente no agrega nada a la predicción la variable educación.

Las razones de ventaja obtenidas para las variables que resultaron significativas en el modelo fueron:

Religión	0.71
Condición de trabajo	1.54
Hijos tenidos vivos	0.77
Edad de la entrevistada	1.04

Que indican que por cada mujer que deseó el último hijo tenido hay la posibilidad de que 1.04 con un año menos lo desearan, de que 0.77 mujeres con un hijo menos lo desearan, de que lo desearan 1.54 mujeres que trabajan por cada una que no trabaja y de que lo desearan 0.71 mujeres católicas por cada mujer no católica.

Resumen de conclusiones.

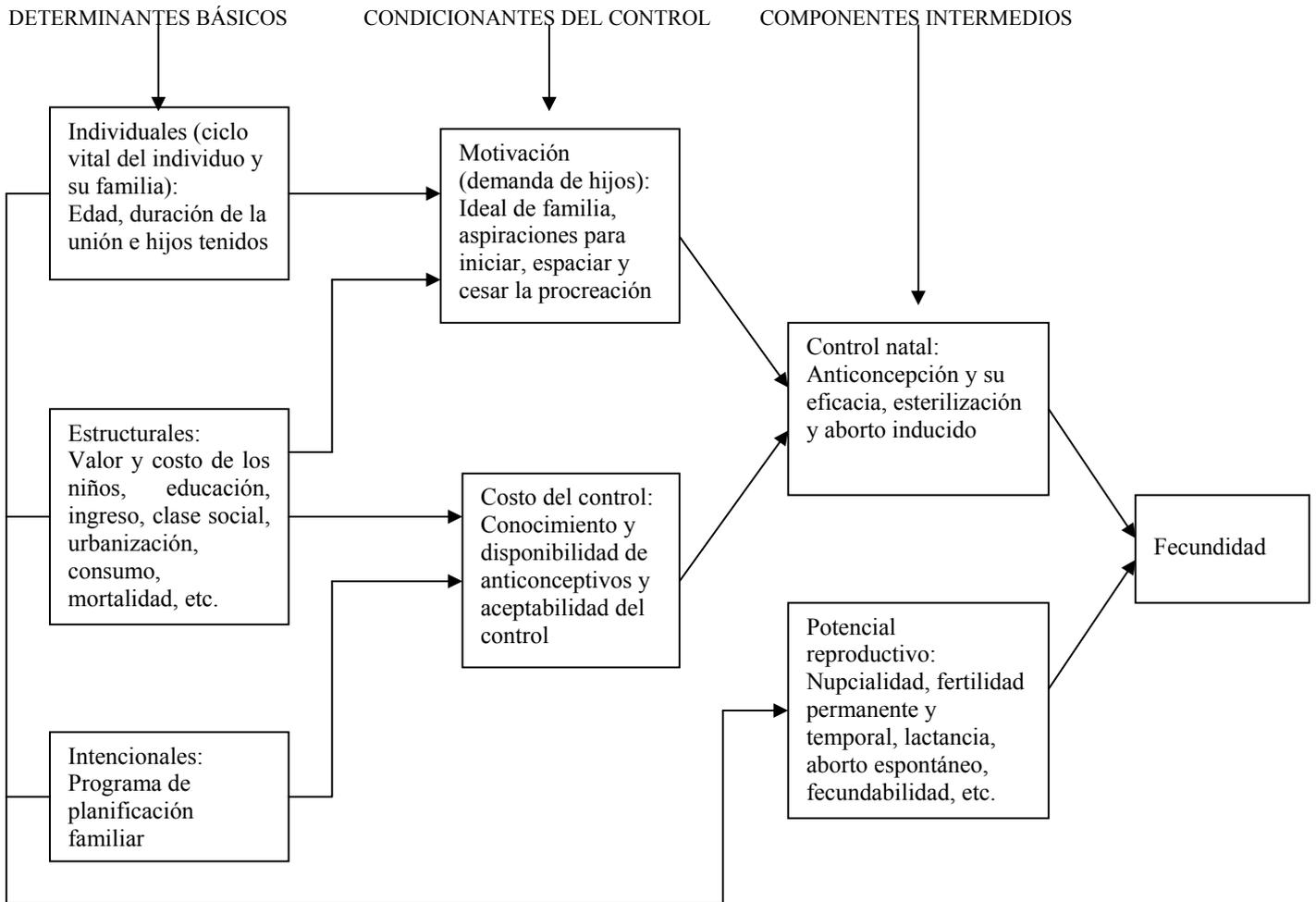
- El número de hijos tenidos no es una variable endógena en los modelos utilizados para estimar el número ideal de hijos, el deseo por un nuevo hijo y el deseo por el último hijo tenido.
- El número ideal de hijos aumenta con el número de hijos tenidos y es más alto para las mujeres católicas que para las que no lo son. No está asociado ni con la condición de trabajo ni con la educación ni con la edad.
- La posibilidad de que una mujer desee un nuevo hijo es menor si se tiene mayor edad, si no se es católica y si se tiene más hijos.
- La posibilidad de que una mujer haya deseado su último hijo es mayor entre las de mayor edad y las que trabajan y menor entre las católicas y entre las que han tenido menos hijos.

Referencias bibliográficas

1. Benefo, KD and Shultz, T.P. Determinants of fertility and child mortality in Cote d'Ivoire and Ghana. Living Standards Measurement Study Working Paper. World Bank. Washington D.C. 1994; 103.
2. Bollen, KA; Guilkey, DK, and Mroz, TA. Binary outcomes and endogenous explanatory variables: test and solutions with an application to the demand for contraceptive use in Tunisia. *Demography*. 1995; 32(1):111-31.
3. Bongaarts, John. A framework for analyzing the proximate determinants of fertility. *Population Development Review*. 1978; 4:105-132.
4. Briscoe, J; Akin, J, and Guilkey, D. People are not passive acceptors of threats to health: endogeneity and its consequences. *International Journal of Epidemiology*. 19:147-153.
5. Conelly, R; DeGraff, DS; Levinson, D, and McCall, B. Tackling endogeneity: alternatives for analysis of women's employment and child care in Brazil. Unpublished. Presented at the Annual Meeting of Population Association of America, New Orleans, Louisiana. 1996; 23.
6. González, M.I., Chen-Mok, M. El número deseado en hijos en Costa Rica: 1993- 1999. *Población y Salud en Mesoamérica*. 2004; 1(2)
7. Gujarati, Damodar. *Econometría* . 4° Edición. México: McGraw-Hill Interamericana; 2003
8. Madrigal, J. Metodología y construcción de un indicador del ingreso familiar. Aplicación a los Censos Nacionales de Vivienda y Población. 1984: Asociación Demográfica Costarricense.
9. Maddala, G.S. *Introducción a la Econometría*. 2° Edición. México: Prentice Hall Hispanoamericana, S.A. 1996
10. Rosero-Bixby; Gómez, M, y Rodríguez, V. Determinantes de la Fecundidad en Costa Rica. Análisis longitudinal de tres encuestas. 1982: Dirección General de Estadística y Censos.
11. Sander W. The effects of women schooling on fertility. *Economic Letters*. 1992 Oct; 40(2):229-33.
12. Stern, S. Estimating family long-term decisions in the presence of endogenous child characteristics. *Journal of Human Resources*. 1995; 30(3):551-80.
13. Tintner, Gerhard. *Econometrics*. New York: John Wiley & Sons; 1963.

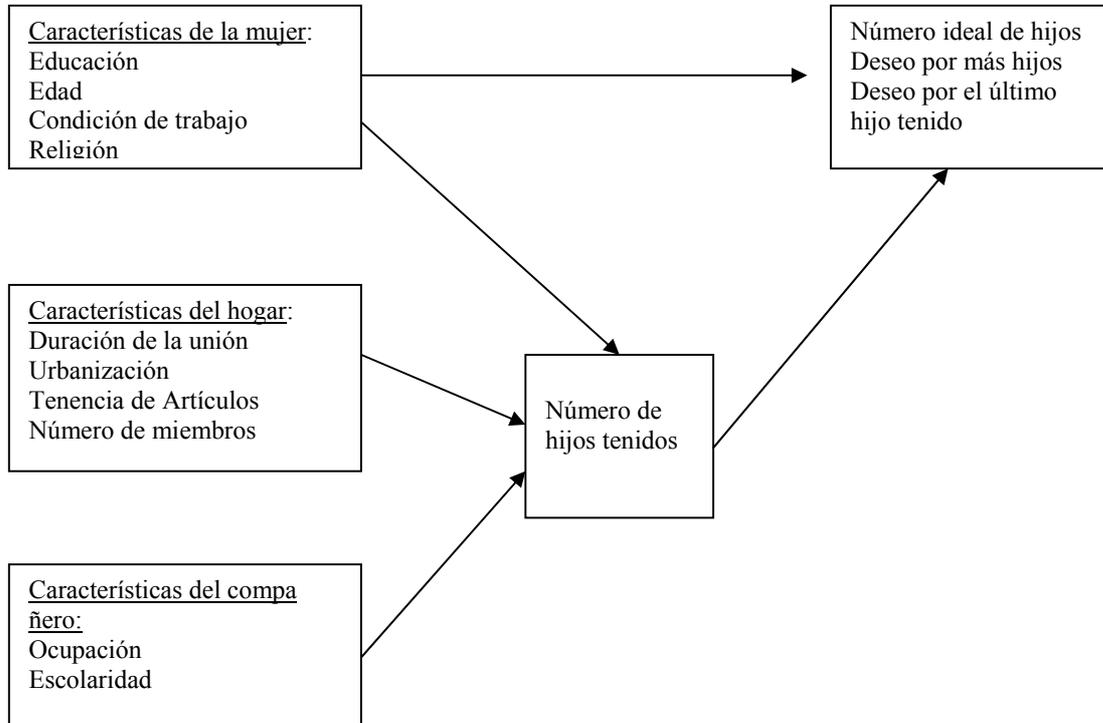
14. Uchudi, M. Spouses' socioeconomic characteristics and fertility differences in sub-Saharan Africa: does spouses' education matter? *Journal of Biosocial Science*. 2001 Oct; 33(4):481-502.
15. Zohoori, Namsar. ¿Es importante la endogeneidad?. Una comparación de los análisis empíricos controlando o no por endogeneidad. Departamento De Nutrición y El Centro De Población De Carolina, Universidad De Carolina Del Norte, Chapel Hill, N.C.

Grafico 1. Determinantes de la Fecundidad Costarricense



Fuente: Rosero-Bixby L., Gómez M. y Rodríguez J. Determinantes de la fecundidad en Costa Rica. Análisis Longitudinal de tres encuestas. Dirección General de Estadística y Censos. 1982

Grafico 2. Marco teórico para la predicción de las preferencias de fecundidad



Cuadro 1. Resultados del análisis de endogeneidad para el número de hijos tenidos en la predicción del número ideal de hijos, del deseo por más hijos y del deseo por el último hijo tenido

Variables independientes	Probabilidades asociadas en la predicción del:		
	Número ideal de hijos	Deseo por más hijos	Deseo por el último hijo tenido
ECUACIÓN 0			
Educación	0,094	0,464	0,719
Condición de trabajo	0,240	0,065	0,003
Religión	0,007	0,003	0,011
Hijos tenidos	0,000	0,000	0,000
Edad	0,290	0,311	0,200
Duración de la unión	0,422	0,088	0,070
Ruralidad	0,508	0,585	0,778
Indice de artefactos	0,230	0,705	0,561
Miembros del hogar	0,217	0,166	0,431
Escolaridad del jefe	0,636	0,535	0,245
Ocupación 1	0,825	0,376	0,318
Ocupación 3	0,882	0,433	0,888
Ocupación 4	0,515	0,175	0,776
Ocupación 5	0,523	0,154	0,408
Ocupación 6	0,294	0,034	0,461
Ocupación 7	0,776	0,241	0,952
Ocupación 8	0,728	0,247	0,080
Ocupación 9	0,935	0,046	0,177
Constante	0,000	0,001	0,459
ECUACIÓN 1			
Educación	0,206	0,646	0,268
Condición de trabajo	0,373	0,194	0,001
Religión	0,022	0,001	0,012
Hijos tenidos	0,000	0,000	0,000
Edad	0,073	0,000	0,000
Constante	0,000	0,000	0,079
ECUACIÓN 2			
Educación	0,000		
Condición de trabajo	0,127		
Religión	0,013		
Edad	0,000		
Duración de la unión	0,000		
Ruralidad	0,820		
Indice de artefactos	0,000		
Miembros del hogar	0,000		
Escolaridad del jefe	0,000		
Constante	0,009		

ECUACIÓN 3

Educación	0,365	0,286	0,112
Condición de trabajo	0,137	0,189	0,003
Religión	0,004	0,009	0,012
Hijos tenidos	0,000	0,000	0,000
Edad	0,360	0,017	0,004
RESIDUOS	0,129	0,125	0,986
Constante	0,000	0,000	0,196
