

ESTRUCTURA DEMOGRÁFICA E INGRESOS.
UN ANÁLISIS EMPÍRICO PARA ARGENTINA

ESTRUCTURA DEMOGRÁFICA E INGRESOS.
UN ANÁLISIS EMPÍRICO PARA ARGENTINA

Mariana Marchionni
Tesis doctoral

Departamento de Economía
Facultad de Ciencias Económicas
Universidad Nacional de La Plata



Marchioni, Mariana

Estructura demográfica e ingresos : un análisis empírico para Argentina -
1a ed. - La Plata : Universidad Nacional de La Plata, 2005.

126 p. ; 25x17 cm.

ISBN 950-34-0337-5

1. Demografía. I. Título

CDD 304.6

Fecha de catalogación: 11/10/2005

A mis padres Isabel y Daniel,
a mi abuela Tita,
a Leo y a nuestros hijos, Santiago y María.

AGRADECIMIENTOS

Quiero agradecer a Walter Sosa Escudero, mi director de tesis. Por sus innumerables e inteligentes ideas y sugerencias. Por ser quien compartió conmigo todo lo que sabe sobre econometría y sobre ser un buen investigador y docente. A mis amigas y compañeras del doctorado, especialmente a Cecilia Rumi por cada una de sus alentadoras palabras y su contagioso optimismo. A los investigadores y docentes del Departamento de Economía de la Facultad de Ciencias Económicas de la Universidad de La Plata, particularmente a Alberto Porto y Mario Szychowski, por sus sugerencias y apoyo. Agradezco también al Programa Mecovi-Argentina (IADB/WB/Cepal) y a la Academia Nacional de Ciencias Económicas por su confianza y por la ayuda financiera que me brindaron a través de sus becas, y a Sebastián Galiani por sus sugerencias.

A mis amigas del alma, Florencia y Sandra, porque a pocas cuadras o a miles de kilómetros estuvieron y están siempre ahí, del otro lado del teléfono o atrás de un mate compartido entre charlas. A mis amigas de toda la vida, Julieta, Lorena, Nadina y Virginia, por haber compartido conmigo la infancia, la adolescencia y el presente.

A Carlos y Zulma, porque a través de ellos he heredado una nueva y maravillosa familia. Por todo el cariño y el apoyo que me dieron tanto en las cuestiones domésticas como en las académicas.

A mi abuela Tita, porque ha sido, es y será siempre uno de los pilares fundamentales de mi vida. Por su oído siempre atento y su contagiosa alegría.

A mi tía Elena, que fue como otra abuela para mí. Por el amor que siempre me dio.

A mis hermanos, Diego, Julieta, Laura y Magdalena. Porque crecimos juntos y el cariño que nos une trasciende continentes y puntos de vista.

A mis padres, Isabel y Daniel, porque con su ejemplo callado han guiado y guían cada uno de mis pasos. Porque me transmitieron todo lo importante en la vida y me enseñaron que esforzarse es tan importante como darse el tiempo para disfrutar.

A Leo, mi esposo, por su amor y su paciencia. Porque con cada consejo hace mi vida más fácil y de mí una mejor persona. Por todas sus inteligentes sugerencias para esta tesis y su apoyo constante durante el proceso.

A Santiago y María, mis hijos, que llenan mi vida de felicidad. Por su ternura y paciencia cuando trabajo. Ir a jugar con ellos fue el principal incentivo para terminar esta tesis.

PRESENTACIÓN

El grado académico máximo que otorga la Universidad Nacional de La Plata es el de Doctor (Ordenanza N° 205/87 del Consejo Superior). Según la Ordenanza, el Doctorado “constituye el desarrollo y perfeccionamiento de un área del conocimiento, cuya universalidad debe atender, en un marco de alto nivel de excelencia académica que permita la obtención de verdaderos aportes originales en el campo elegido. La presentación de una Tesis Doctoral es ineludible para evaluar dichos aportes”.

En 1999 el Departamento de Economía de la Facultad de Ciencias Económicas elaboró un proyecto de creación del Doctorado en Economía como culminación de un proceso que comenzó con la modernización y mejoramiento académico de la carrera de grado (Licenciatura en Economía) y la creación de dos posgrados de alto nivel académico (categorizados Bn y A por la CONEAU). El Doctorado en Economía fue aprobado por el Consejo Académico de la Facultad por Res. N° 129/99 y por la U.N.L.P. el 1° de Septiembre de 2000. En el año 2000 fue presentado como proyecto a la CONEAU que lo acreditó en tal condición. Posteriormente, por Res. N° 348/04 del Ministerio de Educación, Ciencia y Tecnología de la Nación se le dio validez nacional al título.

Las actividades de doctorado comenzaron en Junio de 2000 y comprenden las etapas de creación de conocimientos, transmisión a través de la docencia de grado y posgrado, la extensión y la transferencia a los sectores público y privado.

Luego de más de cuatro años de intensa actividad se han completado dos trabajos de Tesis Doctoral y se han otorgado los dos primeros títulos de Doctor en Economía.

En este volumen se presenta el trabajo de Tesis Doctoral de Mariana Marchionni que fue defendido públicamente el 1° de Julio de 2005 y obtuvo la calificación máxima de 10 puntos.

El Departamento de Economía de la Facultad de Ciencias Económicas y Edulp ponen a disposición de docentes, investigadores, estudiantes y público en general esta Tesis Doctoral que es un aporte original sobre el tema “Estructura demográfica e ingresos. Un análisis empírico para la Argentina”.

Dr. Alberto Porto
Director del Doctorado en Economía
Universidad Nacional de La Plata
La Plata, Octubre de 2005

ÍNDICE

INTRODUCCIÓN.....	1
CAPÍTULO I	
Determinantes demográficos de la distribución del ingreso	9
Apéndice al Capítulo I	31
CAPÍTULO II	
El impacto distributivo de los cambios en las decisiones de fecundidad	41
Apéndice al Capítulo II	65
CAPÍTULO III	
Un modelo de elección binaria con regresor entero endógeno	75
Apéndice al Capítulo III.....	103
COMENTARIOS FINALES	111

INTRODUCCIÓN

...no state has hitherto existed (at least that we have any account of) where the manners were so pure and simple, and the means of subsistence so abundant, that no check whatever has existed to early marriages, among the lower classes, from a fear of no providing well for their families, or among the higher classes, from a fear of lowering their condition in life.

Thomas Malthus

Desde fines del siglo XVIII con la publicación de “Un ensayo sobre el principio de la población” de Thomas Malthus, la interacción entre las cuestiones demográficas y la pobreza ha sido un tema central para la Economía. Ningún análisis que pretenda dar una explicación profunda de la evolución de la distribución del ingreso puede dejar de lado las causas y efectos de cambios en el tamaño y composición de las familias. Pero los factores capaces de explicar los cambios distributivos son múltiples y complejos y, como señalan Atkinson y Bourguignon (2000), aún no existe un marco teórico que permita abordar el análisis de los determinantes de la distribución del ingreso y sus cambios en forma unificada y consistente.

La mayor parte de las investigaciones en esta línea se concentran en estudiar los efectos distributivos de variables vinculadas al mercado laboral. Así se analiza el rol de los retornos a la educación y a la experiencia, y de la brecha salarial entre géneros, entre otros. A diferencia de estos trabajos, esta tesis se aboca al estudio de los factores demográficos como determinantes de la distribución del ingreso, adoptando una perspectiva microeconómica, es decir, se explora el canal demográfico tomando como unidad de análisis al individuo en su contexto familiar.

El enfoque desarrollado a lo largo de esta tesis es predominantemente empírico y se concentra en el caso del Gran Buenos Aires (GBA) –Ciudad de Buenos Aires y su conurbano– durante las tres últimas décadas del siglo XX. Este período fue testigo de profundos cambios distributivos en la Argentina, y en particular en el GBA. Los indicadores de pobreza y desigualdad, luego de mantenerse relativamente bajos y estables durante los años cincuenta y sesenta, comenzaron un ascenso que se acentuó durante la década de 1990. También la estructura demográfica de las familias se transformó significativamente en ese lapso. La fecundidad cayó marcadamente, se generalizó la cohabitación y aumentó la edad de nupcialidad y la incidencia de la

separación y el divorcio. También la participación laboral de la mujer creció como consecuencia de la evolución de la condición femenina a lo largo de estos años.

Ante la contemporaneidad de tan profundas transformaciones demográficas y distributivas, cabe preguntarse acerca de cuáles son las posibles interrelaciones entre ambos fenómenos. En esta tesis se explora una de las direcciones en que esa interdependencia puede operar: el efecto de los cambios demográficos sobre los ingresos y su distribución. Varios interrogantes se abren con respecto a este tema: ¿Cuáles son las características demográficas de los individuos y las familias que pueden afectar a los ingresos familiares, a los individuales y a su distribución? ¿Cuáles son los efectos directos e indirectos de los cambios demográficos? ¿Cuál es la magnitud de esos efectos? ¿Cuál es el impacto sobre la pobreza y cuál sobre la desigualdad en la distribución del ingreso? La investigación realizada en esta tesis busca dar respuesta a estas preguntas.

La tesis consiste en tres trabajos de investigación, que aunque independientes entre sí y autocontenidos, siguen un hilo conductor. La investigación del Capítulo I se concentra en los “Determinantes demográficos de la distribución del ingreso”, y tiene por objetivo presentar evidencia empírica sobre la evolución de distintas variables demográficas y su potencial relación con la tendencia observada de crecimiento de la pobreza y la desigualdad. Este es un trabajo fundamentalmente descriptivo para el que se emplea una serie de datos de corte transversal provenientes de la Encuesta Permanente de Hogares (EPH) que cubre el período 1974-2000.

En el análisis se consideran tres grupos de variables demográficas: tamaño familiar, estructura familiar y patrones de formación de uniones conyugales. En cada caso se presenta evidencia sobre la estructura observada por estrato de ingreso y su evolución a lo largo del tiempo, que permite conjeturar sobre la existencia de tendencias demográficas que pudieron haber contribuido al aumento de la pobreza o de la desigualdad. Sin embargo, sólo se tienen en cuenta los efectos directos de los cambios demográficos. Por ejemplo, al evaluar el impacto distributivo de un aumento en el número de hijos en los hogares más pobres se toman como dados los ingresos totales del hogar, lo que implica suponer —entre otras cosas— que las decisiones laborales no se modifican ante ese cambio en la composición familiar.

En la investigación sobre “El impacto distributivo de los cambios en las decisiones de fecundidad”, presentado en el Capítulo II, se busca cuantificar el efecto de los cambios observados en el número de hijos por hogar sobre la pobreza y la desigualdad, considerando tanto los efectos directos de esos cambios demográficos como los indirectos. La evidencia preliminar del Capítulo I sugiere que la brecha en el número de

hijos entre los estratos superiores e inferiores de la distribución del ingreso se ha ampliado sensiblemente como consecuencia de un aumento en el número de hijos de los hogares más pobres y una reducción en el de los más ricos. El efecto directo de estos cambios sería tanto empobrecedor como desigualador, pero resta conocer su magnitud y la de los efectos indirectos que podrían potenciar o contrarrestar al primero. Con el fin de evaluar en qué medida los cambios en las decisiones reproductivas contribuyeron al aumento registrado en la pobreza y la desigualdad durante las últimas dos décadas en el GBA, se aplica la metodología de microsimulaciones o descomposiciones microeconómicas –ver Bourguignon, Ferreira y Lustig (2004). Esta metodología permite responder la siguiente pregunta contrafáctica: ¿Cuáles habrían sido los cambios distributivos entre dos momentos del tiempo si sólo el comportamiento reproductivo hubiera cambiado en ese período? En particular, manteniendo el resto constante, se analiza el impacto que el cambio en los parámetros que guían la decisión sobre el número de hijos puede haber tenido en la distribución del ingreso familiar equivalente por dos vías: el cambio en el número de adultos equivalentes de cada familia –efecto directo– y la modificación de las decisiones laborales del jefe de hogar y su cónyuge –efecto indirecto.

El aporte metodológico de esta investigación consiste en incorporar un tratamiento más riguroso de las decisiones de fecundidad y su impacto sobre la distribución del ingreso en la literatura de microsimulaciones. La estimación de los parámetros que gobiernan las decisiones de fecundidad y las respuestas en términos de participación laboral a cambios en el tamaño de las familias, constituyen el insumo básico para realizar las descomposiciones microeconómicas. Sin embargo, para mantener el ejercicio en un nivel de complejidad razonable, las decisiones de fecundidad fueron modeladas y estimadas independientemente de las ecuaciones de participación laboral.

El problema de la endogeneidad de las características de los hijos en los modelos de participación laboral femenina –señalado tempranamente por autores como Schultz (1978), Dooley (1982), Moffitt (1984) y Hotz y Miller (1988)– causa la inconsistencia de los estimadores habituales de los parámetros de estos modelos. En la investigación desarrollada en el Capítulo III de esta tesis –“Un modelo de elección binaria con regresor entero endógeno”– se discute una especificación particular para el modelo de participación laboral femenina con fecundidad endógena y se propone un método de estimación en dos etapas que permitiría estimarlo en forma sencilla.

El modelo de interés consiste en una ecuación de elección binaria donde se incluye una variable entera como regresor endógeno. Para la aplicación en cuestión, el modelo de elección binaria representa la decisión de participación laboral de las mujeres, mientras

que la variable entera viene dada por el número de hijos. El estimador propuesto –que surge como una adaptación del método de máxima verosimilitud condicional en dos etapas propuesto por Rivers y Vuong (1988)– provee una alternativa simple a la estrategia adoptada por Weiss (1999), que discute una especificación particular para el modelo en un contexto de información completa. La dificultad computacional de obtener estimadores máximo verosímiles para esa especificación es considerable, de manera que su replicación para encarar sistemáticamente el análisis empírico de problemas que respondan a la misma estructura es improbable o demasiado costosa.

El Capítulo III se diferencia de los dos anteriores tanto por la temática abordada como por la metodología que se aplica. Sin embargo, fue motivado por la evolución de la investigación, en particular por la necesidad de contar con estimadores consistentes de los parámetros del modelo de participación laboral de las mujeres, utilizados en el ejercicio de microsimulaciones. En este último capítulo, entonces, además de proponerse un estimador consistente y de implementación sencilla para ese modelo, se reportan los resultados de estimar ecuaciones de participación laboral de mujeres permitiendo que el número de hijos se determine endógenamente. En esta aplicación empírica se utilizan nuevamente datos para Argentina provenientes de la EPH. La evidencia preliminar encontrada sugiere que la endogeneidad no constituye un problema significativo para este caso, por lo que la estimación en forma independiente de las ecuaciones de participación y de fecundidad sería adecuada.

Referencias

- Atkinson, A. Y Bourguignon, F. (eds.). (2000). *Handbook of Income Distribution*. Elsevier Science, B. V.
- Bourguignon, F., Ferreira, F. y Lustig, N. (eds.) (2004). *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*. New York: Oxford University Press.
- Dooley, M. D. (1982). Labor Supply and Fertility of Married Women: An Analysis with Grouped and Individual Data from the U.S. Census. *Journal of Human Resources*, 17, pp. 499-532.
- Hotz, V.J. y R. A. Miller. (1988). An Empirical Analysis of Life Cycle Fertility and Female Labor Supply. *Econometrica*, 56, pp. 91-118.
- Malthus, Thomas. (1798). *An Essay on the Principle of Population as it affects the future improvement of society: with remarks on the speculations of Mr. Godwin, M. Condorcet, and other writers*. London. Printed for J. Johnson, in St. Paul's Church-yard.
- Moffitt, R. (1984). Profiles of Fertility, Labour Supply and Wages of Married Women: A Complete Life-Cycle Model. *Review of Economic Studies*, 51, pp. 263-278.
- Rivers, D. y Q. H. Vuong. (1988). Limited Information Estimators and Exogeneity Tests for Simultaneous probit Models. *Journal of Econometrics*. Vol 39, pp. 347-366.
- Schultz, T. P. (1978). The Influence of Fertility on Labor Supply of Married Women: Simultaneous Equation Estimates. In R. Ehrenberg (ed.). *Research in Labor Economics*, Vol. 2, JAL Press, pp. 273-351.
- Weiss, A. (1999). A Simultaneous Binary Choice / Count Model with an Application to Credit Card Approvals. In *Cointegration, Causality, and Forecasting: A Festschrift in honour of Clive W. J. Granger*. Oxford University Press.

CAPÍTULO I

DETERMINANTES DEMOGRÁFICOS DE LA DISTRIBUCIÓN DEL INGRESO*

* Versiones preliminares de este trabajo fueron presentadas en los Seminarios de Doctorado y de Economía del Departamento de Economía, Facultad de Ciencias Económicas, Universidad Nacional de La Plata.

Introducción

La estructura demográfica argentina ha sufrido fuertes transformaciones durante las últimas décadas del siglo XX. Caída en la fecundidad, profundo cambio en la dinámica de formación de uniones conyugales, generalización de la cohabitación, aumento de la edad de nupcialidad, aumento en la incidencia de la separación y el divorcio son algunos de los principales fenómenos que caracterizan los cambios en el comportamiento demográfico.

Este período también fue testigo de un considerable aumento en la pobreza y la desigualdad de la distribución del ingreso: los indicadores se mantuvieron relativamente bajos y estables durante las décadas de 1950 y 1960, pero desde mediados de los años setenta presentaron un claro patrón ascendente y especialmente acentuado en la década del noventa. La razón entre el ingreso promedio del decil más rico de la población y el más pobre pasó de 10 a 40 en los últimos 25 años del siglo. De la misma forma la pobreza escaló niveles sin precedentes. A principios de los noventa menos de un 20% de la población era pobre y una década más tarde la pobreza era del 30%. Desde 2001, como consecuencia de la profundización de la crisis, la pobreza subió vertiginosamente, superando al 50% en 2002.¹

Los determinantes de estos cambios distributivos son múltiples y complejos y, como reconocen Atkinson y Bourguignon (2000), aún no existe un marco teórico que permita abordar el análisis de los determinantes de la distribución del ingreso y sus cambios en forma unificada y consistente. Entre los principales determinantes, y quizás los más explorados en la literatura, están los vinculados al mercado laboral. Retornos a la educación, a la experiencia, brecha salarial entre géneros, son algunos de los factores que permiten explicar parte de los cambios distributivos observados.

Desde Malthus con su ensayo sobre el principio de la población (1798), la interacción entre las cuestiones demográficas y la pobreza ha sido un tema central para la Economía. Difícilmente pueda encontrarse una explicación de la evolución del ingreso per capita y de su distribución sin incluir en el análisis las causas y efectos de cambios en el número de miembros y composición de las familias. El objetivo de este trabajo es, justamente, explorar empíricamente los efectos distributivos de cambios en variables

¹ Las estadísticas son para el Gran Buenos Aires –Ciudad de Buenos Aires y su conurbano–, pero patrones similares se presentan en otros aglomerados y para el total del país. Fuente: Centro de Estudios Distributivos, Laborales y Sociales (CEDLAS), Universidad Nacional de La Plata.

demográficas a un nivel microeconómico, para lo que se emplean datos provenientes de la Encuesta Permanente de Hogares (EPH).

El ingreso total y por individuo de una familia depende de sus características demográficas. Por ejemplo, dadas las decisiones de participación laboral y los ingresos de los miembros de un hogar, mayor número de hijos implicará menos ingreso per capita.² Luego, si aumenta la fecundidad en hogares de bajos recursos, aumenta el número de pobres y/o la profundidad de la pobreza caeteris paribus. Además, un perfil demográfico diferente entre familias de distintos estratos socioeconómicos puede afectar la desigualdad en la distribución de ingresos. Siguiendo con el ejemplo, si la fecundidad de las familias más “pobres” aumenta y la de las más “ricas” disminuye, es de esperar un aumento en la desigualdad, manteniendo fijo todo lo demás.

El trabajo se organiza como sigue. La sección 2 argumenta sobre la existencia de relaciones entre los cambios demográficos y los distributivos, insertando esta discusión en la literatura económica. En la sección 3 se caracterizan algunas de las transformaciones del perfil demográfico que operaron desde mediados de la década del setenta, vinculando empíricamente estos cambios con los distributivos. La sección 4 expone algunos comentarios finales y conclusiones.

Las variables demográficas y la distribución del ingreso

Las variables demográficas usualmente aparecen en la literatura económica como un importante factor que contribuye en la determinación de la distribución de ingresos, y por lo tanto a los niveles y dispersión de los mismos que importan para medir pobreza y desigualdad. Sin embargo, citando a Lam (1997) “el rango de temas que se incluye en la discusión es amplio y frecuentemente impreciso”. Entre las variables demográficas que la literatura económica ha vinculado con la distribución de ingresos están la fecundidad, la estructura etaria y por géneros de la población, la estructura familiar y los mecanismos de conformación de uniones conyugales. El efecto de cada una sobre la distribución depende de si se analizan los ingresos individuales o si, por el contrario, la unidad de análisis es la familia, como en el caso del presente trabajo.

Un punto de partida posible para investigar las relaciones entre la distribución del ingreso y las cuestiones demográficas es concentrarse en los determinantes de la porción del ingreso familiar disponible para cada miembro del hogar. La ecuación debajo define

² En este trabajo las palabras hogar y familia se usarán indistintamente.

el ingreso familiar equivalente –en adelante ingreso equivalente– del individuo i del hogar h .

$$y_{ih} = \frac{\sum_{j \in h} \left[\left(\sum_k \alpha_j^k F_j^k w_j^k \right) + T_j \right]}{\sum_{j \in h} a_j} \quad \forall i \in h$$

El numerador de la ecuación representa el ingreso total familiar, compuesto por el ingreso de cada miembro que surge de vender una fracción α de su dotación total F de cada factor productivo k a una tasa o precio w , más las transferencias netas recibidas por cada integrante o por la familia en su conjunto. El ingreso equivalente surge de dividir el ingreso total familiar por el número de adultos equivalentes del hogar, que pondera a cada individuo de acuerdo a sus necesidades calóricas según su edad y sexo. El ingreso familiar per capita se obtiene como un caso particular: cuando los ponderadores a_j son todos iguales a 1 el denominador de la ecuación es igual al número total de miembros del hogar. Para la discusión que sigue se emplea el concepto de ingreso equivalente por brindar una mejor aproximación del nivel de vida individual que el ingreso per capita.³

Sea Y un vector $1 \times P$ de los ingresos equivalentes de los individuos de una población formada por P individuos, es decir $Y = (y_1, y_2, y_3, \dots, y_P)$. Este vector incluye toda la información de la distribución de ingresos que nos interesa, aunque generalmente resulta más útil resumir parte de esa información mediante alguna función $Q(Y)$, donde $Q(\cdot)$ puede ser por ejemplo una medida de disparidad de ingresos como el coeficiente de Gini, o de pobreza como la tasa de incidencia. En este trabajo se discuten distintos canales a través de los cuales las características demográficas podrían afectar los ingresos equivalentes de los individuos y como consecuencia el valor de la función $Q(\cdot)$. Para organizar la discusión, resulta útil emplear la ecuación del ingreso equivalente presentada más arriba.

La fecundidad es la principal variable para explicar diferencias en el número de miembros por familia, y obviamente contribuye a determinar el tamaño familiar que es el denominador de la ecuación del ingreso equivalente. Mayor fecundidad implica

³ Más aún, si se considera la existencia de economías de escala internas al hogar, correspondería que el denominador, ya sea del ingreso per cápita o por adulto equivalente, se eleve a un exponente menor que uno. En este caso se obtiene el ingreso familiar per cápita o por adulto equivalente, ajustado por economías de escala internas al hogar.

menor ingreso equivalente caeteris paribus. Cuando la medida de ingreso relevante es per capita, el tamaño familiar se define simplemente como el número de miembros del hogar, mientras que si lo que interesa es el ingreso equivalente, el tamaño familiar también depende de la composición etaria y por géneros de la familia. Un mayor tamaño familiar puede provenir de (a) un mayor número de miembros dada la estructura etaria y por géneros, (b) una mayor proporción de adultos jóvenes dado el número de miembros y la composición por géneros o (c) una mayor proporción de hombres, dado el número de miembros y su composición etaria. Esto es consecuencia de que los adultos jóvenes de sexo masculino suelen “valer” más adultos equivalentes que otros miembros. Por ejemplo, en términos de adultos equivalentes el ponderador de un niño de 2 años es de 0.50, el de un varón entre 13 y 15 años es de 0.96 y el de una mujer de esa edad es de 0.79.

Dada la fecundidad, las diferencias en la composición etaria y por géneros de un hogar y en el número de miembros también pueden deberse a diferentes estructuras familiares.⁴ Una tipología de los hogares según su estructura distingue entre hogares conyugales – donde cohabitan los cónyuges o al menos uno de los padres con sus hijos solteros– y no conyugales. Los primeros, a su vez, pueden clasificarse en hogares completos –ambos padres presentes– o monoparentales –sólo madre o padre presente–, o en hogares conyugales nucleares o no nucleares –dependiendo de si incluyen otros parientes y no parientes. Además, un hogar conyugal puede estar formado por más de un núcleo conyugal. Así se distinguen núcleos conyugales primarios –núcleos conyugales a los que pertenece el jefe de hogar– y secundarios. Por su parte, entre los hogares no conyugales se destacan los hogares unipersonales, que naturalmente consisten en individuos viviendo solos.⁵

La composición etaria y por géneros del hogar también contribuye a determinar con qué tipo de activos F cuenta ese hogar para vender en el mercado, los retornos a esos factores productivos $-w-$, los incentivos y restricciones que determinan los α , y las transferencias netas T que cada individuo –o el hogar en su conjunto– recibe desde el sector público o privado. Como ejemplo, una mujer que trabaja recibe en el mercado laboral un “retorno a ser mujer” que refleja la brecha salarial entre géneros. Si hay discriminación contra las mujeres, ese retorno es negativo. Si la mujer es jefa de un

⁴ Hay evidencia que la fecundidad también puede depender del género de los hijos si los padres tienen preferencia por una mezcla de géneros en su descendencia. Ver Angrist y Evans (1998) y Cruces y Galiani (2003).

⁵ Para una clasificación exhaustiva de la estructura familiar ver Torrado (1998) y Sosa Escudero y Marchionni (1999).

hogar monoparental, y dado todo lo demás, ese retorno negativo implicaría un menor ingreso equivalente en su hogar que en el de un hogar monoparental con jefe hombre.

Otra dimensión demográfica de fundamental importancia como determinante de la distribución del ingreso es el proceso de formación de uniones conyugales, ya sean legales o consensuales. El fenómeno conocido como homogamia designa la tendencia de hombres y mujeres a formar parejas con compañeros de un mismo entorno social. Este fenómeno responde a la existencia de mecanismos que regulan la elección del cónyuge. Según Torrado (2003) hay fundamentalmente dos factores que explican la existencia de tales mecanismos: la segregación social –barrios privados, escuelas, lugares de trabajo, etc.– y la distribución social de gustos y preferencias que son interiorizados por los individuos.

En la literatura económica los modelos teóricos recrean la homogamia a partir de un juego que tiene lugar en un “mercado de matrimonios”. En este juego, cada jugador es un potencial cónyuge que debe seleccionar de todos los compañeros disponibles el que maximice la utilidad derivada de su situación. Si no hay ninguno capaz de mejorar el status quo, permanecer soltero es la opción elegida.⁶ La utilidad que un individuo puede derivar de un matrimonio depende de las características de quien será su cónyuge y de la calidad de la unión o “amor”. Los ingresos y sus principales determinantes, como la educación y otros atributos generalmente inobservables, se consideran usualmente como características valoradas en el mercado de matrimonios. Por lo general, estos modelos teóricos predicen el mismo tipo de ordenamiento que se observa en la realidad, es decir, una asociación entre individuos de características similares. Individuos con “mejores” atributos –o con mayor dotación de “buenos” atributos– suelen unirse con otros similares, e individuos con “peores” atributos –o con menor dotación de “buenos” atributos– hacen lo mismo.⁷

El ordenamiento de los cónyuges de acuerdo a sus características tiene un efecto potencial sobre la distribución de ingresos en la medida que esas características generen

⁶ Los modelos difieren en el número de veces que el individuo puede ir al mercado de matrimonios, en la posibilidad o no de revertir una decisión de matrimonio mediante el divorcio, en la probabilidad de encontrarse con distintos tipos de individuos en el mercado, etc. Fernández, Guner, and Knowles (2001) presentan un modelo en que cada individuo puede ir dos veces al mercado de matrimonios. La primera vez se encuentra con cualquier tipo de individuos, y la segunda sólo con individuos similares a él como consecuencia de una segregación creciente. Aiyagari, Greenwood, and Guner (2000) y Greenwood, Guner, and Knowles (2003) modelan decisiones de divorcio.

⁷ Fernández, Guner y Knowles (2001) es un ejemplo. Su modelo predice un alto grado de correlación positiva de la educación de los cónyuges. El modelo de Willis (1999) predice el mismo patrón para individuos de altos ingresos. Sin embargo, en los estratos inferiores de ingresos, hay mujeres que prefieren quedarse solteras o incluso ser madres solteras, a casarse con hombres “demasiado pobres”.

retornos en el mercado laboral. Por ejemplo, un ordenamiento en base a la educación de los cónyuges se traduce en una alta correlación positiva entre los niveles educativos de, típicamente, el jefe de hogar y su cónyuge. Con retornos a la educación positivos, y dado todo lo demás –en especial la participación de ambos cónyuges en el mercado laboral–, este comportamiento acentuaría la diferencia de ingresos equivalentes entre hogares de distintos estratos socioeconómicos.⁸

De la discusión anterior queda en evidencia cómo el perfil demográfico podría afectar los ingresos a través del tamaño y estructura familiar, su composición etaria y por géneros, y los patrones que gobiernan la formación de las uniones conyugales. Más aún, en la medida que la evolución de estas variables demográficas difiera por estrato socioeconómico, no sólo habría un efecto sobre los niveles de ingreso sino sobre toda la distribución, por lo que el impacto distributivo de los cambios demográficos se reflejaría tanto en las medidas de pobreza como en las de desigualdad.

Cambios demográficos y su impacto distributivo. Evidencia empírica

En Argentina en general y en el Gran Buenos Aires (GBA) en particular, la distribución del ingreso ha sufrido profundas transformaciones a lo largo de las tres últimas décadas del siglo XX. Tanto la pobreza como la desigualdad han tenido un claro patrón ascendente durante todo el periodo, que se profundizó durante los años noventa. Los niveles de desigualdad y pobreza alcanzados hacia fines de esa década no tuvieron precedentes, siendo comparables solamente con los registrados durante la hiperinflación de 1989 –ver Figuras 1 y 2 en el Apéndice a este capítulo.

Como señalan Gasparini, Marchionni y Sosa Escudero (2001), la identificación de las causas de estas profundas transformaciones distributivas se complica ya que son múltiples los factores que convergen en la explicación. En particular, la década de los noventa fue escenario de importantes reformas económicas que pueden haber operado con distinta fuerza y sentido sobre la distribución del ingreso. Estos cambios económicos incluyeron la liberalización comercial, la adopción de nuevas tecnologías productivas y organizacionales, las privatizaciones, la desregulación, la reforma del

⁸ Kremer (1997) encuentra evidencia que indica que para Estados Unidos la presencia de homogamia en la formación de uniones no tendría un gran efecto sobre la desigualdad en el largo plazo, aunque sí sobre la movilidad social.

sistema de seguridad social, las reformas laborales y la modificación de las políticas fiscal, monetaria y cambiaria.

También se modificaron, en forma autónoma o posiblemente en parte como respuesta a estos cambios económicos, algunos comportamientos de los individuos y las familias que contribuyeron a transformar el perfil demográfico. Caída en la fecundidad, cambios en la dinámica de formación de uniones conyugales, generalización de la cohabitación, aumento de la edad de nupcialidad, aumento en la incidencia de la separación y el divorcio son algunos de los principales fenómenos que caracterizan los cambios demográficos durante las últimas décadas en la Argentina y particularmente en el GBA.⁹

En lo que resta de esta sección se presenta evidencia empírica que documenta algunos de esos cambios demográficos y sus posibles vínculos con el patrón observado de crecimiento sostenido de la desigualdad y la pobreza entre 1974 y 2000. Como fuente de información se utiliza una serie de datos de corte transversal provenientes de la Encuesta Permanente de Hogares (EPH). El análisis se realiza para el GBA –Ciudad de Buenos Aires y su conurbano– que concentra aproximadamente un tercio de la población total del país y para la cual se dispone de información representativa para todo el período. Si bien los datos censales serían la mejor opción para caracterizar el perfil demográfico de una población, enfrentan en el caso argentino limitaciones fundamentales que impedirían llevar adelante los objetivos aquí propuestos. En primer lugar, no se dispone de información sobre ingresos monetarios proveniente de los censos, por lo que un análisis que relacione ingresos familiares con otras características del hogar no sería posible. En segundo lugar, los censos se realizan cada diez años, de manera que entre 1974 y 2000 sólo hubo dos operativos del Censo Nacional de Población y Viviendas, uno en 1981 y el otro en 1991.

Cambios en el comportamiento demográfico claramente pueden provocar y a su vez ser provocados por cambios en las variables económicas, en particular en la distribución del ingreso. En ese sentido un análisis profundo de la relación entre ambos fenómenos requeriría la estimación de modelos de equilibrio general que exceden el alcance de este trabajo. El objetivo aquí se concentra en presentar evidencia sobre la transformación del perfil demográfico y explorar en qué dirección esa transformación podría impactar en la distribución del ingreso, sin ignorar que los cambios distributivos podrían afectar a su vez las características demográficas, contrarrestando o potenciando el primer efecto.

⁹ Estas transformaciones se concentraron especialmente en el GBA el polo más moderno del país y, en general, el primero en experimentar este tipo de cambios. Ver Torrado (2003).

Para ordenar la discusión resulta útil distinguir tres grupos de variables demográficas: (3.1) las que determinan el tamaño de las familias, (3.2) las que determinan la estructura familiar, y (3.3) las que rigen los patrones de formación de uniones conyugales.

3.1. Tamaño de las familias

Como se discutió previamente, cuando se utiliza una escala de equivalencias el tamaño familiar depende tanto del número de miembros como de la estructura etaria y por géneros de los mismos, y estas características demográficas son en gran medida determinadas por la estructura familiar. Tomando como dados todos los demás factores relevantes, un hogar monoparental tendría un integrante menos que un hogar completo, y si se tratara de un hogar con jefa mujer tendría además menos adultos equivalentes que otro con jefe varón. Por lo tanto, la línea divisoria entre lo que es estrictamente tamaño familiar y lo que se relaciona con la estructura familiar es algo confusa. El análisis de este apartado se concentra en el tamaño familiar sin indagar en qué medida este responde a distintas configuraciones de los hogares.

La distribución de los hogares del GBA según el número de miembros se reporta en la Tabla 1 del Apéndice a este capítulo. Las familias más comunes son las de 2, 3 y 4 miembros. En 1974 estos hogares representaban un 68% del total, pero esa participación se contrajo 6 puntos aproximadamente a lo largo del período evidenciando un proceso de reducción del tamaño familiar promedio.

Un fenómeno que sobresale es la considerable expansión de los hogares unipersonales – hogares con un solo miembro. La proporción de hogares unipersonales aumentó aproximadamente 6 puntos durante estos años, desde poco menos del 9% en 1974 a más del 15% en 2000. Estos hogares generalmente se conforman por adultos de edad media a avanzada, especialmente de sexo femenino. La fuerte expansión de este tipo de hogares, sin embargo, se debe fundamentalmente a una mayor proporción de adultos jóvenes viviendo solos, especialmente de sexo masculino. Mientras la gran mayoría de mujeres en hogares unipersonales son de edad media a avanzada –en 2000, un 76% tienen más de 55 años–, más de la mitad de los hombres que viven solos en los noventa tienen menos de 55 años –en 2000 un 61%.¹⁰

Mediante las Tablas 2 y 3 –ver Apéndice– se investiga la existencia de diferencias sistemáticas en el tamaño de los hogares por estrato de ingreso. Para la construcción de la Tabla 2 se ordena a los hogares según su ingreso total familiar, reportándose para

¹⁰ Un análisis descriptivo detallado de estos fenómenos puede verse en Marchionni (2001).

cada quintil el número promedio de miembros por hogar. En la Tabla 3 los quintiles de hogares se definen a partir del ingreso equivalente. En la Tabla 2 se observa una relación positiva entre el tamaño familiar y el ingreso total familiar. Sin embargo, cuando se consideran los recursos potencialmente disponibles por individuo –aproximados por el ingreso equivalente–, la evidencia de la Tabla 3 indica que las familias más “pobres” tienen más miembros que las más “ricas”, sugiriendo que en promedio el ingreso total familiar crece más lentamente que el número de integrantes familiares. En cuanto a la evolución de esta distribución se observa que el contraste de tamaños familiares entre los quintiles extremos de la distribución del ingreso equivalente se ha ido acentuando progresivamente como consecuencia de más miembros en las familias de los quintiles más bajos y menos miembros en las de los más altos. Por ejemplo, mientras que en 1974 las familias del primer quintil de ingresos equivalentes tenían un 57% más de integrantes que las del último quintil, en 2000 las primeras duplicaban el tamaño de las segundas.¹¹ Este comportamiento tendría un efecto empobrecedor y desigualador sobre la distribución de ingresos individuales, lo primero como consecuencia del mayor número de miembros en las familias de los quintiles inferiores y lo segundo por la mayor brecha de tamaño familiar entre quintiles extremos.

Es importante hacer una aclaración sobre la discusión anterior. Encontrar que las familias “pobres” tienden a ser más numerosas que las “ricas” podría ser en parte una consecuencia trivial de ordenar a los hogares por ingreso equivalente, lo que no permite distinguir el efecto del ingreso familiar del efecto del tamaño de la familia en ubicar a un hogar en un determinado quintil. En consecuencia, esta evidencia debe interpretarse cuidadosamente y en todo caso ser tomada como un indicio sobre una posible relación entre el tamaño de las familias y el ingreso individual de sus miembros.

La principal fuente de variaciones en el tamaño familiar es la diferencia en el número de hijos entre familias, es decir, la cantidad de hijos del jefe de hogar explica la mayor parte del aumento observado en el número de miembros por familia –ver Tabla 4 en el Apéndice. Nuevamente, en la Tabla 5 se ordenan los hogares de acuerdo al ingreso equivalente, presentándose el número promedio de hijos del jefe de hogar para cada quintil y considerando dos rangos etarios alternativos de los hijos: menores de 5 años y de 18 años. Para ambos grupos y a lo largo de los años se observa una relación inversa entre el número de hijos y el ingreso equivalente. Una posible explicación de este fenómeno sería la existencia de divergencias en el comportamiento reproductivo por quintiles, en línea con lo que Malthus dio en llamar el problema de la población: las

¹¹ El mismo patrón por quintiles se observa para el número de adultos equivalentes por hogar.

familias “pobres” tienen más hijos. Otra posible explicación sería que, dado todo lo demás y con retornos positivos a la experiencia laboral potencial –relacionada directamente con la edad–, individuos más viejos obtienen mayores ingresos al mismo tiempo que tienen menos probabilidad de tener hijos jóvenes. De esta forma, en los quintiles inferiores se produciría una mayor concentración de familias con padres jóvenes que por un lado perciben ingresos más bajos y por el otro cohabitan con un mayor número de hijos de corta edad.

En cuanto a la evolución de la distribución por quintiles del número de hijos, el patrón es menos claro que cuando se estudia el tamaño total de la familia. Desde la década de 1980, pero principalmente en los años noventa, se produjo una caída generalizada del número promedio de hijos menores de 5 años, algo más acentuada para los hogares con mayores ingresos equivalentes. Si bien este no es el indicador adecuado, podemos pensar que sugiere una tendencia decreciente en la fecundidad y que en ese proceso se observan grupos más rezagados, fenómenos que se verifican cuando se utilizan otras fuentes de información como los censos. Como señala Torrado (2003), el segmento poblacional que todavía presenta altos niveles de fecundidad está compuesto no sólo por los estratos más marginados sino también por algunos sectores de las clases media y obrera.¹² Otra vez se observa una relación negativa entre el ingreso equivalente y el número de menores por hogar al considerar la cantidad de hijos menores de 18 años presentes en el hogar.

En cuanto a la evolución de esta estructura, el promedio de hijos menores de 18 años creció para el 20% más pobre de los hogares y en la última década se contrajo fuertemente para los quintiles superiores, provocando que la brecha entre quintiles extremos se amplíe desde 1980. Estos cambios tenderían a aumentar la pobreza y la desigualdad de la distribución del ingreso equivalente.

Como se mencionó antes, la evidencia discutida sobre la relación entre la distribución del ingreso equivalente y el tamaño familiar debe ser considerada con cautela dado que cuando se ordenan los hogares usando esa medida de ingreso se vuelve algo trivial encontrar que los hogares más pobres son también los más numerosos. Dicho de otra forma, no se puede distinguir la causa que lleva a que un hogar sea más pobre que otro: si son los ingresos familiares bajos o un elevado número de miembros. Este problema

¹² Sin duda esta evidencia abre el interrogante acerca de cuáles son las causas de la brecha en el comportamiento reproductivo entre grupos socioeconómicos diferentes. Entre las dos explicaciones más obvias aparecen diferencias en las preferencias y en el acceso a métodos efectivos de control de la natalidad. Torrado (2003) resalta aquí el rol de los servicios de salud pública al no brindar ni la información ni los elementos que garanticen la materialización de las preferencias por familias más pequeñas. El análisis de estos temas, sin embargo, excede los objetivos del presente trabajo.

de circularidad podría evitarse con un análisis condicional que permita controlar los efectos de la composición familiar y aislarlos de los del tamaño. Supóngase que todas las familias están formadas por padre, madre e hijos, y que los hijos son inactivos. Esto hace que la única fuente de variación del tamaño familiar sea el número de hijos y que el ingreso total familiar quede determinado por la suma de los ingresos de la pareja – desde ahora ingreso conyugal. En un caso como este sería posible analizar la relación entre tamaño e ingreso de la familia, aislando los efectos de la composición.¹³ En la Tabla 6 del Apéndice se presenta evidencia empírica acerca de esta relación. A tal efecto se considera una submuestra de familias con padre y madre presentes en el hogar –hogares con núcleo conyugal completo– en los que la madre tiene entre 30 y 40 años de edad. El grupo analizado aquí incluiría familias con hijos jóvenes y, por lo tanto, probablemente inactivos. Para la construcción de la Tabla 6 los hogares de esta submuestra se agrupan en quintiles según el ingreso conyugal. Se observa el mismo patrón que se describía antes. Los hogares más pobres tienen más hijos, y la brecha entre “pobres” y “ricos” se ha ampliado con el tiempo como consecuencia de un mayor número de hijos entre los primeros y un menor número de hijos entre los últimos. Otra vez, este patrón daría lugar a un aumento y profundización de la pobreza, y a un aumento de la desigualdad en la distribución de ingresos.

Además del número de hijos, otro comportamiento capaz de explicar el tamaño de las familias, especialmente de aquéllas con muchos miembros, es el fenómeno de una familia dentro de otra. Cuando dentro de un hogar existe un núcleo conyugal donde ninguno de los cónyuges es el jefe de hogar se dice que hay un núcleo conyugal secundario. Por una deficiencia de la EPH –que recién se soluciona en parte desde 1998– no es posible identificar la existencia de este tipo de núcleos conyugales. Sin embargo, la presencia en el hogar de yernos o nueras del jefe puede ser útil para detectar al menos algunas de estas familias que conviven en el seno del hogar del jefe. En la Tabla 7 se presenta evidencia que indica que la proporción de hogares que incluyen entre sus miembros a hijos políticos del jefe –y con mucha probabilidad, también a su núcleo conyugal– aumenta considerablemente con el tamaño del hogar y que hacia fines de los años noventa esa proporción había crecido considerablemente para los hogares más numerosos.

Finalmente, en la Tabla 8 se presenta una aproximación al efecto distributivo del cambio en el tamaño familiar documentado en este apartado. En el primer panel se presenta el cociente entre el ingreso equivalente promedio del quintil 5 y el quintil 1 y

¹³ Notar que en este caso no hay otros contribuyentes al ingreso familiar que los padres, y por lo tanto mayor número de miembros no implica un potencial mayor número de generadores de ingresos.

en el segundo el coeficiente de Gini del ingreso equivalente promedio entre quintiles. En la columna (i) se reportan los indicadores observados para cada año. Las cifras que aparecen en las demás columnas son el resultado de un ejercicio de simulación. Por ejemplo, en la columna (ii) se calculan ambos indicadores pero suponiendo que el número promedio de adultos equivalentes se mantiene fijo en los niveles de 1974. Es decir, a cada hogar del primer quintil –de ingreso equivalente– se le asigna el número de adultos equivalentes promedio del primer quintil –de ingreso equivalente– que se obtiene a partir de la EPH de 1974. En la columna (iii) se mantiene el tamaño familiar en los valores de 1980 y así sucesivamente.

La tabla muestra resultados interesantes. Como puede verse en la columna (i) –y también fue discutido más arriba– este período se caracteriza por un aumento de la desigualdad que se acelera durante la década de 1990. Estos cambios se suavizan al mantener fijo el tamaño familiar, lo que indica un efecto no menor de los cambios relativos en el número de adultos equivalentes por quintiles sobre la distribución del ingreso. De los más de 9 puntos de aumento en el cociente de ingresos observado entre quintiles extremos en el período 1974-2000, un 66% sería atribuible a los cambios en el tamaño familiar, lo que constituye un porcentaje muy significativo.¹⁴ En el caso del coeficiente de Gini, los cambios en el número de adultos equivalentes explicarían un 73% del aumento observado en la desigualdad del ingreso equivalente entre quintiles entre esos años. Si se considera la década del noventa, la evolución del tamaño de las familias lograría explicar un 61% y 36% del crecimiento de esos indicadores de desigualdad, respectivamente.

El ejercicio detrás de la Tabla 8 es naturalmente muy rudimentario, aunque es indicativo de fenómenos presumiblemente relevantes. Uno de los problemas del ejercicio es que se mantienen constantes los ingresos familiares en la simulación. Es de esperar que el cambio del número de miembros en un hogar genere cambios en las conductas laborales de sus miembros.¹⁵ La dirección de las mismas, sin embargo, no es obvia. Si bien un mayor número de niños implica por un lado la necesidad de mayores recursos, lo cual estimula el trabajo de todos los miembros en condiciones de hacerlo, por otro lado

¹⁴ Considerando las seis simulaciones realizadas, el cociente entre los ingresos del quinto y primer quintil creció entre 1974 y 2000 un promedio de 3.1 puntos –aproximadamente un 34% del cambio observado entre esos años. Este sería el aumento en la desigualdad atribuible a todos los demás factores que afectan la distribución del ingreso equivalente.

¹⁵ En el trabajo de investigación que se presenta en el Capítulo II de esta tesis se aplica una metodología que permite tener en cuenta estos efectos indirectos de los cambios demográficos conjuntamente con los directos.

familias con más niños reducen la probabilidad de la madre de trabajar y la de los hijos de educarse, realimentándose el proceso de empobrecimiento.¹⁶

3.2. Estructura familiar

Los cambios descriptos en el tamaño familiar sugieren una profunda transformación en las estructuras que adoptan las familias. En la Tabla 9 del Apéndice se presenta evidencia sobre esta transformación. Allí se distinguen 3 categorías de hogares: (1) hogares completos –son los que tienen jefe y cónyuge presentes–, (2) hogares monoparentales –en los que está ausente el cónyuge del jefe pero hay hijos solteros presentes–, y (3) hogares no conyugales –constituidos por parientes o no parientes que no forman un núcleo conyugal, a diferencia de los dos primeros grupos formados por hogares conyugales.

Los hogares conyugales completos representaban alrededor del 75% del total de hogares del GBA en 1974, y pese a que a lo largo del período considerado perdieron 10 puntos de participación, siguen siendo la estructura familiar más común. Compensando en parte la contracción de este grupo, la proporción de hogares monoparentales creció llamativamente. En 1974 sólo un 7% de los hogares se constituía a partir de núcleos conyugales con ausencia de uno de los cónyuges. Esa participación creció más de 5 puntos hasta alcanzar casi el 13% en 2000. A pesar de esta expansión de los hogares monoparentales, la participación de los hogares conyugales –completos más monoparentales– cayó en 5 puntos.

Otra dimensión de la estructura familiar que sufrió una profunda transformación en estos años es la relacionada con el género del jefe de hogar –ver Tabla 10 del Apéndice. La gran mayoría de los hogares conyugales son completos y con jefe hombre –siempre por encima del 80% del total de hogares conyugales. En segundo lugar se ubican los hogares monoparentales con jefa mujer. En general los hogares monoparentales tienen por jefe a una mujer y las mujeres jefas lo son generalmente de hogares monoparentales. La proporción –sobre el total de hogares conyugales– de hogares monoparentales con jefas mujeres creció casi 7 puntos en el GBA entre 1974 y 2000. Este cambio es

¹⁶ El efecto estimado de la fecundidad sobre la participación laboral de las mujeres es por lo general negativo, aunque cuando se tiene en cuenta la naturaleza endógena del comportamiento reproductivo el efecto es más pequeño. Ver sobre este punto Angrist y Evans (1998) y para Argentina Cruces y Galiani (2003). Los efectos de distintos tamaños y composiciones familiares sobre la educación de los hijos está documentado para Argentina en Sosa Escudero y Marchionni (2000), donde también se presenta evidencia empírica del efecto negativo que tiene la cantidad de hermanos sobre las probabilidades de educarse.

esencialmente el reflejo de la fuerte expansión de los hogares monoparentales discutida en el párrafo anterior. La proporción de hogares completos con jefe hombre –la estructura familiar tradicional– cayó 10 puntos a lo largo del período. Esta caída se explica en parte por la menor participación de hogares completos en el total de hogares, pero también por una mayor disposición de los hogares completos de reconocer como jefe a la cónyuge mujer –notar el aumento de participación de hogares completos con jefa mujer en el total de hogares conyugales.

Resta explorar el vínculo que podría existir entre la estructura familiar y la distribución de ingresos individuales. En la Tabla 11 se reporta la tasa de incidencia de la pobreza para distintas conformaciones familiares. En 1974, el tipo de estructura familiar con mayor proporción de pobres eran los hogares monoparentales. Aproximadamente un 13% de los individuos que formaban parte de este tipo de hogares tenían un ingreso equivalente inferior a la línea de la pobreza, proporción que más que duplicaba la tasa de incidencia para el total del GBA. Para el mismo año, la tasa de incidencia para los hogares no conyugales rondaba el 8% y el 5% para las familias conyugales completas. Esta estructura se modificó considerablemente en los años siguientes. Durante la década de 1980 y hasta principios de los años noventa, las tasas de pobreza para los hogares completos y monoparentales fueron similares entre sí y levemente superiores a la tasa de incidencia total del aglomerado GBA. Esta estructura se mantuvo, acompañando al crecimiento de la pobreza hasta comienzos de los noventa. En esa década, sin embargo, la tasa de incidencia para los hogares monoparentales creció más rápido, midiendo alrededor de 5 puntos más que para los hogares completos y entre 13 y 15 puntos más que para el total del GBA. La fuerte expansión de los hogares monoparentales durante los años noventa coincide entonces con el crecimiento –en términos absolutos y relativos– de la proporción de individuos pobres viviendo en este tipo de hogares. Si bien este punto requiere un análisis más profundo que se deja para investigaciones futuras, una posible explicación serían las diferentes posibilidades al alcance de uno y otro tipo de familias para amortiguar el impacto sobre los ingresos familiares provocado por el crecimiento del desempleo que caracterizó a estos años.

3.3. Formación de uniones conyugales

En este apartado se presenta evidencia empírica que permite describir el comportamiento en torno a la formación de las uniones conyugales que dan origen a los hogares. Interesa particularmente explorar el grado de asociación entre cónyuges de las características capaces de generar retornos en el mercado laboral para evaluar el impacto de la homogamia sobre la distribución de ingresos. En ese sentido sería

relevante evaluar los niveles de ingresos de cada cónyuge previo a la formación de la pareja. Sin embargo, esta estrategia no es viable por no disponerse de un panel de los individuos que finalmente constituyen núcleos conyugales. Lo que sí puede observarse a partir de la EPH son los ingresos individuales del hombre y la mujer una vez que la familia ha sido formada. Pero dada la interdependencia o simultaneidad en las decisiones laborales de los miembros de una familia, los ingresos ex post no son una buena medida para diagnosticar fenómenos de homogamia. Por ejemplo, para los hogares completos del GBA, Gasparini *et al.* (2004) encuentran que la participación laboral del jefe de hogar tendría un efecto negativo sobre la decisión de participación de los cónyuges que son mayormente mujeres. Luego, es común observar hombres que trabajan cuyas esposas tienen ingresos laborales nulos, lo que no da ninguna información sobre la potencial productividad de estas en el mercado laboral.

Una alternativa que controla en parte este problema de endogeneidad es utilizar la educación en lugar de los ingresos. La educación es un proceso de largo plazo, con un origen –y en muchos casos también un final– previo al de la constitución de la pareja. Por esto la educación de un individuo es más independiente de las características del nuevo hogar que su ingreso. Con retornos a la educación positivos, y dado todo lo demás, un individuo con educación alta tendrá ingresos mayores.

En la EPH se reporta el máximo nivel educativo alcanzado por los individuos: primaria incompleta, primaria completa, secundaria incompleta, secundaria completa, superior o universitaria incompleta y superior o universitaria completa. En las celdas de la Tabla 12 se reporta la proporción de parejas –en hogares completos y sobre el total de estos hogares– formadas por un varón con el nivel educativo dado por la fila y una mujer con el nivel educativo dado por la columna. La información de los seis paneles de la Tabla 12 se resume en la Tabla 13. Aproximadamente en el 50% de los hogares completos el nivel educativo del jefe de hogar coincidía con el de su cónyuge en 1974, porcentaje que se contrajo algo a lo largo del período analizado indicando cierta disminución del grado de homogamia vinculada a la educación formal. Esto implicaría una reducción en la correlación de los niveles educativos dentro de las parejas a lo largo del período, lo que potencialmente podría tener un efecto igualador sobre la distribución de ingresos. Cuando se consideran cónyuges cuya educación difiere como máximo en un nivel, la proporción sube en 35 puntos aproximadamente.¹⁷

¹⁷ Evidencia similar puede encontrarse en Gasparini, Marchionni y Sosa Escudero (2001) y en Torrado (2003).

Por otra parte la homogamia presentaba cierta asimetría por género que ha ido desapareciendo con el tiempo. En 1974, en aproximadamente el 34% de los hogares la mujer tenía menos educación que el hombre, mientras sólo en el 17% de las parejas la mujer era la más educada. En los años noventa ambos porcentajes se encontraban alrededor del 26%.

Por último, se exploraron posibles patrones de formación de uniones conyugales en base a características típicamente inobservables para el analista pero claramente observables para los que participan del mercado matrimonial. Se estimaron ecuaciones de Mincer para los jefes de hogar y sus cónyuges, y se computaron los errores estimados de esas regresiones.¹⁸ En la literatura es usual interpretar a los residuos de estas regresiones como la remuneración correspondiente a los factores inobservables para el analista pero que contribuyen en la determinación de los ingresos laborales. En términos de la ecuación del ingreso equivalente presentada al comienzo de la sección 2, el residuo sería una estimación de $F^k w^k$ donde el k-ésimo factor productivo no es observable. Una denominación habitual que en la literatura se ha dado a este factor es el de *habilidad innata*.

Entre los regresores de las ecuaciones de Mincer se incluyeron variables tradicionales como edad, educación y género, y se realizaron estimaciones separadas para jefes de hogar y cónyuges siguiendo la especificación de Gasparini *et al.* (2004). Se utilizó el método de Heckman Máximo Verosímil para controlar por la posible autoselección en la muestra de individuos ocupados.

Una alta correlación positiva entre cónyuges de los residuos estimados podría interpretarse como evidencia sobre la existencia de un ordenamiento positivo en base a estas características inobservables. Con esta premisa se analizó la correlación entre los residuos de las regresiones de jefes de hogar y sus cónyuges. La correlación en 1980 y 1986 es del 23%, del 19% en 1992 y del 33% en 1998 y 2000.¹⁹ Si bien los valores son inferiores a los que a priori se esperaban, se observa un fuerte incremento de la correlación hacia fines de la última década. Esto ejercería un potencial efecto

¹⁸ Las ecuaciones de Mincer son modelos de precios hedónicos donde se estudian los determinantes de los ingresos laborales individuales en función de variables de capital humano, típicamente la experiencia laboral y la educación, y características del empleo. Su nombre se debe al trabajo pionero de Mincer (1974).

¹⁹ Por falta de disponibilidad de datos relacionados al ingreso laboral, no se hicieron los cálculos correspondientes a 1974.

desigualador, contrario a lo que ocurría al analizar asociación por educación entre cónyuges.²⁰

Comentarios finales

En este trabajo se documentó la evolución de distintas variables demográficas entre 1974 y 2000 en el GBA y se discutieron los posibles efectos de esos cambios sobre la pobreza y la desigualdad en la distribución del ingreso.

Con relación al tamaño de las familias, se observa que a medida que crece el ingreso equivalente, el número total de miembros y de hijos jóvenes del jefe de hogar disminuye. Esta evidencia es consistente con el comportamiento conocido como “problema de la población” a partir del pionero ensayo de Malthus. En cuanto a la evolución de esta estructura, se observa que con el transcurso del tiempo el contraste de tamaños familiares entre los quintiles extremos de la distribución del ingreso equivalente se ha ido acentuando progresivamente como consecuencia de más miembros en las familias de los quintiles más bajos y menos miembros en las de los más altos. Este comportamiento contribuiría a aumentar tanto la pobreza como la desigualdad.

La estructura de las familias también se transformó fuertemente durante este período, destacándose el marcado aumento de la participación de los hogares monoparentales con jefa mujer. En la década de 1990, esta expansión coincidió con la de la proporción de individuos pobres viviendo en este tipo de hogares.

Por último, se exploraron posibles patrones de formación de uniones conyugales en base a características de los cónyuges valoradas en el mercado laboral. Se observa que el nivel educativo entre los esposos presentaba una fuerte correlación, aunque aparentemente el grado de homogamia ha disminuido algo en los últimos años.

La evidencia aquí presentada constituye una primera aproximación al estudio de las relaciones entre los fenómenos demográficos y la distribución del ingreso desde una perspectiva microeconómica. En este trabajo sólo se ha explorado el canal que va desde los cambios demográficos hacia la distribución del ingreso y se han discutido

²⁰ Sin embargo estas cifras deben ser tomadas con extrema cautela. Si bien se controla por el problema de selección al realizar las estimaciones, al calcular las correlaciones sólo se consideran parejas en las que ambos cónyuges están empleados y reportan ingresos positivos, ya que para los demás no se cuenta con una medida de los residuos. Este problema de muestra truncada podría generar sesgos en los coeficientes de correlación.

únicamente los efectos directos o parciales de esos cambios. Entre otros puntos, queda para investigaciones futuras estudiar la determinación simultánea de ambos fenómenos –demográficos y distributivos–, la consideración de los efectos indirectos de los cambios demográficos en la distribución del ingreso y la cuantificación del impacto distributivo de estos cambios. En la investigación que se presenta en el Capítulo II de esta tesis se realiza un avance sobre los dos últimos puntos.

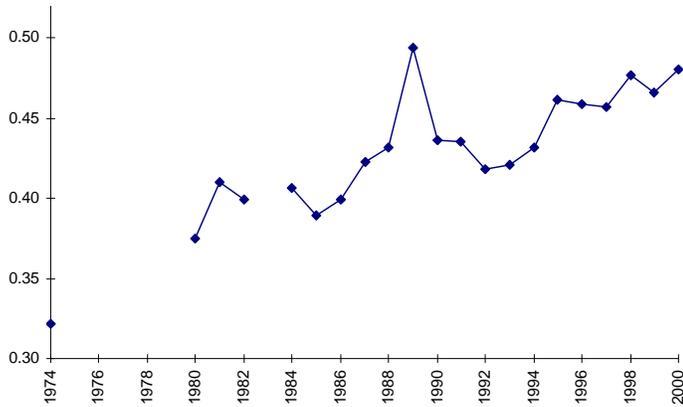
Referencias

- Aiyagari, S. Rao, Jeremy Greenwood and Nezih Guner. (2000). On the State of the Union. *Journal of Political Economy*. 108:2, pp. 213-244.
- Angrist, J. D. and W. N. Evans (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *American Economic Review*. Vol.88, No. 3, pp. 451-477.
- Atkinson, A. Y Bourguignon. F. (eds.). (2000). *Handbook of Income Distribution*. Elsevier Science, B. V.
- Attanasio, O. Y Székely, M. (2003) *The Family in Flux: household Decision-Making in Latin America*. Research Department Technical Series, Latin American Research Network, IDB, Washington, DC.
- Becker, Gary. (1974). A Theory of Marriage. In *Economics of the Family*, edited by T. W. Schultz, Chicago: University of Chicago Press, pp. 299-344.
- CEDLAS. Estadísticas distributivas: www.depeco.econo.unlp.edu.ar/cedlas.
- Cruces, G. and S. Galiani (2003). Generalizing the Causal Effect of Fertility on Female Labor Supply. Mimeo.
- Fernández, R., Nezih Guner y John Knowles. (2001). Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality. *NBER Working Papers* 8580. National Bureau of Economic Research.
- Gasparini, Leonardo, Mariana Marchionni y Walter Sosa Escudero. (2004). Characterization of inequality changes through microeconomic decompositions. The case of Greater Buenos Aires. En Bourguignon, F., F. Ferreira and N. Lustig (eds.). *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*. New York: Oxford University Press.
- Gasparini, Leonardo, Mariana Marchionni y Walter Sosa Escudero. (2001). *Distribución del ingreso en Argentina: perspectivas y efectos sobre el bienestar*. Premio Fulvio Salvador Pagani 2001. Fundación Arcor-Triunfar.
- Greenwood, Jeremy, Nezih Guner and John Knowles. (2003). More on Marriage, Fertility, and the Distribution of Income. *International Economic Review*. Vol. 44, pp. 827-862.
- Knowles, John. (1999). Social Policy, Equilibrium Poverty, and Investment in Children. Mimeo. Philadelphia, Pennsylvania: University of Pennsylvania, Department of Economics.
- Kremer, M. (1997) How much does sorting increase inequality?. *Quarterly Journal of Economics*. Vol. 112, N° 1.
- Lam, D. (1997). Demographic Variables and Income Inequality. En *Handbook of Population and Family Economics*, Vol. 1B, Capítulo 18.
- Malthus, Thomas. (1798). *An Essay on the Principle of Population as it affects the future improvement of society: with remarks on the speculations of Mr. Godwin, M. Condorcet, and other writers*. London. Printed for J. Johnson, in St. Paul's Church-yard.

- Marchionni, Mariana. (2001). Perfil demográfico. Argentina y Gran Buenos Aires. En *Características demográficas y pobreza en la Argentina y El Salvador*. Banco Interamericano de Desarrollo, Diálogo Regional de Política de la Red para la Reducción de la Pobreza y la Protección Social.
- Mincer, J. (1974). *Schooling, experience, and earnings*. National Bureau of Economic Research. Columbia University Press. New York.
- Sosa Escudero, Walter y Mariana Marchionni. (2000). Household Structure, Gender and the Economic Determinants of School Attendance in Argentina. *Poor People in a Rich Country*, Vol. II, Background paper N° 7. World Bank.
- Torrado, Susana. (1998). *Familia y diferenciación social. Cuestiones de método*. Eudeba.
- Torrado, Susana. (2003). *Historia de la Familia en la Argentina Moderna. (1870-2000)*. Ediciones de La Flor.

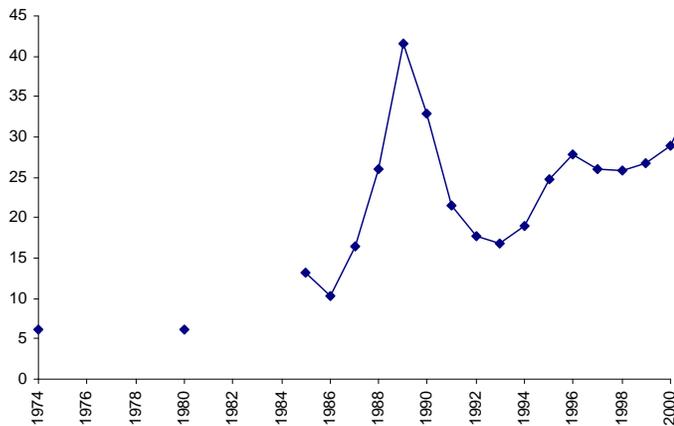
Apéndice al Capítulo I

Figura 1
Coeficiente de Gini del ingreso familiar equivalente.
GBA 1974-2000



Fuente: Gasparini, Marchionni y Sosa Escudero (2001) sobre la base de la EPH.

Figura 2
Tasa de incidencia de la pobreza para la distribución del ingreso familiar equivalente
GBA 1974-2000



Fuente: Gasparini, Marchionni y Sosa Escudero (2001) sobre la base de la EPH.

Tabla 1
Porcentaje de hogares por número de miembros por hogar

número de miembros por hogar	1974	1980	1986	1992	1998	2000
1	8.9	10.3	10.9	13.0	16.1	15.2
2	23.1	23.0	23.3	22.4	23.0	22.0
3	23.0	20.8	20.0	18.6	18.7	19.4
4	22.6	22.0	21.5	20.1	19.9	19.8
5	12.7	12.8	12.7	13.3	11.3	12.1
6	4.9	6.5	6.3	7.0	6.1	5.1
7	2.4	2.3	3.0	2.7	2.5	3.2
8 o más	2.4	2.4	2.3	3.0	2.6	3.3
Total	100.0	100.0	100.0	100.0	100.0	100.0

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Tabla 2
Miembros por hogar por quintil de ingreso total familiar

quintiles de ingreso total familiar	1974	1980	1986	1992	1998	2000
1	2.2	2.1	2.2	2.1	2.5	2.8
2	3.1	3.0	3.0	3.3	3.3	3.3
3	3.5	3.5	3.5	3.7	3.4	3.6
4	3.7	3.8	3.8	3.8	3.5	3.6
5	4.3	4.2	4.0	3.9	3.6	3.6
Total	3.4	3.3	3.3	3.4	3.3	3.4

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Nota: quintiles de hogares

Tabla 3
Miembros por hogar por quintil de ingreso equivalente

quintiles de ingreso equivalente	1974	1980	1986	1992	1998	2000
1	4.1	4.1	4.2	4.4	4.7	4.8
2	3.7	3.5	3.7	3.3	3.6	3.8
3	3.3	3.1	2.9	3.6	3.0	3.2
4	3.0	3.1	3.0	3.0	2.8	2.8
5	2.6	2.9	2.7	2.6	2.4	2.3
Total	3.4	3.3	3.3	3.4	3.3	3.4

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Nota: quintiles de hogares

Tabla 4

Proporción de hijos solteros del jefe en el total de miembros del hogar

número de miembros por hogar	1974	1980	1986	1992	1998	2000
1	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2	5.8%	5.6%	8.6%	9.4%	9.3%	9.2%
3	28.9%	29.6%	31.6%	32.2%	34.7%	33.5%
4	42.2%	44.5%	45.4%	45.6%	45.9%	46.2%
5	47.0%	47.2%	50.9%	51.3%	53.5%	52.4%
6	47.6%	50.0%	52.1%	53.0%	52.1%	52.1%
7	59.2%	48.7%	58.1%	54.8%	55.8%	55.0%
8 o más	57.6%	53.0%	58.2%	54.9%	55.4%	56.0%
Total	31.4%	32.2%	34.7%	35.5%	35.4%	35.6%

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Tabla 5

Hijos del jefe por hogar por quintiles de ingreso equivalente

A. Menores de 5 años

quintiles de ingreso equivalente	1974	1980	1986	1992	1998	2000
1	0.59	0.64	0.61	0.62	0.55	0.53
2	0.47	0.40	0.38	0.35	0.33	0.35
3	0.30	0.23	0.19	0.25	0.19	0.23
4	0.17	0.25	0.21	0.22	0.13	0.16
5	0.16	0.22	0.18	0.19	0.12	0.16
Total	0.34	0.35	0.32	0.32	0.26	0.28

B. Menores de 18 años

quintiles de ingreso equivalente	1974	1980	1986	1992	1998	2000
1	1.74	1.71	1.86	1.97	2.02	2.04
2	1.29	1.13	1.27	1.08	1.13	1.23
3	0.84	0.73	0.72	1.06	0.68	0.70
4	0.56	0.69	0.66	0.69	0.53	0.55
5	0.38	0.56	0.61	0.53	0.36	0.40
Total	0.96	0.96	1.02	1.07	0.95	0.98

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Nota: quintiles de hogares .

Tabla 6
Número de hijos por hogar por quintil de ingreso conyugal

quintiles de ingreso conyugal	1974	1980	1986	1992	1998	2000
1	2.37	2.57	2.56	2.79	3.07	3.00
2	2.20	2.26	2.51	2.67	2.66	2.93
3	2.27	2.18	2.53	2.24	2.31	2.33
4	2.03	2.10	2.15	2.37	1.98	1.95
5	2.02	2.08	2.18	2.06	1.69	1.64
Total	2.18	2.24	2.39	2.42	2.34	2.37

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Nota: quintiles de hogares. Se consideran sólo hogares con madre y padre presentes, cuya madre tiene entre 30 y 40 años de edad.

Tabla 7
Proporción de hogares con yerno o nuera por tamaño de hogar

número de miembros por hogar	1974	1980	1986	1992	1998	2000
1	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%
3	0.4%	0.7%	1.4%	1.2%	0.9%	0.3%
4	3.4%	2.1%	3.1%	2.4%	3.1%	3.7%
5	6.7%	8.1%	6.4%	6.3%	5.7%	5.6%
6	12.0%	9.7%	10.4%	11.5%	16.6%	11.6%
7	7.6%	14.1%	12.2%	14.4%	13.5%	16.7%
8 o más	20.3%	26.0%	19.3%	22.5%	30.5%	32.1%
Total	3.0%	3.2%	3.2%	3.4%	3.5%	3.6%

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Tabla 8
Desigualdad observada y simulada

	cada año	adultos equivalentes de					
		1974	1980	1986	1992	1998	2000
Cociente 5to/1er quintil	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
1974	5.60	5.60	7.39	7.84	8.16	10.43	10.57
1980	7.04	5.20	7.04	7.35	7.66	9.78	9.91
1986	7.98	5.63	7.51	7.98	8.29	10.60	10.74
1992	9.62	6.25	8.33	8.84	9.62	11.75	11.91
1998	13.89	7.14	9.53	10.11	10.52	13.89	13.62
2000	14.85	7.58	10.11	10.72	11.16	14.26	14.85
Gini entre quintiles							
1974	0.314	0.314	0.371	0.380	0.391	0.421	0.415
1980	0.356	0.287	0.355	0.358	0.370	0.401	0.394
1986	0.374	0.300	0.360	0.374	0.379	0.410	0.404
1992	0.410	0.316	0.378	0.386	0.410	0.428	0.421
1998	0.453	0.337	0.395	0.405	0.414	0.453	0.436
2000	0.457	0.351	0.409	0.418	0.427	0.455	0.457

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Nota: quintiles de hogares. La distribución relevante es la del ingreso equivalente.

Tabla 9

Porcentaje de hogares de acuerdo a su estructura

Familias según su estructura	1974	1980	1986	1992	1998	2000
Completas	75.5%	75.2%	73.0%	71.3%	64.9%	65.4%
Monoparentales	7.3%	7.9%	9.8%	10.0%	12.2%	12.8%
No Conyugales	17.2%	17.0%	17.2%	18.7%	22.9%	21.9%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Tabla 10

Porcentaje de hogares por sexo del jefe y estructura del hogar sobre total de hogares conyugales

Sexo del jefe / estructura familiar	1974	1980	1986	1992	1998	2000
hombre / completa	89.9%	89.5%	87.2%	86.1%	81.9%	80.2%
hombre / monoparental	1.3%	1.7%	2.1%	2.3%	2.7%	2.3%
mujer / completa	1.3%	1.0%	1.1%	1.6%	2.3%	3.5%
mujer / monoparental	7.5%	7.7%	9.7%	10.0%	13.2%	14.0%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Tabla 11

Tasa de incidencia de la pobreza por estructura del hogar

Familias según su estructura	1974	1980	1986	1992	1998	2000
Completas	5.2%	6.4%	10.8%	18.6%	27.3%	30.4%
Monoparentales	13.5%	6.8%	10.7%	18.1%	33.7%	34.5%
No Conyugales	7.7%	1.8%	6.2%	9.8%	12.7%	14.7%
Total	6.0%	6.0%	10.4%	17.7%	26.6%	29.4%

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Tabla 12
Asociación de los cónyuges por nivel educativo

A. 1974

nivel educativo del cónyuge varón	nivel educativo de la cónyuge mujer						Total
	Primario incompleto	Primario completo	Secundario incompleto	Secundario completo	Universitario incompleto	Universitario completo	
Primario incompleto	18.6	7.0	0.9	0.1	0.1	0.0	26.7
Primario completo	9.1	22.4	2.8	2.2	0.0	0.1	36.68
Secundario incompleto	1.8	8.6	2.9	1.7	0.5	0.2	15.68
Secundario completo	0.3	3.8	2.0	3.5	0.2	0.5	10.26
Universitario incompleto	0.2	0.9	0.8	1.8	1.2	0.5	5.36
Universitario completo	0.1	0.9	0.5	2.1	0.7	1.2	5.31
Total	30.2	43.5	9.9	11.4	2.6	2.5	100

B. 1980

nivel educativo del cónyuge varón	nivel educativo de la cónyuge mujer						Total
	Primario incompleto	Primario completo	Secundario incompleto	Secundario completo	Universitario incompleto	Universitario completo	
Primario incompleto	14.7	6.6	1.4	0.4	0.0	0.0	23.19
Primario completo	8.5	23.4	3.2	2.3	0.4	0.2	37.85
Secundario incompleto	1.8	7.3	3.5	2.8	0.7	0.4	16.49
Secundario completo	0.2	3.4	1.9	3.9	0.6	0.3	10.26
Universitario incompleto	0.2	1.1	0.7	1.5	1.4	0.5	5.28
Universitario completo	0.2	1.0	0.7	2.9	0.5	1.7	6.93
Total	25.4	42.8	11.4	13.8	3.6	3.0	100

C. 1986

nivel educativo del cónyuge varón	nivel educativo de la cónyuge mujer						Total
	Primario incompleto	Primario completo	Secundario incompleto	Secundario completo	Universitario incompleto	Universitario completo	
Primario incompleto	10.2	6.8	1.3	0.8	0.0	0.0	19
Primario completo	6.5	21.3	4.2	2.7	0.2	0.3	35.04
Secundario incompleto	1.9	7.8	4.4	4.0	0.4	0.5	18.89
Secundario completo	0.7	3.8	1.9	5.4	0.5	0.9	13.06
Universitario incompleto	0.0	0.8	0.7	1.8	1.0	1.0	5.3
Universitario completo	0.0	0.8	0.6	2.7	1.3	3.3	8.7
Total	19.2	41.2	12.9	17.4	3.4	5.9	100

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Nota: Los porcentajes son sobre el total de hogares con núcleo conyugal primario completo.

*Tabla 12 (cont.)
Asociación de los cónyuges por nivel educativo*

D. 1992

nivel educativo del cónyuge varón	nivel educativo de la cónyuge mujer						Total
	Primario incompleto	Primario completo	Secundario incompleto	Secundario completo	Universitario incompleto	Universitario completo	
Primario incompleto	7.0	5.7	0.7	0.2	0.0	0.1	13.7
Primario completo	5.2	22.9	4.5	3.5	0.3	0.4	36.64
Secundario incompleto	1.1	6.6	5.0	3.5	0.7	0.6	17.43
Secundario completo	0.6	2.8	2.0	8.0	1.2	1.8	16.32
Universitario incompleto	0.1	0.6	0.5	2.0	1.3	1.8	6.2
Universitario completo	0.0	0.7	0.6	3.1	1.4	3.9	9.71
Total	13.9	39.3	13.3	20.1	4.8	8.6	100

E. 1998

nivel educativo del cónyuge varón	nivel educativo de la cónyuge mujer						Total
	Primario incompleto	Primario completo	Secundario incompleto	Secundario completo	Universitario incompleto	Universitario completo	
Primario incompleto	5.6	4.4	1.0	0.3	0.0	0.1	11.36
Primario completo	3.6	19.7	6.1	2.9	0.6	0.4	33.17
Secundario incompleto	1.1	6.9	5.7	3.5	1.3	0.9	19.47
Secundario completo	0.2	2.9	2.8	7.4	1.5	2.1	16.87
Universitario incompleto	0.0	0.8	1.1	2.2	1.4	2.0	7.39
Universitario completo	0.0	0.4	0.8	2.5	1.7	6.5	11.74
Total	10.5	35.0	17.5	18.7	6.5	11.8	100

F. 2000

nivel educativo del cónyuge varón	nivel educativo de la cónyuge mujer						Total
	Primario incompleto	Primario completo	Secundario incompleto	Secundario completo	Universitario incompleto	Universitario completo	
Primario incompleto	4.5	4.2	0.7	0.5	0.1	0.1	10.14
Primario completo	4.4	18.7	5.2	2.7	0.8	0.8	32.61
Secundario incompleto	1.0	6.1	6.2	3.9	1.0	1.3	19.53
Secundario completo	0.3	3.4	2.9	7.1	1.1	2.6	17.42
Universitario incompleto	0.1	0.6	0.4	2.0	1.7	2.3	7.06
Universitario completo	0.1	0.6	0.7	2.7	1.8	7.4	13.24
Total	10.3	33.6	16.1	19.0	6.4	14.6	100

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Nota: Los porcentajes son sobre el total de hogares con núcleo conyugal primario completo.

Tabla 13

Porcentaje de hogares según características educativas de los cónyuges

Nivel educativo de los cónyuges	1974	1980	1986	1992	1998	2000
igual nivel	49.8	48.6	45.5	47.9	46.3	45.6
difieren como máximo en un nivel	84.2	82.0	81.2	81.7	81.0	79.5
la mujer tiene menos educación que el hombre	33.5	31.7	31.1	27.2	26.8	27.0
el hombre tiene menos educación que la mujer	16.7	19.7	23.4	24.9	26.9	27.4

Fuente: elaboración propia sobre la base de la EPH, onda octubre.

Nota: Los porcentajes son sobre el total de hogares con núcleo conyugal primario completo.

CAPÍTULO II

EL IMPACTO DISTRIBUTIVO DE LOS CAMBIOS EN LAS DECISIONES DE FECUNDIDAD*

* Otra versión de este trabajo en colaboración con Leonardo Gasparini se encuentra publicada en los Anales de la XXXVII Reunión Anual de la Asociación Argentina de Economía Política, Tucumán 2002. Versiones preliminares de este trabajo fueron presentadas en los Seminarios de Doctorado y de Economía del Departamento de Economía de la Universidad Nacional de La Plata, Seminario de Economía de la Universidad de San Andrés, 3ra. reunión de la Network Inequality and Poverty – Capítulo Argentino (IADB–World Bank–LACEA, Universidad Torcuato Di Tella, 2002) y 8th. Annual Meeting de la Latin American and Caribbean Economic Association (Puebla, México 2003). Parte de esta investigación fue financiada por una beca del Programa Mecovi–Argentina (IADB–World Bank–Cepal).

Introducción

En la investigación realizada para el Capítulo I de esta tesis se presentó evidencia sobre los cambios demográficos que operaron en las últimas décadas en Argentina y se discutieron los posibles efectos de estos cambios sobre la distribución del ingreso.²¹ Para algunas de las variables demográficas consideradas se observa una evolución diferencial por estrato de ingreso. Este es el caso, por ejemplo, del patrón seguido por el número de hijos en el hogar: la brecha en el número de hijos entre los estratos superiores e inferiores de la distribución del ingreso se ha ensanchado sensiblemente. Mientras que el número promedio de hijos menores de 16 años en los hogares del quintil más pobre de la distribución del ingreso familiar equivalente creció de 2.6 a 2.8 entre 1980 y 2000, ese promedio cayó de 1.3 a 0.8 en los hogares del quintil de mayores ingresos.

Los canales a través de los cuales los cambios demográficos pueden afectar distintos rasgos de la distribución del ingreso fueron también discutidos en el Capítulo I. *Caeteris paribus*, un aumento en el número de hijos de las familias pobres y de aquellas marginalmente por encima de la línea de la pobreza incrementa el nivel y grado de la pobreza. Por su parte, un cambio diferencial en las decisiones de fecundidad entre estratos de ingreso, como el señalado en el párrafo anterior, puede incrementar el nivel de desigualdad.²² Sin embargo, para el análisis realizado en ese capítulo sólo se tuvieron en cuenta los efectos directos de los cambios demográficos. Por ejemplo, al evaluar el impacto distributivo de un aumento en el número de hijos en los hogares de menores recursos se mantenía constante el ingreso total familiar del hogar, lo que implica suponer que las decisiones laborales no se modifican ante un cambio de este tipo en la composición familiar.

El objetivo de esta investigación, que constituye el Capítulo II de esta tesis, es evaluar en qué medida los cambios en las decisiones de fecundidad contribuyeron al aumento registrado en la pobreza y la desigualdad durante las últimas dos décadas en el Gran Buenos Aires –Ciudad de Buenos Aires y su conurbano–, considerando tanto los efectos

²¹ Ver también Gasparini (2002) y Marchionni (2002) para una documentación de estos cambios.

²² La expresión “decisiones de fecundidad” se utiliza por comodidad a lo largo de este trabajo. Por un lado el número de hijos en el hogar es un indicador imperfecto de fecundidad. Por el otro, y pese a que se habla de “decisiones”, para este trabajo resulta irrelevante si la fecundidad ocurre como consecuencia de una decisión racional y libre de la pareja o no.

directos como los indirectos de esos cambios demográficos. Con este fin se aplica la metodología de microsimulaciones sugerida por Bourguignon, Ferreira y Lustig (2004). El aporte metodológico del presente artículo consiste en incorporar un tratamiento más riguroso de las decisiones de fecundidad y su impacto sobre la distribución del ingreso en la literatura de microsimulaciones.

La estimación de los parámetros que gobiernan las decisiones de fecundidad y las respuestas en términos de participación laboral a cambios en la cantidad de hijos de las familias, constituyen el insumo básico para llevar a cabo las microsimulaciones. Usando esos parámetros puede generarse una distribución contrafáctica de ingresos, que es la que habría existido en el año t si las decisiones de fecundidad y los consecuentes ajustes en la participación laboral se hubieran regido por los parámetros estimados para el año t' .

La especificación adoptada en este trabajo supone que el número de hijos sigue un proceso de Poisson, y sus parámetros se estiman consistentemente usando el modelo de regresión de Poisson. En cuanto a los ingresos laborales horarios y las horas trabajadas, se supone que se determinan simultáneamente en un modelo de equilibrio del mercado laboral. Para la estimación de los parámetros de este modelo se usa una estrategia consistente: la ecuación de ingresos se estima por el método de Heckman máximo verosímil y la de horas trabajadas por el método Tobit.

Con los parámetros así estimados se procede con las simulaciones. Sobre la distribución de ingresos simulada en el año t usando parámetros del año t' se computan distintos indicadores de pobreza y desigualdad y se los compara con los indicadores observados para el año t . La diferencia entre ambos conjuntos de indicadores es señal del impacto distributivo de los cambios en las decisiones de fecundidad.

Los resultados de aplicar esta metodología a datos de la Encuesta Permanente de Hogares del Gran Buenos Aires sugieren que durante la década del ochenta el aumento en el número de hijos de las familias de ingresos bajos y medios-bajos incrementó en una magnitud no despreciable la pobreza y la desigualdad. Este efecto operó fundamentalmente de manera directa, a través del aumento en el número de miembros del hogar, y en menor medida de manera indirecta, a través de la reducción de las horas trabajadas de miembros adultos como consecuencia del mayor número de hijos en el hogar. Durante los noventa el tamaño de las familias cayó moderadamente y de forma aproximadamente uniforme entre estratos de ingreso, lo que generó un efecto ligeramente reductor de la pobreza, sin alterar significativamente la desigualdad. Finalmente, la relación negativa entre las horas trabajadas por el cónyuge del hogar y el

número de hijos se fue debilitando con el tiempo, lo cual parece haber contribuido a una reducción de la pobreza y la desigualdad.

El resto del trabajo se organiza de la siguiente manera. En la sección 2 se discuten los canales de influencia de las características demográficas sobre la distribución del ingreso y se presenta evidencia preliminar. En la sección 3 se describe la metodología de descomposiciones microeconómicas y en la 4 la estrategia de estimación. Los resultados se presentan y analizan en la sección 5, mientras que en la sección 6 se desarrollan las conclusiones.

Fecundidad y distribución del ingreso: vínculos y evidencia preliminar

Tanto la distribución del ingreso como la estructura demográfica se han modificado sensiblemente en el Gran Buenos Aires (GBA) en las últimas dos décadas.²³ En esta sección se presenta evidencia sobre ambos fenómenos y se discuten brevemente las posibles vinculaciones.

Sea D_t la distribución del ingreso familiar equivalente (y) entre los N individuos de una economía en el momento t

$$(1) \quad D_t = \{y_{1t}, y_{2t}, \dots, y_{Nt}\} \text{ con}$$

$$(2) \quad y_{it} = \frac{Y_{ht}}{A_{ht}} \quad \forall i \in h$$

donde Y_h denota el ingreso total de la familia h y A_h el total de adultos equivalentes en ese hogar.²⁴

Las Figuras 1 y 2 del Apéndice al Capítulo I –ver página 31– muestran la evolución de dos dimensiones de la distribución del ingreso familiar equivalente para el GBA desde mediados de los años setenta. Las figuras reflejan fenómenos conocidos: tanto la

²³ Existe evidencia que sugiere que estos cambios también se manifestaron en el ámbito nacional. Ver por ejemplo Marchionni (2002) y Gasparini, Marchionni y Sosa Escudero (2001). Este trabajo, sin embargo, se restringe al Gran Buenos Aires dada una mayor disponibilidad temporal de información de hogares.

²⁴ Las palabras hogar y familia son utilizadas indistintamente en este trabajo.

pobreza como la desigualdad han aumentado sensiblemente en el GBA durante las dos últimas décadas.²⁵

El ingreso familiar equivalente del individuo i , utilizado como *proxy* de su nivel de bienestar, se determina en el hogar del cual es miembro, y es afectado por las decisiones de fecundidad en ese hogar. El aumento en el número de hijos incrementa el denominador en (2) y por ende, manteniendo constante el resto, reduce el ingreso equivalente de todos los miembros del hogar.²⁶ Adicionalmente, el número de hijos influye sobre las decisiones de participación de algunos miembros del hogar, típicamente de las madres, modificando las horas trabajadas o la probabilidad de ser económicamente activas, lo cual afecta el numerador de (2).

Los patrones de fecundidad no son homogéneos por grupos de ingreso. Desde Malthus, el llamado problema de la población hace referencia esencialmente al mayor número de hijos en las familias pobres.²⁷ Este fenómeno está reflejado en las Tablas 1 a 4 –ver Apéndice–, que muestran el número promedio de hijos menores de 16 años para aquellos hogares con jefes entre 25 y 45 años clasificados según el ingreso familiar equivalente, el ingreso total de la pareja, el ingreso laboral horario del jefe y el nivel educativo del jefe. Estas tablas sugieren no sólo que las familias más carenciadas –en términos de ingresos, salarios y nivel educativo– tienden a tener más hijos, sino que la brecha en el patrón de comportamiento reproductivo entre estratos sociales se ha ensanchado sensiblemente en las últimas dos décadas. Este ensanchamiento fue acelerado entre 1980 y 1992 cuando la fecundidad de los estratos bajos y medios aumentó y la de los estratos altos se redujo, y más moderado entre 1992 y 2000 cuando la tasa de fecundidad de toda la población se redujo, aunque a tasas menores en las familias de menores recursos.²⁸

La evidencia presentada hasta el momento indica que en las últimas décadas el GBA experimentó cambios significativos en la distribución del ingreso y en las decisiones de fecundidad. ¿Existe algún vínculo entre ambos fenómenos? Naturalmente, esta es una

²⁵ La información sobre evolución de la distribución del ingreso en el Gran Buenos Aires es abundante. Ver Altimir *et al.* (2000), Gasparini *et al.* (2001) y Llach y Montoya (1999), entre otros.

²⁶ Si bien el ingreso equivalente cae, no está claro cómo es afectado el bienestar de todos los miembros del hogar cuando aumenta el número de hijos. En particular, es probable que la utilidad de quienes tomaron la decisión de fecundidad no caiga. Esta discusión, si bien muy relevante, es de difícil solución y excede el marco de los típicos estudios de distribución del ingreso, al que pertenece este trabajo.

²⁷ Ver Anand y Morduch (1999).

²⁸ Como se señaló en el Capítulo I de esta tesis, el número de hijos en el hogar es sólo una medida aproximada de la fecundidad, pero la más cercana que puede definirse a partir de la EPH.

pregunta difícil de contestar tanto teórica como empíricamente. Las decisiones de fecundidad y las decisiones laborales que determinan los ingresos posiblemente se determinan simultáneamente, por lo que únicamente la estimación de un modelo de equilibrio general puede echar luz sobre el vínculo profundo entre esos fenómenos.

Este artículo sigue un camino más modesto. La pregunta que intenta responder es cuáles habrían sido los cambios distributivos entre dos momentos del tiempo si sólo el comportamiento reproductivo hubiera cambiado en ese período. En particular, manteniendo el resto constante, se analiza el impacto que el cambio en los parámetros que guían la decisión sobre el número de hijos puede haber tenido en la distribución del ingreso familiar equivalente por dos vías: el cambio en el número de adultos equivalentes de cada familia y el cambio en las decisiones laborales del jefe de hogar y su cónyuge.

Este ejercicio debe ser interpretado de la siguiente manera. Si los cambios observados en los patrones de fecundidad son el resultado de cambios en parámetros que sólo afectan esa decisión, entonces el ejercicio provee una estimación del impacto distributivo del cambio en esos parámetros. Por el contrario, si los cambios observados en el comportamiento reproductivo son el resultado de cambios en parámetros que afectan tanto a la fecundidad como a otras variables determinantes del ingreso equivalente –por ejemplo, la conformación familiar, las decisiones educativas y laborales, etc.–, entonces el ejercicio debe interpretarse como indicativo del signo y magnitud del efecto que tales cambios en los parámetros tienen sobre la distribución del ingreso únicamente a través del canal demográfico. En este último caso es necesario tener presente que ese impacto distributivo puede ser reforzado o contrarrestado a través de otros canales que el ejercicio ignora.

Un nuevo hijo, en principio, incrementa el denominador de (2). Sin embargo, cuando este hijo crece puede aportar al numerador de (2) al incorporarse a la fuerza laboral y compartir sus ingresos con el resto de la familia.²⁹ De hecho, la decisión de tener un hijo puede estar afectada por las perspectivas de ese aporte a los ingresos familiares. Idealmente, para incorporar estas consideraciones el análisis debería centrarse en el impacto de las decisiones reproductivas sobre la distribución del ingreso permanente, en lugar del ingreso corriente. Desafortunadamente, no sólo se enfrentan dificultades analíticas y conceptuales para ese tipo de análisis, sino fundamentalmente informativas: casi todas las encuestas de hogares, entre ellas la EPH, sólo relevan el ingreso corriente de las personas.

²⁹ Alternativamente puede abandonar el hogar y transferir fondos a sus padres.

Descomposiciones microeconómicas

En esta sección se detalla la metodología de descomposiciones microeconómicas o microsimulaciones bosquejada en la sección anterior. Esta metodología, iniciada por Blinder (1973) y Oaxaca (1973) para estudiar diferencias en las medias de dos distribuciones, fue ampliada por Juhn, Murphy y Pierce (1993) para considerar diferencias en toda la distribución, no necesariamente en las medias. Más recientemente, Bourguignon, Ferreira y Lustig (2004) generalizaron el enfoque, permitiendo su aplicación a formas funcionales diversas, no necesariamente lineales.³⁰ Gasparini, Marchionni y Sosa Escudero (GMS) (2004) aplican una variante de esta metodología al caso argentino.³¹ El presente trabajo utiliza el marco de GMS, extendiéndolo para considerar las decisiones de fecundidad y su impacto sobre la pobreza y la desigualdad. El aporte metodológico de este artículo es incorporar una modelización más cuidadosa de las decisiones reproductivas y su impacto distributivo en la literatura de microsimulaciones, donde estas decisiones eran o bien ignoradas, o tratadas marginalmente.³²

De acuerdo a la ecuación (2), el ingreso familiar equivalente del individuo i en el momento t (y_{it}) es el cociente entre el ingreso total familiar del hogar y el número de integrantes –en adultos equivalentes. El ingreso total familiar Y_{ht} es la suma de los ingresos laborales y no laborales –supráíndices L y NL respectivamente.

$$(3) \quad Y_{ht} = \sum_{j \in h} (Y_{jt}^L + Y_{jt}^{NL})$$

³⁰ Otras aplicaciones de esta metodología pueden encontrarse en Bourguignon, Fournier y Gurgand (2001) y Bourguignon, Ferreira y Lustig (eds.) (2004).

³¹ Altimir *et al.* (2000) y Menéndez y González Rozada (2001) realizan ejercicios con metodologías semejantes aplicados a la Argentina.

³² Una excepción es Ferreira y Leite (2002), quienes incluyen decisiones de fecundidad mediante un modelo logístico multinomial. Estos autores están interesados fundamentalmente en los efectos distributivos de los cambios en la estructura educativa de la población. Estos cambios repercuten en las decisiones de fecundidad, y el cambio en el número de hijos afecta al ingreso equivalente por las vías directa e indirecta descriptas más arriba.

Se asume que los ingresos no laborales están dados exógenamente. El ingreso laboral es el producto del ingreso laboral horario (w) por el número de horas trabajadas (L).

$$(4) \quad Y_{it}^L = w_{it} \cdot L_{it}$$

Tanto w como L dependen de características observables (X_{it}) y no observables (ε_{it}) para el analista, y de los vectores de parámetros β_t y λ_t que determinan la forma en que las características observables afectan a los ingresos horarios y a las horas trabajadas, respectivamente. De esta forma, el ingreso laboral del individuo i puede representarse por la ecuación (5).

$$(5) \quad Y_{it}^L = Y(X_{it}, \varepsilon_{it}; \beta_t, \lambda_t)$$

El tamaño de la familia –en adultos equivalentes– es determinado por ciertas características observables del hogar (Z_{ht}), ciertas características inobservables (e_{ht}) y el vector de parámetros η_t que describe las decisiones de fecundidad. Formalmente,

$$(6) \quad A_{ht} = A(Z_{ht}, e_{ht}; \eta_t)$$

Combinando (2), (3), (5) y (6) se obtiene

$$(7) \quad y_{it} = F\left(\left\{X_{jt}\right\}_{j \in h}, Z_{ht}, \left\{\varepsilon_{jt}\right\}_{j \in h}, e_{ht}, \left\{Y_{jt}^{NL}\right\}_{j \in h}; \beta_t, \lambda_t, \eta_t\right) \quad \forall i \in h$$

donde X y Z pueden tener variables en común. Cambiando todos o algunos de los argumentos de la ecuación (7) es posible simular ingresos. Por ejemplo, la ecuación (8) representa el ingreso equivalente que un individuo del hogar h habría tenido en el momento t si los parámetros que determinan las decisiones de fecundidad (η) hubieran sido los del momento t' y todo lo demás quedara fijo en sus valores al momento t .

$$(8) \quad y_{it}(\eta_{t'}) = F\left(\left\{X_{jt}\right\}_{j \in h}, Z_{ht}, \left\{\varepsilon_{jt}\right\}_{j \in h}, e_{ht}, \left\{Y_{jt}^{NL}\right\}_{j \in h}; \beta_t, \lambda_t, \eta_{t'}\right) \quad \forall i \in h$$

La distribución entre individuos del ingreso familiar equivalente simulado a través de este ejercicio –distribución simulada– viene dada por la ecuación (9).

$$(9) \quad D_t(\eta_{t'}) = \{y_{1t}(\eta_{t'}), \dots, y_{N_t}(\eta_{t'})\}$$

El objetivo del ejercicio propuesto es comparar las distribuciones (1) y (9) en función de algún indicador distributivo, como por ejemplo algún indicador de pobreza o de desigualdad.

$$(10) \quad I(D_t(\eta_{t'})) - I(D_t)$$

donde $I(D_t)$ es un indicador de alguna dimensión de la distribución D_t . Supongamos que $I(\cdot)$ representa la tasa de incidencia de la pobreza. En ese caso la diferencia entre la pobreza simulada y la real se interpreta como la contribución de un cambio de los parámetros η entre t y t' al verdadero cambio observado en la pobreza entre esos dos momentos. Esta diferencia –y la correspondiente al medir desigualdad– será referida como *efecto tamaño-directo*. Este efecto capta el impacto distributivo de las consecuencias que el cambio en las decisiones reproductivas tiene sobre el ingreso familiar equivalente a través de la vía más directa: la modificación del denominador en (2).

Naturalmente, cambios en η pueden repercutir también en el numerador de (2). Específicamente, las horas trabajadas por un individuo pueden depender del número de hijos en el hogar.³³ Formalmente, puede considerarse que las horas trabajadas vienen dadas por

$$(11) \quad L_{it} = l(X_{Lit}^{-H}, H_{it}, \varepsilon_{it}^L; \lambda_t^{-H}, \lambda_t^H)$$

donde X_L incluye sólo las variables de X en la ecuación (5) que determinan las horas trabajadas y no necesariamente el ingreso horario, H es el número de hijos del individuo i que viven en el hogar y el supraíndice $-H$ indica que el vector excluye la variable H – en el caso de X – o el coeficiente λ^H asociado con la misma –en el caso de λ . ε^L son características inobservables que contribuyen a la determinación de las horas trabajadas.

³³ Este es típicamente el caso de las mujeres (ver Killingsworth y Heckman (1986) y Blundell y MaCurdy (1999) para resúmenes de la evidencia sobre este punto). Se asume que el salario horario no depende del número de hijos.

Reemplazando (11) en (7) se tiene que el ingreso familiar equivalente del individuo i – y, por supuesto, de los demás individuos que integran su hogar– puede escribirse como

$$(12) \quad y_{it} = F\left(\left\{X_{jt}^{-H}, H_{jt}\right\}_{j \in h}, Z_{ht}, \left\{\varepsilon_{jt}\right\}_{j \in h}, e_{ht}, \left\{Y_{jt}^{NL}\right\}_{j \in h}; \beta_t, \lambda_t^{-H}, \lambda_t^H, \eta_t\right) \quad \forall i \in h$$

La ecuación (13) representa el ingreso equivalente que un individuo del hogar h habría tenido en el momento t si las decisiones de participación en el mercado laboral hubieran cambiado en respuesta al número simulado de hijos usando los parámetros de fecundidad en t' , manteniendo todo lo demás fijo en sus valores al momento t , inclusive el número de adultos equivalentes del hogar.

$$(13) \quad y_{it}(H_{it'}) = F\left(\left\{X_{jt}^{-H}, H_{jt'}\right\}_{j \in h}, Z_{ht}, \left\{\varepsilon_{jt}\right\}_{j \in h}, e_{ht}, \left\{Y_{jt}^{NL}\right\}_{j \in h}; \beta_t, \lambda_t^{-H}, \lambda_{t'}^H, \eta_t\right) \quad \forall i \in h$$

Con esta simulación puede calcularse el impacto distributivo del cambio en las horas trabajadas en respuesta al cambio en el número de hijos entre los años t y t' , manteniendo el resto fijo en sus valores al momento t , incluido el tamaño familiar en el denominador de la ecuación de ingresos equivalentes. Este efecto será referido como *efecto tamaño-horas*.

Se realiza una tercera descomposición, que surge de computar la diferencia entre la distribución observada en t y la simulada para ese año si los ingresos se determinaran según la ecuación (14), donde se reemplaza λ^H en t por su valor en t' .

$$(14) \quad y_{it}(\lambda_{t'}^H) = F\left(\left\{X_{jt}^{-H}, H_{jt}\right\}_{j \in h}, Z_{ht}, \left\{\varepsilon_{jt}\right\}_{j \in h}, e_{ht}, \left\{Y_{jt}^{NL}\right\}_{j \in h}; \beta_t, \lambda_t^{-H}, \lambda_{t'}^H, \eta_t\right) \quad \forall i \in h$$

El parámetro λ^H mide el efecto de un cambio en el número de hijos sobre las horas trabajadas por un individuo. Nótese que a diferencia de los dos efectos anteriores, los cambios en λ^H no reflejan cambios de corte netamente demográfico, sino más bien cambios en la forma de tomar decisiones laborales en función de variables demográficas, o en la forma como el mercado laboral reacciona frente a individuos con ciertas características demográficas. El impacto distributivo de estos cambios en los parámetros de la ecuación de horas trabajadas es denominado *efecto parámetros-horas*.

A partir de estas tres descomposiciones se computan dos efectos acumulativos. El *efecto tamaño-total* acumula los efectos *tamaño-directo* y *tamaño-horas*, agregando el impacto

sobre la distribución del ingreso del cambio en los parámetros que guían la decisión de fecundidad. El efecto total incluye adicionalmente al *efecto parámetros-horas*.

Resumiendo,

- i. *Efecto tamaño-directo*: cambios en el tamaño familiar con ingreso familiar constante.
- ii. *Efecto tamaño-horas*: las horas trabajadas responden al cambio simulado en el número de hijos pero se mantiene constante el número de adultos equivalentes del hogar.
- iii. *Efecto parámetros-horas*: cambia el coeficiente del número de hijos en la ecuación de horas trabajadas (λ^H).
- iv. *Efecto tamaño-total*: (i)+(ii)
- v. *Efecto total*: (i)+(ii)+(iii)

Estrategia de estimación

Para realizar las simulaciones descritas en la sección anterior es necesario estimar modelos de fecundidad y de ingresos laborales. En esta sección se detalla la metodología para estimar los parámetros que guían la decisión sobre el número de hijos, la decisión de participación en el mercado laboral y la determinación de los salarios horarios.

4.1. Decisiones de fecundidad

Según la teoría económica las decisiones de fecundidad surgen de un proceso de optimización de los padres mediante el cual evalúan los beneficios de tener hijos contra los costos de oportunidad asociados a su crianza. La evaluación de estos beneficios y costos depende de ciertas características propias del hogar y de cada uno de los cónyuges. Las decisiones de fecundidad, entonces, pueden ser representadas por la ecuación (15).

$$(15) \quad H_h = H(Z_h, e_h; \eta)$$

donde, igual que antes, H es el número de hijos, Z es un vector de características observables del hogar y e incluye todas las características del hogar no observables para el analista pero que influyen sobre el comportamiento reproductivo de esa familia.

Para la estimación de este modelo se supone que

$$(16) \quad H_h \sim \text{Poisson}(\mu_h) \quad \text{con } \mu_h = E(H_h|Z_h) = \exp(Z_h \eta)$$

De esta forma,

$$(17) \quad \text{Prob}(H_h = H_0) = \frac{\exp(-Z_h \eta)(Z_h \eta)^{H_0}}{H_0!} \quad \text{para } H_0 = 0, 1, 2, \dots$$

Este es el modelo de regresión de Poisson, a partir del cual es posible estimar consistentemente los parámetros η por el método de máxima verosimilitud. Puede demostrarse que la propiedad de consistencia de los estimadores máximo verosímiles de η se mantiene siempre que la verdadera distribución sea cualquiera de la familia de lineales exponenciales –a la que pertenece la distribución de Poisson– y que la media condicional en (16) esté correctamente especificada.³⁴ Los estimadores máximo verosímiles de η –que por simplicidad también son denotados por η –, se utilizan para las simulaciones descritas en la sección 3.

El número de hijos del hogar h en el año t simulado usando los parámetros estimados para el año t' viene dado por (18).

$$(18) \quad H_{ht}(\eta_{t'}) = F_{t'|Z_{ht}}^{-1} \circ F_{t|Z_{ht}}(H_{ht})$$

donde $F_{t|Z_{ht}}(\cdot)$ es la función que da el ranking relativo de su argumento en la distribución del año t condicional en las características observables Z_{ht} . En este caso $F_{t|Z_{ht}}(\cdot)$ es la función de probabilidad acumulada de una variable con distribución Poisson con

³⁴ Un supuesto algo más realista, y que se ha utilizado en la literatura, es que los hijos siguen una distribución binomial negativa –ver por ejemplo, Rao, Mazumdar, Waller y Li (1973), Hamdan (1975), y Wooldridge (2000). Sin embargo, en el presente trabajo se emplea el modelo de Poisson por dos motivos: 1) como ya se dijo más arriba, los estimadores siguen siendo consistentes y 2) para los hogares completos –que representan, dependiendo el año, entre el 63% y el 82% del total de hogares en la muestra– no puede rechazarse la hipótesis nula que la distribución de hijos por hogar es Poisson versus una Binomial Negativa –tipo NB2– sobre la base de un test de razón de verosimilitud. Para más detalles sobre el modelo de Poisson y el Binomial Negativo, ver Cameron y Trivedi (1998). También se empleó un modelo de decisiones secuenciales donde el jefe –y su cónyuge si lo hubiere– deciden entre tener o no tener hijos, y luego el número de hijos. Los resultados obtenidos no difieren significativamente de los que surgen con la especificación de Poisson.

parámetro $\exp(Z_h \eta)$. De esta manera, se obtiene para cada hogar en el año t , el número simulado de hijos usando los parámetros que gobernaban el comportamiento reproductivo en el año t' .

El mecanismo (18) resulta apropiado para simular hijos cuando se considera la relevancia de los inobservables en la determinación de las decisiones de fecundidad. Dos hogares con las mismas características observables Z que tienen un número de hijos diferente, claramente difieren en sus características inobservables e , aunque la predicción del número esperado de hijos para ambos hogares sería la misma e igual a $\exp(Z\eta)$. Como el objetivo es simular cambios en el número de hijos como consecuencia únicamente de cambios en los parámetros η , es necesario mantener fijos los factores inobservables. Con ese fin, se caracteriza a cada hogar según el cuantil que ocupa en la distribución de hijos del año t . Si el número de hijos del hogar h en el momento t lo ubica en el cuantil q_h , entonces $F_{t|Z_{ht}}(H_{ht})=q_h$. El número simulado de hijos para el hogar h usando los parámetros del momento t' es aquél que lo ubicaría en el cuantil q_h pero de la distribución de hijos con parámetros $\eta_{t'}$.³⁵

Dada la definición de ingreso familiar equivalente, es necesario transformar el número de hijos en adultos equivalentes, lo cual idealmente implica considerar su estructura etaria y de género. Dada la dificultad en incluir estas dimensiones en el análisis, se optó por aplicar un ajuste sencillo. Específicamente, se transforma proporcionalmente al número de hijos simulado multiplicándolo por el cociente entre el número de adultos equivalentes que representan los hijos y el número de hijos del hogar h en el año t .³⁶ El supuesto implícito detrás de este ajuste es que si bien el número de hijos cambia al realizar la simulación, su estructura etaria y de género se mantiene igual a la de los hijos observados en el momento t .

Una vez simulado el número de hijos en adultos equivalentes, se construye el ingreso equivalente del individuo i del hogar h en el año t , simulado mediante los parámetros de fecundidad del año t' . Para su cálculo se emplea la ecuación (19).

³⁵ Si bien el supuesto sobre la distribución de los hijos en (16) no es crucial para la consistencia de los estimadores, aquí se hace evidente que no es irrelevante a efectos de la simulación. La distribución binomial negativa tiene mayor varianza que una de Poisson con la misma media. Sin embargo, como se mencionó anteriormente, para la gran mayoría de los hogares –entre un 63% y un 82%– no puede rechazarse la hipótesis nula de equidispersión en la distribución de los hijos.

³⁶ Para los hogares que no tenían hijos en el año t pero tienen hijos según la simulación, el factor de proporcionalidad utilizado es el cociente entre el número medio de hijos en adultos equivalentes y el número medio de hijos en toda la muestra.

$$(19) \quad y_{it}(\eta_{it}) = \frac{Y_{ht}}{A_{ht} - HA_{ht} + HA_{ht}(\eta_{it})} \quad \forall i \in h \text{ en } t$$

donde en el denominador, al número total de adultos equivalentes del hogar en el año t , se le resta el número de adultos equivalentes representados por los hijos que realmente hay en esa familia en el momento t (HA_{ht}) y se le suma el número simulado ($HA_{ht}(\eta_{it})$) con el procedimiento antes descrito. El ingreso equivalente simulado por (19), que es la forma explícita que adopta la ecuación (8), se utiliza en el cálculo del *efecto tamaño-directo*.

4.2. Decisiones laborales

Para computar los efectos *tamaño-horas* y *parámetros-horas* es necesario estimar los parámetros λ de la ecuación de horas trabajadas (11), para lo que se adopta la misma especificación y estrategia que en GMS. El ingreso laboral del individuo i en el momento t es el producto de su salario horario (w) y de las horas trabajadas (L) –ver ecuación (4). Ambas variables se determinan a partir del siguiente modelo que describe, en su forma reducida, el equilibrio en el mercado laboral.

$$(20) \quad \ln w_{it}^* = X_{1it} \beta_t + \varepsilon_{it}^W$$

$$(21) \quad L_{it}^* = X_{2it} \lambda_t^{-H} + H_{it} \lambda_t^H + \varepsilon_{it}^L$$

$$\text{con } w_{it} = w_{it}^* \quad \text{y} \quad L_{it} = L_{it}^* \quad \text{si} \quad L_{it}^* > 0$$

$$w_{it} = 0 \quad \text{y} \quad L_{it} = 0 \quad \text{si} \quad L_{it}^* \leq 0$$

$$(\varepsilon_{it}^W, \varepsilon_{it}^L) \sim N(0,0, \sigma_W^2, \sigma_L^2, \rho)$$

donde w_{it}^* y L_{it}^* son variables latentes, no observables. Siguiendo a GMS, esta especificación corresponde a un *Tobit Type III* según la clasificación de Amemiya (1985). Es posible estimar consistentemente los parámetros del modelo (20)-(21) de la siguiente forma:³⁷ (i) estimar (20) por el método de Heckman máximo verosímil,

³⁷ Esta estrategia, si bien consistente, no es eficiente. En relación a este punto caben dos aclaraciones, que oportunamente señalan GMS: (a) esta alternativa es computacionalmente más simple que la aplicación de

usando una versión censurada de (21) como ecuación de selección, donde en lugar de las horas trabajadas se usa un indicador que sólo diferencia individuos que trabajan de los que no, y (ii) estimar (21) por el método Tobit. De esta forma se obtienen estimaciones de los parámetros λ^H y λ^L de la ecuación de horas trabajadas, necesarios para realizar las simulaciones para el cómputo de los efectos *tamaño-horas* y *parámetros-horas*.

Aunque nuestros estimadores de λ^H podrían ser sesgados a causa de la potencial endogeneidad de la fecundidad en la ecuación de horas trabajadas, las consecuencias de este sesgo en el *efecto tamaño-horas* calculado sobre la pobreza serían despreciables ya que la evidencia disponible indica que este sesgo es pequeño para mujeres en los quintiles más bajos de la distribución de ingresos.³⁸ Angrist y Evans (1998) construyen un estimador de variables instrumentales explotando las preferencias de los padres por una mezcla de géneros en su descendencia. Usando datos para USA encuentran que los estimadores de mínimos cuadrados tienden a sobrestimar el efecto de la fecundidad sobre la participación laboral de mujeres con al menos dos hijos, pero que el sesgo es menos severo para mujeres pobres y con menores niveles educativos. Además, Cruces y Galiani (2003) muestran que estas preferencias por la mezcla de géneros también están presentes en Argentina y que la mezcla de géneros resulta ser también un instrumento válido cuando se usan datos de este país. Sin embargo, ellos no encuentran diferencias significativas entre las estimaciones obtenidas usando ese instrumento y las tradicionales de mínimos cuadrados.³⁹

Del modelo (20)-(21) también se obtienen estimaciones de los errores para los individuos que trabajan $-\varepsilon^W$ y ε^L , que se interpretan como el efecto conjunto de los inobservables sobre el ingreso horario y las horas trabajadas, respectivamente, y de σ_w^2 , σ_L^2 y ρ . Para calcular los efectos *tamaño-horas* y *parámetros-horas* es necesario simular las horas trabajadas de todos los individuos, para lo que se requieren estimaciones de los coeficientes λ y de los errores ε^L de la ecuación (21). Para los individuos que en t no trabajan, las horas trabajadas observadas son nulas, por lo que no se cuenta para ellos con estimaciones del término de error. Usando el mismo

un procedimiento de información completa y (b) la pérdida de eficiencia no es necesariamente grande para un tamaño de muestra dado.

³⁸ Ya que el *efecto parámetros-horas* evalúa el impacto distributivo de cambios en λ^H , las consecuencias del sesgo son menos claras.

³⁹ Dado que el sesgo parece ser más grande para mujeres altamente educadas, podrían surgir resultados interesantes sobre la desigualdad en la distribución de ingresos que no son captados por la especificación aquí considerada.

procedimiento que GMS se simulan para estos individuos los inobservables ε^W y ε^L muestreando aleatoriamente pares de errores de la distribución implícita en el modelo (20)-(21), consistentes con la decisión de participación observada en t .

Los parámetros β estimados de la ecuación (20) y la estimación de ε^W obtenida por el procedimiento anterior se usan para imputar los ingresos horarios de los individuos que no trabajaban en t pero que al emplear los parámetros de t' tienen horas trabajadas simuladas positivas.

Resultados

En esta sección se presentan los resultados de aplicar la metodología descrita en las dos secciones anteriores con el fin de evaluar el posible vínculo entre cambios en las decisiones de fecundidad y distribución del ingreso.

La estimación del modelo de fecundidad de la ecuación (17) se realiza separadamente para los hogares completos –organizados alrededor de un jefe y su cónyuge– e incompletos o monoparentales –el jefe del hogar no tiene cónyuge. En ambos casos la variable dependiente se define como el número de hijos menores de 16 años presentes en el hogar. En un intento de reflejar lo mejor posible las verdaderas decisiones de fecundidad la muestra se limita a los hogares con jefe mayor de 25 años y menor de 45. Las Tablas 5 y 6 reportan los resultados de la estimación de los modelos de fecundidad para los años 1980, 1986, 1992, 1998 y 2000 en el GBA.⁴⁰ Es interesante notar el incremento en el tiempo del valor absoluto de los coeficientes de las *dummies* que indican nivel educativo superior completo de los padres. Este incremento es indicio del ensanchamiento en las brechas de fecundidad entre estratos de ingreso.

Las ecuaciones de ingresos horarios (20) y de horas trabajadas (21) se estiman para jefes y cónyuges en forma separada. Por simplicidad, se supone que las decisiones de participación de los cónyuges dependen del ingreso de los jefes, pero las de éstos son

⁴⁰ A continuación se describen las abreviaturas de los nombres de las variables usadas en las tablas 5 a 10. *edad* y *edad2* son la edad y su cuadrado en años; *pric*, *seci*, *secc*, *supi* y *supc* son *dummies* que toman el valor 1 si el individuo tiene máximo nivel educativo primario completo, secundario incompleto, secundario completo, superior incompleto y superior completo, respectivamente. Para la ecuación de fecundidad de hogares completos, las variables de edad y categorías educativas se distinguen con la letra “m” o “p” según se trate de la madre o del padre. *jefamu* vale 1 si el hogar tiene jefa mujer; *divorc*, *casado* y *viudo* son indicadores de estado civil (el primero es una abreviatura para divorciado o separado, el segundo agrupa individuos casados o unidos); *hombre* vale 1 para los varones; *hijos16* es el número de hijos menores de 16 años y *mujerh16* es igual a *hijos16* si la jefa es mujer y cero en caso contrario; *concorre* vale 1 para los individuos que reciben actualmente educación formal en cualquier nivel; *ijj* es el ingreso individual total del jefe de hogar.

independientes de los ingresos o participación de los primeros.⁴¹ Ambas ecuaciones de horas trabajadas incluyen como variable independiente al número de hijos menores de 16 años. Se supone que las decisiones de participación de los demás miembros del hogar no dependen del número de hijos del jefe. En las Tablas 7 y 8 se reportan las estimaciones del modelo de ingresos horarios para jefes y cónyuges, respectivamente, mientras que en las Tablas 9 y 10 se presentan los resultados de las estimaciones de los modelos de horas trabajadas. Hay dos fenómenos que resulta particularmente interesante notar y que se desprenden de las Tablas 9 y 10. El primero, es que las madres son las que ajustan su participación laboral como respuesta a un cambio en el número de hijos. El número de hijos sólo es significativo en la ecuación de horas de los jefes cuando interactúa con un indicador de madre –variable *mujerh16*–, mientras que en la ecuación para cónyuges, que son prácticamente todas mujeres, es significativo por sí sólo. El segundo fenómeno tiene que ver con la magnitud de la respuesta de las horas trabajadas de las madres ante un aumento en el número de hijos. El vínculo entre las decisiones de horas trabajadas de los cónyuges y la presencia de hijos menores de 16 años en el hogar se ha debilitado considerablemente entre 1986 y 2000, mientras que para las madres jefas de hogar se ha fortalecido desde 1980.

Con la estimación de estos parámetros es posible implementar la metodología descrita en la sección 3. El impacto de los cambios demográficos es evaluado sobre dos dimensiones de la distribución del ingreso: la pobreza y la desigualdad. La pobreza se mide mediante el indicador FGT –Foster, Greer y Thorbecke (1984)– con tres parámetros alternativos: 0, 1 y 2. En el primer caso el FGT coincide con la tasa de incidencia de la pobreza, en el segundo con el indicador de brecha de la pobreza y en el tercero con la llamada profundidad de la pobreza. En todos los ejercicios se utiliza la línea oficial de la pobreza computada por el INDEC. Para captar desigualdad se utilizan varios indicadores alternativos usuales en la literatura.⁴² Todos los indicadores de pobreza y desigualdad se calculan sobre la distribución entre individuos del ingreso familiar equivalente.

La Tabla 11 muestra los resultados de las micros simulaciones para el GBA en el período 1980-2000 en términos de la tasa de incidencia de la pobreza. La tabla se lee de la manera siguiente. Entre 1980 y 1986 en el GBA la tasa de incidencia de la pobreza creció 4.74 puntos porcentuales en la submuestra de hogares con jefe entre 25 y 45 años con respuestas completas y consistentes de salarios y horas trabajadas. El promedio del

⁴¹ Esta especificación secuencial es similar a la utilizada en Bourguignon *et al.* (2001).

⁴² Aquí sólo se presentan los resultados para el coeficiente de Gini. Los resultados para otros indicadores de desigualdad son en general semejantes y están disponibles para el lector interesado.

efecto tamaño-directo cambiando el año base entre 1980 y 1986 es 1.35.⁴³ A grandes rasgos, esto implica que si sólo los parámetros que guían las decisiones de fecundidad hubieran cambiado entre 1980 y 1986, y si el consiguiente cambio en el número de hijos sólo hubiera modificado el denominador de (2) no afectando los ingresos totales familiares, entonces la tasa de pobreza se habría incrementado en 1.35 puntos porcentuales en el período. El impacto del cambio en las decisiones de fecundidad a través de la modificación de las decisiones de participación laboral –*efecto tamaño-horas*– también tiene signo positivo, aunque su valor promedio es muy cercano a cero.⁴⁴ Algo superior resulta ser el *efecto parámetros-horas*: el cambio en los parámetros que vinculan las decisiones de horas trabajadas con el número de hijos menores dan cuenta, entre 1980 y 1986, de un aumento en la tasa de incidencia de la pobreza de medio punto porcentual. El *efecto tamaño-total*, que resume los efectos que operan a través del cambio en el número de hijos, tiene un valor de 1.40, lo cual representa un 30% del cambio observado en la pobreza en el período, mientras que el *efecto total* –que incluye el efecto *parámetros-horas*– da cuenta de un 40% del cambio real.

A lo largo de las últimas dos décadas la pobreza creció en la submuestra analizada en más de 23 puntos. Este notable incremento fue alimentado por múltiples factores económicos y sociales. Los factores demográficos no parecen haber sido ajenos a este proceso. El ejercicio muestra que el incremento en el tamaño familiar de los hogares cercanos a la línea de la pobreza puede dar cuenta, por sí sólo, de alrededor de 3 de los 23 puntos de incremento de la tasa de pobreza entre 1980 y 2000. Este impacto del cambio en las decisiones de fecundidad se dio esencialmente en el período 1980-1992 y perdió fuerza en los noventa. Mientras que el *efecto tamaño-directo* da cuenta de casi el 30% del incremento de la pobreza entre 1980 y 1986, ese valor cae al 8% entre 1986 y 1992 y al 3% entre 1992 y 2000.

El *efecto tamaño-horas* ha sido positivo hasta 1992 y negativo en el resto del período analizado, aunque los valores estimados son menores comparados con los del *efecto tamaño-directo*, y posiblemente no significativos. Los valores son en general mayores, en valor absoluto, para el *efecto parámetros-horas*. Este efecto resulta desde 1986 reductor de la pobreza.

⁴³ Se reporta el promedio ya que las simulaciones no son independientes del año base elegido –*path dependence*. Intuitivamente, esto ocurre porque los mismos cambios en los coeficientes se aplican a dos poblaciones diferentes, con distribuciones distintas de las características observables e inobservables.

⁴⁴ Idealmente, se debería practicar un ejercicio de test de hipótesis para determinar la significatividad estadística de cada resultado. Este ejercicio implica varias complicaciones por lo que se deja para una próxima etapa de la investigación.

En las Tablas 12 y 13 se reportan los resultados de computar la brecha de la pobreza –el FGT(1)– y la profundidad de la misma –el FGT(2). Los resultados cualitativos son aproximadamente semejantes a los de la Tabla 11.

El impacto del cambio en las decisiones demográficas sobre la desigualdad es resumido en la Tabla 14 donde se reportan cambios en el coeficiente de Gini. El *efecto tamaño-directo* resulta positivo durante todo el período, aunque su relevancia difiere marcadamente entre subperíodos. Mientras que entre 1980 y 1992 este efecto representa casi el 40% del cambio observado en el Gini, entre 1992 y 2000 representa sólo un 4%. Los efectos *tamaño-horas* y *parámetros-horas* resultan menores, particularmente en los años ochenta.

Una interpretación plausible de los resultados presentados es la siguiente. Durante la década de 1980 el número de hijos aumentó en las familias de ingresos bajos y medios-bajos y se redujo en el resto. Este fenómeno incrementó tanto el número de pobres como el grado de severidad de la pobreza, y generó un aumento en la desigualdad. Los resultados sugieren que estos efectos, si bien no han sido los dominantes, dan cuenta de una no insignificante fracción del incremento en la pobreza y la desigualdad experimentado entre 1980 y 1992. El mayor número de hijos en familias de ingresos bajos y medios-bajos puede haber conducido a algunas madres a dejar el trabajo o trabajar menos horas. Este efecto, sin embargo, parece haber sido menor: su impacto sobre la pobreza y la desigualdad resulta en general muy bajo.

Durante los noventa el tamaño de las familias pobres y cercanas a la línea de la pobreza de la submuestra analizada cayó moderadamente en el GBA –salvo entre 1998 y 2000–, lo que generó un efecto reductor de la pobreza aunque de magnitud menor. Dado que la caída en el tamaño familiar fue generalizada a través de los estratos de ingresos, la desigualdad no se vio mayormente alterada. La reducción en el número de hijos puede haber alentado esencialmente a algunas madres a trabajar, lo cual implicó una caída en la pobreza. El efecto, sin embargo, no parece haber sido cuantitativamente relevante ni sobre la pobreza ni sobre la desigualdad.

A partir de mediados de los ochenta la relación negativa entre el número de hijos y la decisión de participación laboral de los cónyuges se fue debilitando gradualmente. La salida al mercado laboral de madres con hijos, que años antes hubieran permanecido inactivas, naturalmente puede afectar la distribución. Los resultados de las tablas 11 a 14 indican que este movimiento hacia el mercado laboral redujo la pobreza y la desigualdad, aunque nuevamente el impacto parece haber sido cuantitativamente menor.

Comentarios finales

A lo largo de las últimas dos décadas, la pobreza y la desigualdad aumentaron fuertemente en la Argentina. Al mismo tiempo, operaron profundas transformaciones demográficas en el seno de las familias. Este trabajo estudia empíricamente la relación entre los cambios en las decisiones de fecundidad, y en las decisiones laborales ligadas al número de hijos, y los cambios en la distribución del ingreso familiar equivalente.

El trabajo concluye que si bien los fenómenos demográficos no parecen ocupar un lugar central en la explicación de los cambios distributivos ocurridos en la Argentina, han contribuido de manera no despreciable en esas transformaciones distributivas. El incremento del tamaño familiar en hogares de ingresos bajos y medios-bajos contribuyó de manera significativa al aumento de la pobreza y la desigualdad en los ochenta. La reversión de esta tendencia para los hogares de la muestra durante los años noventa tuvo un efecto despreciable sobre ambos. El debilitamiento en la conexión entre horas trabajadas y número de hijos para las mujeres cónyuges tuvo en el período analizado un efecto reductor de la pobreza y la desigualdad, pero también de magnitud menor.

Con este trabajo se ha incorporado explícitamente la discusión de cuestiones demográficas a la literatura de microdescomposiciones con el objetivo de cuantificar el impacto distributivo de cambios en las decisiones de fecundidad. Estas fueron modeladas y estimadas independientemente de las ecuaciones de participación laboral. Queda para futuras investigaciones la utilización de especificaciones más realistas que permitan la determinación conjunta de la fecundidad y la situación de empleo. En la investigación desarrollada en el Capítulo III de esta tesis se propone un método de estimación en dos etapas que permitiría estimar en forma sencilla un modelo de participación laboral con fecundidad endógena.

Finalmente, otras dimensiones demográficas podrían incluirse al análisis de descomposiciones microeconómicas. Resultaría particularmente interesante cuantificar el efecto sobre la pobreza y la desigualdad de los cambios en los patrones de formación de uniones conyugales discutidos en el Capítulo I de esta tesis.

Referencias

- Altimir, O., Beccaria, L. y González Rozada, M. (2000). La evolución de la distribución del ingreso familiar en la Argentina. *Serie de Estudios en Finanzas Públicas* 3, La Plata.
- Amemiya, T. (1985). *Advanced Econometrics*, Harvard University Press.
- Anand, S. y Morduch, J. (1999). Poverty and the population problem. En De Santis y Livi-Bacci (eds.). *Population and Poverty in the Developing World*, Oxford University Press.
- Angrist, J. D. and W. N. Evans (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *American Economic Review*. Vol.88, No. 3, pp. 451-477.
- Blinder, A. (1973). Wage discrimination: reduced form and structural estimate. *The Journal of Human Resources* VIII (4), 436-453.
- Blundell, R. y MaCurdy, T. (1999). Labor supply: a review of alternative approaches. En Ashenfelter and Card (eds.) *Handbok of Labor Economics*. North Holland, vol. 3A, cap. 27.
- Bourguignon, F., Ferreira, F. y Lustig, N. (eds.) (2004). *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*. New York: Oxford University Press.
- Bourguignon, F., Fournier, M. y Gurgand, M. (2001). Fast development with a stable income distribution: Taiwan, 1979-94. *Review of Income and Wealth* 47 (2), 139-163.
- Cameron, A. C. y Trivedi, P. K. (1998). *Regression Analysis of Count Data*. Cambridge University Press.
- Cruces, G. and S. Galiani (2003). Generalizing the Causal Effect of Fertility on Female Labor Supply. Mimeo.
- Ferreira, F. y Leite, P. (2002). Educational Expansion and Income Distribution. A Micro-Simulation for Ceará. Mimeo.
- Foster, J, Greer, J. y Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica* 52.
- Gasparini, L. (2002). Perfil demográfico de pobres y nuevos pobres. En *Características demográficas y pobreza en la Argentina*. Red para la Reducción de la Pobreza y la Protección Social, Banco Interamericano de Desarrollo.
- Gasparini, L., Marchionni, M. y W. Sosa Escudero (2001). *La Distribución del Ingreso en la Argentina. Perspectivas y Efectos sobre el Bienestar*. Premio Fulvio Salvador Pagani 2001, Fundación Arcor-Triunfar.
- Gasparini, Leonardo, Mariana Marchionni y Walter Sosa Escudero. (2004). Characterization of inequality changes through microeconomic decompositions. The case of Greater Buenos Aires. En Bourguignon, F., F. Ferreira and N. Lustig (eds.). *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*. New York: Oxford University Press.

- Hamdan, M. A.. (1975). Correlation between the numbers of two types of children when the family size distribution is zero-truncated negative binomial. *Biometrics*. 31, pp. 765-766.
- Juhn, C., Murphy, K. y Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy* 101, June.
- Killingsworth, M. y Heckman, J. (1986). Female labor supply: a survey. En Ashenfelter and Layard (eds.) *Handbok of Labor Economics*. North Holland, vol. 1, cap. 2
- Llach, J. y Montoya, S. (1999). *En pos de la equidad. La pobreza y la distribución del ingreso en el área Metropolitana del Buenos Aires: diagnóstico y alternativas de política*. IERAL, Buenos Aires.
- Marchionni, M. (2001). Perfil demográfico de la Argentina. En *Características demográficas y pobreza en la Argentina*. Red para la Reducción de la Pobreza y la Protección Social, Banco Interamericano de Desarrollo.
- Menéndez, A. y González Rozada, M. (2001). Unemployment and inequality. Mimeo.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor market. *International Economic Review* 14 (3), 693-709.
- Rao, B. R., Mazumdar, S., Waller, J., y Li, C.C. (1973). Correlation between the numbers of two types of children in a family. *Biometrics* 29, pp. 271-279.
- Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

Apéndice al Capítulo II

Tabla 1

Promedio de hijos menores de 16 años en la muestra por quintiles de hogares de la distribución del ingreso familiar equivalente

Quintiles	1980	1986	1992	1998	2000
1	2.62	2.79	2.91	2.94	2.79
2	2.05	2.12	2.25	1.85	1.97
3	1.57	1.68	1.86	1.48	1.54
4	1.21	1.41	1.31	1.03	1.01
5	1.26	1.08	0.90	0.71	0.76
Total	1.74	1.82	1.85	1.60	1.61

Fuente: Elaboración propia en base a datos de la EPH.
La muestra incluye hogares con jefes entre 25 y 45 años.

Tabla 2

*Promedio de hijos menores de 16 años por quintiles de hogares de la distribución del ingreso total de los cónyuges**

Quintiles	1980	1986	1992	1998	2000
1	1.74	1.93	2.07	2.07	1.92
2	1.69	1.81	2.15	1.79	1.89
3	1.91	1.94	1.89	1.55	1.60
4	1.74	1.74	1.69	1.40	1.42
5	1.63	1.65	1.44	1.20	1.23
Total	1.74	1.82	1.85	1.60	1.61

Fuente: Elaboración propia en base a datos de la EPH. La muestra incluye hogares con jefes entre 25 y 45 años.
* En el caso de los hogares monoparentales, es el ingreso total del jefe de hogar.

Tabla 3

Promedio de hijos menores de 16 años

por quintiles de hogares de la distribución del ingreso laboral horario del jefe de hogar

Quintiles	1980	1986	1992	1998	2000
1	1.67	1.89	2.07	2.00	1.81
2	1.87	1.83	2.10	1.80	1.88
3	1.61	1.83	1.82	1.61	1.58
4	1.82	1.75	1.77	1.46	1.59
5	1.74	1.79	1.47	1.15	1.21
Total	1.74	1.82	1.85	1.60	1.61

Fuente: Elaboración propia en base a datos de la EPH. La muestra incluye hogares con jefes entre 25 y 45 años.

Tabla 4

Promedio de hijos menores de 16 años por educación del jefe de hogar

Educación del jefe	1980	1986	1992	1998	2000
Primaria incompleta	2.20	2.38	2.84	2.41	2.70
Primaria completa	1.81	1.87	2.08	1.96	2.02
Secundaria incompleta	1.45	1.82	1.94	1.78	1.78
Secundaria completa	1.65	1.68	1.63	1.35	1.38
Superior incompleta	1.46	1.44	1.19	1.02	1.04
Superior completa	1.55	1.44	1.23	0.96	1.01
Total	1.74	1.81	1.85	1.60	1.61

Fuente: Elaboración propia en base a datos de la EPH. La muestra incluye hogares con jefes entre 25 y 45 años.

Tabla 5

Estimación del modelo de fecundidad. Hogares completos con jefe entre 25 y 45 años.

Modelo de Regresión de Poisson.

Variable dependiente: número de hijos menores de 16 años.

	1980	1986	1992	1998	2000
Edadm	0.114 (3.11)**	0.191 (5.39)**	0.146 (3.51)**	0.223 (5.50)**	0.188 (5.22)**
edadm2	-0.002 (3.44)**	-0.003 (5.57)**	-0.002 (3.69)**	-0.003 (5.87)**	-0.003 (5.48)**
Edadp	0.256 (3.85)**	0.151 (2.58)**	0.181 (2.59)**	0.102 (1.45)	0.171 (2.53)*
edadp2	-0.003 (3.66)**	-0.002 (2.44)*	-0.002 (2.53)*	-0.001 (1.20)	-0.002 (2.37)*
Pricm	-0.080 (1.16)	-0.172 (2.50)*	-0.121 (1.38)	-0.107 (1.00)	-0.131 (1.27)
Secim	-0.317 (3.46)**	-0.230 (2.78)**	-0.261 (2.46)*	-0.073 (0.65)	-0.162 (1.45)
Seccm	-0.291 (2.78)**	-0.312 (3.67)**	-0.343 (3.14)**	-0.314 (2.46)*	-0.419 (3.44)**
Supim	-0.224 (1.51)	-0.494 (3.33)**	-0.478 (2.96)**	-0.426 (2.70)**	-0.410 (2.79)**
Supcm	-0.466 (2.56)*	-0.322 (2.54)*	-0.380 (2.72)**	-0.430 (2.85)**	-0.502 (3.71)**
Pricp	-0.132 (1.91)	-0.209 (3.08)**	-0.194 (2.20)*	-0.282 (2.93)**	-0.183 (1.85)
Secip	-0.225 (2.50)*	-0.193 (2.57)*	-0.238 (2.34)*	-0.315 (3.04)**	-0.272 (2.52)*
Seccp	-0.091 (0.80)	-0.223 (2.52)*	-0.301 (2.63)**	-0.464 (3.88)**	-0.395 (3.35)**
Supip	-0.236 (1.75)	-0.288 (2.58)*	-0.480 (3.20)**	-0.565 (3.72)**	-0.422 (2.78)**
Supcp	-0.068 (0.49)	-0.177 (1.49)	-0.382 (2.63)**	-0.481 (3.18)**	-0.391 (2.74)**
Jefamu	0.031 (0.09)	0.021 (0.08)	-0.146 (0.60)	-0.184 (1.26)	-0.479 (2.74)**
Constante	-5.511 (5.16)**	-4.797 (5.06)**	-4.400 (3.85)**	-4.488 (3.87)**	-5.210 (4.71)**
Observac.	834	1042	698	804	810

Valor absoluto de los estadísticos z entre paréntesis. * significativo a un nivel del 5%; ** significativo a un nivel de 1%

Tabla 6
Estimación del modelo de fecundidad.
Hogares monoparentales con jefe entre 25 y 45 años
Modelo de Regresión de Poisson.
Variable dependiente: número de hijos menores de 16 años.

	1980	1986	1992	1998	2000
Edad	0.172 (0.73)	0.380 (1.97)*	0.434 (2.01)*	0.609 (4.03)**	0.554 (3.33)**
edad2	-0.002 (0.74)	-0.005 (1.97)*	-0.007 (2.21)*	-0.009 (4.24)**	-0.008 (3.51)**
Pric	0.086 (0.34)	-0.211 (1.04)	-0.546 (2.44)*	-0.085 (0.44)	-0.845 (3.23)**
Seci	0.007 (0.02)	-0.384 (1.73)	-0.758 (2.48)*	-0.388 (1.94)	-1.255 (4.40)**
Secc	-0.017 (0.05)	-0.881 (3.21)**	-0.854 (3.38)**	-0.945 (3.58)**	-1.441 (4.95)**
Supi	-0.247 (0.52)	-0.719 (2.23)*	-0.959 (2.56)*	-0.989 (3.60)**	-1.879 (5.64)**
Supc	-0.459 (0.73)	-1.366 (3.69)**	-1.276 (4.27)**	-1.402 (5.08)**	-2.059 (6.50)**
Divorc	1.128 (4.19)**	1.066 (3.38)**	0.906 (4.26)**	0.981 (6.02)**	1.090 (6.89)**
Viudo	1.146 (3.76)**	0.890 (5.34)**	1.038 (3.53)**	1.098 (5.34)**	0.853 (3.57)**
Jefamu	1.093 (3.40)**	0.686 (3.40)**	1.370 (4.72)**	1.383 (6.91)**	1.057 (5.55)**
Constante	-4.838 (1.16)	-7.405 (2.14)*	-7.933 (2.05)*	-11.366 (4.25)**	-9.539 (3.22)**
Observac.	148	202	171	292	300

Valor absoluto de los estadísticos z entre paréntesis.* significativo a un nivel del 5%; ** significativo a un nivel de 1%

Tabla 7

Estimación de la ecuación de ingresos horarios. Jefes de hogar entre 25 y 45 años

Método de Heckman Máximo Verosímil.

Variable dependiente: logaritmo del ingreso laboral horario.

	1980	1986	1992	1998	2000
Edad	0.053 (1.27)	0.095 (2.53)*	0.056 (1.21)	-0.008 (0.19)	0.063 (1.46)
edad2	-0.001 (1.06)	-0.001 (2.25)*	-0.001 (0.96)	0.000 (0.53)	-0.001 (1.19)
Pric	0.202 (3.90)**	0.149 (2.62)**	0.206 (2.78)**	0.244 (2.90)**	0.029 (0.31)
Seci	0.432 (7.33)**	0.395 (6.93)**	0.317 (3.98)**	0.427 (4.95)**	0.173 (1.78)
Secc	0.797 (11.09)**	0.535 (8.16)**	0.646 (7.99)**	0.770 (8.63)**	0.518 (5.31)**
Supi	0.942 (11.45)**	0.911 (12.62)**	0.990 (10.59)**	0.989 (10.02)**	0.816 (7.67)**
Supc	1.395 (16.34)**	1.234 (16.59)**	1.238 (14.10)**	1.487 (16.03)**	1.259 (12.27)**
Hombre	0.182 (1.82)	0.211 (3.00)**	0.183 (3.02)**	0.183 (3.34)**	0.145 (2.61)**
Constant	7.162 (9.75)**	-1.954 (2.87)**	-0.671 (0.83)	0.301 (0.39)	-0.681 (0.89)
Observac.	982	1244	869	1096	1110
Obs.Cens.	27	51	39	87	95
Log Lik.	-891.31	-1172.06	-823.22	-1153.84	-1179.95
Rho	-0.11	-0.24	0.75	0.46	0.42
LR rho=0	0.78	0.72	0.01	0.05	0.08

Valor absoluto de los estadísticos z entre paréntesis.* significativo a un nivel del 5%; ** significativo a un nivel de 1%

LR rho=0 indica el p-valor para un test de Razón de Verosimilitud de independencia entre las ecuaciones de salarios y selección.

Tabla 8

Estimación de la ecuación de ingresos horarios. Cónyuges.

Hogares con jefes entre 25 y 45 años.

Método de Heckman Máximo Verosímil.

Variable dependiente: logaritmo del ingreso laboral horario.

	1980	1986	1992	1998	2000
Edad	0.126 (2.76)**	0.033 (0.83)	-0.008 (0.19)	0.074 (1.99)*	0.074 (1.73)
edad2	-0.002 (2.56)*	-0.000 (0.71)	0.000 (0.34)	-0.001 (1.73)	-0.001 (1.73)
Pric	-0.053 (0.45)	0.076 (0.68)	-0.238 (1.65)	0.235 (1.38)	-0.254 (1.07)
Seci	0.114 (0.85)	0.234 (1.73)	-0.085 (0.55)	0.401 (2.42)*	-0.210 (0.87)
Secc	0.683 (5.42)**	0.539 (4.73)**	0.268 (1.78)	0.662 (3.92)**	0.122 (0.49)
Supi	0.964 (5.79)**	0.722 (4.70)**	0.532 (2.96)**	0.811 (4.27)**	0.585 (2.23)*
Supc	1.470 (7.74)**	1.023 (7.10)**	0.579 (3.35)**	1.250 (6.54)**	0.991 (3.56)**
Hombre	0.629 (1.80)	-0.043 (0.15)	0.310 (1.53)	0.448 (2.91)**	0.178 (1.09)
Constant	5.631 (6.81)**	-0.771 (1.14)	0.948 (1.22)	-0.999 (1.48)	-0.518 (0.66)
Observac.	834	1042	698	804	810
Obs.Cens.	605	716	432	488	493
Log Lik.	-620.70	-868.98	-612.54	-725.71	-741.41
Rho	0.53	0.27	0.16	0.31	0.27
LR rho=0	0.03	0.24	0.43	0.18	0.24

Valor absoluto de los estadísticos z entre paréntesis. * significativo a un nivel del 5%; ** significativo a un nivel de 1%

LR rho=0 indica el p-valor para un test de Razón de Verosimilitud de independencia entre las ecuaciones de salarios y selección.

Tabla 9

Estimación de la ecuación de horas trabajadas. Jefes de hogar entre 25 y 45 años.

Método Tobit.

Variable dependiente: horas trabajadas

	1980	1986	1992	1998	2000
Edad	-0.790 (0.62)	1.558 (1.37)	1.191 (0.81)	1.827 (1.15)	1.653 (1.08)
edad2	0.011 (0.63)	-0.022 (1.35)	-0.020 (0.95)	-0.028 (1.26)	-0.026 (1.21)
Pric	1.220 (0.79)	3.241 (1.98)*	2.865 (1.24)	9.074 (3.05)**	12.525 (4.02)**
Seci	0.091 (0.05)	0.139 (0.08)	5.578 (2.22)*	8.429 (2.75)**	16.164 (5.02)**
Secc	0.306 (0.14)	3.272 (1.69)	2.611 (1.02)	9.210 (2.88)**	14.032 (4.28)**
Supi	-0.924 (0.35)	1.724 (0.73)	4.259 (1.36)	8.515 (2.28)*	15.651 (4.11)**
Supc	-6.072 (2.38)*	-0.731 (0.35)	4.179 (1.49)	11.086 (3.35)**	18.573 (5.47)**
Hombre	17.487 (5.36)**	7.499 (3.00)**	7.195 (2.32)*	9.639 (3.48)**	9.881 (3.81)**
Casado	-0.790 (0.33)	-5.713 (4.11)**	1.118 (0.45)	2.750 (1.26)	6.542 (3.15)**
hijos16	0.240 (0.51)	0.192 (0.48)	0.479 (0.91)	0.967 (1.65)	0.825 (1.50)
mujerh16	-1.938 (1.30)	-4.228 (3.11)**	-4.231 (3.53)**	-5.524 (4.72)**	-6.059 (4.62)**
Concurre	-3.745 (0.84)	-9.688 (2.83)**	-12.808 (3.27)**	-11.212 (2.84)**	-6.028 (1.67)
Constante	47.578 (2.14)*	15.029 (0.75)	20.852 (0.81)	-1.217 (0.04)	-7.751 (0.29)
Observac.	982	1244	869	1096	1110

Valor absoluto de los estadísticos t entre paréntesis. * significativo a un nivel del 5%; ** significativo a un nivel de 1%

Tabla 10

Estimación de la ecuación de horas trabajadas. Cónyuges.

Hogares con jefe entre 25 y 45 años.

Método Tobit.

Variable dependiente: horas trabajadas.

	1980	1986	1992	1998	2000
Edad	5.625 (2.35)*	3.052 (1.42)	4.906 (2.32)*	0.188 (0.10)	2.597 (1.54)
edad2	-0.071 (2.06)*	-0.034 (1.11)	-0.057 (1.92)	0.004 (0.16)	-0.030 (1.28)
Pric	-11.258 (1.83)	-10.838 (1.92)	1.471 (0.21)	-9.314 (1.22)	15.300 (1.65)
Seci	-2.770 (0.37)	-15.286 (2.31)*	3.870 (0.50)	1.037 (0.13)	27.153 (2.90)**
Secc	11.483 (1.49)	-2.448 (0.40)	14.573 (1.92)	6.796 (0.86)	35.447 (3.68)**
Supi	36.847 (3.19)**	13.352 (1.44)	19.457 (1.92)	9.416 (0.98)	38.990 (3.70)**
Supc	56.118 (4.92)**	27.132 (3.40)**	36.891 (4.25)**	37.534 (4.40)**	61.484 (6.11)**
Hombre	73.822 (3.22)**	49.921 (2.69)**	48.813 (3.94)**	45.247 (6.09)**	46.776 (6.04)**
hijos16	-8.772 (4.72)**	-11.189 (7.46)**	-8.452 (5.78)**	-6.319 (4.61)**	-5.774 (3.93)**
Concurre	-22.819 (1.23)	-7.944 (0.57)	6.354 (0.45)	8.022 (0.88)	-3.468 (0.43)
lij	-0.000 (3.82)**	-0.008 (1.92)	-0.015 (4.25)**	-0.004 (1.72)	-0.007 (2.84)**
Constant	-104.657 (2.59)**	-52.614 (1.48)	-88.321 (2.41)*	-12.098 (0.38)	-79.120 (2.65)**
Observations	834	1042	698	804	810

Valor absoluto de los estadísticos t entre paréntesis. * significativo a un nivel del 5%; ** significativo a un nivel de 1%

Tabla 11

Cambios en la tasa de incidencia –FGT(0)– real y simulada

Cambios en la tasa de incidencia real y simulada							
Efectos	1980-1986	1986-1992	1992-1998	1998-2000	1980-1992	1992-2000	1980-2000
real	4.74	6.90	12.63	-0.90	11.65	11.73	23.38
tamaño-directo	1.35	0.58	-1.25	0.89	2.91	0.32	3.02
tamaño-horas	0.07	0.11	-0.19	-0.01	0.35	-0.03	0.00
parámetros-horas	0.47	-0.61	-0.89	0.00	-0.22	-0.61	-0.77
tamaño-total	1.40	0.74	-1.24	0.89	3.51	0.44	3.28
efecto total	1.88	0.07	-2.00	0.89	3.43	-0.04	2.71

Tabla 12

Cambios en la brecha de la pobreza –FGT(1)– real y simulada

Cambios en la brecha de la pobreza real y simulada							
Efectos	1980-1986	1986-1992	1992-1998	1998-2000	1980-1992	1992-2000	1980-2000
real	2.33	1.90	7.39	-0.21	4.23	7.18	11.41
tamaño-directo	0.31	0.08	-0.30	0.41	1.00	0.16	1.48
tamaño-horas	0.25	-0.21	-0.02	0.03	0.15	0.03	0.16
parámetros-horas	0.12	0.33	-0.35	-0.02	-0.01	-0.35	-0.13
tamaño-total	0.58	0.34	-0.31	0.44	1.13	0.20	1.65
efecto total	0.28	0.01	-0.64	0.43	1.12	-0.18	1.43

Tabla 13

Cambios en la profundidad de la pobreza –FGT(2)– real y simulada

Cambios en la profundidad de la pobreza real y simulada							
Efectos	1980-1986	1986-1992	1992-1998	1998-2000	1980-1992	1992-2000	1980-2000
real	1.47	0.80	4.66	-0.30	2.27	4.36	6.63
tamaño-directo	0.15	0.23	-0.13	0.27	0.42	0.07	0.76
tamaño-horas	0.04	0.06	0.00	0.02	0.10	0.02	0.14
parámetros-horas	0.16	-0.20	-0.21	0.00	0.00	-0.21	-0.02
tamaño-total	0.19	0.29	-0.13	0.29	0.51	0.08	0.89
efecto total	0.33	0.10	-0.33	0.29	0.51	-0.15	0.80

Tabla 14

Cambios en el coeficiente de Gini real y simulado

Cambios en el gini real y simulado							
Efectos	1980-1986	1986-1992	1992-1998	1998-2000	1980-1992	1992-2000	1980-2000
real	2.52	2.30	5.89	0.65	4.82	6.55	11.37
tamaño-directo	0.73	0.72	0.00	0.23	1.88	0.27	2.13
tamaño-horas	-0.12	0.05	-0.14	0.16	-0.01	0.13	-0.03
parámetros-horas	0.00	-0.08	-0.21	-0.02	0.01	-0.15	-0.22
tamaño-total	0.61	0.76	-0.14	0.40	1.86	0.39	2.10
efecto total	0.79	0.66	-0.22	0.26	1.93	0.13	2.08

CAPÍTULO III

UN MODELO DE ELECCIÓN BINARIA CON REGRESOR ENTERO ENDÓGENO*

* Esta investigación fue en parte financiada por una beca de la Academia Nacional de Ciencias Económicas. Agradezco a Sebastián Galiani por sus sugerencias, a William Greene por cedermme gentilmente los datos para la aplicación empírica y a Laura Ripani por su colaboración. Cualquier error es de mi exclusiva responsabilidad.

1. Introducción

Varias aplicaciones empíricas discutidas en la literatura pueden representarse mediante modelos de elección binaria con regresores endógenos enteros. Un ejemplo viene dado por los modelos de participación laboral femenina que incluyen como variable explicativa al número de hijos. En el Capítulo II de esta tesis se utilizó una especificación que suponía que las decisiones de fecundidad eran independientes de las de participación laboral de las mujeres. Autores como Schultz (1978), Dooley (1982), Moffitt (1984) y Hotz y Miller (1988), entre otros, señalaron tempranamente el problema de la endogeneidad de las características de los hijos en estos modelos de participación laboral femenina. La endogeneidad puede deberse a la presencia de factores inobservables que afectan tanto la fecundidad como las decisiones laborales de las mujeres, o a la determinación simultánea de ambas variables.

Otro ejemplo es el que discute Weiss (1999), que consiste en un modelo sobre los determinantes de la aprobación de solicitudes para tarjetas de crédito, donde una de las variables explicativas es el número de veces que el solicitante mantuvo un saldo deudor en otras cuentas de crédito. La endogeneidad de esta variable entera puede justificarse en este caso por la posible existencia de factores inobservables que afectan tanto el uso de fuentes de crédito por parte de los demandantes de tarjetas como la decisión de los encargados de evaluar las solicitudes.

Los métodos estándar para estimar modelos de elección binaria, como los estimadores máximo verosímiles de los modelos probit y logit, resultan inconsistentes ante la presencia de regresores endógenos. En la literatura los avances se han concentrado mayormente en la estimación de modelos de elección binaria en los que la variable explicativa endógena es continua. Una especificación particularmente conveniente es la propuesta por Rivers y Vuong (1988) considerando un modelo estructural para la variable binaria de interés y modelos lineales en forma reducida para los regresores endógenos. El estimador de Rivers y Vuong es máximo verosímil condicional y se obtiene mediante un procedimiento en dos etapas. En la primera etapa se estiman por mínimos cuadrados los modelos para las variables endógenas y en la segunda etapa se estima por máxima verosimilitud un modelo probit, que incluye los errores de estimación de la etapa previa como regresores adicionales. Comparado con otros estimadores propuestos en la literatura, el estimador de Rivers y Vuong se desempeña

favorablemente en términos de performance en muestras chicas, utilidad para realizar tests de hipótesis y simplicidad en su implementación.⁴⁵

En el espíritu de la estrategia de Rivers y Vuong, este trabajo propone un estimador en dos etapas para un modelo de elección binaria en el que la variable explicativa endógena es entera en lugar de continua. Este método provee una alternativa simple a la estrategia adoptada por Weiss (1999), que discute una especificación particular para el modelo de elección binaria con un regresor endógeno entero en un contexto de máxima verosimilitud. Weiss supone un modelo binomial negativo para la variable entera, lo que introduce una gran complejidad en la estructura del modelo simultáneo. La dificultad computacional de obtener estimadores máximo verosímiles para esa especificación es considerable, por lo que resulta deseable evaluar otras alternativas de implementación más sencilla.

La alternativa de estimación que se explora en el presente trabajo surge de considerar que los modelos para variables enteras son en esencia modelos de regresión exponencial. En los modelos de regresión estándar para este tipo de variables, como el poisson y el binomial negativo, se supone generalmente que la esperanza condicional de la variable entera es exponencial. Si Y es la variable entera, el supuesto implica que $E[Y|X]=\exp(X\beta)$ donde X es un vector de variables explicativas y β el vector de parámetros asociado a X . Esta especificación de la esperanza condicional es coherente con un modelo de regresión exponencial para la variable entera donde el término de error puede ser aditivo con media cero –Modelo Exponencial Aditivo– o bien multiplicativo con media uno –Modelo Exponencial Multiplicativo. Para cada una de estas dos especificaciones se propone un estimador máximo verosímil condicional que se obtiene mediante un procedimiento en dos etapas, y que permite estimar consistentemente los parámetros del modelo de elección binaria bajo una amplia gama de supuestos acerca de la distribución de los errores en la ecuación de la variable endógena, como así también evaluar directamente la hipótesis de exogeneidad del regresor entero. Además de ser de muy sencilla aplicación, la evidencia de Monte Carlo sugiere un buen desempeño de este estimador en muestras chicas.

El resto del trabajo se organiza como sigue. En la sección 2 se presenta el modelo de interés y se discute el modelo tratado en Weiss (1999). En la sección 3 se propone un estimador máximo verosímil condicional para el modelo de elección binaria cuando la especificación del regresor endógeno entero es aditiva y se discuten otras

⁴⁵ Si la conveniencia computacional no se tiene en cuenta, naturalmente es preferible el estimador máximo verosímil.

especificaciones que han sido estudiadas en la literatura. En la sección 4 se presenta evidencia de experimentos de Monte Carlo que permiten evaluar la performance en muestras chicas de ese estimador. La sección 5 discute la especificación multiplicativa para la variable endógena y propone otro estimador en dos etapas. En la sección 6 se presentan dos aplicaciones empíricas. La primera es el ejemplo de aprobación de solicitudes para tarjetas de crédito estudiado por Weiss (1999) en base a un modelo de Greene (1994). La segunda aplicación consiste en un modelo de participación laboral de mujeres que incluye entre las variable explicativas al número de hijos menores de 6 años. Finalmente, en la sección 7 se concluye.

2. El modelo

El interés de este trabajo se centra en un modelo de elección binaria con un regresor entero endógeno. Una especificación general para la forma estructural de este modelo viene dada por las ecuaciones (22) y (23).

$$(22) \quad Y_{1i}^* = \delta Y_{2i} + X_{1i}\beta_1 + u_{1i} \quad \text{y} \quad Y_{1i} = 1[Y_{1i}^* > 0]$$

con $i = 1, \dots, n$

$$(23) \quad E[Y_{2i} | X_{2i}] = \exp(X_{2i}\beta_2)$$

Y_{1i}^* es una variable latente no observable de la que sólo se conoce su signo a través del indicador Y_{1i} , y Y_{2i} es una variable entera y potencialmente endógena con esperanza condicional exponencial, supuesto consistente con una amplia gama de modelos. X_{1i} y X_{2i} son, respectivamente, vectores de $1 \times k_1$ y $1 \times k_2$ regresores exógenos cuyo primer elemento es la unidad.

La especificación discutida por Weiss (1999) para el modelo simultáneo de elección binaria con regresor entero endógeno –en adelante *modelo de Weiss*– responde a la estructura general dada por las ecuaciones (22)-(23). En el modelo de Weiss, la variable entera viene determinada por las ecuaciones (24) y (25) debajo. La esperanza condicional de la variable entera –ecuación (24)– tiene la típica forma exponencial pero depende de un término inobservable e independiente de X_{2i} , representado por v_{2i} .⁴⁶ Este término indica la presencia de heterogeneidad no observada afectando los propios

⁴⁶ En Mullahy (1997) se discuten estimadores consistentes para el modelo de la esperanza condicional de una variable entera en presencia de correlación entre los regresores observables e inobservables.

parámetros de la distribución condicional de la variable entera. La ecuación (25) implica que, condicional en la realización de v_{2i} , la variable entera sigue una distribución poisson con parámetro λ_i .⁴⁷

$$(24) \quad E[Y_{2i} | X_{2i}, v_{2i}] \equiv \lambda_i = \exp(X_{2i}\beta_2 + v_{2i})$$

$$(25) \quad P(Y_{2i} = y_2 | v_{2i}) \equiv Po(\lambda_i, y_2) = \frac{\lambda_i^{y_2} \exp(-\lambda_i)}{y_2!} \quad \text{con } y_2 = 0, 1, 2, \dots$$

Cuando v_{2i} sigue una distribución exponencial gamma (EG), es posible encontrar una forma cerrada para la distribución marginal de la variable entera. Después de la normalización $E[\exp(v_{2i})]=1$, necesaria para la identificación de (24), la p.d.f de v_{2i} viene dada por

$$(26) \quad v_{2i} \sim EG(\alpha) = \frac{\alpha^\alpha}{\Gamma(\alpha)} \exp(-\alpha \exp(v_{2i})) \exp(\alpha v_{2i})$$

donde α es el parámetro de la función gamma asociada a esta distribución y la $\text{Var}[\exp(v_{2i})]=1/\alpha$. La especificación dada por las ecuaciones (24), (25) y (26) corresponde al modelo binomial negativo para la variable entera, donde $E[Y_{2i}|X_{2i}] = \exp(X_{2i}\beta_2)$ y $\text{Var}[Y_{2i}|X_{2i}] = \exp(X_{2i}\beta_2)[1 + \exp(X_{2i}\beta_2)/\alpha]$.

La representación adoptada por Weiss para incorporar la endogeneidad de Y_{2i} en la ecuación de elección binaria, viene dada por:

$$(27) \quad u_{1i} = \varphi h(v_{2i}) + e_{1i} \quad \text{con } h(v_{2i}) = \Phi^{-1}[G(v_{2i})]$$

donde $G(\cdot)$ y $\Phi(\cdot)$ son las funciones de distribución acumulada de v_{2i} y normal estándar respectivamente, con lo que $h(\cdot)$ es normal estándar. Además, supone que la distribución

⁴⁷ La incorporación de heterogeneidad no observable en la determinación de la media condicional de Y_{2i} a través del término v_{2i} permite romper con el supuesto de equidispersión –igualdad entre esperanza y varianza condicional de la variable entera– al que el modelo de poisson está atado.

de e_{1i} condicional en (Y_{2i}, v_{2i}, X_{1i}) es normal con media cero y varianza $\sigma_{e_1}^2 = 1 - \varphi^2$, y por lo tanto la distribución condicional de u_{1i} es normal estándar.

La obtención de estimadores máximo verosímiles de los parámetros del modelo de Weiss implica una dificultad computacional considerable, por lo que resulta conveniente evaluar la hipótesis nula de exogeneidad de la variable entera, es decir $\varphi=0$, antes de encarar la estimación conjunta del modelo. Si la endogeneidad no fuera relevante la ecuación de elección binaria tendría la estructura de un modelo probit estándar, pudiéndose estimar como es usual por máxima verosimilitud independientemente de la ecuación de la variable entera.⁴⁸ Weiss desarrolla un test para evaluar $H_0: \varphi=0$ versus la hipótesis alternativa $\varphi \neq 0$ basado en Smith y Blundell (1986).⁴⁹

El modelo de Weiss logra acomodar los dos rasgos fundamentales del modelo bajo consideración: la endogeneidad en un modelo de elección binaria y la naturaleza entera de la variable explicativa endógena. Sin embargo, la búsqueda de estimadores completamente eficientes en ese contexto implica un procedimiento computacionalmente muy complejo que requiere, entre otras cosas, de la aplicación de procesos de integración numérica, de manera que su replicación para encarar sistemáticamente el análisis empírico de problemas que respondan a la misma estructura es improbable o demasiado costosa.

⁴⁸ Otra situación en que la endogeneidad no sería relevante es cuando $\delta=0$. Sin embargo, si $\varphi \neq 0$ la distribución de u_{1i} podría no ser normal ya que depende de la especificación de $h(\cdot)$ y de la distribución de e_{1i} –ver ecuación (27). En ese caso el estimador máximo verosímil del modelo probit resultaría inconsistente para el modelo de elección binaria, por lo que sería necesario estimar el modelo conjunto para determinar $h(v_{2i})$, la distribución de v_{2i} y si $\delta=0$.

⁴⁹ Sustituyendo (27) en el modelo de elección binaria (22), y reemplazando $h(v_{2i})$ por el residuo generalizado que se define como

$$\begin{aligned} \hat{h}(v_{2i}) &= E[h(v_{2i}) | Y_{2i} = y_2] \Big|_{(\hat{\beta}_1, \hat{\alpha})} \\ &= \frac{1}{P(Y_{2i} = y_2)} \int h(v_{2i}) Po(\hat{\lambda}_i, y_2) EG(v_{2i}, \hat{\alpha}) dv_{2i} \end{aligned}$$

se obtiene

$$Y_{1i}^* = \delta Y_{2i} + X_{1i} \beta_1 + \varphi \hat{h}(v_{2i}) + e_{1i}$$

que responde a la especificación de un modelo probit estándar, pudiéndose estimar de la forma usual por máxima verosimilitud. El estadístico de prueba de este test es la razón t del coeficiente φ , cuya distribución asintótica bajo la hipótesis nula es normal estándar. Weiss también muestra cómo evaluar la hipótesis nula de exogeneidad de Y_{2i} usando un test de multiplicador de Lagrange (LM). Ambos tests son asintóticamente equivalentes bajo la hipótesis nula, pero el que se basa en Smith y Blundell es más fácil de implementar.

En este trabajo se propone una estructura alternativa para el modelo de elección binaria con un regresor endógeno entero, basada en supuestos no más arbitrarios que los adoptados por Weiss pero de estimación considerablemente más simple. Nuevamente partimos de la especificación general del modelo dada por las ecuaciones (22) y (23). Como señala Mullahy (1997), la especificación exponencial para la media condicional implica que los modelos para variables enteras puedan considerarse esencialmente como modelos de regresión exponencial. En ese contexto, la única característica fundamental de la variable dependiente Y_{2i} sería la no negatividad y no el hecho de ser entera. Hay básicamente dos modelos de regresión exponencial consistentes con (23). El primero, al que llamaremos Modelo Exponencial Aditivo (MEA), viene dado por (28). El segundo, Modelo Exponencial Multiplicativo (MEM), viene dado por (29) y se caracteriza por un tratamiento simétrico de las variables observables (X_{2i}) y las inobservables (η_{2i}).

Modelo Exponencial Aditivo – MEA:

$$(28) \quad Y_{2i} = \exp(X_{2i}\beta_2) + u_{2i} \quad \text{con } E[u_{2i}|X_{2i}] = 0$$

Modelo Exponencial Multiplicativo – MEM:

$$(29) \quad Y_{2i} = \exp(X_{2i}\beta_2 + \eta_{2i}) \quad \text{con } E[\exp(\eta_{2i})|X_{2i}] = 1$$

Los modelos (28) y (29) son observacionalmente equivalentes a menos que se adopten supuestos adicionales sobre los momentos de orden superior de los términos de error.⁵⁰ En las secciones 3 y 4 se discute el modelo de elección binaria de la ecuación (22) cuando la especificación del regresor endógeno entero corresponde al MEA. Luego, en la sección 5, se estudia el mismo modelo cuando la variable endógena se determina a partir del MEM.

3. Estimación en dos etapas del modelo aditivo

Dada la considerable dificultad computacional involucrada en la obtención de estimadores completamente eficientes de nuestro modelo, resulta de interés evaluar

⁵⁰ Notar que ambos modelos son consistentes con la esperanza condicional de Y_{2i} dada en (2) pero dan lugar a diferentes varianzas condicionales de Y_{2i} . Ver Wooldridge (1992).

otras alternativas más simples y cuya pérdida de eficiencia en muestras chicas no sea necesariamente importante.

En esta sección se propone un estimador consistente para los parámetros del modelo de elección binaria cuando se adopta el modelo MEA para el regresor endógeno (ver ecuación (28)). La endogeneidad es consecuencia de la correlación entre la variable entera, Y_{2i} , y el término de error de la ecuación de elección binaria (22), u_{1i} . Esta correlación se incorpora al modelo mediante la ecuación (30), que supone que u_{1i} y u_{2i} satisfacen:

$$(30) \quad u_{1i} = \theta u_{2i} + \varepsilon_{1i}$$

Sustituyendo (30) en (22) se obtiene:

$$(31) \quad Y_{1i}^* = \delta Y_{2i} + X_{1i} \beta_1 + \theta u_{2i} + \varepsilon_{1i}$$

Supongamos que se observa un muestra *iid* (Y_{1i}, Z_i) donde $Z_i \equiv (Y_{2i}, X_{1i}, u_{2i})$. Si la distribución de ε_{1i} condicional en Z_i es normal estándar, el modelo responde a la especificación de un probit estándar, de manera que los parámetros del modelo de elección binaria pueden ser estimados consistentemente por máxima verosimilitud de la manera usual.

En el caso relevante para nosotros, sin embargo, u_{2i} no es observable. Los estimadores máximo verosímiles del modelo probit de Y_{1i} en Y_{2i} y X_{1i} serán inconsistentes a menos que se adopten fuertes supuestos sobre la distribución de u_{2i} y la posible dependencia entre u_{2i} y Y_{2i} . En este contexto se propone un estimador máximo verosímil condicional que se obtiene mediante un procedimiento en dos etapas que sigue en espíritu la estrategia de Rivers y Voung (1988) –en adelante RV. Llamaremos a este estimador 2EMEA por tratarse de un estimador en dos etapas para el modelo en que la variable endógena responde a la especificación MEA.

Manteniendo el supuesto en (28) que $E[u_{2i}|X_{2i}]=0$ y suponiendo que se observa una muestra *iid* $(Y_{1i}, X_{1i}, Y_{2i}, X_{2i})$, el siguiente procedimiento permite estimar consistentemente los parámetros δ , β_1 y θ .

- 1° etapa.** Estimar (28) por mínimos cuadrados no lineales (NLS). Obtener estimadores consistentes de u_{2i} definidos como $\hat{u}_{2i} \equiv Y_{2i} - \exp(X_{2i} \hat{\beta}_2)$. Aquí la propiedad de consistencia depende de la aditividad de u_{2i} y del supuesto $E[u_{2i}|X_{2i}]=0$.
- 2° etapa.** Estimar por máxima verosimilitud un modelo probit de Y_{1i} en Y_{2i} , X_{1i} y \hat{u}_{2i} como variables explicativas. Esta etapa consiste básicamente en reemplazar u_{2i} por un estimador consistente obtenido en la etapa previa. La propiedad de consistencia de los estimadores de δ , β_1 y θ depende de la normalidad de la distribución de ε_{1i} condicional en Z_i y de la correcta especificación de la media condicional de Y_{1i}^* .

El estimador 2EMEA es entonces consistente y asintóticamente normal bajo una amplia gama de posibles distribuciones del término de error de la ecuación de la variable endógena, por lo que es posible realizar tests t sobre los coeficientes individuales del modelo de elección binaria usando esa distribución asintótica. En particular, la hipótesis nula de exogeneidad de la variable entera puede evaluarse directamente mediante un test de significatividad individual sobre el coeficiente de los residuos de la primera etapa en la ecuación de la variable binaria. Si no puede rechazarse la hipótesis nula de exogeneidad, el modelo de elección binaria respondería a la especificación del modelo probit estándar y podría estimarse por máxima verosimilitud independientemente de la ecuación de la variable entera.

Con respecto a la matriz de varianzas y covarianzas del estimador 2EMEA, el estimador usual del probit que se obtiene en la segunda etapa resulta inconsistente porque ignora la variabilidad muestral de $\hat{\beta}_2$. En este trabajo se emplean técnicas de bootstrap para obtener estimaciones de los errores estándar de los coeficientes, dejando para investigaciones futuras la derivación analítica de la matriz de varianzas y covarianzas del estimador en dos etapas.

Otras especificaciones relacionadas discutidas en la literatura

Más arriba se señalaba que el estimador 2EMEA surge como una adaptación del método de máxima verosimilitud condicional en dos etapas propuesto por RV. El modelo estudiado por RV consiste en una ecuación estructural para la variable binaria de interés, que depende de un conjunto de regresores endógenos continuos representados por ecuaciones lineales en forma reducida. En la primera etapa del procedimiento se estiman por MCO los modelos de las variables endógenas y en la segunda etapa se estima un modelo probit estándar, incluyendo los errores de estimación de la etapa

previa como regresores adicionales. Comparado con otros estimadores para modelos de elección binaria con regresores endógenos continuos, el estimador de RV se desempeña favorablemente en términos de performance en muestras chicas, utilidad para realizar tests de hipótesis y simplicidad en su implementación.

Como señalan RV, si bien el supuesto de normalidad conjunta en la distribución condicional de los errores es estándar para este tipo de modelos, no es necesario para la consistencia del estimador en dos etapas que proponen, así como tampoco lo es para el estimador 2EMEA. La consistencia sólo requiere que la distribución condicional de u_{1i} sea normal con media lineal en u_{2i} y varianza constante.

Otro estimador similar al de RV es el sugerido por Lee (1981) –probit con variables instrumentales (IVP). Este estimador también se obtiene mediante un procedimiento en dos etapas, pero se diferencia del de RV en que en la segunda etapa no se estima la forma estructural sino la forma reducida del modelo de elección binaria. Usando la notación para nuestro modelo, la segunda etapa del estimador IVP implicaría estimar un modelo probit estándar de Y_{1i} en X_{2i} , X_{1i} y el residuo estimado por mínimos cuadrados ordinarios en la primera etapa. Según la evidencia de Monte Carlo presentada por RV, el IVP se comporta algo peor en muestras chicas que el estimador de RV, mientras que no es posible establecer un ordenamiento en cuanto a su eficiencia asintótica.

Una especificación de nuestro modelo pero “cabeza abajo” es la que proponen Windmeijer y Santos Silva (1997). Puede pensarse que el modelo (22)-(23) surge como una de las dos triangularizaciones posibles del siguiente modelo simultáneo más general:

$$(32) \quad Y_{1i}^* = \delta Y_{2i} + X_{1i}\beta_1 + u_{1i} \quad \text{y} \quad Y_{1i} = 1[Y_{1i}^* > 0]$$

$$(33) \quad E[Y_{2i} | X_{2i}] = \exp(\gamma Y_{1i} + X_{2i}\beta_2)$$

En el modelo (32)-(33) no puede darse que δ y γ sean simultáneamente distintos de cero. Esto violaría el requisito de consistencia $\text{Prob}(Y_{1i}=1) + \text{Prob}(Y_{1i}=0) = 1$ (ver Maddala (1983, p.118) y Blundell y Smith (1994)). Por lo tanto la endogeneidad en este sistema no puede operar en ambas direcciones. El modelo relevante será entonces uno con variable dependiente entera y regresor endógeno binario ($\delta=0$) o, alternativamente, otro con variable dependiente binaria y regresor endógeno entero ($\gamma=0$), como el discutido hasta ahora.

Windmeijer y Santos Silva se concentran en la otra triangularización, es decir, suponen que en la ecuación de la variable entera hay un regresor binario potencialmente endógeno, empleando alternativamente las especificaciones MEA y MEM para representar la variable entera. Estos autores discuten un estimador basado en el método generalizado de momentos (GMM). Intuitivamente, el procedimiento consiste en instrumentar la variable binaria endógena mediante la probabilidad predicha obtenida al estimar la ecuación (32), que responde a la estructura de un modelo probit estándar, por máxima verosimilitud. Sin embargo, pese a que los modelos MEA y MEM son observacionalmente indistinguibles, los instrumentos que resultan adecuados bajo una especificación generalmente dejan de serlo bajo la otra. La aplicación empírica que presentan Windmeijer y Santos Silva consiste en un modelo de determinación del número de visitas al médico en el último mes, que incluye entre los regresores un indicador binario sobre la percepción del propio estado de salud del individuo. En este caso puede pensarse que la endogeneidad surge como consecuencia de errores de medición de la variable binaria –autorreportada– posiblemente correlacionados con los determinantes inobservables de las visitas al médico.

4. Evidencia de Monte Carlo

En esta sección se presentan los resultados de algunos experimentos de Monte Carlo diseñados para evaluar las propiedades del estimador 2EMEA en muestras chicas. Si bien este es el principal objetivo, en cada caso se presentan también los resultados de usar el estimador máximo verosímil de un modelo probit estándar ignorando el problema de endogeneidad. La comparación de los resultados obtenidos por ambos métodos permite cuantificar la severidad de este problema cuando se utiliza ese estimador inconsistente.

Las simulaciones se basan en un modelo de elección binaria con dos regresores, uno de los cuales, Y_{2i} , es potencialmente endógeno y responde a la especificación MEA (ecuación (28)). El modelo viene dado por las ecuaciones (34) a (36).

$$(34) \quad Y_{1i}^* = \delta Y_{2i} + b_{11} + b_{12}x_{1i} + u_{1i} \quad \text{y} \quad Y_{1i} = 1[Y_{1i}^* > 0]$$

$$(35) \quad Y_{2i} = \exp(b_{21} + b_{22}x_{2i}) + u_{2i}$$

$$(36) \quad u_{1i} = \theta u_{2i} + \varepsilon_{1i}$$

Para las simulaciones se utilizaron los siguientes valores de los parámetros: $\delta=1$, $b_{11}=-2$, $b_{12}=1$, $b_{21}=-1$ y $b_{22}=1$.⁵¹ Cada simulación se basó en una muestra aleatoria de 1000 observaciones y se realizaron 500 replicaciones. La variable exógena x_{1i} se generó a partir de una distribución normal estándar, mientras que x_{2i} se obtuvo independientemente a partir de una distribución uniforme definida en el intervalo $[0, 12^{1/2}]$. Los valores para ε_{1i} y u_{2i} se generaron a partir de distribuciones normales independientes con medias cero y varianzas iguales a uno y σ_2^2 , respectivamente. Las observaciones u_{1i} se generaron usando (36), con lo que $E[u_{1i}]=0$, $\text{Var}[u_{1i}] \equiv \sigma_1^2 = 1 + \theta^2 \sigma_2^2$ y $\theta = \rho \sigma_1 / \sigma_2$, donde $\rho \equiv \text{Corr}[u_{1i}, u_{2i}]$. Cuando $\rho \neq 0$ o, equivalentemente, $\theta \neq 0$, Y_{2i} es endógena en la ecuación (34) y los estimadores máximo verosímiles de un probit estándar de Y_{1i} en Y_{2i} y x_{1i} son inconsistentes para los parámetros –reescalados usando σ_1 – del modelo de elección binaria. Este problema se agudiza a medida que ρ aumenta en valor absoluto, lo que, dada la varianza de u_{2i} , también implica mayor valor absoluto de θ .⁵²

En primer lugar se realiza un experimento para estudiar el efecto de aumentos en la correlación entre los errores de la ecuación de elección binaria (u_{1i}) y de la variable entera (u_{2i}). La varianza de u_{1i} aumenta como consecuencia del aumento en la correlación porque $\sigma_1^2 = 1/(1-\rho^2)$, mientras que la varianza de u_{2i} se mantiene fija ($\sigma_2^2 = 1$) a lo largo de los experimentos. Se usan cinco conjuntos alternativos de parámetros cuyos valores se presentan en las primeras tres columnas de la Tabla 1 –ver Apéndice al final de este capítulo. Estos valores de los parámetros permiten evaluar situaciones con correlaciones extremas como así también varios casos intermedios. Mientras la correlación aumenta desde cero a casi 0.9, el desvío estándar de u_1 más que se duplica, pasando de 1 a 2.24. Los resultados de Monte Carlo de estimar el modelo (34)-(36) con estos parámetros se presentan en las restantes columnas de la Tabla 1.

Naturalmente el sesgo del probit aumenta marcadamente con la correlación. Por ejemplo, la raíz cuadrada del error cuadrático medio (RMSE) del coeficiente de la

⁵¹ El valor de b_{22} sólo se modifica para provocar cambios en la bondad del ajuste del modelo estimado en la primera etapa. Ver Tabla 2.

⁵² El signo de θ no afecta los resultados de los experimentos de Monte Carlo aquí presentados pero sí los reportados por RV. Esta diferencia surge cuando se reescalan las estimaciones para poder comparar los distintos métodos. En el presente trabajo los estimadores del modelo probit estándar se reescalan usando $\text{SD}(u_{1i}) = (1 + \theta^2 \sigma_2^2)^{1/2}$, mientras que RV usan $(1 + (\theta + \delta)^2 \sigma_2^2)^{1/2}$, que es el desvío estándar de la forma reducida del probit y que claramente no sólo depende de la magnitud sino también del signo de θ .

variable endógena se triplica al pasar de una correlación nula a una moderada de 0.5, y casi se triplica nuevamente cuando la correlación alcanza el valor de 0.9. El sesgo para estos valores de correlación equivale a una sobreestimación del orden del 1.5%, 18% y 53%, respectivamente, del valor real del coeficiente ($\delta=1$). Comportamientos similares se observan para los otros coeficientes.

Es de esperar que el sesgo del estimador 2EMEA también se acentúe a medida que la correlación aumenta. Intuitivamente, al reemplazar u_{2i} por su estimación existe parte del verdadero u_{2i} que continúa excluyéndose del modelo. Esa omisión produce un sesgo que se agudiza a medida que crece la correlación –y consecuentemente θ . A pesar de esto, es notable de esta tabla la buena performance del estimador 2EMEA, incluso para los niveles más altos de correlación. Por ejemplo, con una correlación de 0.9, el estimador 2EMEA sobrestima al verdadero coeficiente de la variable endógena en sólo un 2.2% y el RMSE del probit es 6 veces mayor que el RMSE del estimador 2EMEA, que sólo aumenta un 24% al pasar de una correlación nula a una de 0.9.

El buen desempeño del estimador 2EMEA que se desprende de la Tabla 1 en parte se debe a que la bondad del ajuste del modelo exponencial – R^2 implícito en el modelo no lineal y que por simplicidad será referido desde ahora como R^2 – estimado en la primera etapa es muy alto. Intuitivamente, cuanto menor es la bondad del ajuste que surge de estimar la ecuación (36) por NLS, más se asemejan los residuos de esa regresión a la verdadera Y_{2i} , agravándose el problema de omitir u_{2i} y estimar el modelo como si se tratara de un modelo probit estándar. En relación al estimador 2EMEA, cuanto menor sea el R^2 del modelo de la primera etapa, más difícil será la identificación de δ y θ en la segunda, demorando la convergencia del estimador y aumentando las medidas de sesgo y RMSE para un tamaño de muestra dado.

Para explorar este punto y su efecto sobre el desempeño del estimador en muestras chicas, se realizan una serie de experimentos de Monte Carlo para distintos ajustes del modelo MEA, manteniendo constante la correlación entre u_{1i} y u_{2i} , y por ende también la propia varianza de u_{1i} . Dada la distribución de x_{2i} y los parámetros b_{21} y b_{22} , el R^2 del modelo de la primera etapa aumenta cuando cae la varianza del término de error de ese modelo, es decir cuando cae σ_2^2 . Por ejemplo, para los experimentos presentados en la Tabla 1, la varianza de u_{2i} es relativamente baja ($\sigma_2^2=1$) y el ajuste promedio muy alto (95%). Por otra parte, también θ aumenta cuando cae σ_2^2 , por lo que la varianza de u_{2i} no sólo afecta el ajuste del modelo para la variable endógena sino que también modifica la relación entre u_{1i} y u_{2i} .

Una forma de provocar cambios en el R^2 sin modificar θ es mediante cambios en la varianza del término exponencial de la ecuación (36). Esta es la estrategia que se sigue

en este ejercicio. Se realizan siete experimentos con una correlación entre moderada y alta de $\rho=0.7$, $\sigma_2^2=1$ y distintos valores de b_{22} que permiten que el R^2 promedio varíe entre un 12% y un 99%. Los valores de los parámetros empleados y los resultados de Monte Carlo de estimar el modelo (34)-(36) con estos parámetros se presentan en la Tabla 2 del Apéndice.

En la Tabla 2 puede verse cómo mejora la performance del estimador 2EMEA a medida que lo hace el ajuste del modelo estimado en la primera etapa. Cuando el ajuste es muy bajo, por ejemplo 12% en las últimas dos filas de la tabla, el desempeño de los estimadores 2EMEA, especialmente para estimar el coeficiente vinculado a la variable endógena $-\delta$, es muy pobre. Esto se manifiesta en la magnitud del sesgo, pero aún más en los RMSE, considerablemente superiores a los del probit a causa del explosivo aumento de la varianza del estimador en 2 etapas.

Para los experimentos previos los errores del modelo de la variable endógena, u_{2i} , se supusieron normales. Como se remarcó en la sección 3, la normalidad de u_{2i} no es condición necesaria para la consistencia del estimador 2EMEA. A continuación se realiza una serie de experimentos de Monte Carlo con el objetivo de evaluar la performance de este estimador ante cambios en la correlación entre los errores cuando los correspondientes a la ecuación de la variable endógena se distribuyen uniformemente en el intervalo $(-1.73, 1.73)$, de manera que como en las Tablas 1 y 2, $\sigma_2^2=1$. En la Tabla 3 del Apéndice se explora el efecto de aumentos en la correlación de 0 a 0.9, manteniendo constante el R^2 promedio del modelo estimado en la primera etapa. En la Tabla 4 se explora el efecto de cambios en el ajuste de ese modelo para una correlación entre moderada y alta de $\rho=0.7$. Como era de esperar, los resultados son similares a los obtenidos para errores normales en la ecuación de la variable endógena – ver Tablas 1 y 2.

Por último, se presenta un experimento que tiene por objetivo evaluar, en muestras chicas, la potencia del test de exogeneidad de Y_{2i} en el modelo de elección binaria, para distintos valores del parámetro θ , y por ende de la correlación entre los términos de error. La hipótesis nula de exogeneidad ($H_0: \theta=0$) se evalúa mediante el estadístico de prueba z , definido como $z \equiv \hat{\theta} / \hat{sd}(\hat{\theta})$, cuya distribución asintótica bajo la hipótesis nula es $N(0,1)$. Se computa la frecuencia relativa de rechazo de H_0 utilizando para el contraste esa distribución considerando distintos niveles de significatividad asintótica y bondades de ajuste del modelo estimado en la primera etapa. En cada caso se estima el error estándar del coeficiente estimado $\hat{\theta}$ mediante técnicas de bootstrap, usando 100 submuestras para cada una de las 500 replicaciones del experimento de Monte Carlo.

En la Tabla 5 se presentan los resultados de este ejercicio. En las primeras tres columnas se reportan los valores de los parámetros empleados para cada experimento. En las demás columnas se muestran los valores promedio en 500 replicaciones del coeficiente estimado $\hat{\theta}$ y de su desvío estándar estimado, las frecuencias relativas de rechazo de la hipótesis nula de exogeneidad y el p-valor asociado al estadístico de Kolmogorov-Smirnov (K-S) para evaluar la hipótesis de normalidad de la distribución empírica de $\hat{\theta}$. La hipótesis nula de exogeneidad se evalúa para tres niveles de significatividad asintótica alternativos y considerando siempre una hipótesis alternativa bilateral. La primera fila de la tabla corresponde al caso en que la hipótesis nula de exogeneidad es cierta, por lo que las frecuencias relativas de rechazo que se reportan constituyen el error tipo I del test. Si en lugar de tratarse de la distribución asintótica del estadístico z , la normal estándar fuera su distribución exacta, el error tipo I coincidiría con el correspondiente nivel de significatividad del test. Para los demás casos la frecuencia relativa de rechazo de H_0 mide la potencia del test.

Se desprende de la Tabla 5 que aún para correlaciones entre bajas y medias –alrededor de $\rho=0.4$ – la potencia del test es muy alta. Por su parte, la probabilidad de rechazar una hipótesis nula de exogeneidad cuando es cierta se acerca mucho al nivel de significatividad asintótico del test, en especial para el caso con menor R^2 .

Según los resultados de aplicar el test de normalidad de Kolmogorov-Smirnov, no puede rechazarse la normalidad de la distribución empírica del estimador 2EMEA de θ para los niveles de significatividad habituales –la única excepción aparece en la última fila de la tabla.

5. Un modelo multiplicativo para la variable endógena

Hasta aquí el trabajo se ha concentrado en la especificación aditiva –MEA– del modelo de regresión para la variable endógena, representado por la ecuación (28). Como se señaló anteriormente, haciendo los supuestos necesarios sobre la esperanza del término de error, tanto el MEA como el MEM –Modelo Exponencial Multiplicativo– son consistentes con $E[Y_{2i} | X_{2i}] = \exp(X_{2i}\beta_2)$ y observacionalmente indistinguibles entre sí.

En esta sección se discute brevemente el modelo de elección binaria cuando el regresor entero endógeno viene determinado por el MEM (ecuación (29)). Asimismo se propone otro estimador máximo verosímil condicional en dos etapas que resulta consistente bajo

esta nueva especificación, y que al igual que el estimador 2EMEA es de muy sencilla aplicación. El modelo viene dado por

$$(37) \quad Y_{1i}^* = \delta Y_{2i} + X_{1i} \beta_1 + u_{1i} \quad \text{y} \quad Y_{1i} = 1[Y_{1i}^* > 0]$$

$$(38) \quad Y_{2i} = \exp(X_{2i} \beta_2 + \eta_{2i}) \quad \text{con} \quad E[\exp(\eta_{2i})|X_{2i}] = 1$$

$$(39) \quad \mu_{1i} = \theta (\eta_{2i} - M) + \varepsilon_{1i}$$

La diferencia con el MEA es que el término de error en la ecuación (38) recibe un tratamiento simétrico al de las variables observables X_{2i} . La ecuación (39) introduce la endogeneidad de Y_{2i} en el modelo de elección binaria.⁵³ Se supone que ε_{1i} es normal estándar condicional en Y_{2i} , X_{1i} y η_{2i} , con lo que la distribución condicional de u_{1i} resulta normal con media cero y varianza $\sigma_1^2 = 1 + \theta^2 \sigma_\eta^2$.

Nuevamente, si $\theta=0$ los parámetros –reescalados– de la ecuación (37) pueden estimarse consistentemente mediante el estimador máximo verosímil de un probit estándar de Y_{1i} en Y_{2i} y X_{2i} . Cuando $\theta \neq 0$ tanto este método como el estimador 2EMEA resultan inconsistentes. El siguiente procedimiento en dos etapas permite estimar consistentemente los parámetros δ , β_1 y θ cuando el regresor endógeno entero en el modelo de elección binaria sigue la especificación MEM.

1° etapa. Linealizar la ecuación (38) aplicando logaritmos y estimarla por el método de mínimos cuadrados ordinarios (MCO). Obtener los residuos de esa regresión. Estos residuos son estimadores consistentes de $\eta_{2i} - M$.

2° etapa. Estimar por máxima verosimilitud un modelo probit estándar de Y_{1i} en Y_{2i} , X_{2i} y los residuos obtenidos en la etapa anterior.

Al linealizar la ecuación para la variable entera en la primera etapa, surge un problema con este procedimiento debido a la alta frecuencia de ceros que suele presentar este tipo de variables. Una opción sería restringir la estimación a la submuestra para la cual

⁵³ Notar que el término $\eta_{2i} - M$ cumple aquí un rol similar al de la función $h(\cdot)$ en la ecuación (27) del modelo de Weiss (1999). La necesidad de encontrar una expresión cerrada para la distribución de Y_{2i} es lo que hace que Weiss utilice $h(v_{2i})$ con v_{2i} exponencial gamma y no $\eta_{2i} - M$ con η_{2i} normal. Ver Weiss (1999), pág. 432-433.

$Y_{2i} > 0$. Otra alternativa consistiría en reemplazar el valor de la variable entera para las observaciones con $Y_{2i} = 0$ por un valor positivo arbitrariamente pequeño y estimar el modelo para toda la muestra.

A continuación se presentan los resultados de un experimento de Monte Carlo diseñado para evaluar la performance en muestras chicas del estimador 2EMEM –2 Etapas Modelo Exponencial Multiplicativo. Las simulaciones para el experimento de Monte Carlo se basan en el modelo (40)-(42).

$$(40) \quad Y_{1i}^* = \delta Y_{2i} + b_{11} + b_{12}x_{1i} + u_{1i} \quad \text{y} \quad Y_{1i} = 1[Y_{1i}^* > 0]$$

$$(41) \quad Y_{2i} = \exp(b_{21} + b_{22}x_{2i} + \eta_{2i}) \quad \text{con} \quad E[\exp(\eta_{2i})|x_{2i}] = 1$$

$$(42) \quad \mu_{1i} = \theta (\eta_{2i} - M) + \varepsilon_{1i}$$

En general, se emplearon los mismos valores de los parámetros que en los ejercicios de la sección 4, es decir $\delta = 1$, $b_{11} = -2$, $b_{12} = 1$, $b_{21} = -1$ y $b_{22} = 1$. Al igual que antes cada simulación se basó en una muestra aleatoria de 1000 observaciones y se realizaron 500 replicaciones. x_{1i} y ε_{1i} se generaron, independientemente, a partir distribuciones normal estándar, mientras que para x_{2i} se usó una distribución uniforme definida en el intervalo $[0, 12^{1/2}]$. Los valores para η_{2i} se generaron a partir de una distribución normal con media $M < 0$ y varianza $\sigma_\eta^2 = -2M$, de manera que $\exp(\eta_{2i})$ sigue una distribución log-normal que satisface $E[\exp(\eta_{2i})|X_{2i}] = 1$. De la ecuación (42), se sigue que la distribución de u_{1i} tiene media igual a cero y varianza igual a $\sigma_1^2 = 1 + \theta^2 \sigma_2^2$, y que $\theta = \rho \sigma_1 / \sigma_2$, donde $\rho \equiv \text{Corr}[u_{1i}, \eta_{2i}]$.

En primer lugar se fijó $M = -0.02$, que implica un R^2 –promedio en 500 replicaciones– del 96% para el modelo lineal estimado en la primera etapa por MCO, y se repitió el experimento para distintos niveles de correlación entre los residuos. Se estimó el modelo (40)-(42) usando el estimador máximo verosímil del modelo probit estándar, el estimador 2EMEA y el método 2EMEM. Los resultados de este ejercicio se presentan en la Parte A de la Tabla 6 –ver Apéndice–, mientras que en la Parte B se muestran los correspondientes a $M = -1$ con un ajuste promedio del modelo de la primera etapa del 33%. Los resultados sugieren una buena performance del estimador 2EMEM tanto en términos absolutos como relativos, aún cuando hay considerable endogeneidad en el modelo o el ajuste de la primera etapa es sólo moderado.

6. Dos aplicaciones empíricas

En esta sección se presentan dos aplicaciones empíricas que responden a la estructura de un modelo de elección binaria con regresores enteros y endógenos. El primer ejemplo tiene el interés de ser el discutido por Weiss (1999), mientras que del segundo surgió la motivación para realizar esta investigación.

Ejemplo empírico 1: Aprobación de solicitudes para tarjetas de crédito

La primera aplicación consiste en un modelo de aprobación de solicitudes para obtener tarjetas de crédito. Los datos fueron empleados originalmente por Greene (1994) y luego por Weiss (1999), quién estimó el modelo por Máxima Verosimilitud empleando la especificación presentada en la sección 2.

La variable dependiente del modelo de elección binaria indica si la solicitud para una nueva tarjeta de crédito fue o no aprobada. Se supone que esta decisión es afectada por una variable entera (*Major Derogatory Reports - MDR*), que representa el número de veces que el solicitante mantuvo una deuda en alguna cuenta de crédito por más de sesenta días. La muestra cuenta con información sobre 1319 solicitudes, de las cuales 1023 fueron aprobadas.

Las empresas de tarjetas de crédito tercerizan la evaluación de las solicitudes. Los equipos que finalmente se encargan de esta instancia han automatizado el proceso que permite obtener la calificación crediticia del solicitante –que a su vez determinará si la solicitud es o no aprobada–, pero ese es un proceso que se mantiene en secreto. Greene (1994) encuentra que el número de *MDR* es una variable altamente significativa para explicar la aprobación de solicitudes para obtener una nueva tarjeta de crédito, a la vez que sugiere que podría ser apropiado un modelo simultáneo. La existencia de factores inobservables que afecten tanto al uso del crédito por parte de los individuos como a las decisiones sobre la aprobación de sus solicitudes daría lugar a la endogeneidad de *MDR* en la ecuación de elección binaria. Los modelos de elección binaria y para la variable entera –*MDR*– en Weiss son básicamente iguales a los que presenta Greene (1994), excepto porque Weiss incluye *MDR* y MDR^2 como regresores del primero. La definición de las variables independientes se da en la Tabla 7 del Apéndice, donde se mantienen las denominaciones originales de las mismas.

Además del modelo binomial negativo para la variable endógena, Weiss explora otras especificaciones alternativas, en particular un modelo binomial negativo con exceso de ceros –*zero-inflated negative binomial*. En la aplicación de aprobación de solicitudes

evalúa comparativamente ambos modelos y concluye que no hay evidencia contundente que permita elegir uno por sobre el otro.

En la Tabla 8 se presentan los resultados de estimar el modelo de aceptación de solicitudes para tarjetas de crédito usando alternativamente el estimador del modelo probit estándar, el estimador del modelo de Weiss –se reportan los resultados que aparecen en Weiss (1999)–, el estimador 2EMEA y el estimador 2EMEM. El resultado más interesante es la llamativa similitud de los coeficientes estimados de MDR y MDR^2 usando los estimadores máximo verosímiles del modelo de Weiss y el estimador 2EMEM. En ambos casos también se rechaza la hipótesis nula de exogeneidad de MDR en el modelo de elección binaria, aunque las estimaciones para θ difieren sustancialmente. En cuanto a las demás variables, los intervalos de confianza de los coeficientes estimados por máxima verosimilitud a partir del modelo de Weiss siempre contienen a los estimados por 2EMEM, y viceversa.

Por su parte los resultados de utilizar el estimador 2EMEA difieren sustancial y sospechosamente de los demás. Si bien también se rechaza la exogeneidad de MDR , las estimaciones del coeficiente de esta variable, de su cuadrado y del coeficiente θ no están contenidas en los intervalos de confianza para las estimaciones de Weiss ni para las que surgen usando 2EMEM. Lo mismo puede decirse para muchas de las variables explicativas independientes.

Determinar cuál de las tres especificaciones del modelo endógeno mejor se adapta a los datos escapa a los objetivos de este trabajo, por lo que se deja para futuras investigaciones.

Ejemplo empírico 2: El efecto del número de hijos en la participación laboral femenina

En la literatura económica se han estudiado extensamente las influencias de los hijos en diferentes decisiones del hogar como las relativas a la oferta laboral, demanda de bienes, patrones de consumo intertemporal, etc. El objetivo de la aplicación que aquí se presenta es explorar el efecto que los hijos tienen sobre la decisión de participación laboral de las mujeres.

Como señalan Nakamura y Nakamura (1992), la evidencia empírica sugiere que el impacto de los hijos sobre la oferta laboral de sus madres se da especialmente a través de la propensión a trabajar más que de las horas trabajadas, lo que justifica concentrarse en los determinantes de la probabilidad que una mujer trabaje. La variable dependiente

del modelo de elección binaria correspondiente indicaría la situación laboral –1 si la mujer trabaja y 0 en caso contrario.⁵⁴

El efecto de los hijos sobre la participación laboral de las mujeres surge como consecuencia de la necesidad de dedicar tiempo y otros recursos a la crianza. Esto da lugar a un efecto directo sobre la participación laboral corriente que difiere de acuerdo a la edad y cantidad de hijos, y a la frecuencia de nacimientos, entre otros factores. La evidencia empírica sugiere que es fundamentalmente la presencia de hijos de menos de 6 años lo que afecta la situación laboral de las madres.^{55, 56}

Originalmente, las variables relativas a los hijos se incluían directamente en las ecuaciones de participación laboral de mujeres. Sin embargo, y como ha sido discutido extensamente en la literatura, es altamente probable que el número de hijos, así como otras *proxies* para captar el efecto de los hijos, sea endógeno a la decisión de participación laboral de las mujeres. Esa endogeneidad puede surgir tanto por la presencia de factores inobservables que afectan ambas variables o por la simultaneidad en su determinación. Sea cual fuere la causa, la estimación consistente del efecto de los hijos sobre la participación laboral requiere controlar este problema de endogeneidad.

El método de variables instrumentales (IV) es seguramente la herramienta más usada para evitar los problemas causados por la endogeneidad. En su versión más simple, el método IV consiste en la estimación de una ecuación auxiliar para la variable endógena –primera etapa de la estimación–, de donde se obtienen las predicciones que se usan para reemplazar –instrumentar– luego a la variable endógena en el modelo de interés. Para el ejemplo en cuestión, el método IV consistiría en la estimación de un modelo del número de hijos, por ejemplo, y el posterior reemplazo del número de hijos por su predicción en el modelo de participación laboral de mujeres.

Este método soluciona el problema de endogeneidad sólo cuando los regresores de la ecuación auxiliar estimada en la primera etapa son exógenos en el modelo de interés.

⁵⁴ En el Capítulo II se modeló la participación de laboral medida por las horas trabajadas para poder realizar las microsimulaciones. En cambio, basándose en la justificación del párrafo, la aplicación aquí presentada estudia la participación laboral como un fenómeno binario.

⁵⁵ Ver por ejemplo Schultz (1978) y Bowen y Finegan (1969, pág. 98).

⁵⁶ Además de los efectos directos, los hijos afectan la situación laboral presente de sus madres por otras vías que operan sobre un horizonte temporal más extenso. Estos efectos indirectos son consecuencia de que factores como la acumulación de experiencia laboral son afectados a lo largo del tiempo en respuesta a los efectos directos del pasado, presente y futuro. Los efectos indirectos mayormente se reflejan en el salario al que una mujer puede aspirar y al tipo de elección ocupacional que realiza. Sobre este punto ver Nakamura y Nakamura (1992).

Para el caso de participación laboral femenina, nada garantiza que las variables capaces de explicar el número de hijos –edad de la mujer, nivel educativo, nivel educativo de su cónyuge, etc.– sean exógenas, lo que podría agravar en lugar de aliviar el problema de correlación entre el término de error y los regresores de la ecuación de participación laboral. Otro problema de este método, es que su desempeño depende crucialmente del ajuste del modelo estimado en la primera etapa y que para ecuaciones que explican el número de hijos el ajuste suele ser muy bajo –Nakamura y Nakamura (1992).

La presunción de que los determinantes del número de hijos, por ejemplo, están también correlacionados con características inobservables que afectan la participación laboral de las mujeres, es lo que llevó a que numerosos investigadores buscaran a lo largo de los años instrumentos exógenos para el número de hijos. Rosenzweig y Wolpin (1980 a, b) instrumentaron el número de hijos explotando la naturaleza exógena de los nacimientos de mellizos. Más recientemente, Angrist y Evans (1998) mostraron que la probabilidad de que una mujer tenga un tercer hijo es mayor cuando los dos primeros son del mismo sexo. Este fenómeno estaría reflejando una preferencia por la “mezcla de géneros” de los hijos, dando lugar a una variación exógena en la fecundidad de mujeres con al menos dos hijos. A partir de esta evidencia, Angrist y Evans instrumentaron el número de hijos en las ecuaciones de participación laboral para mujeres con al menos dos hijos, usando la combinación de géneros en los dos primeros nacimientos.⁵⁷ Con ambos instrumentos –mellizos y mezcla de sexos– se encuentra que al ignorar la endogeneidad del número de hijos se sobrestima su efecto negativo sobre la participación laboral de las mujeres.

Un problema con este tipo de instrumentos es que son aplicables sólo a un subgrupo de mujeres. Por ejemplo, para el instrumento usado por Rosenzweig y Wolpin se consideran sólo mujeres con al menos un hijo y para el de Angrist y Evans mujeres con al menos dos hijos.

En la aplicación que aquí se presenta el modelo de interés viene representado por una ecuación de participación laboral de mujeres. La variable dependiente de esta ecuación es el indicador binario *trabaja* que vale 1 para mujeres que trabajan y cero en caso contrario. Como regresores en este modelo se incluyen edad, nivel educativo, situación laboral del esposo, indicador de asistencia actual a establecimiento educativo y número de hijos menores 6 años. La ecuación auxiliar modela al número de hijos de menos de 6

⁵⁷ Posteriormente, Cruces y Galiani (2003) generalizaron la aplicación de este instrumento usando datos para Argentina y México.

años en función de la edad y educación de la madre, de la educación del padre y del tipo de unión entre ellos –legal o consensual.⁵⁸

Los datos provienen de la Encuesta Permanente de Hogares (EPH) y se utiliza información para los 29 aglomerados disponibles de todo el país correspondiente al mes de octubre de 2002. La EPH presenta dos limitaciones que condicionan la definición de las variables de interés. En primer lugar, no es posible identificar relaciones madre-hijo que no involucren a una madre que sea jefa de hogar o cónyuge del jefe. En el caso de que la mujer sea cónyuge del jefe del hogar, tampoco se sabe si los hijos del jefe también son hijos de ella –por ejemplo, casos de segundos matrimonios. En segundo lugar, como la EPH sólo recaba información sobre los individuos que habitan en el hogar al momento de la entrevista, no es posible reconstruir el número total de hijos de una mujer sino solamente el número de hijos que actualmente viven con ella dentro del mismo hogar.⁵⁹

Consecuentemente, la población de interés está compuesta por mujeres que son jefas de hogar o cónyuges del jefe, en unión legal o consensual con sus cónyuges –o *casadas* para abreviar. A partir de este grupo de mujeres se definen dos muestras. La *muestra completa* incluye mujeres casadas de 21 a 46 años –un total de 7750 mujeres– y la *muestra joven* se compone de mujeres casadas de 21 a 35 años –4051 mujeres.⁶⁰ La variable endógena del modelo –*hijos6*– se define como el número de niños menores de 6 años que habitan en el hogar y son hijos de la mujer o de su cónyuge.

El modelo de participación laboral de mujeres se estima para las dos muestras usando alternativamente tres estimadores: el estimador MV de un modelo probit estándar, el estimador 2EMEA y el estimador 2EMEM. En los dos últimos casos se reportan los errores estándar obtenidos mediante técnicas de bootstrap. Los resultados se presentan en la Tabla 9 del Apéndice. Cuando se usa el estimador 2EMEA, es decir, se emplea la especificación aditiva para el modelo del número de hijos y se lo estima por NLS en la primera etapa, se encuentra que el efecto de *hijos6* sobre la participación laboral de las

⁵⁸ Tanto en la ecuación auxiliar como en el modelo de participación laboral se incluyen términos cuadráticos para edad y controles geográficos –cinco *dummies* regionales. La educación se incorpora mediante seis indicadores: primaria incompleta (omitida), primaria completa, secundaria incompleta, secundaria completa, superior incompleta y superior completa. Naturalmente esta especificación está sujeta a las críticas al primero de los enfoques de IV discutidos más arriba.

⁵⁹ Una fuente de información para Argentina que no sufre de estas limitaciones es el Censo de Población y Viviendas que se realiza cada 10 años. Los datos del último censo realizado en 2001 aún no están disponibles.

⁶⁰ El rango etario de la muestra completa es empleado por ejemplo por Nakamura y Nakamura (1985), mientras que Angrist y Evans (1998) estiman el modelo para mujeres entre 21 y 35 –muestra joven.

mujeres no es significativo y que no puede rechazarse la hipótesis nula de exogeneidad de esa variable. Los resultados de emplear el estimador 2EMEM, por el contrario, indican un efecto negativo y significativo del número de hijos sobre la probabilidad de trabajar de las mujeres. La magnitud del efecto es similar para las dos muestras y la diferencia con respecto al efecto estimado por un probit estándar no es estadísticamente significativa, aunque para la muestra completa el efecto estimado por 2EMEM es más negativo. En el único caso en que parece haber evidencia de endogeneidad del número de hijos es cuando se emplea el estimador 2EMEM sobre la muestra completa.

Nuevamente, determinar cuál de las especificaciones se adapta mejor a los datos demandaría un análisis que escapa a los objetivos de este trabajo.

7. Conclusiones

En este trabajo se discutió la estimación de un modelo de elección binaria donde una de las variables explicativas es endógena y entera. El único trabajo –al menos del único que tengo conocimiento– que se ha concentrado en la estimación de un modelo con esta misma estructura es Weiss (1999), que propone una especificación particular para este modelo en un contexto de máxima verosimilitud. La búsqueda de estimadores completamente eficientes conlleva la necesidad de adoptar un conjunto de supuestos ciertamente arbitrarios y una considerable dificultad computacional.

En este trabajo se exploró una alternativa de estimación consistente y de más sencilla aplicación para un modelo que responde a la misma estructura general que el modelo de Weiss, es decir, un modelo de elección binaria con regresor entero endógeno. La especificación particular que se adoptó –no más arbitraria que la propuesta por Weiss– surge de considerar que los modelos para variables enteras son en esencia modelos de regresión exponencial. Se propusieron así dos estimadores máximo verosímiles condicionales que se obtienen en dos etapas: el estimador 2EMEA cuando la especificación del término de error en el modelo de la variable entera es aditiva y el 2EMEM cuando es multiplicativa. Estos estimadores son consistentes aún cuando los errores en la ecuación de la variable entera no son normales y su cómputo es muy simple a partir de cualquiera de los paquetes econométricos disponibles.

Los experimentos de Monte Carlo, por su parte, mostraron un buen desempeño –en términos de sesgo y RMSE– de estos estimadores en muestras chicas, aún con grados de endogeneidad considerables y ante distribuciones de los errores en la ecuación de la variable entera significativamente alejadas de la normal. También los experimentos de

Monte Carlo mostraron en qué medida la bondad del ajuste del modelo de la variable endógena es un importante determinante del desempeño del estimador en dos etapas.

Por último, se presentaron dos aplicaciones empíricas. La primera es el mismo ejemplo estudiado por Weiss (1999): un modelo de aprobación de solicitudes para tarjetas de crédito. Al comparar resultados, sorprende la similitud de los obtenidos por Weiss a partir del estimador máximo verosímil de su modelo con los que surgen de emplear el estimador en dos etapas 2EMEM, considerablemente más fácil de implementar. La segunda aplicación empírica consiste en un modelo de participación laboral femenina donde se considera que el número de hijos menores de 6 años podría ser endógeno. Sin embargo, no se encuentra evidencia contundente que permita rechazar la hipótesis nula de exogeneidad de esa variable y consecuentemente los coeficientes estimados usando el estimador MV de un probit estándar prácticamente no difieren de los obtenidos por los métodos en dos etapas.

Varios aspectos quedan pendientes para investigaciones futuras. La formalización de algunos de los resultados obtenidos experimentalmente y la derivación analítica de las matrices de varianzas y covarianzas de los estimadores en dos etapas –estimadas aquí empleando técnicas de bootstrap- es un paso natural, a la luz de los auspiciosos resultados experimentales de este trabajo.

Referencias

- Angrist, J. D. and W. N. Evans (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *American Economic Review*. Vol.88, No. 3, pp. 451-477.
- Bowen, W. G. and T. A. Finegan. (1969). *The Economics of Labor Force Participation*. Princeton University Press.
- Browning, M. (1992). Children and Household Economic Behavior. *Journal of Economic Literature*. Vol. XXX, pp. 1434-1475.
- Cameron, A. C. and P. K. Trivedi (1998). *Regression Analysis of Count Data*. Cambridge University Press.
- Cruces, G. and S. Galiani (2003). Generalizing the Causal Effect of Fertility on Female Labor Supply. Mimeo.
- Dooley, M. D. (1982). Labor Supply and Fertility of Married Women: An Analysis with Grouped and Individual Data from the U.S. Census. *Journal of Human Resources*, 17, pp. 499-532.
- Greene, W. H. (1994). Accounting for excess zeros and sample selection in poisson and negative binomial regression models. *Working Paper EC-94-10*, Department of Economics, New York University.
- Hotz, V.J. and R. A. Miller. (1988). An Empirical Analysis of Life Cycle Fertility and Female Labor Supply. *Econometrica*, 56, pp. 91-118.
- Lee, L. F. (1981). Simultaneous equation models with discrete and censored dependent variables. En: C. Manski y McFadden, eds. *Structural analysis of discrete data with economic applications*. MIT Press, Cambridge, MA.
- Maddala, G. S (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge University Press.
- Moffitt, R. (1984). Profiles of Fertility, Labour Supply and Wages of Married Women: A Complete Life-Cycle Model. *Review of Economic Studies*, 51, pp. 263-278.
- Mullahy, J. (1997). Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior. *The Review of Economics y Statistics*. Vol. 79, No. 4.
- Nakamura, A. and M. Nakamura (1985). *The Second Paycheck: A Socioeconomic Analysis of Earnings*. Academic Press.
- Nakamura, A. and M. Nakamura (1992). The Econometrics of Female Labor Supply and Children. *Econometric Reviews*, Vol. 11, No. 1, pp. 1-71
- Rivers, D. and Q. H. Vuong. (1988). Limited Information Estimators and Exogeneity Tests for Simultaneous probit Models. *Journal of Econometrics*. Vol 39, pp. 347-366.
- Rosenzweig, M. R. and K. I. Wolpin (1980a). Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment. *Econometrica*. Vol. 48, No. 1, pp. 227-240.

- Rosenzweig, M. R. and K. I. Wolpin (1980b). Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models. *Journal of Political Economy*. Vol. 88, pp. 328-348.
- Schultz, T. P. (1978). The Influence of Fertility on Labor Supply of Married Women: Simultaneous Equation Estimates. In R. Ehrenberg (ed.). *Research in Labor Economics*, Vol. 2, JAL Press, pp. 273-351.
- Schultz, T. P. (1978). The Influence of Fertility on Labor Supply of Married Women: Simultaneous Equation Estimates". En R. Ehrenberg (ed.). *Research in Labor Economics*. Vol. 2, JAL Press, pp. 273-351.
- Smith, R. and Richard W. Blundell (1986). An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply. *Econometrica*, Vol. 54, No.3, pp. 679-685.
- Weiss, A. (1999). A Simultaneous Binary Choice / Count Model with an Application to Credit Card Approvals. In *Cointegration, Causality, and Forecasting: A Festschrift in honour of Clive W. J. Granger*. Oxford University Press.
- Windmeijer, F. A. G. and J. M. C. Santos Silva (1997). Endogeneity in Count Data Models: An Application to Demand for Health Care. *Journal of Applied Econometrics*, Vol. 12, pp. 281-294.
- Wooldridge, J. M. (1997). Quasi-Likelihood Methods for Count Data. *Handbook of Applied Econometrics*, Vol. 2, Chapter 8.
- Wooldridge, J. M. (1992). Some Alternatives to the Box-Cox Regression Model. *International Economic Review*. Vol. 33, No. 4, pp. 935-955.

Apéndice al Capítulo III

Tabla 1
Efecto de aumentos en la correlación

Parámetros			Estimador	δ			b_{11}			b_{12}			R^2
θ	ρ	σ_2		Coef.	Sesgo	RMSE	Coef.	Sesgo	RMSE	Coef.	Sesgo	RMSE	
0.0	0.0	1.0	Probit	1.015	0.015	0.003	-2.031	-0.031	0.007	1.020	0.020	0.004	-
			2EMEA	1.017	0.017	0.003	-2.035	-0.035	0.007	1.022	0.022	0.004	0.95
0.1	0.1	1.0	Probit	1.044	0.044	0.004	-2.074	-0.074	0.007	1.019	0.019	0.004	-
			2EMEA	1.017	0.017	0.003	-2.033	-0.033	0.007	1.020	0.020	0.004	0.95
0.5	0.4	1.0	Probit	1.179	0.179	0.009	-2.305	-0.305	0.015	1.042	0.042	0.005	-
			2EMEA	1.017	0.017	0.003	-2.034	-0.034	0.007	1.021	0.021	0.004	0.95
1.0	0.7	1.0	Probit	1.348	0.348	0.016	-2.629	-0.629	0.029	1.081	0.081	0.006	-
			2EMEA	1.022	0.022	0.004	-2.046	-0.046	0.008	1.024	0.024	0.005	0.95
2.0	0.9	1.0	Probit	1.530	0.530	0.024	-3.086	-1.086	0.050	1.095	0.095	0.007	-
			2EMEA	1.022	0.022	0.004	-2.052	-0.052	0.010	1.026	0.026	0.005	0.95

Notar: los valores presentados en las filas rotuladas *probit* corresponden a las estimaciones obtenidas usando el estimador máximo verosímil del modelo probit estándar. A fines comparativos se reescalan esos coeficientes multiplicándolos por el desvío estándar de $u_1 = (1-\rho^2)^{1/2}$.

Tabla 2
Efecto del ajuste del modelo estimado en la primera etapa

Parámetros				Estimador	δ			b_{11}			b_{12}			R^2
θ	ρ	σ_2	b_{22}		Coef.	Sesgo	RMSE	Coef.	Sesgo	RMSE	Coef.	Sesgo	RMSE	
1.0	0.7	1.0	1.5	Probit	1.345	0.345	0.016	-2.631	-0.631	0.030	1.085	0.085	0.007	-
				2EMEA	1.027	0.027	0.004	-2.059	-0.059	0.009	1.031	0.031	0.005	0.998
1.0	0.7	1.0	1.0	Probit	1.348	0.348	0.016	-2.629	-0.629	0.029	1.081	0.081	0.006	-
				2EMEA	1.022	0.022	0.004	-2.046	-0.046	0.008	1.024	0.024	0.005	0.952
1.0	0.7	1.0	0.8	Probit	1.434	0.434	0.020	-2.755	-0.755	0.035	1.097	0.097	0.006	-
				2EMEA	1.018	0.018	0.003	-2.037	-0.037	0.007	1.018	0.018	0.004	0.861
1.0	0.7	1.0	0.6	Probit	1.997	0.997	0.045	-3.381	-1.381	0.062	1.227	0.227	0.011	-
				2EMEA	1.013	0.013	0.005	-2.036	-0.036	0.008	1.022	0.022	0.004	0.671
1.0	0.7	1.0	0.4	Probit	2.614	1.614	0.073	-3.723	-1.723	0.078	1.373	0.373	0.017	-
				2EMEA	1.004	0.004	0.011	-2.022	-0.022	0.010	1.020	0.020	0.004	0.422
1.0	0.7	1.0	0.2	Probit	2.832	1.832	0.082	-3.586	-1.586	0.072	1.427	0.427	0.020	-
				2EMEA	0.898	-0.102	0.038	-1.965	0.035	0.021	1.017	0.017	0.004	0.227
1.0	0.7	1.0	0.01	Probit	2.863	1.863	0.084	-3.391	-1.391	0.063	1.432	0.432	0.020	-
				2EMEA	2.353	1.353	2.428	-2.467	-0.467	0.926	1.015	0.015	0.004	0.124

Notar: los valores presentados en las filas rotuladas *probit* corresponden a las estimaciones obtenidas usando el estimador máximo verosímil del modelo probit estándar. A fines comparativos se reescalan esos coeficientes multiplicándolos por el desvío estándar de $u_1 = (1-\rho^2)^{1/2}$.

Tabla 3
Errores uniformes en la ecuación de la variable endógena

Parámetros			Estimador	δ			b_{11}			b_{12}			R^2
θ	ρ	σ_2		Coef.	Sesgo	RMSE	Coef.	Sesgo	RMSE	Coef.	Sesgo	RMSE	
0.0	0.0	1.0	Probit	1.012	0.012	0.003	-2.025	-0.025	0.007	1.017	0.017	0.004	-
			2EMEA	1.014	0.014	0.003	-2.028	-0.028	0.007	1.019	0.019	0.004	0.95
0.1	0.1	1.0	Probit	1.045	0.045	0.004	-2.077	-0.077	0.008	1.017	0.017	0.004	-
			2EMEA	1.015	0.015	0.003	-2.031	-0.031	0.007	1.019	0.019	0.004	0.95
0.5	0.4	1.0	Probit	1.193	0.193	0.009	-2.320	-0.320	0.016	1.031	0.031	0.005	-
			2EMEA	1.019	0.019	0.004	-2.040	-0.040	0.008	1.022	0.022	0.005	0.95
1.0	0.7	1.0	Probit	1.360	0.360	0.017	-2.624	-0.624	0.029	1.030	0.030	0.005	-
			2EMEA	1.023	0.023	0.004	-2.050	-0.050	0.008	1.023	0.023	0.005	0.95
2.0	0.9	1.0	Probit	1.527	0.527	0.024	-3.039	-1.039	0.047	0.967	-0.033	0.006	-
			2EMEA	1.021	0.021	0.004	-2.051	-0.051	0.010	1.020	0.020	0.005	0.95

Notar: los valores presentados en las filas rotuladas *probit* corresponden a las estimaciones obtenidas usando el estimador máximo verosímil del modelo probit estándar. A fines comparativos se reescalan esos coeficientes multiplicándolos por el desvío estándar de $u_i = (1-\rho^2)^{1/2}$.

Tabla 4
Ajuste del modelo de la variable endógena con errores uniformes

Parámetros				Estimador	δ			b_{11}			b_{12}			R^2
θ	ρ	σ_2	b_{22}		Coef.	Sesgo	RMSE	Coef.	Sesgo	RMSE	Coef.	Sesgo	RMSE	
1.0	0.7	1.0	1.5	Probit	1.357	0.357	0.017	-2.621	-0.621	0.030	1.037	0.037	0.006	-
				2EMEA	1.028	0.028	0.005	-2.063	-0.063	0.010	1.035	0.035	0.006	0.99
1.0	0.7	1.0	1.0	Probit	1.360	0.360	0.017	-2.624	-0.624	0.029	1.030	0.030	0.005	-
				2EMEA	1.023	0.023	0.004	-2.050	-0.050	0.008	1.023	0.023	0.005	0.95
1.0	0.7	1.0	0.8	Probit	1.429	0.429	0.020	-2.722	-0.722	0.033	1.034	0.034	0.004	-
				2EMEA	1.018	0.018	0.004	-2.041	-0.041	0.008	1.019	0.019	0.004	0.86
1.0	0.7	1.0	0.6	Probit	2.112	1.112	0.050	-3.502	-1.502	0.068	1.195	0.195	0.010	-
				2EMEA	1.014	0.014	0.005	-2.038	-0.038	0.008	1.021	0.021	0.004	0.67
1.0	0.7	1.0	0.4	Probit	2.666	1.666	0.075	-3.755	-1.755	0.079	1.374	0.374	0.018	-
				2EMEA	1.014	0.014	0.011	-2.029	-0.029	0.011	1.019	0.019	0.004	0.42
1.0	0.7	1.0	0.2	Probit	2.841	1.841	0.083	-3.596	-1.596	0.072	1.427	0.427	0.020	-
				2EMEA	0.920	-0.080	0.040	-1.981	0.019	0.022	1.017	0.017	0.004	0.23
1.0	0.7	1.0	0.0	Probit	2.870	1.870	0.084	-3.403	-1.403	0.064	1.433	0.433	0.020	-
				2EMEA	-1.909	-2.909	3.429	-0.916	1.084	1.324	1.016	0.016	0.004	0.12

Notar: los valores presentados en las filas rotuladas *probit* corresponden a las estimaciones obtenidas usando el estimador máximo verosímil del modelo probit estándar. A fines comparativos se reescalan esos coeficientes multiplicándolos por el desvío estándar de $u_1 = (1-\rho^2)^{1/2}$.

Tabla 5
Potencia del test de exogeneidad

				$R^2 = 0.67$ (promedio en 500 replicaciones)					$R^2 = 0.95$ (promedio en 500 replicaciones)					
Parámetros			Estimaciones (promedio en 500 reps.)		Frec. rel. de rechazo de $H_0: \theta=0$ Significatividad asintótica			K-S** p-valor	Estimaciones (promedio en 500 reps.)		Frec. rel. de rechazo de $H_0: \theta=0$ Significatividad asintótica			K-S** p-valor
θ	ρ	σ_2	θ	Bootstrapped SD(θ)	1%	5%	10%		θ	Bootstrapped SD(θ)	1%	5%	10%	
0.0	0.0	1.0	0.0067	0.089	0.012	0.046	0.096	0.837	0.0018	0.077	0.004	0.036	0.098	0.531
0.1	0.1	1.0	0.1086	0.091	0.078	0.204	0.336	0.987	0.1022	0.078	0.090	0.274	0.394	0.758
0.5	0.4	1.0	0.5084	0.103	0.992	0.998	1.000	0.988	0.5081	0.086	1.000	1.000	1.000	0.737
1.0	0.7	1.0	1.0168	0.128	1.000	1.000	1.000	0.623	1.0119	0.106	1.000	1.000	1.000	0.656
2.0	0.9	1.0	2.0362	0.204	1.000	1.000	1.000	0.801	2.0492	0.178	1.000	1.000	1.000	0.029

*Test de normalidad Kolmogorov-Smirnov. Los p-valores corresponden a la hipótesis nula que postula que la distribución empírica del estimador de θ es normal.

Tabla 6
 Parte A: Modelo Multiplicativo $-R^2$ alto

Parámetros				Estimador	δ			b_{11}			b_{12}			R^2
θ	ρ	M	σ_2		Coef.	Sesgo	RMSE	Coef.	Sesgo	RMSE	Coef.	Sesgo	RMSE	
0.0	0.0	-0.02	0.20	Probit	1.013	0.013	0.003	-2.025	-0.025	0.006	1.016	0.016	0.004	-
				2EMEA	1.015	0.015	0.003	-2.030	-0.030	0.006	1.018	0.018	0.004	0.96
				2EMEM	1.014	0.014	0.003	-2.029	-0.029	0.006	1.018	0.018	0.004	0.96
0.5	0.1	-0.02	0.20	Probit	1.019	0.019	0.003	-2.041	-0.041	0.006	1.016	0.016	0.004	-
				2EMEA	1.029	0.029	0.003	-2.041	-0.041	0.006	1.017	0.017	0.004	0.96
				2EMEM	1.014	0.014	0.003	-2.028	-0.028	0.006	1.018	0.018	0.004	0.96
1.0	0.2	-0.02	0.20	Probit	1.025	0.025	0.003	-2.057	-0.057	0.007	1.015	0.015	0.004	-
				2EMEA	1.043	0.043	0.004	-2.048	-0.048	0.007	1.012	0.012	0.004	0.96
				2EMEM	1.014	0.014	0.003	-2.027	-0.027	0.006	1.018	0.018	0.004	0.96
2.0	0.4	-0.02	0.20	Probit	1.038	0.038	0.004	-2.090	-0.090	0.007	1.014	0.014	0.004	-
				2EMEA	1.067	0.067	0.005	-2.049	-0.049	0.007	0.992	-0.008	0.004	0.96
				2EMEM	1.015	0.015	0.003	-2.029	-0.029	0.007	1.019	0.019	0.004	0.96
3.0	0.5	-0.02	0.20	Probit	1.050	0.050	0.004	-2.127	-0.127	0.008	1.014	0.014	0.004	-
				2EMEA	1.085	0.085	0.005	-2.033	-0.033	0.007	0.961	-0.039	0.004	0.96
				2EMEM	1.016	0.016	0.003	-2.033	-0.033	0.007	1.020	0.020	0.004	0.96
5.0	0.7	-0.02	0.20	Probit	1.073	0.073	0.005	-2.196	-0.196	0.011	1.012	0.012	0.004	-
				2EMEA	1.096	0.096	0.006	-1.948	0.052	0.007	0.875	-0.125	0.007	0.96
				2EMEM	1.020	0.020	0.004	-2.040	-0.040	0.008	1.021	0.021	0.004	0.96

Parte B: Modelo Multiplicativo - R^2 intermedio

Parámetros				Estimador	δ			b_{11}			b_{12}			R^2
θ	ρ	M	σ_2		Coef.	Sesgo	RMSE	Coef.	Sesgo	RMSE	Coef.	Sesgo	RMSE	
0.0	0.0	-1	1.41	Probit	1.017	0.017	0.003	-2.025	-0.025	0.005	1.018	0.018	0.004	-
				2EMEA	1.019	0.019	0.003	-2.024	-0.024	0.006	1.021	0.021	0.004	0.18
				2EMEM	1.016	0.016	0.004	-2.026	-0.026	0.006	1.020	0.020	0.004	0.33
0.1	0.1	-1	1.41	Probit	1.088	0.088	0.005	-2.124	-0.124	0.008	1.026	0.026	0.004	-
				2EMEA	1.082	0.082	0.005	-2.056	-0.056	0.006	1.021	0.021	0.004	0.18
				2EMEM	1.018	0.018	0.004	-2.030	-0.030	0.006	1.023	0.023	0.004	0.33
0.5	0.6	-1	1.41	Probit	1.501	0.501	0.023	-2.708	-0.708	0.032	1.126	0.126	0.007	-
				2EMEA	1.303	0.303	0.014	-2.051	-0.051	0.007	0.978	-0.022	0.004	0.18
				2EMEM	1.017	0.017	0.004	-2.031	-0.031	0.007	1.021	0.021	0.004	0.33
1.0	0.8	-1	1.41	Probit	2.187	1.187	0.054	-3.637	-1.637	0.074	1.299	0.299	0.015	-
				2EMEA	1.546	0.546	0.025	-1.921	0.079	0.010	0.910	-0.090	0.006	0.18
				2EMEM	1.018	0.018	0.005	-2.040	-0.040	0.008	1.026	0.026	0.005	0.33
2.0	0.9	-1	1.41	Probit	3.615	2.615	0.117	-5.397	-3.397	0.152	1.475	0.475	0.023	-
				2EMEA	2.016	1.016	0.047	-1.634	0.366	0.021	0.780	-0.220	0.011	0.18
				2EMEM	1.025	0.025	0.007	-2.052	-0.052	0.011	1.022	0.022	0.005	0.33

Notar: los valores presentados en las filas rotuladas *probit* corresponden a las estimaciones obtenidas usando el estimador máximo verosímil del modelo probit estándar. A fines comparativos se reescalan esos coeficientes multiplicándolos por el desvío estándar de $u_1 = (1-\rho^2)^{1/2}$.

Tabla 7

Variables independientes- modelo de aprobación de solicitudes para tarjetas de crédito

Income		Ingreso, en diez miles de \$
Age		Edad, en años
Curadd	*	Número de años viviendo en la actual dirección
Exp-inc	**	Gasto promedio mensual dividido por el ingreso anual
Avgexp	**	Gasto mensual promedio con tarjeta de crédito
Ownrent	*	Variable binaria que indica propiedad de la casa
Selfempl	*	Variable binaria, =1 si trabajador por cuenta propia
Depndt	*	Número de personas a cargo del solicitante (dependientes)
Inc-per	*	Ingreso por dependiente
Major		Variable binaria, =1 si tiene una tarjeta de crédito principal
Active	*	Número de cuentas de crédito activas
Accounts	*	Número de cuentas abiertas.

Se mantienen los nombres de las variables que emplea Weiss (1999).

* Se incluye sólo en la ecuación de elección binaria. ** Se incluye sólo en la ecuación de *MDR*. Las demás variables se incluyen en ambos modelos.

Tabla 8
Estimación del modelo de aprobación de solicitudes para tarjetas de crédito.

	MV probit estándar	MV modelo de Weiss*	2EMEA**	2EMEM**
Constante	0.496 (0.196)	0.618 (0.175)	1.233 (1.069)	0.555 (0.231)
Income	0.110 (0.057)	0.089 (0.048)	0.061 (0.290)	0.111 (0.066)
Age	-0.007 (0.005)	-0.004 (0.005)	0.014 (0.025)	-0.003 (0.006)
Curadd	0.000 (0.001)	0.000 (0.001)	-0.002 (0.002)	0.000 (0.001)
Ownrent	0.257 (0.110)	0.194 (0.103)	0.127 (0.322)	0.242 (0.111)
Selfempl	-0.344 (0.168)	-0.264 (0.142)	0.268 (0.510)	-0.342 (0.175)
Depndt	-0.134 (0.073)	-0.106 (0.064)	-0.114 (0.167)	-0.126 (0.078)
Inc-per	0.001 (0.077)	0.01 (0.063)	0.118 (0.218)	0.015 (0.095)
Major	0.265 (0.110)	0.235 (0.104)	0.573 (0.753)	0.248 (0.117)
Active	-0.047 (0.028)	-0.032 (0.016)	-0.050 (0.062)	-0.047 (0.041)
Accounts	0.119 (0.030)	0.087 (0.019)	0.128 (0.069)	0.119 (0.043)
MDR	-0.944 (0.146)	-1.6 (0.099)	-3.327 (1.199)	-1.704 (0.413)
MDR2	0.046 (0.047)	0.095 (0.021)	0.024 (0.117)	0.109 (0.032)
Res. 1° etapa	- -	0.887 (0.069)	2.721 (1.074)	0.095 (0.039)
Obs..	1319	1319	1319	1319
Log-Lik.	-486.36	-1485.73	-95.73	-478.14
Pseudo R2	0.307	ND	0.864	0.319
R2 1° etapa	-	-	0.196	0.023

Errores estándar entre paréntesis. * Resultados obtenidos por Weiss (1999). Errores estándar basados en el estimador BHHH/OPG de la matriz de información. ** Errores estándar obtenidos por técnicas de bootstrap (1000 replicaciones). ND: no disponible en el trabajo de Weiss.

Tabla 9
Estimación del modelo de participación laboral de mujeres

	Mujeres casadas - Todos los Aglomerados - 2002					
	21 a 35 años			21 a 46 años		
	(1)	(2)	(3)	(4)	(5)	(6)
	Probit	2EMEA	2EMEM	Probit	2EMEA	2EMEM
Constante	-2.091 (1.084)	-1.453 (1.193)	-2.080 (1.084)	-1.775 (0.377)	-2.064 (0.791)	-1.729 (0.391)
Edad	0.139 (0.077)	0.136 (0.078)	0.140 (0.077)	0.104 (0.022)	0.111 (0.028)	0.107 (0.023)
Edad cuadrado	-0.002 (0.001)	-0.002 (0.001)	-0.002 (0.001)	-0.001 (0.000)	-0.001 (0.0003)	-0.002 (0.0003)
Prim. Comp.	0.056 (0.094)	0.014 (0.113)	0.054 (0.095)	0.100 (0.061)	0.108 (0.067)	0.094 (0.061)
Sec. Inc.	0.175 (0.094)	0.101 (0.123)	0.172 (0.095)	0.162 (0.062)	0.177 (0.072)	0.152 (0.063)
Sec. Comp.	0.251 (0.094)	0.127 (0.145)	0.247 (0.097)	0.287 (0.062)	0.310 (0.086)	0.274 (0.062)
Sup. Inc.	0.464 (0.108)	0.274 (0.196)	0.457 (0.111)	0.457 (0.075)	0.493 (0.118)	0.436 (0.076)
Sup. Comp.	1.230 (0.104)	1.076 (0.176)	1.225 (0.105)	1.289 (0.068)	1.310 (0.086)	1.276 (0.068)
Empleo esposo	-0.482 (0.061)	-0.484 (0.061)	-0.483 (0.065)	-0.396 (0.042)	-0.396 (0.043)	-0.399 (0.044)
Asiste	-0.390 (0.092)	-0.400 (0.091)	-0.389 (0.096)	-0.210 (0.074)	-0.206 (0.074)	-0.207 (0.075)
Hijos6	-0.149 (0.022)	-0.473 (0.272)	-0.161 (0.036)	-0.103 (0.018)	-0.018 (0.209)	-0.154 (0.031)
Res. 1° etapa		0.328 (0.274)	0.003 (0.008)		-0.087 (0.211)	0.012 (0.006)
Obs.	4051	4051	4051	7750	7750	7750
Log. Lik.	-2529.80	-2528.73	-2529.70	-4907.06	-4906.92	-4904.90
Pseudo R2	0.09	0.10	0.09	0.09	0.09	0.09
R2 1° Etapa		0.63	0.04		0.56	0.22

Errores estándar entre paréntesis. Los de las columnas (2), (3), (5) y (6) fueron obtenidos por técnicas de bootstrap (1000 replicaciones).

COMENTARIOS FINALES

Esta tesis se concentró en estudiar empíricamente la relación entre los ingresos y las características demográficas de los hogares, resaltando el importante papel que estas últimas juegan en la determinación de la distribución del ingreso y consecuentemente en las mediciones de pobreza y desigualdad.

Creo que son dos las principales contribuciones esta tesis. En primer lugar, poner en el centro de la discusión la relevancia cualitativa de las cuestiones demográficas como determinantes de la distribución del ingreso, que típicamente no son tenidas en cuenta en la literatura económica especializada. En segundo lugar, cuantificar el efecto distributivo de las variables demográficas.

Cada uno de los tres capítulos que componen esta tesis abordaron la misma temática desde distintos ángulos, empleando metodologías diversas. En el Capítulo I se presentó evidencia empírica para Argentina sobre la evolución de distintas variables demográficas y su potencial relación con el comportamiento de los indicadores distributivos de pobreza y desigualdad. En particular, se observó que durante las últimas décadas el contraste de tamaños familiares entre quintiles extremos de la distribución del ingreso se ha ido acentuando progresivamente como consecuencia de una expansión del tamaño promedio de las familias de los quintiles más bajos y una contracción para las de los quintiles superiores, con un potencial efecto directo empobrecedor y desigualador.

A partir de la hipótesis del impacto distributivo de estos cambios demográficos durante el periodo de análisis, en el Capítulo II se buscó cuantificar la magnitud de ese efecto directo así como también evaluar la importancia de los efectos indirectos producidos fundamentalmente como consecuencia del ajuste de las decisiones de participación laboral de algunos miembros del hogar. Con ese objetivo se aplicó la metodología de microdescomposiciones econométricas o microsimulaciones, donde a los modelos de determinación de ingresos usuales se incorporaron los determinantes demográficos a través de la modelización de la fecundidad de las familias. Del análisis surgió que si bien los fenómenos demográficos no constituyen la principal explicación de los cambios distributivos que operaron durante las últimas décadas del siglo XX en Argentina, dan cuenta de una parte no despreciable de esos cambios, fundamentalmente durante los años ochenta y vía el canal demográfico directo.

Por último, el Capítulo III se concentró en explorar alternativas de estimación consistentes para una especificación particular del modelo de participación laboral femenina. Considerando que el número de hijos y la decisión de una mujer de trabajar o no se determinan simultáneamente, el modelo debe ser capaz de acomodar dos rasgos

particulares de las variables involucradas: una variable dependiente binaria y un regresor endógeno entero. En este capítulo se explotó el hecho de que la mayoría de los modelos para variables enteras son esencialmente modelos de regresión exponencial. Concentrándose en este aspecto del modelo, y no en la naturaleza entera de la variable endógena, se propusieron estimadores en dos etapas de implementación considerablemente más simple que otras alternativas discutidas en la literatura. Asimismo, esta estrategia de estimación provee un test directo para evaluar la hipótesis nula de exogeneidad. De la evaluación a través de experimentos de Monte Carlo, surgió fuerte evidencia sobre el buen desempeño de los estimadores propuestos.

