

Avances de Investigación

Salud y desarrollo humano

Número de hermanos, orden de nacimiento
y resultados educativos en la niñez:
evidencia en Perú

Víctor Saldarriaga Lescano

1



Avances de Investigación 1

**Número de hermanos, orden de nacimiento
y resultados educativos en la niñez:
evidencia en Perú**

Víctor Saldarriaga Lescano *

* Investigador Asistente del Grupo de Análisis para el Desarrollo (GRADE). El autor agradece la valiosa asesoría de Juan José Díaz y las oportunas sugerencias de Jorge Agüero y Fernando Fernández durante la elaboración del presente estudio. Esta investigación fue realizada con el financiamiento de la beca otorgada a investigadores junior por Think Tank Initiative a través de GRADE. Las opiniones vertidas en este documento, junto con cualquier error u omisión, son de responsabilidad exclusiva del autor. Información para contacto: vsaldarriaga@grade.org.pe.

La serie Avances de Investigación que inaugura el Grupo de Análisis para el Desarrollo (GRADE) busca difundir los resultados en proceso de los estudios que realizan sus investigadores. En concordancia con los objetivos de la institución, su propósito es realizar investigación académica rigurosa con un alto grado de objetividad, para estimular y enriquecer el debate, el diseño y la implementación de políticas públicas.

Las opiniones y recomendaciones vertidas en estos documentos son responsabilidad de sus autores y no representan necesariamente los puntos de vista de GRADE ni de las instituciones auspiciadoras.

Esta publicación se llevó a cabo con la ayuda de una subvención del Centro Internacional de Investigaciones para el Desarrollo, Canadá, bajo la iniciativa Think Tank.

Lima, diciembre de 2011

© Grupo de Análisis para el Desarrollo, GRADE
Av. Grau 915, Barranco, Lima, Perú
Teléfono: 2479988
Fax: 2471854
www.grade.org.pe

Directora de Investigación: Lorena Alcázar
Revisión de texto y cuidado de edición: Carolina Teillier, Alicia Infante
Asistente de edición: Paula Pino V.
Diseño de carátula: Elena González
Diagramación: Impresiones y Ediciones Arteta E.I.R.L.
Cajamarca 239 C, Barranco, Lima, Perú. Teléfono: 247-4305 / 265-5146

Índice

Resumen	7
1. Introducción	9
2. Marco conceptual	13
3. Estrategia de identificación	25
4. Datos y variables utilizadas	33
5. Validez de los instrumentos	41
6. Resultados	47
7. Robustez	57
8. Resultados adicionales	61
9. Posibles mecanismos	65
10. Conclusiones	71
Referencias bibliográficas	75
Tablas	81

RESUMEN

El presente estudio investiga el efecto del número de hijos y el orden de nacimiento en la asistencia a la escuela, en la probabilidad de retraso escolar y en el año o grado normativo para la edad en niños de 6 a 14 años de edad. La estrategia de identificación se basa en el uso de variables instrumentales, tomando los eventos de nacimientos múltiples y preferencias por balance de sexo de los hijos como potenciales instrumentos de identificación. Utilizando datos de las Encuestas Demográficas y de Salud Familiar (ENDES) de los años 1996, 2000, 2004-2008 y 2009, se encuentra que un hijo adicional en familias afectadas por eventos de nacimientos múltiples incrementa la probabilidad de asistencia a la escuela de los hijos mayores entre 3,5 y 4,5 puntos porcentuales. Este efecto es mayor en niños residentes en zonas rurales, en los varones y en aquellos niños que reportan vivir en hogares donde ambos padres se encuentran presentes. No obstante, el efecto es menor para niños con edades más cercanas a los 6 años. No se encuentran efectos significativos en la probabilidad de retraso escolar ni tampoco en el año o grado normativo para la edad de los niños, lo cual sugiere que no existe un vínculo causal entre el tamaño familiar y el progreso escolar de los niños. Adicionalmente, se encuentran efectos diferenciados de acuerdo al orden de nacimiento de los hijos.

1. INTRODUCCIÓN

La interrogante sobre cómo el tamaño familiar afecta la redistribución de los recursos al interior de las familias es una de las más perdurables en la ciencia económica. Durante los últimos cuarenta años, muchos economistas y otros científicos sociales han desarrollado una serie de planteamientos teóricos referentes a la manera en que las familias ajustan su presupuesto ante cambios en las preferencias por fecundidad. El estudio más influyente en este campo de la economía de la familia es el modelo teórico de la interacción entre la cantidad y la calidad de los hijos¹ propuesto por Becker y Lewis en 1973, el cual sugiere que una reducción en el número deseado de hijos debería incrementar la inversión de los padres por cada hijo, de manera tal que la familia en su conjunto acumule mayor capital humano.

Recientemente se han llevado a cabo muchos estudios para mostrar evidencia empírica de que el tamaño familiar tiene un efecto negativo en la inversión en educación, salud y consumo de bienes en general de cada hijo. A pesar de que la gran mayoría de estudios encuentra evidencia de una relación inversa, dos estudios recientes han cuestionado la existencia de una interacción entre la cantidad y calidad de los hijos (Black *et al.* [2005] y Angrist *et al.* [2010]). Los resultados

1 El título original del artículo de Becker y Lewis es “*On the Interaction Between the Quantity and Quality of Children*”. A lo largo del documento se utilizará la traducción exacta del título del artículo, denominando al modelo original de Becker y Lewis como el modelo de “interacción entre la cantidad y calidad de los hijos”. Se recomienda que se entienda el concepto de “calidad” como todo atributo o característica propia de los individuos que genera cierto grado de atracción en los mercados en que participan a lo largo de su ciclo de vida. Un ejemplo específico de este tipo de mercados es el mercado laboral.

encontrados en ambos estudios han sido objeto de motivación para que cada vez más economistas se interesen en el tema y han abierto una discusión en torno a la mecánica de la asignación de recursos al interior de las familias condicionada en el tamaño familiar.

A pesar del gran número de estudios empíricos realizados, pocos de estos se han hecho en países en desarrollo, por lo cual no existe evidencia de cómo invierten los padres en sus hijos ante cambios en el tamaño familiar en países donde los niveles de ingresos son menores, existen restricciones de crédito para un gran porcentaje poblacional, las tasas de fecundidad aún se encuentran, en promedio, por encima de los dos hijos por familia y el sistema educativo no se encuentra totalmente desarrollado. Estos factores mencionados llevan a pensar que es en estos países donde la interacción entre la cantidad y calidad de los hijos debiera de tener mayor importancia al interior de las familias y, por consiguiente, el interés en explorar la mecánica de la asignación de los recursos en el hogar debiera de ser mayor.

En el presente estudio el objetivo principal es responder a la pregunta sobre cómo afecta un hijo adicional a la inversión en educación y los resultados educativos de los hijos mayores en las familias para el caso específico del Perú. La motivación por responder a esta interrogante puede dividirse de acuerdo a su relevancia en la economía de la familia y en temas de interés político. En el primer caso, la motivación radica en mostrar evidencia a favor o cuestionar la existencia de una interacción entre cantidad y calidad formulada en el modelo original de Becker y Lewis, al menos en un país donde el ingreso medio es aún relativamente bajo. Por el lado del interés político, esta investigación permite discutir de qué manera podrían verse afectados los logros educativos (y por consiguiente los estándares de vida) de las generaciones futuras ante cambios demográficos en el país; y por otro lado, abre el debate acerca de la relevancia de la estructura familiar en la focalización de los beneficiarios de los programas de transferencias condicionadas para el caso del Perú.

Para responder a la pregunta de interés, la estrategia de identificación seguida se basa en el uso de variables instrumentales. De manera concreta, se utilizan indicadores de nacimientos múltiples

y de mismo sexo de los dos primeros hijos como fuentes de variación exógena en el número de hijos que tienen las familias, para, de este modo, identificar a través de una estimación por mínimos cuadrados en dos etapas el efecto causal que tiene el número de hijos en variables de inversión (asistencia a la escuela) y resultados educativos (probabilidad de que el niño se encuentre en retraso escolar y edad normativa para el grado o año de educación básica) de niños entre 6 y 14 años de edad. Los datos utilizados en el análisis empírico provienen de las Encuestas Demográficas y de Salud Familiar de los años 1996, 2000, 2004-2008 y 2009. La ventaja de estas encuestas es la riqueza en información acerca del historial de nacimientos que ha tenido la madre y la posibilidad de identificación de los niños presentes en el hogar con la madre entrevistada.

Dada la estrategia de identificación basada en el uso de variables instrumentales y la información que brindan las encuestas utilizadas, los objetivos secundarios del presente estudio son, en primer lugar, validar el uso de los eventos de nacimientos múltiples y preferencias por balance de sexo de los hijos como instrumentos de identificación; y en segundo lugar, depurar el efecto asociado exclusivamente al número de hijos de aquellos factores que surgen por defecto al aumentar el tamaño familiar, como lo son el orden de nacimiento y la distancia temporal entre nacimientos contiguos de los hijos.

De manera resumida, las estimaciones por mínimos cuadrados ordinarios indican que un mayor número de hijos tiene efectos negativos en la probabilidad de asistencia a la escuela, efectos positivos en la probabilidad de retraso escolar y efectos negativos en la edad normativa para el año o grado de educación básica en los hijos mayores de las familias. Estos resultados son consistentes con la hipótesis de restricción de recursos que plantea el modelo de Becker y Lewis.

Sin embargo, al efectuar la estimación por mínimos cuadrados en dos etapas se encuentra que un hijo adicional en familias afectadas por eventos de nacimientos múltiples tiene un efecto positivo que varía entre 3,5 y 4,5 puntos porcentuales en la probabilidad de asistencia a la escuela de los dos hermanos mayores condicional en el número de

hijos que ha tenido la madre. Estos resultados son estadísticamente significativos al menos al 95% y son robustos aun al agregar nuevas variables de control y también ante cambios en la muestra utilizada. No se encuentran efectos significativos en la probabilidad de retraso escolar ni tampoco en la edad normativa para el grado del niño, lo cual es consistente con investigaciones previamente realizadas (Black *et al.* [2005], Cáceres-Delpiano [2006], Agüero y Marks [2008] y Angrist *et al.* [2010]) que sugieren que no existe un efecto causal del tamaño familiar en variables de resultados educativos. Al utilizar el indicador de *mismo sexo de los dos primeros hijos* como instrumento de identificación, los resultados obtenidos parecen ser más confusos. No obstante, no es posible rechazar la hipótesis nula de instrumentos débiles utilizando este instrumento, lo cual implica que las estimaciones en dos etapas en base a las preferencias por balance de sexo de los hijos no son consistentes. Finalmente, los resultados encontrados sugieren la existencia de efectos heterogéneos de acuerdo al orden de nacimiento de los hijos.

El estudio se organiza como sigue. La segunda sección presenta el modelo teórico propuesto por Becker y Lewis y la discusión detallada de los hallazgos en los estudios empíricos previamente realizados. La tercera sección presenta la estrategia de identificación utilizada, seguida por la cuarta sección que describe los datos, la composición de la muestra y las variables utilizadas para el análisis empírico. La quinta sección analiza la validez de los instrumentos utilizados en la estimación por dos etapas. La sexta sección presenta los resultados obtenidos en las regresiones. La séptima y la octava presentan los resultados heterogéneos de acuerdo a distintos subgrupos poblacionales y resultados adicionales utilizando como alternativa otro tipo de variables dependientes, respectivamente. La novena sección observa los posibles mecanismos por los cuales se pueden dar los efectos positivos del tamaño familiar en la asistencia a la escuela. Finalmente, la décima sección expone las conclusiones del documento.

2. MARCO CONCEPTUAL

2.1. Consideraciones teóricas

Para establecer la relación existente entre la calidad y cantidad de los hijos, considérese el siguiente modelo propuesto por Becker y Lewis [1973], donde la familia maximiza su utilidad en función del número de hijos (n), la calidad deseada por hijo (q) y un conjunto agregado de bienes de consumo (y):

$$U = U(n, q, y), \quad (1)$$

donde se asume que cada argumento tiene un efecto positivo en la utilidad de la familia. La restricción de ingreso de la familia puede escribirse como sigue:

$$R = \pi_n n + \pi_q q + \pi_{qn} + \pi_y y, \quad (2)$$

donde R es el ingreso total de la familia. En cuanto a los precios, el componente π_n denota el precio asociado exclusivamente al número de hijos, como los costos de prácticas contraceptivas y cuidado prenatal. De manera similar, π_q es el precio asociado únicamente a la calidad, no a la cantidad. Este componente tiene los atributos de un “bien público” al interior de la familia; es el caso de la vestimenta, libros, computadoras, características de los padres y todo tipo de bienes que pueden ser fácilmente transmitidos y utilizados entre los hijos sin existir costos adicionales a la adquisición de los mismos. Por otro lado, π_y representa el precio del conjunto agregado de bienes

consumidos por la familia. Finalmente, el componente π representa el precio unitario de la calidad deseada para cada hijo. Puede pensarse en este precio unitario asociado a la calidad deseada para cada hijo como la suma ponderada de los precios de todos los bienes que potencian los atributos de los hijos (salud, educación, nutrición, entre otros), tal como lo caracterizan Rosenzweig y Wolpin [1980].

En el modelo de cantidad y calidad propuesto por Becker y Lewis, se asume que los padres invierten equitativamente entre el número de hijos el presupuesto total asignado al capital humano de su progenie. Dicho de otro modo, los padres invierten el mismo monto en calidad sin importar el orden de nacimiento o la distancia temporal existente entre los nacimientos de cada hijo. Asumiendo una solución interior en la optimización restringida de la familia, las condiciones de primer orden pueden representarse de la siguiente manera:

$$\frac{\partial U}{\partial n} = \lambda(\pi_n + \pi q) = \lambda p_n; \quad \frac{\partial U}{\partial y} = \lambda(\pi_q + \pi n) = \lambda p_q; \quad \frac{\partial U}{\partial y} = \lambda \pi_y = \lambda p_y, \quad (3)$$

donde λ es el multiplicador de Lagrange (utilidad que reporta cada unidad monetaria adicional a la familia) en la ecuación de optimización restringida. La idea central del modelo se basa en la forma en que los precios sombra (el costo asociado a un incremento unitario en la cantidad de un bien manteniendo constante la cantidad demandada de los demás bienes) de la cantidad y calidad se encuentran relacionados. El precio sombra de la cantidad de hijos (p_n) se encuentra positivamente relacionado con q , la calidad deseada para los hijos, y el precio sombra de la calidad de los hijos (p_q) se encuentra directamente relacionado con n , el número de hijos. La interpretación económica de la interacción entre la cantidad y la calidad de los hijos es que la calidad deseada es más costosa a medida que el número de hijos aumenta, pues el incremento en la calidad deberá repartirse entre un mayor número de hijos. Similarmente, el costo asociado a tener un mayor número de hijos aumenta a medida que la calidad deseada por hijo sea mayor.

Dos implicancias directas se pueden recoger del modelo. En primer lugar, debido a que el precio sombra asociado al aumento de la cantidad se encuentra relacionado positivamente con la calidad deseada, un aumento en la calidad deseada para cada hijo (manteniendo n , q , π_n , y π_q , constantes) resultaría en una menor *demanda* por hijos. En segundo lugar, por el mismo hecho de que el precio sombra asociado al aumento de la calidad se encuentra relacionado directamente con la cantidad, un aumento en el número de hijos (todos los precios y cantidades constantes) resultaría en un aumento del precio sombra asociado a la calidad deseada y, por consiguiente, una disminución en la demanda por calidad.

Además, existe también un efecto resultante de variaciones de los precios de la cantidad y la calidad. A modo de ejemplo, uno puede suponer que el precio asociado a la cantidad, π_n , aumenta exógenamente debido a una mejora tecnológica en los métodos contraceptivos. Dejando todo lo demás constante, este aumento en el precio asociado a la cantidad resultaría en un aumento inmediato del precio sombra de la cantidad, tal como lo predice $\partial U / \partial n$ en (3), lo cual tendría como consecuencia una reducción del número de hijos que desean las familias. Bajo la misma lógica, una disminución del precio asociado a la calidad deseada, π_q , que puede darse como consecuencia, por ejemplo, de un aumento exógeno de la educación de la madre, resultaría en una disminución del precio sombra de la calidad deseada, tal como predice $\partial U / \partial q$ en (3), lo cual tendría como consecuencia inmediata una mayor demanda de calidad y una menor *demanda* de hijos². En cualquiera de los dos escenarios descritos, esta disminución resultante de la demanda de hijos implica, al mismo

2 Nótese que la educación de los padres muestra una correlación inversa con el número de hijos en la familia. Esta correlación negativa que describe el modelo explica la relación negativa usualmente observada entre la educación y la fertilidad de la mujer (ver DeTray [1974] y Ben-Porath [1973]). El argumento utilizado para sustentar la relación negativa entre la educación y la fertilidad de la mujer se basa en que las madres más educadas tienen una mayor valoración por la educación de sus hijos (invierten más en capital humano). Es por este mismo motivo que se observa también que los padres más educados (lo cual guarda también una relación positiva con la riqueza de la familia) tienen hijos con altos niveles de educación.

tiempo, un aumento de la demanda por calidad, de manera tal que se da un efecto de “retroalimentación” entre las demandas de ambos bienes.

Un escenario extremo podría incluir una variación en el nivel general de los precios, lo cual puede deberse, por ejemplo, a un aumento general en el nivel de salarios. En tal caso, suponiendo que todos los precios aumentan en la misma proporción en que aumentan los salarios y el número de hijos deseados, y la calidad deseada por hijo y el patrón de consumo de las familias se mantienen invariantes, la elección entre cantidad y calidad de los hijos podría resolverse al analizar los precios relativos de ambos bienes. De manera particular, Becker y Lewis asumen que $\pi_n n > \pi_q q$. Más aún, los autores son puntuales al señalar que la cantidad de hijos es un sustituto más cercano, en comparación con la calidad deseada para cada hijo, de los bienes consumidos por la familia, de tal manera que un incremento general del ingreso resultaría en un aumento de la calidad, un aumento del consumo y una disminución del número de hijos en la familia. El argumento que los autores utilizan a favor de su supuesto se basa en que el precio asociado al aumento del tamaño familiar no tiene únicamente un componente nominal, sino que también introduce un costo asociado al tiempo asignado al cuidado pre y post natal de los hijos, lo cual puede tener una mayor valoración para la familia en términos de asignación de tiempo a lo largo del ciclo de vida. Este componente “temporal” es lo que hace que el precio asociado al número de hijos sea relativamente mayor al precio asociado a la calidad deseada por cada hijo. Para este caso específico, una variación en el nivel general de los precios como consecuencia de un aumento en el nivel salarial de la población, tendría como consecuencia una disminución de la *demanda* de hijos y, consiguientemente, un aumento en la calidad deseada.

De manera general, debido a la forma en que interactúan los precios sombra de la cantidad y calidad, un aumento en la demanda de cualquiera de ambos bienes (como consecuencia de una reducción en el precio asociado a dicho bien, por ejemplo) tiene como resultado una disminución en la demanda del otro bien, dejando constante el precio

y la cantidad de los demás bienes consumidos por el hogar. Por este motivo, cualquier análisis empírico que no considere la determinación conjunta que describe el modelo de interacción entre cantidad y calidad podría confundir el “efecto verdadero” de cómo el tamaño familiar afecta la inversión que los padres hacen en cada hijo y los resultados educativos de los mismos.

La siguiente sección describe la estrategia de identificación para determinar la existencia (y dirección) del efecto causal del número de hermanos sobre la inversión y desempeño de los niños. De manera puntual, esta estrategia se basa en movimientos exógenos en el número de hijos, independientes de las preferencias de los padres y de los precios asociados a la cantidad de hijos y la calidad deseada para cada hijo.

2.2. Evidencia empírica previa

En el modelo original de Becker y Lewis, un incremento en el tamaño familiar (como consecuencia de un aumento en los costos de los métodos contraceptivos, por ejemplo) reduce la calidad de cada hijo debido a que un aumento en la cantidad incrementa el precio sombra de la calidad deseada para cada hijo. Se han aportado dos hipótesis para explicar la existencia de la relación negativa entre tamaño familiar e inversión y resultados educativos de los niños. La primera hipótesis corresponde a la de “dilución de recursos” (Powell y Steelman [1990] y van Ejick y de Graaf [1995]), la cual argumenta que el resultado en el desarrollo de los hijos depende directamente de los insumos de tiempo y recursos asignados por sus padres. Así, mientras más hijos haya en la familia, menos cantidad de recursos estarán disponibles para cada hijo y, por consiguiente, el desarrollo de los niños se verá afectado negativamente. La segunda hipótesis corresponde al “modelo de confluencia” (Zajonc y Markus [1975], Zajonc [1976], Zajonc, Markus y Markus [1979] y Zajonc y Mulally [1997]), la cual predice que el desarrollo intelectual de los niños se ve afectado por el grado de madurez intelectual, que se define por

la edad promedio de cada familia. En esta línea, las familias con un mayor número de hijos en edades tempranas tienen un menor grado de madurez intelectual, lo cual afecta negativamente la educación de los niños.

Usualmente, los estudios empíricos han basado su estrategia de identificación en el uso de variables instrumentales para resolver el problema de endogeneidad, a modo de constatar si un aumento exógeno del número de hijos en la familia tiene repercusiones negativas en la inversión en la calidad y los resultados educativos de los hijos mayores. Diversos estudios usan eventos de nacimientos múltiples o preferencias por balance de sexo de los hijos como fuentes de variación exógena en el número de hijos. La determinación “natural” que caracteriza a los eventos de nacimientos múltiples y al sexo de los hijos justifica su uso como potenciales instrumentos para la identificación.

Rosenzweig y Wolpin [1980] encuentran que los niños provenientes de familias con nacimientos múltiples en zonas rurales de la India tienen, en promedio, resultados educativos menos favorables que los niños provenientes de familias donde no se registraron este tipo de nacimientos³. Cáceres-Delpiano [2006] realiza estimaciones en dos etapas en una muestra de niños de Estados Unidos, utilizando eventos de nacimientos múltiples como instrumento de identificación. El autor encuentra efectos negativos del número de hijos en la probabilidad de asistencia a escuelas privadas y efectos positivos de esta misma variable en la probabilidad de que los hijos compartan habitaciones. Ello podría implicar que los padres no dejan de invertir totalmente en la educación de los hijos, sino que (1) reemplazan los bienes de mayor calidad por sustitutos más cercanos o (2) disminuyen la inversión en cada hijo, invirtiendo en bienes que pueden ser compartidos entre los hijos. El autor encuentra también

3 Los autores utilizan la desviación estándar de los años de educación con respecto a la media para la edad. A pesar de que los nacimientos múltiples suponen variaciones exógenas en el tamaño familiar, la variable causal de interés que utilizan los autores es la razón del número de niños que han nacido de embarazos múltiples con respecto al total de niños en las familias. Esta variable introduce efectos confusos debido a que, por construcción, el denominador (número de hijos en la familia) es determinado endógenamente por los padres.

que el tamaño familiar no tiene efectos significativos en variables como el retraso escolar, años de educación y embarazo adolescente, las cuales guardan mayor relación con el progreso escolar y resultados favorables en la etapa de adultez de los hijos. Siguiendo la misma estrategia, Li, Zhang y Zhu [2007] utilizan datos provenientes del Censo Poblacional de China y encuentran una relación negativa entre el tamaño familiar y la asistencia a la escuela y el nivel educativo alcanzado de los niños. Los autores señalan que esta relación negativa es aún mayor en zonas rurales, donde el sistema educativo es más pobre comparado con el de zonas urbanas⁴.

Un instrumento alternativo de identificación se basa en las preferencias de las familias por balance de sexo de los hijos. Este instrumento fue introducido por Angrist y Evans [1998] en un estudio realizado para explicar la participación de las mujeres en la fuerza laboral a partir de variaciones exógenas en la fecundidad. Siguiendo la misma estrategia de identificación, Conley y Glauber [2006] utilizan datos provenientes de los censos de Estados Unidos de 1980 y 1990 para investigar el efecto del tamaño familiar en la probabilidad de asistencia a escuelas privadas y la probabilidad de retraso escolar en los niños, y encuentran que un hijo adicional en la familia reduce significativamente la probabilidad de que los hijos asistan a escuelas privadas; no obstante, no encuentra resultados significativos en indicadores de retraso escolar. Lee [2007] explora las preferencias de los padres por los hijos varones para explicar el número de niños en familias de Corea del Sur, y halla efectos negativos del tamaño familiar en la inversión en educación de los hijos.

En contraste, dos estudios recientes hechos por Black *et al.* [2005] y Angrist *et al.* [2010] han cuestionado la existencia de la interacción entre cantidad y calidad formulada por Becker y Lewis [1973]. Siguiendo un procedimiento de estimación por variables

4 Efectos negativos del tamaño familiar en resultados educativos de los niños han encontrado también por Blake [1981], Hanushek [1992], Iacovou [2001] y Patrinos y Psacharopoulos [1997], este último para el caso del Perú. No obstante, estos estudios asumen que el número de hijos es exógeno en la familia luego de controlar por características de los padres, orden de nacimiento y características de los niños.

instrumentales, Black *et al.* [2005] utilizan tanto los eventos de nacimientos múltiples como el sexo de los dos primeros hijos como instrumentos de identificación aplicados a una muestra de niños de Noruega. Los autores concluyen que no existen efectos del tamaño familiar en la escolaridad de los niños y que la relación negativa encontrada entre el número de hijos y los años de educación de los individuos estaría capturando efectos asociados al orden de nacimiento y no los efectos exclusivos del tamaño familiar. Aplicando la misma estrategia de identificación a una muestra de niños de Israel, Angrist *et al.* [2010] no encuentran evidencia de la existencia de una interacción entre la cantidad y calidad deseada por cada hijo.

Otros estudios que utilizan instrumentos distintos a los comúnmente usados en la literatura son los de Qian [2006] y Agüero y Marks [2008]. Qian [2006] explota la flexibilización de la política de control poblacional en China como instrumento de identificación. En particular, en 1978 el gobierno de China autorizó a las familias a tener un hijo adicional en algunas zonas rurales del país en caso de que el hijo primogénito fuese mujer⁵, y encuentra que un hijo adicional en las familias afectadas por esta autorización del gobierno chino incrementa en casi 16 puntos porcentuales la probabilidad de matrícula del hijo mayor en la escuela⁶. Los argumentos utilizados para defender la relación positiva encontrada se basan en (1) la existencia de una ganancia por economías de escala al interior de la familia donde los hijos comparten los bienes materiales, (2) el beneficio social de los hijos mayores con la llegada de hermanos menores al jugar el rol de maestros, lo cual refuerza su entendimiento de las materias escolares, y (3) la inversión en bienes que pueden ser fácilmente transferibles entre los hijos. De otro lado, Agüero y Marks

5 La política de Un Solo Hijo en China dictamina que las familias no pueden tener más de un solo hijo tanto en zonas urbanas como en zonas rurales del país. Una excepción a esta ley se da en el caso de nacimientos múltiples en las familias.

6 Debido a que la estrategia de identificación seguida implica que el hijo primogénito debe ser mujer para que la familia sea afectada por la autorización de tener un hijo adicional, los resultados obtenidos por Qian [2006] deben interpretarse como un aumento en la probabilidad de matrícula en la escuela únicamente para las hijas mayores.

[2008] utilizan la infertilidad de la madre después del primer parto como instrumento de identificación. La infertilidad de la madre se define como la imposibilidad de poder concebir un hijo después de un año de relaciones sexuales regularmente realizadas sin utilizar método contraceptivo alguno y, por tal motivo, este instrumento puede ser utilizado como un límite superior del número de hijos que puede concebir la familia. Los resultados de los autores señalan que luego de instrumentalizar por infertilidad de la madre posterior al primer parto, no existen resultados estadísticamente significativos en variables asociadas a la educación de los niños.

Sin embargo, algunos estudios presentan limitaciones en cuanto a la estructura de datos utilizados para el análisis empírico y la ausencia de una discusión apropiada acerca de otros mecanismos por los cuales el aumento en el tamaño familiar puede afectar la educación de los niños.

En primer lugar, algunos de los estudios realizados incluyen a individuos adultos dentro de su muestra de análisis (Black *et al.* [2005]). La inclusión de individuos adultos en la muestra puede introducir sesgo en la estimación del parámetro de interés en cuanto a que los individuos pueden invertir en su propio capital humano durante la etapa de adultez, haciendo que las variables dependientes tengan error de medida. A pesar de la posibilidad de capturar efectos de largo plazo del tamaño familiar en la educación de los hijos al incluir a individuos adultos en la muestra utilizada, el modelo original de Becker y Lewis describe la inversión de los padres durante la etapa de niñez de su progenie y, por tanto, un análisis más fino debería incluir únicamente a niños como unidad de análisis, debido a que corresponde a la etapa en la que la inversión en capital humano es una decisión exclusiva de los padres.

En segundo lugar, un aumento del tamaño familiar podría afectar el comportamiento de las familias por canales alternativos. Uno de estos canales es la decisión de las madres de salir de la fuerza laboral ante un aumento exógeno del número de hijos. No obstante, Iacovou [2000] y Jaramillo y Montalva [2010], este último para el caso del Perú, no encuentran efectos significativos de un aumento

exógeno en el número de hijos en la oferta laboral de la madre⁷. Otros posibles resultados originados por la fertilidad han sido también documentados por la literatura económica, como es el caso del trabajo infantil (Rosenzweig y Evenson [1977] y Levy [1985]) y la ruptura marital de los padres (Dahl y Moretti [2004] y Lundberg [2005]).

Tercero, los instrumentos utilizados para la identificación podrían alterar por sí mismos el comportamiento de los padres (Rosenzweig y Wolpin [2000]). Al respecto, Rosenzweig y Zhang [2009] señalan que los nacimientos múltiples pueden afectar la reasignación de los recursos al interior del hogar debido a que los niños que son gemelos o mellizos presentan un menor peso al nacer, haciendo que los padres inviertan más en aquellos niños con mejores dotaciones iniciales⁸ (niños que nacen de partos singulares). Bajo esta hipótesis, el uso de nacimientos múltiples como instrumento de identificación introduciría un sesgo positivo en los resultados educativos de los hijos mayores. Los autores concluyen que controlar por el peso al nacer de los niños remueve el sesgo introducido por la reasignación de los recursos en las familias ante la presencia de nacimientos múltiples y encuentran que la probabilidad de que el hijo mayor asista a la escuela es 14 puntos porcentuales menos en familias con dos hijos en comparación con familias con un solo hijo, manteniendo invariantes las dotaciones iniciales de los hijos.

Por último, existen otro tipo de factores que surgen por defecto al aumentar el tamaño de la familia; concretamente, el orden de nacimiento de los hijos y la distancia temporal entre nacimientos. El orden de nacimiento es importante debido al ciclo económico de la familia al momento del nacimiento de alguno de los hijos. Así, los hijos que nacen en condiciones económicas favorables para la familia

7 Cáceres-Delpiano [2006] encuentra efectos negativos del tamaño familiar sobre la probabilidad de que la madre trabaje.

8 Las dotaciones iniciales comprenden todo tipo de características propias de los individuos que los diferencian del resto de los individuos al interior de una comunidad. Ejemplos de estas características pueden ser la habilidad o las deficiencias físicas o mentales que un individuo pueda tener al momento de nacer. Una mayor profundización con respecto al efecto de las dotaciones iniciales en la inversión por cada hijo, puede verse en Becker y Tomes [1976] y Behrman, Pollak y Taubman [1982].

pueden tener más ventaja en comparación con los hijos que nacen en condiciones opuestas, en cuanto a que los padres tienen una mayor capacidad de inversión en estos hijos. La literatura referente al orden de nacimiento señala también algunas otras hipótesis en las que se mencionan: la dilución de los recursos de la familia, efectos culturales de la familia y efectos biológicos de la madre (ver Hauser y Sewell [1985], Birdsall [1991], Hanushek [1992], Ejrnæs y Pörtner [2004], Black *et al.* [2005] y Dammert [2010]). En cuanto a la distancia temporal entre los nacimientos, Powell y Steelman [1990] señalan que la “densidad” del número de niños contemporáneos puede ser perjudicial para el desarrollo de los hijos en la medida en que puede generar un conflicto en la asignación de recursos por parte de los padres⁹.

En las siguientes secciones se describe la estrategia de identificación y los datos utilizados para una correcta identificación del efecto causal del número de hijos en la inversión y resultados educativos de los niños.

9 También se han investigado otros efectos, como la composición de género de los hijos (Powell y Steelman [1990], Butcher y Case [1994], Hauser y Kuo [1997] y Conley [2000]).

3. ESTRATEGIA DE IDENTIFICACIÓN

El interés central es determinar la existencia de un efecto causal del aumento del tamaño familiar tanto en la inversión por hijo como también en los resultados educativos de los niños. En caso de que exista dicho efecto, uno podría preguntarse, en primer lugar, si la dirección del efecto es congruente con la interacción formulada en el modelo teórico de Becker y Lewis; y, en segundo lugar, en qué magnitud afectaría un hijo adicional a la inversión y los resultados educativos de los niños al interior de las familias.

Formalmente, la relación existente entre el tamaño familiar (número de hijos en la familia) y la inversión y desempeño educativo del niño puede escribirse de la siguiente manera:

$$q_{ik} = \alpha + \gamma n_{ik} + X_{ik} \beta + \varepsilon_{ik}$$

En la ecuación anterior, q_i es una variable de inversión o resultado educativo del i -ésimo niño en una familia donde la madre ha tenido al menos k partos, n_{ik} es el número de hijos en la familia (número de hermanos que tiene el niño más uno), X_{ik} es un conjunto de variables de control y ε_{ik} es un término de error que captura todos los demás factores omitidos. Adicionalmente, α , γ y β representan un conjunto de parámetros de interés que se quiere estimar, donde β es un vector de dimensión $r \times 1$. De manera concreta, el interés del análisis empírico se centra en determinar el parámetro γ , el cual se asume que es común para toda la población y se asocia al número de hijos en la familia. Si se cumple que $Cov(n_{ik}, \varepsilon_{ik}) = 0$, entonces γ identifica el efecto causal de n en q .

Sin embargo, la estimación de una regresión lineal usando mínimos cuadrados ordinarios (MCO) produciría resultados sesgados del parámetro de interés γ debido a la determinación conjunta explicada por el modelo de cantidad y calidad. Se pueden mencionar cuatro canales distintos por los cuales el tamaño familiar y el término de error en la ecuación anterior se encuentran correlacionados. En principio, las familias podrían tener una valoración alta por la educación y al mismo tiempo tener preferencias por un número reducido de hijos, lo cual no es un efecto causal propiamente dicho. En segundo lugar, las tasas de fecundidad han ido disminuyendo en el tiempo, tanto en zonas urbanas como en zonas rurales del país, mientras que la tendencia de la cobertura y el acceso educativo ha ido en aumento, especialmente en la educación básica primaria, donde puede decirse que la asistencia a la escuela alcanza la totalidad. En este caso, una regresión lineal podría capturar efectos seculares en las preferencias por fecundidad y cobertura educativa en vez del efecto causal de interés. Tercero, los padres pueden decidir aumentar el tamaño de la familia en función de la calidad del hijo primogénito. En este sentido, si los padres pueden identificar correctamente la habilidad innata del primogénito, entonces los padres con hijos menos hábiles pueden restringir el número de hijos que desean tener a futuro debido a que anticiparán una “menor calidad” de su progenie y preferirán compensar las *carencias* del primer hijo. Finalmente, como ha sido mencionado en la sección anterior, es predecible que las familias con padres más educados tengan un menor número de hijos. Ello implica que, dada la relación positiva existente entre la educación y el ingreso, las familias más ricas son las que tienen familias menos extensas, por lo cual los resultados favorables tanto en la inversión como en el desempeño educativo de los hijos podrían darse como consecuencia del bienestar de la familia y no por el tamaño de la misma.

Consecuentemente, es necesario un método alternativo para estimar γ . Una alternativa es encontrar una fuente de variación exógena en el número de hijos en la familia para una correcta identificación del efecto causal. Esta fuente de variación exógena permite el uso de variables instrumentales (VI) como metodología

para estimar correctamente γ . Para hacer uso de VI deben cumplirse dos supuestos fundamentales: (1) el instrumento utilizado, m_{ik} , no debe estar correlacionado con el término de error de la ecuación estructural ($Cov(m_{ik}, \varepsilon_{ik}) = 0$) y (2) debe existir correlación entre el instrumento y la variable causal endógena ($Cov(m_{ik}, n_{ik}) \neq 0$)¹⁰. En este estudio se utilizan dos instrumentos distintos que permiten identificar correctamente el efecto causal del tamaño familiar en la inversión y resultados educativos de los hijos.

En primer lugar, Ronsenzweig y Wolpin [1980] analizan la posibilidad de utilizar los nacimientos múltiples (existencia de gemelos o mellizos) como instrumento del número de niños en la familia. La existencia de un nacimiento múltiple al momento de la concepción resulta en un aumento exógeno de un hijo adicional en la familia. Además, los nacimientos múltiples son determinados al azar por la naturaleza y no existe ninguna relación con factores que podrían estar afectando a la inversión y los resultados educativos de los niños por otros canales. De manera intuitiva, los embarazos múltiples se asemejan a una subvención con respecto al precio asociado exclusivamente al número de hijos (π_n). El costo en términos nominales de dicha subvención puede relacionarse a los gastos en esfuerzos contraceptivos que realizarán los padres en los períodos posteriores al nacimiento múltiple: padres con preferencias por un menor número de hijos tenderán a realizar mayores esfuerzos contraceptivos, por lo cual el efecto de la “subvención” ha sido mayor en dichas familias¹¹.

10 Existen, por supuesto, restricciones adicionales, tanto en la estructura de la muestra utilizada como en la relación del instrumento con la variable explicativa endógena, que deben cumplirse para asegurar la correcta identificación del efecto causal. Para mayores referencias, ver Heckman [1997], Deaton [2009], Heckman y Urzúa [2009], Imbens [2010] y Leamer [2010].

11 Los nacimientos múltiples son más comunes en mujeres a las cuales se les han realizado tratamientos de fertilidad o se les han administrado drogas para poder concebir. Si los nacimientos múltiples identificados en la muestra que será utilizada en este estudio guardan relación con tratamientos de fertilidad en las mujeres, el uso de dicho instrumento podría quedar invalidado en la estrategia de identificación que se sigue. La encuesta utilizada no permite identificar si las mujeres han seguido este tipo de tratamientos que posibilitan la

En este caso, el parámetro hallado por VI corresponderá a un “efecto local” tal como lo describen Imbens y Angrist [1994]. Esto debido a que la población afectada por el estímulo del instrumento (nacimientos múltiples) podría tener características distintas en comparación con la población no afectada. No obstante, la aleatoriedad que caracteriza al evento de nacimientos múltiples puede hacer suponer que no existe una selección *ex-ante* de la población afectada por este tipo de embarazos¹².

En esta línea, el evento de nacimientos múltiples podrá ser utilizado para comparar, por ejemplo, cómo cambian la inversión y los resultados educativos del niño mayor comparando el evento de un nacimiento múltiple (gemelos o mellizos) con respecto a un nacimiento singular en el segundo parto de la madre. La misma comparación es posible para los dos hijos mayores al comparar el evento de un nacimiento múltiple con respecto a nacimientos singulares en el tercer parto de la madre. De manera general, es posible utilizar el evento de los nacimientos múltiples en el *k-ésimo* parto para todo hijo nacido

concepción. En el Perú, alrededor del 1,5% de mujeres en edades de 15 a 49 años reportan ser infértiles, y el acceso a tratamientos de fertilidad es bastante costoso (alrededor de US\$ 8 mil) para el caso de la fecundación in vitro. Por estos motivos, resulta plausible asumir que las mujeres infértiles que pueden acceder a tratamientos de fertilidad no representan un grupo considerable en la población de mujeres en edad fértil y, por consiguiente, los resultados obtenidos no estarían siendo sesgados por selección en la muestra.

Para corroborar este supuesto, se realizan estimaciones incluyendo en la muestra únicamente a los niños concebidos antes del año 2000. Los resultados no difieren sistemáticamente de los reportados en las tablas al final del documento y, aunque no se muestran, se pueden facilitar a pedido del lector. Alternativamente, la quinta sección revisa cómo los métodos contraceptivos hormonales (píldoras, inyecciones, espumas, dispositivos intrauterinos, etc.) podrían alterar el ciclo ovulatorio de la mujer y con ello afectar la incidencia de nacimientos múltiples.

- 12 Heckman [1997] observa el rol de la heterogeneidad de la población y la sensibilidad de cómo puede verse modificado el comportamiento de los individuos ante cambios inducidos por las VI. Por ejemplo, los individuos con preferencias por mayor fecundidad tienden a tomar drogas que aumentan su fertilidad, lo cual, se ha comprobado, tiene un efecto positivo en los embarazos múltiples. Un ejemplo contrario puede ser la existencia de abortos selectivos ante el acontecimiento de embarazos múltiples en la familia. Este tipo de cambios podrían invalidar el supuesto de exogeneidad en la manera en que la población que acepta tener un hijo adicional por el estímulo del instrumento es una población, en sí misma, seleccionada (tiene preferencias por un mayor número de hijos).

antes de dicho parto. La idea general se basa en comparar a dos niños (o pares) equivalentes pero que difieran en el número de hermanos que tiene cada uno.

Sin embargo, para incrementar la precisión de los resultados, la muestra que será utilizada incluirá únicamente a aquellos niños para los cuales la madre presenta entre dos y cuatro partos ($k = 2,3,4$). Restringir la muestra para los hijos cuyas madres han tenido entre dos y cuatro nacimientos presenta dos ventajas: (1) asegura que las preferencias por número de hijos sean iguales en familias con nacimientos múltiples y singulares, y (2) se evita que el evento de nacimientos múltiples se presente debido a un mayor número de nacimientos reportados por la madre. Asimismo, restringir el análisis para hijos que han nacido antes del k -ésimo nacimiento tiene la ventaja de: (1) evitar el problema de selección que surge debido a que las familias que han tenido hijos después de un nacimiento múltiple pueden diferir de aquellas que deciden tener más hijos después de un nacimiento singular, y (2) evitar el sesgo “hacia abajo” que se introduce en el orden de nacimiento como consecuencia de un nacimiento múltiple.

En segundo lugar, es posible argumentar que, independientemente de las preferencias por el número de hijos, la calidad deseada y los precios asociados a cada bien, las familias presentan preferencias por un balance de género de sus hijos. Al respecto, Ben-Porath y Welch [1976] documentan que en el censo de Estados Unidos de 1970, alrededor del 56% de familias cuyos dos primeros hijos eran del mismo sexo reportaban tener un tercer hijo. En el Perú, cerca del 50% de las madres cuyos dos primeros hijos son del mismo sexo reportan tener al menos un hijo más. Intuitivamente, el hecho de tener a los dos primeros hijos del mismo sexo, ambos varones o mujeres, implica explotar las preferencias por balance de sexo de los hijos que tienen las familias. En el modelo teórico descrito, este hecho implica un movimiento exógeno en la utilidad marginal de la familia con respecto al número de hijos ($\partial U/\partial n$), pues el sexo de los hijos es determinado al azar por la naturaleza, haciendo que las familias sean más propensas a tener un tercer hijo. En esta línea, un indicador de

“mismo sexo” de los dos primeros hijos puede ser también utilizado como potencial instrumento.

No obstante, el uso de dicho instrumento requiere un supuesto importante: el hecho de que los dos primeros hijos tengan el mismo sexo únicamente afectará las preferencias por el número de hijos que desea tener la familia (todas las demás variables se mantienen constantes)¹³. En este caso, el parámetro hallado se interpretará también como un “efecto local” en la medida en que la población tratada (los que tienen un hijo adicional) se caracteriza justamente por tener preferencias por balance de género en su progeñe.

Debido a que las preferencias por balance de sexo inducen a tener un tercer hijo, este instrumento podrá ser utilizado únicamente en la muestra de niños que reportan tener al menos un hermano adicional, es decir, en los niños provenientes de familias en las que la madre reporta haber tenido dos o más partos.

Habiendo discutido la intuición que valida el uso de cada uno de los instrumentos propuestos y los efectos que, potencialmente, cada instrumento puede capturar, la primera etapa del modelo de VI puede escribirse como sigue:

$$n_{ik} = \varphi + \delta m_{ik} + X_{ik} \beta + u_{ik}$$

En esta ecuación, la variable m_{ik} es una variable dicotómica que toma el valor de 1 si el nacimiento de orden k es múltiple, y de 0 en caso contrario, cuando se utiliza como instrumento la presencia de nacimientos múltiples en la familia; y, alternativamente, tomará el valor de 1 cuando los dos primeros hijos son del mismo sexo, y de 0 en cualquier otro caso, cuando se utilice el indicador de mismo sexo como instrumento del número de hijos. El parámetro δ mide el efecto de los nacimientos múltiples o de la composición del género de

13 Es factible suponer que tener hijos del mismo sexo reduce el precio asociado a la calidad, π_q , en la medida en que los bienes utilizados por los hijos pueden ser fácilmente transferibles (como es el caso de la vestimenta). De ser este el caso, el uso del sexo de los dos primeros hijos estaría violando las restricciones de exclusión requeridas para el uso de variables instrumentales y, por consiguiente, se introduciría sesgo en el parámetro de interés.

los hijos en el número de hijos que tienen las familias. El vector X_{ik} es el conjunto total de variables de control que se utilizan en la ecuación estructural y u_{ik} es un término de error que captura todos los demás factores omitidos.

El uso de nacimientos múltiples como instrumento de identificación requiere a su vez controlar por la edad de la madre tanto en el primer parto como en el k -ésimo parto en la primera etapa de la regresión. Esto debido a que la probabilidad de nacimientos múltiples es creciente con respecto a la edad de la madre, como consecuencia de que la frecuencia de irregularidades en el ciclo ovulatorio de la mujer es mayor a medida que aumenta la edad. Para el caso de la composición por género de los hijos, Angrist y Evans [1998] controlan por el sexo del primer hijo (un indicador de género) a modo de capturar efectos heterogéneos entre hombres y mujeres.

4. DATOS Y VARIABLES UTILIZADAS

4.1. Datos

A partir de la Encuesta Demográfica y de Salud Familiar (ENDES) de los años 1996, 2000, 2004-2008 y 2009 se construye un *pool* de datos que será utilizado para el análisis empírico de este estudio. La ventaja que tiene la ENDES en cuestiones de composición familiar es que incluye un módulo de historial de nacimientos de mujeres entre 15 y 49 años de edad, el cual contiene información acerca del sexo de los hijos nacidos vivos, tipo de parto (singular o múltiple), fecha de cada nacimiento, entre otros datos de interés, sin considerar si los hijos se encuentran o no presentes en el hogar. Asimismo, es posible identificar a los niños presentes en el hogar con sus respectivas madres, lo cual permite añadir información completa acerca de la composición familiar de los niños. Por otro lado, el uso de cortes transversales repetidos, además de añadir observaciones para aumentar la precisión en las estimaciones, permite recoger efectos de trayectorias temporales tanto en el acceso a la educación como en las preferencias por fecundidad de las familias.

La ENDES tiene representatividad a nivel nacional y cuenta con dos cuestionarios específicos, el cuestionario del hogar y el cuestionario individual de mujeres. El cuestionario del hogar incluye preguntas sobre características de la vivienda y características de todos los individuos que residen en la vivienda (miembros y no miembros del hogar). A partir del cuestionario del hogar se selecciona de manera aleatoria a las mujeres de 15 a 49 años (rango de edad fértil) que responderán el cuestionario individual de mujeres. La cantidad de mujeres seleccionadas es proporcional al número de mujeres que

residen en la vivienda y estas pueden ser o no miembros del hogar. El criterio utilizado para la selección aleatoria es la fecha de nacimiento de las mujeres. El cuestionario individual de las mujeres incluye preguntas de antecedentes de las mujeres (lugar de nacimiento, lugar de residencia previa, educación, etc.), historial de partos, antecedentes de fecundidad de la madre, información acerca de la nupcialidad, preferencias por fecundidad, violencia doméstica, entre otros módulos informativos.

Para el análisis empírico se utilizan dos bases de datos: la base de mujeres y la base de niños. La base de mujeres es utilizada básicamente para propósitos de validez de los instrumentos. Por otro lado, la base de niños es utilizada para el análisis del efecto del tamaño familiar en variables de inversión y resultados educativos de los niños.

En el contexto de la validez de instrumentos, la base de mujeres se dividirá en cuatro muestras distintas de acuerdo al instrumento que se quiera validar. Así, para probar la validez de las preferencias de los padres por balance de sexo de sus hijos se utiliza el total de mujeres entrevistadas que alguna vez tuvieron por lo menos un hijo nacido vivo, sin considerar si sigue vivo o falleció después de su nacimiento. El propósito de utilizar esta muestra es verificar si existe o no selección del sexo de los hijos por parte de los padres. Esta muestra se identificará como la “muestra completa” de mujeres.

De otro lado, para probar la aleatoriedad en el evento de nacimientos múltiples, la base de mujeres se divide en tres muestras adicionales. La primera muestra la componen las mujeres que, como mínimo, han tenido dos partos. La segunda muestra la componen las mujeres que han tenido como mínimo tres partos. La última muestra queda compuesta por las mujeres que han tenido como mínimo cuatro partos. Estas muestras serán identificadas como las muestras de mujeres con “2 a más hijos”, “3 a más hijos” y “4 a más hijos”, respectivamente. El motivo de dividir a las mujeres que han tenido dos a más hijos de acuerdo al número de partos se debe a la estrategia de identificación descrita en la sección anterior. De manera concreta, el interés se basa en corroborar la aleatoriedad del evento de nacimientos múltiples en cada parto que ha tenido la mujer. Además,

debido a que la muerte de algún hijo no solamente reduce el tamaño familiar, sino que también sesga “hacia abajo” los efectos de orden de nacimiento de los hijos menores y podría estar correlacionado con variables tales como la educación de la madre y el ingreso del hogar, se excluye de estas tres muestras a las mujeres que reportan haber tenido, por lo menos, un hijo que nació vivo pero que falleció después de su nacimiento. Los detalles de la composición de las muestras de “2 a más hijos”, “3 a más hijos” y “4 a más hijos” y las estadísticas descriptivas de cada una de estas muestras, así como también los de la muestra completa de mujeres, se muestran en las tablas 1 y 3.

La segunda base es la que se utiliza para la estimación del efecto del tamaño familiar en la inversión y los resultados educativos de los niños. Debido a que se quiere estimar el impacto del tamaño familiar en variables educativas, el análisis queda restringido exclusivamente a niños que se encuentran en la transición de educación básica, la cual, por norma, incluye a niños en edades de 6 a 17 años. Sin embargo, la encuesta no provee la variable identificadora de la madre del niño a partir de los 15 años en adelante. Por este motivo el análisis queda restringido a niños de 6 a 14 años de edad. Adicionalmente, es necesario que la madre del niño haya sido entrevistada para poder adherir los datos completos del número de hermanos que tiene el niño (número de hijos que tiene la madre). Este filtro de entrevista de la madre excluye por defecto a todos los niños cuya madre ha fallecido o no se encuentra presente en el hogar al momento de la entrevista¹⁴.

La base de niños se divide en tres muestras distintas de acuerdo al número de partos que reporte la madre del niño. Así, la muestra de “2 a más hijos” la componen todos los niños miembros del hogar que reporten ser los primeros en el historial de nacimientos de la madre (hijos mayores) y cuya madre ha tenido como mínimo dos

14 Considerando los problemas potenciales que pueden surgir por selección debido a la presencia de la madre, se realizan pruebas de diferencia de medias para verificar si ambos grupos (los niños con madre presente y aquellos cuya madre está ausente del hogar) difieren sistemáticamente en características individuales, estado educacional y características del hogar. No se encuentran resultados significativos entre ambas muestras.

partos en su vida. La muestra de “3 a más hijos” la componen todos los niños miembros del hogar que reporten ser los primeros o segundos en el historial de nacimientos de la madre y cuya madre ha tenido como mínimo tres partos en su vida. Finalmente, la muestra de “4 a más hijos” la componen los niños miembros del hogar que reporten ser los primeros, segundos o terceros en el historial de nacimientos de la madre y cuya madre ha tenido como mínimo cuatro partos en su vida. Cinco restricciones adicionales son necesarias para una correcta identificación: (1) se excluye a los niños para los cuales la madre reporta haber dado a luz un hijo que nació con vida pero falleció después de su nacimiento; (2) se excluye a los niños que han nacido en partos múltiples (es decir, no hay gemelos o mellizos en ninguna de las muestras); (3) se excluye a los niños para los cuales la madre reporta que el segundo, tercer o cuarto parto, según corresponda, es de trillizos o cuatrillizos; (4) se excluye de la muestra a los niños que no tengan datos completos en el módulo de información educacional; y (5), para el caso de las muestras de “3 a más hijos” y “4 a más hijos”, es necesario que todos los hijos mayores se encuentren presentes en el hogar al momento de la entrevista¹⁵. La composición y las estadísticas descriptivas de las tres muestras de niños se detallan en las tablas 2 y 4.

Al comparar las características de la madre en la muestra de niños (tabla 4) y las características de las mujeres entrevistadas (tabla 3) condicionando en el número de partos que ha tenido la madre (mujer), se pueden observar ligeras diferencias en el número de hijos promedio, la edad en que la madre o mujer tuvo al primer hijo y el porcentaje de nacimientos múltiples en el cuarto parto. Estas diferencias, aunque no son significativas, aparecen principalmente por la restricción de muestra por fallecimiento de algún hijo en la familia y por la condición de que todos los hijos mayores deben estar presentes en el hogar según corresponda en la muestra. Estas restricciones implican que son las madres más jóvenes -y por ello, con un menor número de hijos- las que están siendo recogidas en la muestra de niños. Tal como ha sido

15 Esta restricción no altera de manera significativa los resultados obtenidos. Las regresiones que incluyen a los hijos provenientes de familias cuyos hijos mayores no están en el hogar no se presentan en este documento; no obstante, se pueden ser facilitar a pedido del lector.

mencionado en el pie de página, estas restricciones no afectan los resultados de las estimaciones.

4.2. Variables utilizadas

Las variables dependientes utilizadas se dividen en dos categorías: variables de inversión, que corresponden exclusivamente a las decisiones de los padres, y variables de resultados educativos (progreso escolar), que resultan de la inversión de los padres más un componente de “esfuerzo” de los niños.

En la categoría de variables de inversión se considera que la asistencia a la escuela es un buen indicador de inversión en capital humano de los hijos que realizan los padres. De este modo, la variable dependiente tomará el valor de 1 si el niño reporta que se encuentra actualmente asistiendo a la escuela, y 0 en caso contrario. Por otro lado, se consideran dos variables de inversión adicionales: una dicotómica que indica si los niños comparten o no habitaciones en la vivienda y una dicotómica que indica si el estado civil de la madre corresponde a la categoría de “divorciada” o “separada”¹⁶. El efecto del número de hijos en la probabilidad de que los niños compartan habitaciones y en la probabilidad de que la madre esté divorciada será analizado en la sección 8.

En la categoría de resultados educativos se consideran dos indicadores de retraso escolar. En primer lugar, se construye un indicador de la edad normativa para el año o grado más alto al que ha llegado el niño. Debido a que el historial de nacimientos de la madre permite

16 Para construir la variable dependiente que reporta si los niños comparten o no una habitación, primero se calcula el número de habitaciones “disponibles” en la vivienda, que se obtiene restando el número de miembros del hogar mayores de edad al número total de habitaciones utilizadas para dormir en la vivienda, considerando la posibilidad de que los padres y otras parejas que habitan en la vivienda puedan compartir una misma habitación. Nótese que este indicador podría ser negativo si el hacinamiento en la vivienda fuera muy alto. En esos casos, el número de habitaciones “disponibles” queda censurado en cero. La variable dependiente (indicador de habitaciones compartidas) toma el valor de 1 si el número de niños en la familia es mayor al número de habitaciones disponibles y 0 en caso contrario.

identificar el mes y el año en que nació el niño, es posible construir una regla de correspondencia con respecto a la edad en la que se empieza la escuela. Como regla general, los niños empiezan la escuela a los 6 años de edad. No obstante, si el niño nació en julio o meses posteriores, entonces habrá empezado la escuela a los 7 años de edad. En este contexto, los niños que reportan tener 8 años de edad y estar en el primer año de primaria tendrán un año de retraso en el indicador de edad normativa. La misma lógica se cumple para edades subsiguientes¹⁷. Por tal motivo es posible cuantificar el efecto del tamaño familiar tanto en la probabilidad de retraso escolar (margen extensivo), que tomará el valor de 1 si el niño tiene al menos un año de retraso, como en los años de retraso escolar (margen intensivo).

La variable explicativa de interés en la ecuación estructural es el número de hijos en la familia. Dependiendo de la muestra utilizada, esta variable tendrá un mínimo de dos hijos, tres hijos ó cuatro hijos. En las tres muestras, esta variable no excede de once hijos por familia. El vector de variables de control incluye la edad y el sexo del niño (1 si es hombre), la edad de la madre al primer parto, la edad de la madre al *k*-ésimo parto (dependiendo de la muestra en la que se efectúe la estimación), los años de educación de la madre, la edad del jefe de hogar, el sexo del jefe de hogar, variables dicotómicas por quintil de riqueza del hogar, un indicador de ámbito geográfico (1 si es rural), variables dicotómicas por departamento de residencia y variables dicotómicas por año de la encuesta. Agregar efectos fijos por año de la encuesta permite recoger efectos asociados a la tendencia en preferencias por fecundidad como también trayectorias educativas en el tiempo.

Dados los objetivos secundarios de la investigación, se incluyen también indicadores de orden de nacimiento como controles adicionales. La inclusión de este tipo de controles incrementa la precisión de los resultados debido a que el efecto del número de

17 Bajo esta regla de correspondencia, el indicador de edad normativa alcanzará un mínimo de -14, el cual corresponde a los niños que tienen 14 años de edad y nunca han asistido a la escuela, y un máximo de 0, el cual corresponde a los niños de cualquier edad que reportan no tener retraso escolar.

hijos queda depurado del efecto de orden que surge por defecto al aumentar tamaño familiar. Con este objetivo, se utilizarán dos tipos de indicadores de orden de nacimiento. En primer lugar, se introducen variables dicotómicas que indican el orden en que nació el niño. En segundo lugar, un indicador de orden de nacimiento que varía en el rango $[0,1]$ se construirá de la siguiente manera:

$$o_i = \frac{k_i - 1}{n_i - 1},$$

de tal modo que el indicador tome el valor de 0 para el caso del hijo mayor y 1 para el caso del hijo menor en la familia. Esta construcción permite ponderar el orden de nacimiento de acuerdo al número de hijos existentes en la familia, dando más peso a las familias con un menor número de hijos. No obstante, la introducción de ambos indicadores en el análisis empírico puede realizarse únicamente en los niños provenientes de familias que han tenido más de dos nacimientos ($k = 3,4$), debido a la composición de la muestra. Finalmente, se introducirán también en las regresiones efectos de distancia temporal entre nacimientos, a modo de incrementar la precisión de las estimaciones. Los efectos que surgen por distancia temporal serán capturados por variables dicotómicas que indican si la distancia temporal entre nacimientos contiguos es menor a dos años, entre tres y cuatro años o más de cuatro años.

5. VALIDEZ DE LOS INSTRUMENTOS

El objetivo de esta sección es revisar si existen canales alternativos al número de hijos en la familia por los cuales los nacimientos múltiples y el sexo de los hijos pueden afectar la inversión y los resultados educativos de los niños. De manera general, el interés se centra en determinar la existencia de selección por fecundidad en las familias que reportan haber tenido hijos gemelos o mellizos y, por otro lado, detectar la existencia de preferencias por sexo de los hijos en las familias.

Para empezar, la tabla A.1 del Anexo A muestra las diferencias en los promedios de las variables características de los niños de acuerdo a los eventos de nacimientos múltiples y mismo sexo de los dos primeros hijos. La segunda, tercera y cuarta columnas de la tabla muestran los promedios de las características de los primogénitos cuya madre no tuvo un nacimiento múltiple en el segundo parto (singular), los promedios de los primogénitos cuya madre tuvo un nacimiento múltiple en el segundo parto (múltiple) y el estadístico t asociado a la diferencia de los promedios, respectivamente. La quinta, sexta y séptima columnas presentan los mismos descriptivos para la muestra de tres a más hijos. La octava, novena y décima columnas presentan las comparaciones para la muestra de cuatro a más hijos y las últimas tres columnas de la tabla presentan las comparaciones diferenciando a los primogénitos de acuerdo al sexo de su segundo hermano (sexo distinto y mismo sexo).

Para la muestra de dos a más hijos se encuentran diferencias en el sexo del niño, sexo del jefe de hogar, quintil de riqueza (quintil inferior), edad de la madre al segundo parto y años de educación de la madre. Un hallazgo interesante tiene que ver con las diferencias en la distancia temporal con el nacimiento subsiguiente: los niños

que reportan tener hermanos gemelos o mellizos suelen estar más distanciados en edad que aquellos con hermanos nacidos en partos singulares. Diferencias similares se encuentran al analizar la muestra de tres a más hijos. En el caso de la muestra de cuatro a más hijos, los niños con hermanos nacidos en partos múltiples parecen ser bastante diferentes a los niños con hermanos nacidos en partos singulares. Estas diferencias sugieren que es necesaria la inclusión de las características del niño, características del hogar y características de la madre como variables de control en el análisis empírico cuando se utiliza a los nacimientos múltiples como instrumento de identificación para suprimir cualquier heterogeneidad observable.

Para el caso del indicador de mismo sexo de los dos primeros hijos, la comparación de promedios sugiere que no hay diferencias significativas entre aquellos niños cuya madre reporta el mismo sexo de sus dos primeros hijos y aquellos niños cuya madre reporta sexos distintos de sus dos primeros hijos.

5.1. Nacimientos múltiples

El uso de VI como estrategia de identificación requiere que el evento de nacimientos múltiples como potencial instrumento no esté correlacionado con características no observables en la familia. Debido a que los embarazos múltiples son, por naturaleza, aleatorios entre las familias, este tipo de correlaciones no pueden hallarse de manera directa. Alternativamente, es posible relacionar la probabilidad de nacimientos múltiples en el segundo, tercer y cuarto parto con características observables como la educación de la madre y el quintil de riqueza del hogar. Por otro lado, es posible considerar la probabilidad de existencia de transmisiones “genéticas” de fertilidad en las familias, por lo cual se verifica si la probabilidad de nacimientos múltiples puede determinarse por la fertilidad de la madre de la mujer. Esto es posible debido a que la ENDES cuenta con un módulo de historial de nacimientos de la madre de la mujer entrevistada (módulo de mortalidad materna). Este módulo reporta

también el mes y año de nacimiento de los hermanos de la mujer entrevistada, lo cual permite identificar la existencia de nacimientos múltiples en el historial de nacimientos de la madre. Finalmente, se verifica si el uso de métodos contraceptivos hormonales puede modificar la ovulación de la mujer, tal que pueda tener un efecto en la probabilidad de nacimientos múltiples.

Las tablas A.2, A.3 y A.4 del Anexo A muestran los resultados obtenidos en las pruebas F de significancia individual y conjunta de las variables explicativas mencionadas en el párrafo anterior, en la probabilidad de nacimientos múltiples para el segundo, tercer y cuarto parto, respectivamente. Se agregan en las regresiones controles adicionales como la edad de la mujer al primer nacimiento, la edad de la mujer al *k*-ésimo parto y variables dicotómicas por año de encuesta. Todas las pruebas F de significancia conjunta para el caso de educación de la madre y quintil de riqueza del hogar y significancia individual para el caso de nacimientos múltiples de la madre y uso de métodos contraceptivos hormonales, no pueden rechazar la hipótesis nula de que los coeficientes obtenidos en las regresiones son estadísticamente iguales a cero. Dados los resultados obtenidos en las pruebas de hipótesis, es posible interpretar que no existe selección o correlación del instrumento de nacimientos múltiples con alguna de las características de la mujer y del hogar señaladas anteriormente.

Sin embargo, el evento de nacimientos múltiples puede afectar por otros canales a los niños. De manera concreta, se considera la existencia de al menos dos canales por los cuales el niño puede verse afectado ante el nacimiento de hermanos gemelos o mellizos: la distancia temporal entre el nacimiento de sus hermanos y las dotaciones iniciales de sus hermanos. En el caso de distancia temporal entre nacimientos, Black *et al.* [2005] encuentran que los hijos mayores tienen menos años de educación si la distancia temporal entre el nacimiento de los dos niños nacidos inmediatamente después del hijo mayor es menor. Al extrapolar esto al caso de los gemelos o mellizos, uno esperaría que este efecto negativo sea mayor debido a que la distancia temporal entre hermanos gemelos o mellizos es igual a cero (la mínima distancia temporal). Por otro lado, como ha

sido mencionado por Rosenzweig y Zhang [2009], los niños que son gemelos o mellizos presentan un menor peso al nacer. Ello implica que, en un escenario de refuerzo de las dotaciones iniciales de los hijos, los padres invertirán más en la educación de los hijos que no son gemelos o mellizos.

A modo de probar ambas hipótesis, se explora si la educación de los niños se ve afectada negativamente con respecto a la inversa de la distancia temporal entre el nacimiento de sus dos hermanos nacidos inmediatamente después¹⁸. Se agregan en la regresión controles adicionales como la edad del niño, el sexo del niño y variables dicotómicas por año de la encuesta. Los coeficientes resultantes son todos negativos y las pruebas F rechazan la hipótesis nula de que los coeficientes son estadísticamente iguales a cero, lo cual implica que sí existe un efecto negativo de la distancia temporal entre el nacimiento de los dos hermanos nacidos inmediatamente después y la educación del hijo mayor. Esto es particularmente interesante debido a que, al utilizar nacimientos múltiples como instrumento de identificación se introduciría un sesgo negativo en la variable causal endógena, de tal modo que, de existir un efecto negativo entre el número de hijos y la asistencia a la escuela o el retraso escolar, entonces los coeficientes resultantes de la estimación por mínimos cuadrados en dos etapas (MC2E) serían un límite inferior del efecto de la cantidad en la calidad.

De otro lado, para probar si los padres refuerzan las dotaciones iniciales de sus hijos, se analiza de qué manera afectan los nacimientos múltiples directamente a la educación de los hijos mayores; es decir, se estima la forma reducida de MC2E, pero agregando como variables de control el número de hijos, edad del niño, sexo del niño, edad de la madre al primer parto, edad de la madre al k -ésimo parto y variables dicotómicas por año de la encuesta, de tal modo que si existiera algún otro efecto del evento de nacimientos múltiples que no esté siendo capturado por el número de hijos en la familia, el coeficiente asociado

18 Para esta regresión se utiliza únicamente la muestra de tres a más hijos y se excluye a todos los niños cuya madre reporta haber tenido nacimientos múltiples en el tercer parto. La unidad de análisis es el hijo mayor.

a esta variable debería de resultar estadísticamente distinto de cero. Como se esperaba, las pruebas F de significancia individual para cada muestra no rechazan la hipótesis nula de que los coeficientes asociados a nacimientos múltiples son iguales a cero en las tres muestras (ver tabla A.5 del Anexo A). Por ello, se interpreta que no existe un efecto que favorece a los hijos mayores debido a reasignación de los recursos al interior de las familias.

En resumen, es posible concluir que no existe sesgo de selección por preferencias de fecundidad en las familias en ninguna de las tres muestras ni existe sesgo por reasignación de recursos del hogar debido al menor peso al nacer de los gemelos o mellizos. No obstante, las regresiones efectuadas señalan que los eventos de nacimientos múltiples introducen un sesgo negativo en la asistencia a la escuela y un sesgo positivo en los indicadores de retraso escolar de los hijos mayores, lo cual debe tomarse en cuenta en el análisis posterior. Esto implica que las estimaciones por MC2E, utilizando el instrumento de nacimientos múltiples, estarían sesgadas hacia encontrar efectos negativos del número de hijos en la familia por defecto.

5.2. Preferencias por balance de sexo

Si las familias tienen preferencias por algún género específico de sus hijos, entonces el uso del indicador de mismo sexo como instrumento de identificación quedaría invalidado en el análisis empírico. A modo de verificar la existencia de preferencias por género se explora de qué manera afecta la razón de hijos varones nacidos vivos con respecto al total de hijos nacidos vivos a la probabilidad de haber tenido algún aborto y, por otro lado, al número de hijos que nacieron vivos pero que murieron antes de cumplir el primer año de edad (mortalidad infantil). Si existiera selección por el sexo de los hijos, puede esperarse que la variable explicativa de interés sea estadísticamente distinta de cero. Adicionalmente, para permitir efectos diferenciados por ámbito geográfico, se divide la muestra de acuerdo al ámbito urbano o rural. Además, se agrega un término de interacción entre la razón de hijos

varones nacidos vivos con respecto al total de hijos nacidos vivos y una dicotómica que toma el valor de 1 si la mujer indica haber vivido en una ciudad durante su niñez. Esta interacción permite recoger efectos diferenciados de acuerdo al ámbito de residencia durante la niñez y al ámbito de residencia actual, a modo de identificar la persistencia de efectos “culturales” en las familias.

La tabla A.6 del Anexo A muestra los resultados de las pruebas F de significancia individual de los coeficientes resultantes. Las pruebas F indican que todos los coeficientes hallados son estadísticamente iguales a cero. Estos resultados son evidencia de que no existen preferencias por el género de los hijos en las familias, tanto en zonas urbanas como también en zonas rurales. Ello implica que no existe sesgo por selección *ex-ante* del sexo de los hijos y, por lo tanto, el uso del indicador de mismo sexo de los dos primeros hijos como instrumento de identificación es válido en el análisis empírico.

6. RESULTADOS

6.1. Tamaño familiar

A partir del análisis de diferencias en los promedios de las características de los niños comparando de acuerdo a los eventos de nacimientos múltiples y mismo sexo de los dos primeros hijos es posible obtener las primeras etapas de la estimación por MC2E.

Al observar la variación generada por los nacimientos múltiples, estos eventos aumentan entre 0,51, 0,56 y 0,48 el número de hijos para las familias con dos a más hijos, tres a más hijos y cuatro a más hijos, respectivamente. Para el caso del indicador de mismo sexo de los dos primeros hijos, la primera etapa muestra una variación de 0,03 hijos adicionales. Todos los estimados son estadísticamente significativos al menos al 95% de significancia.

La tabla 5 presenta los coeficientes obtenidos en la primera etapa de estimación por MC2E de los eventos de nacimientos múltiples y mismo sexo de los dos primeros hijos como variables explicativas del número de hijos en la familia controlando por el vector completo de características sociodemográficas de los niños. El efecto de los nacimientos múltiples (segunda columna) se divide de acuerdo al número de partos que ha tenido la madre: dos a más hijos (filas 1 a 3), tres a más hijos (filas 4 a 6) y cuatro a más hijos (filas 7 a 9). La tercera columna reporta los coeficientes obtenidos con el indicador de mismo sexo de los dos primeros hijos. Los errores estándar agrupados a nivel de la madre se reportan entre paréntesis y los estadísticos R2 de bondad de ajuste se reportan entre corchetes.

En el caso de nacimientos múltiples, los coeficientes puntuales se encuentran en alrededor de 0,80 para las tres muestras. La inter-

pretación de los coeficientes obtenidos es que el evento de nacimientos múltiples en el segundo, tercer y cuarto orden afecta en 0,83, 0,85 y 0,78 hijos adicionales¹⁹, respectivamente, en la familia²⁰. Estos resultados son consistentes con el hecho de que los nacimientos de gemelos o mellizos incrementan de manera exógena un miembro adicional en la familia. Asimismo, los estadísticos *t* resultantes son lo suficientemente grandes, tal que es posible rechazar la hipótesis de instrumentos débiles al considerar el evento de nacimientos múltiples como instrumento de identificación. La prueba de instrumentos débiles de Cragg y Donald corrobora este resultado.

Al comparar entre las tres muestras es posible observar que el efecto de los nacimientos múltiples es mayor en el tercer parto, en comparación con el segundo parto. Sin embargo, el efecto del cuarto parto es menor que aquel del tercer e incluso del segundo parto. El resultado de que los nacimientos múltiples en el tercer parto son mayores a los del segundo parto guarda relación con el hecho de que la muestra de dos a más hijos incluye algunas familias para las cuales el tamaño familiar deseado no se ve afectado por el evento de nacimientos múltiples. Para las familias con al menos dos hijos, los nacimientos múltiples en el segundo parto afectan únicamente el período en el que tendrán a los hijos subsiguientes. Sin embargo, cuando la muestra queda restringida a familias con al menos tres hijos, el evento de nacimientos múltiples en el tercer parto tiende a afectar más el tamaño familiar, haciendo que la familia se acerque (o incluso pase) el tamaño familiar inicialmente deseado. Por otro lado, el menor efecto encontrado en la muestra de cuatro a más hijos puede relacionarse al hecho de que estas familias tienen, por defecto, preferencias por un mayor tamaño familiar.

19 Los resultados obtenidos son similares a aquellos obtenidos por Cáceres-Delpiano [2006] y Black *et al.* [2005].

20 Debido a que la ocurrencia de nacimientos múltiples se encuentra positivamente correlacionado con la edad de la madre (Rosenzweig y Wolpin [1980]) y dado que el promedio de edad al segundo, tercer y cuarto parto de las madres incluidas en la muestra es menor a 30 años, el efecto de los nacimientos múltiples en el número de hijos es mayor al controlar por edad de la madre al primer parto y edad de la madre al *k*-ésimo parto en comparación con las primeras etapas mostradas en la tabla A.1.

Para el caso del indicador de mismo sexo, se estima que en familias con al menos dos hijos, el hecho de que los dos hijos mayores sean del mismo sexo aumenta el tamaño familiar en 0,04 hijos. Aun cuando el parámetro resultante es similar al resultado obtenido por Angrist *et al.* [2010] (0,07; EE= 0,01) y es estadísticamente significativo al 99%, el estadístico t resultante no es lo suficientemente grande como para rechazar la hipótesis de instrumentos débiles²¹. El problema con los instrumentos débiles es que introducen sesgo en la segunda etapa de la estimación por MC2E, incluso aún más que las estimaciones por MCO. Para evitar errores de inferencia, no se comentarán los resultados obtenidos en la segunda etapa de estimación mediante el uso del indicador de mismo sexo como instrumento. No obstante, los coeficientes resultantes serán reportados en las tablas que prosiguen a lo largo del documento.

6.2. Inversión en educación y resultados educativos de los niños

De manera general, las formas reducidas mostradas en la tabla A.1 evidencian que los eventos de nacimientos múltiples en el segundo y tercer parto de la madre tienen un efecto positivo de alrededor de 4 puntos porcentuales en la probabilidad de asistir a la escuela, tanto para los primogénitos en la muestra de dos a más hijos, como para los dos hijos mayores en la muestra de tres a más hijos. En la muestra de cuatro a más hijos, a pesar de que el evento de un nacimiento múltiple en el cuarto parto de la madre tiene un efecto positivo en la probabilidad de asistencia escolar de los tres hijos mayores, este no es estadísticamente significativo. De otro lado, los eventos de nacimientos múltiples no generan una variación estadísticamente significativa en la probabilidad de retraso escolar ni tampoco en el año o grado normativo para la edad del niño en ninguna de las tres muestras. Para el caso de las preferencias por balance de género de

21 La prueba de instrumentos débiles de Cragg y Donald señala que los parámetros resultantes al utilizar el indicador de mismo sexo como instrumento de identificación presentan más de 10% de sesgo con respecto al parámetro poblacional.

los hijos, se encuentra que el indicador de mismo sexo de los dos primeros hijos no genera una variación estadísticamente significativa en ninguna de las tres variables de resultados educativos consideradas.

Los estimadores de Wald se pueden hallar dividiendo las formas reducidas entre las primeras etapas mostradas en la tabla A.1. La tabla 6 muestra los resultados por MCO (sin incluir variables de control) y los estimadores de Wald para cada uno de los resultados educativos considerados, diferenciando de acuerdo al instrumento utilizado. Al analizar los resultados obtenidos por MCO se encuentra que un hijo adicional en la familia disminuye entre 1 y 2 puntos porcentuales la probabilidad de asistencia a la escuela, aumenta entre 12 y 15 puntos porcentuales la probabilidad de retraso escolar y disminuye entre 0,29 y 0,38 años de retraso escolar de los hijos mayores dependiendo del tamaño inicial de la familia. En contraparte, la estimación por VI muestra signos opuestos a los obtenidos por MCO. En la muestra de dos a más hijos se encuentra que un hijo adicional en la familia aumenta en 7,4 puntos porcentuales la probabilidad de asistencia del hijo primogénito a la escuela. En familias con tres a más hijos y cuatro a más hijos los coeficientes puntuales resultan ser 0,066 y 0,021, respectivamente. No obstante, el coeficiente resultante en la muestra de cuatro a más hijos es estadísticamente no significativo. Al considerar los indicadores de retraso escolar y año o grado normativo para la edad del niño, los resultados son significativos únicamente en la muestra de tres a más hijos. En esta muestra, un hijo adicional en la familia disminuye la probabilidad de retraso escolar de los dos primeros hijos en 22 puntos porcentuales y tiene un efecto positivo de 0,496 años en el año o grado normativo para la edad.

Las tablas 7, 8 y 9 muestran los resultados de la estimación por MCO y de la segunda etapa de la estimación por MC2E del número de hijos en las variables dependientes de asistencia a la escuela, probabilidad de retraso escolar y edad normativa para el grado escolar, respectivamente, controlando por el vector completo de características sociodemográficas de los niños.

Los coeficientes obtenidos de las estimaciones por MCO (segunda columna de cada tabla) muestran la existencia de una relación nega-

tiva entre el número de hijos en la familia y las variables de inversión y resultados educativos, tal como lo predice el modelo de Becker y Lewis. Interesantemente, este efecto es creciente (en términos absolutos) con respecto al tamaño familiar: mientras más hijos haya en la familia, el efecto de un hijo adicional tendrá un efecto cada vez más negativo en la asistencia a la escuela y el progreso escolar de los hijos mayores. Así, para el caso de asistencia a la escuela, se encuentra que cada hijo adicional en la familia disminuye la probabilidad de asistencia del hijo mayor en 0,4 puntos porcentuales en familias con al menos dos hijos. En el caso de familias con tres a más hijos y cuatro a más hijos, cada hijo adicional disminuye la probabilidad de asistencia a la escuela de los dos hijos y tres hijos mayores en 0,7 y 1,4 puntos porcentuales, respectivamente. Todos los resultados son significativos al menos al 95%.

Del mismo modo, los resultados obtenidos por MCO predicen que cada hijo adicional aumenta entre 4 y 6 puntos porcentuales la probabilidad de retraso escolar de los hijos mayores, dependiendo del tamaño familiar inicial. Al analizar el margen intensivo, un hijo adicional en la familia tiene como consecuencia un retraso de entre 0,12 y 0,16 años de educación en los hijos mayores, condicional en la edad de los niños. Todos los resultados encontrados para los indicadores de retraso escolar, tanto en el margen extensivo como en el margen intensivo, son significativos al 99%.

Contrariamente a estos resultados, los coeficientes obtenidos al estimar por MC2E (tercera columna de cada tabla) utilizando el evento de nacimientos múltiples como instrumento de identificación, muestran signos opuestos a los resultados obtenidos por MCO, sugiriendo que, de manera general, un hijo adicional en familias con dos a más hijos afectadas por el evento de nacimientos múltiples tiene un efecto positivo en la probabilidad de asistencia a la escuela y efectos negativos en el margen extensivo e intensivo de retraso escolar de los hijos mayores²².

Al analizar el efecto del número de hijos en la probabilidad de asistencia a la escuela, en la muestra de dos a más hijos se obtiene que

22 Resultados similares han sido encontrados por Qian [2006] y Agüero y Marks [2008].

un hijo adicional en familias afectadas por nacimientos múltiples en el segundo parto de la madre tiene un efecto positivo de 3,6 puntos porcentuales en la probabilidad de asistencia a la escuela del hijo mayor, con una significancia estadística del 95%. En esta misma línea, en las muestras de tres a más hijos y cuatro a más hijos se encuentran efectos positivos de 4,6 puntos porcentuales y 0,8 puntos porcentuales en la probabilidad de asistencia a la escuela de los hijos mayores, respectivamente. En la muestra de tres a más hijos el coeficiente resulta ser estadísticamente significativo al 99%; sin embargo, en la muestra de cuatro a más hijos no se puede rechazar la hipótesis nula de que el coeficiente hallado sea estadísticamente igual a cero.

En el caso de la probabilidad de retraso escolar (margen extensivo), los coeficientes puntuales obtenidos utilizando el evento de nacimientos múltiples como instrumento de identificación son decrecientes, variando entre 0,002 y -0,223 de acuerdo al tamaño familiar inicial. Para las tres muestras, las pruebas de significancia individual no rechazan la hipótesis nula de que el coeficiente asociado al número de hijos en la familia es estadísticamente igual a cero. Similarmente, los coeficientes puntuales obtenidos en la estimación por MC2E utilizando como indicador de resultado el número de años de retraso en educación que tiene el niño condicional a su edad (margen intensivo) varían entre -0,034 y 0,16. No obstante, al igual que los resultados del número de hijos en la probabilidad de retraso escolar, estos coeficientes resultan ser estadísticamente no significativos. De acuerdo con los resultados obtenidos en los indicadores de resultados escolares, es posible concluir que un hijo adicional en familias afectadas por eventos de nacimientos múltiples no tiene efecto alguno, tanto en la probabilidad de retraso como en los años de retraso educativo de acuerdo a la edad de los hijos mayores de la familia.

Considerando los resultados obtenidos, la discusión se centra exclusivamente en el efecto del número de hijos en la probabilidad de asistencia a la escuela de los hijos mayores, variable que se define como la más cercana a un indicador de inversión en la calidad de cada hijo. Así, debido a que el análisis empírico se ha realizado considerando

únicamente a los hijos mayores de la familia, es necesario verificar si estos resultados encontrados están capturando otros efectos que surgen al incrementar el número de hijos en la familia; concretamente, el orden de nacimiento y la distancia temporal entre los nacimientos de los hijos.

6.3. Orden de nacimiento y distancia temporal

Como se ha mencionado en la sección de evidencia empírica previa, existen efectos adicionales que surgen al incrementar el número de hijos en la familia. Estos efectos son el orden de nacimiento de cada hijo y la distancia temporal existente entre nacimientos contiguos. Para verificar que los resultados obtenidos utilizando el evento de nacimientos múltiples como instrumento de identificación en la estimación por MC2E no estén capturando efectos de orden de nacimiento y espaciado temporal, se incluyen indicadores que permitan mantener constantes ambos efectos en las regresiones. Específicamente, se agregan variables dicotómicas de acuerdo al orden de nacimiento para las muestras de tres a más y cuatro a más hijos, siendo el hermano mayor (orden de nacimiento igual a 1) la categoría base, y variables dicotómicas de acuerdo a la distancia temporal existente entre el nacimiento del niño evaluado y el nacimiento que continúa (nacimiento del hermano o hermanos menores). Los indicadores de distancia temporal se clasifican de la siguiente manera: menos de dos años, de dos a cuatro años y cuatro a más años, siendo la dicotómica de “menos de 2 años” la categoría base. Alternativamente a las dicotómicas que diferencian por orden de nacimiento, se realizan estimaciones que incluyan el orden de nacimiento relativo; esto es, tomando en cuenta el número de hijos que tiene la familia.

La tabla 10 muestra los resultados de la primera etapa de estimación por MC2E utilizando los eventos de nacimientos múltiples como instrumento de identificación e incluyendo los indicadores de orden de nacimiento y distancia temporal entre nacimientos como variables de control en la ecuación estructural. La segunda y tercera

columna se diferencian en la manera en que el indicador de orden de nacimiento ha sido introducido en la ecuación estructural: la segunda columna incluye dicotómicas de acuerdo al orden de nacimiento para las muestras de tres a más y cuatro a más hijos, mientras que la tercera columna introduce el orden de nacimiento relativo del hijo en la familia. Al comparar los coeficientes reportados en la tabla 10 con los coeficientes reportados en la tabla 6, no se encuentran diferencias significativas entre ambas especificaciones. Estos resultados implican que los coeficientes obtenidos en la estimación por MC2E en la primera etapa se mantienen robustos incluso al introducir controles adicionales que surgen por defecto al aumentar el número de hijos en la familia.

La tabla 11 muestra los coeficientes obtenidos en la segunda etapa de estimación por MC2E del número de hijos en la probabilidad de asistencia a la escuela. En la muestra de dos a más hijos, las estimaciones por MCO se mantienen estables al introducir indicadores de distancia temporal entre nacimientos. Del mismo modo, los coeficientes obtenidos mediante MC2E no se ven afectados por la introducción de la distancia temporal entre nacimientos como control adicional. Nótese que los coeficientes de la tercera y cuarta columnas son exactamente los mismos ya que en esta muestra no es posible introducir efectos de orden de nacimiento debido a que, por construcción, incluye únicamente al hijo mayor de la familia. De manera general, se encuentra que un hijo adicional en familias con dos a más hijos con nacimientos múltiples en el segundo parto de la madre tiene un efecto positivo de 3,5 puntos porcentuales (99% de significancia estadística) en la probabilidad de asistencia a la escuela del hijo mayor. De otro lado, las muestras de tres a más hijos y cuatro a más hijos sí permiten agregar indicadores de orden de nacimiento. Así, en la muestra de tres a más hijos se encuentra que el efecto de un hijo adicional, como consecuencia de nacimientos múltiples en el tercer parto de la madre, tiene un efecto positivo que varía entre 4,4 y 4,7 puntos porcentuales, dependiendo del indicador de orden de nacimiento utilizado, en la probabilidad de que los dos hijos mayores asistan a la escuela. Estos resultados parecen ser muy similares a los resultados mostrados en la tabla 5, lo cual da

cuenta de la robustez de los resultados obtenidos aun al agregar más controles en la regresión. Similarmente, los coeficientes obtenidos son estadísticamente significativos al 99%. Para la muestra de cuatro a más hijos no se encuentran diferencias significativas en los coeficientes al incluir los controles adicionales de orden de nacimiento y distancia temporal entre nacimientos en la ecuación estructural.

Analizando los indicadores de orden de nacimiento en la muestra de tres a más hijos y cuatro a más hijos, es notable la existencia de efectos diferenciados entre hermanos al interior de una misma familia. Así, en la muestra de tres a más hijos es posible notar que el segundo hijo tiene 1,7 puntos porcentuales menos de probabilidad de asistencia a la escuela en comparación con el hijo mayor. En la muestra de cuatro a más hijos, el segundo hijo tiene 2,5 puntos porcentuales menos de probabilidad de asistencia a la escuela, mientras que el tercer hijo tiene 2,9 puntos porcentuales menos de probabilidad de asistencia a la escuela en comparación con el hijo mayor de la familia. Estos resultados, en ambas muestras, son significativos al 99%. A pesar de ello, el indicador alternativo de orden de nacimiento que corrige por el número de hijos en la familia muestra ser estadísticamente no significativo para ninguna de las muestras. Este resultado puede darse debido a la construcción de este indicador: el número de hijos en la familia es endógenamente determinado. Por tanto, incluir dicha variable al mismo tiempo que se incluye el número de hijos como variable endógena causal, puede generar confusión en la interpretación de los parámetros. Bajo este escenario, los coeficientes asociados a las variables dicotómicas que indican el orden de nacimiento del niño en análisis son más fiables. Esta evidencia se contradice con el supuesto del modelo teórico de Becker y Lewis, quienes asumen que los padres invierten de manera equitativa en la calidad de cada uno de sus hijos; y más aún, resulta ser distinta a los resultados obtenidos por Black *et al.* [2005], quienes argumentan que los coeficientes hallados, asociados al número de hijos en la familia, estarían capturando únicamente el orden de nacimiento de los hijos.

Las variables dicotómicas que diferencian de acuerdo a la distancia temporal entre nacimientos de los hijos presentan resultados más

confusos. En la muestra de dos a más hijos los coeficientes asociados a las dicotómicas de distancia entre nacimientos resultan ser positivos y estadísticamente significativos al 99%, lo que significa que una mayor distancia temporal entre los nacimientos es favorable para cada hijo. Sin embargo, al introducir efectos de orden de nacimiento en las muestras de tres a más y cuatro a más hijos, no es posible rechazar la hipótesis de que los coeficientes asociados a la distancia temporal entre nacimientos contiguos sean estadísticamente iguales a cero. Estos resultados, no obstante, deben ser tratados con cautela debido a que la distancia temporal entre uno y otro nacimiento es endógenamente determinada por los padres.

Las tablas 12 y 13 muestran los coeficientes obtenidos del número de hijos en la probabilidad de retraso escolar y los años de retraso escolar condicional en la edad del niño, respectivamente, en la segunda etapa de estimación por MC2E, utilizando como instrumento de identificación los eventos de nacimientos múltiples e incluyendo el orden de nacimiento y distancia temporal entre nacimientos como controles adicionales. De manera general, estos resultados no presentan diferencias significativas en comparación con los resultados de las tablas 7 y 8, los cuales no controlan por orden de nacimiento ni distancia temporal entre nacimientos.

7. ROBUSTEZ

En la tabla 14 se comprueba qué tan sensibles son los resultados obtenidos al estimar el efecto del número de hijos en la probabilidad de asistencia a la escuela para distintos subgrupos de las muestras de dos a más hijos y tres a más hijos en la familia, para las cuales se obtienen resultados estadísticamente significativos en el análisis previamente realizado.

La segunda y tercera columnas de la tabla 14 reportan los coeficientes obtenidos al estimar por MC2E utilizando el evento de nacimientos múltiples como instrumento de identificación de acuerdo al ámbito urbano y rural, respectivamente. Esto debido a que el mayor número de escuelas, la mayor cercanía a la escuela y la mayor probabilidad de acceso al crédito en las zonas urbanas, en comparación con las zonas rurales del país, podrían introducir efectos heterogéneos de acuerdo al ámbito geográfico de residencia. Los resultados obtenidos sugieren que el efecto de un hijo adicional en la familia como consecuencia de nacimientos múltiples en el segundo y tercer parto de la madre es mayor para los hijos mayores de las familias residentes en zonas rurales, en comparación con los hijos mayores de las familias residentes en zonas urbanas del país. Estos resultados son válidos tanto en la muestra de dos a más hijos como en la muestra de tres a más hijos y son significativos al 99%. Nótese que el coeficiente obtenido es mayor incluso al controlar por orden de nacimiento en la muestra de tres a más hijos en comparación con la muestra de dos a más hijos para familias residentes en zonas rurales del país. Esto es particularmente interesante debido a que, al observar los resultados para familias residentes en zonas urbanas, los coeficientes obtenidos en ambas muestras son muy parecidos, lo cual sugiere que los padres

invierten más en la educación de los hijos mayores en zonas rurales ante un aumento exógeno del número de hijos en la familia.

En la cuarta y quinta columnas de la tabla 14, la muestra se divide de acuerdo al sexo de los niños. Hacer esta distinción puede ser interesante para analizar la existencia de preferencias sesgadas en función del sexo de los hijos en las familias. Para los niños, un hijo adicional en la familia, como consecuencia de nacimientos múltiples, tiene un efecto positivo de 5 puntos porcentuales en la probabilidad de asistencia a la escuela, y es igual para la muestra de familias con dos a más hijos y la muestra de familias con tres a más hijos (ambos resultados son significativos al 99%). Los coeficientes resultantes para las niñas son positivos pero menores a los de los niños. En el caso de las niñas, un hijo adicional en la familia tiene un efecto de 3,4 a 3,8 puntos porcentuales más de probabilidad de asistencia a la escuela dependiendo del tamaño inicial de la familia (los resultados en ambas muestras son significativos al 99%). Estos resultados pueden ser evidencia de que sí existen preferencias en la inversión en educación de acuerdo al sexo de los hijos mayores ante un aumento exógeno del número de hijos en la familia.

En adición, se divide también la muestra de acuerdo a la presencia del padre en el hogar. Esto es posible debido a que la ENDES incluye una variable identificadora del padre de los niños²³. Como se esperaba, los coeficientes obtenidos son todos positivos (y significativos al 99%) y mayores en familias en donde el padre se encuentra presente en el hogar. Este mayor efecto, como consecuencia de una figura paterna presente, puede deberse a las ganancias por división del trabajo (actividades remuneradas y actividades concernientes al hogar) que tienen las familias en donde ambos padres están presentes.

Finalmente se hace la distinción por distancia temporal entre nacimientos de los hijos. Para ello, la muestra incluye únicamente a

23 Debido a que se quiere medir el efecto de una figura paterna presente en la familia, no se hacen distinciones acerca de si los niños nacidos de una misma madre tienen padres distintos. Para esos casos se considera en la muestra de "padre presente" a los niños para los cuales al menos un hermano (un hijo en la familia) reporte que su padre se encuentra presente en el hogar.

aquellos niños para los cuales la diferencia entre el año de su nacimiento y el año del nacimiento del hermano que les sigue en orden es menor a dos años. Esta distinción es particularmente interesante debido a que los costos de reasignación de los recursos del hogar pueden ser mayores mientras mayor sea la distancia temporal entre edades de los hijos. A modo de ejemplo, es más costoso para los padres que un hijo deje de asistir a la escuela cuando está cursando los últimos años de educación secundaria, en comparación con que el hijo esté cursando los primeros años de educación primaria, debido a que los padres ya han invertido dinero, tiempo y otros recursos de la familia en la formación educacional del hijo mayor. Por tal motivo, si los costos de reasignación de recursos son menores mientras menor sea la distancia entre edades de los hijos, se esperaría que el efecto positivo encontrado de un hijo adicional en la familia en la probabilidad de asistencia a la escuela tienda a cero mientras menos sea la distancia temporal entre nacimientos contiguos.

La columna 9 de la tabla 14 reporta los coeficientes resultantes al efectuar esta restricción. Los coeficientes puntuales son 0,049 y 0,031 para las muestras de dos a más hijos y tres a más hijos respectivamente (ambos significativos al menos al 95%), los cuales son muy parecidos a los coeficientes reportados en las tablas 5 y 10. Más aún, la columna 10 de la tabla 14 agrega la restricción de que la edad de los niños debe ser igual o menor a 10 años. Los coeficientes resultan ser estadísticamente significativos al 95% en ambas muestras, aunque menores a los reportados en la columna 9, lo cual implica que no es únicamente la distancia temporal entre nacimientos, sino también la edad de los hijos, lo que determina la magnitud del efecto de un hijo adicional en la familia en la probabilidad de asistencia a la escuela de los hijos mayores.

8. RESULTADOS ADICIONALES

La tabla 15 reporta los coeficientes obtenidos por MCO y MC2E del número de hijos en la probabilidad de que los niños compartan una habitación y la probabilidad de que la madre se haya divorciado.

Debido a que las habitaciones de la vivienda pueden ser vistas como un “bien público” para los hijos en caso de ser compartidas, los padres podrían compensar el efecto de un mayor tamaño familiar dividiendo la inversión, inicialmente asignada para cada hijo, en este tipo de bienes que son fácilmente transmitidos entre los hijos. Esto es más factible en la medida en que los costos por transferencia de bienes entre los hijos sean menores. Por ejemplo, los niños de un mismo sexo o de edades similares podrían compartir habitaciones, juegos, vestimenta e incluso el tiempo que los padres dedican al acompañamiento en las actividades escolares de los hijos. Esto implica, al mismo tiempo, que mientras los niños compartan características similares, podrían existir ganancias por economías de escala al interior de las familias, atenuando el efecto negativo en la inversión por cada hijo que surge como consecuencia de cambios exógenos del tamaño familiar.

La segunda columna de la tabla 15 muestra los coeficientes obtenidos de la estimación por MCO del número de hijos en la probabilidad de que los niños compartan habitaciones. Como se esperaba, estos coeficientes son positivos, pero estadísticamente significativos (al 99%) únicamente en las muestras de dos a más hijos y tres a más hijos. Esto podría deberse a que la muestra de cuatro a más hijos incluye familias más numerosas, lo cual implica que es bastante probable que los miembros de esas familias ya se encuentren compartiendo habitaciones. La tercera columna de la tabla 15 reporta los resultados

obtenidos de la estimación por MC2E. En este caso, el efecto del número de hijos en la probabilidad de que los niños compartan la habitación es estadísticamente significativo únicamente en la muestra de dos a más hijos, implicando que en familias con al menos dos hijos, el efecto de un hijo adicional como consecuencia de nacimientos múltiples en el segundo parto de la madre aumenta en 6,4 puntos porcentuales la probabilidad de que los hijos compartan habitaciones. Este resultado, sin embargo, podría tener una explicación bastante lógica: esta muestra incluye a familias que tienen por lo menos dos hijos y que pueden residir en viviendas que no requieren mayor espacio, como es el caso de los departamentos. En estas familias, una variación exógena en el tamaño familiar implica que los niños necesariamente tendrán que compartir habitaciones, más aún si la vivienda cuenta en total con tres o menos habitaciones utilizadas para dormir. Contrariamente, las muestras de tres a más hijos y cuatro a más hijos incluyen familias que, por tener un mayor número de miembros, residen en viviendas más amplias o es más probable que los miembros de la familia compartan dormitorios.

Por otra parte, la estabilidad conyugal de los padres puede verse también como una inversión de tiempo asignado a la crianza conjunta de los hijos. Así, si los padres valoran altamente la calidad de sus hijos y consideran que la estabilidad marital tiene efectos positivos en la formación del capital humano de su progenie, entonces el divorcio resultaría más costoso para los padres. Algunos estudios (Manski *et al.* [1992] y Haverman y Wolfe [1995]) han documentado que los hijos de padres divorciados tienen menores logros educativos que los hijos de familias “intactas”. Al mismo tiempo, es factible pensar que un mayor número de hijos en la familia eleva los costos asociados al divorcio, disminuyendo la probabilidad de que los padres se separen. Por consiguiente, un mayor número de hijos puede tener efectos positivos en la estabilidad marital de los padres, lo cual, a su vez, puede favorecer los resultados educativos de los hijos.

La cuarta y quinta columnas de la tabla 15 reportan los coeficientes obtenidos al estimar por MCO y MC2E el efecto del número de hijos en la probabilidad de que la madre se haya divorciado o separado,

respectivamente. Los resultados de la estimación por MCO son negativos y estadísticamente significativos al menos al 95% en las tres muestras. Estos resultados implican que la llegada de un hijo adicional a la familia disminuye entre 1.2 y 3 puntos porcentuales la probabilidad de que la madre se divorcie. Este resultado, no obstante, es decreciente (en términos absolutos) con respecto al tamaño familiar inicial, implicando que la magnitud del efecto de un hijo adicional en la familia, en la probabilidad de que la madre se divorcie será menor a medida que aumenta el tamaño familiar.

Por el contrario, los coeficientes obtenidos al estimar por MC2E son estadísticamente iguales a cero en las tres muestras. Estos resultados sugieren que los coeficientes obtenidos por MCO podrían estar capturando efectos que pueden estar correlacionados con preferencias por un mayor número de hijos y al mismo tiempo un vínculo marital más sólido, como es el caso de creencias religiosas o rechazo a métodos contraceptivos en las familias. Del mismo modo, si la hipótesis que sugiere que la estabilidad conyugal de los padres tiene efectos positivos en los resultados educativos de los niños fuera cierta, entonces los resultados obtenidos en la estimación por MC2E del número de hijos en la probabilidad de divorcio de la madre serían consistentes con los resultados obtenidos en las tablas 8, 9, 12 y 13, que muestran que una variación exógena del tamaño familiar no tiene efecto alguno, tanto en el margen extensivo como en el margen intensivo asociados al retraso escolar.

9. POSIBLES MECANISMOS

La principal pregunta que surge a partir de los resultados encontrados es qué puede estar explicando la relación causal positiva existente entre el número de hijos y la asistencia a la escuela de los hijos mayores en la familia y, al mismo tiempo, la no existencia de un vínculo causal entre el tamaño familiar y los resultados educativos de los niños. En esta sección se intentará abordar algunos mecanismos que pueden estar relacionados con los resultados encontrados en este estudio.

En el caso de los resultados educativos obtenidos por los niños, una posible explicación viene dada por una ganancia en productividad del hogar debido a economías de escala que se generan al aumentar el tamaño familiar. Bajo esta hipótesis, existen dos posibles escenarios. En primer lugar, es posible argumentar que los niños mayores tienen una mayor ganancia en términos de madurez debido a las nuevas responsabilidades adquiridas como consecuencia de la llegada de un hijo adicional. En segundo lugar, es factible pensar que los hijos mayores se benefician con un aumento del tamaño familiar debido a que pueden jugar el rol de maestros o referencia más próxima de sus hermanos menores. Esto no solo implica que los hijos mayores puedan tomar conciencia de ello y tener una actitud más responsable y madura, sino que pueden reforzar también las materias aprendidas en la escuela transmitiendo sus conocimientos a los hijos menores de la familia.

Alternativamente, si los hijos mayores pueden notar la restricción financiera y de otros recursos del hogar debido a la llegada de un hijo adicional a la familia (por ejemplo, pueden darse cuenta de que sus padres dedican más tiempo al cuidado de los hijos menores), es posible que quieran captar la atención de sus padres asignando un mayor

esfuerzo en la escuela (obteniendo mejores calificaciones). Dado que los resultados educativos son consecuencia tanto de la inversión en educación como del esfuerzo asignado por el niño, es posible pensar que la menor inversión en la educación de los hijos puede verse compensada por un mayor esfuerzo en el aprendizaje del niño.

Con respecto a la asistencia a la escuela, la literatura económica provee de algunas hipótesis en referencia a la reasignación de recursos que hacen los padres al aumentar el tamaño de la familia.

Un argumento frecuentemente utilizado por los economistas se basa en que los padres pueden reducir los gastos de su progeñie invirtiendo en “bienes públicos” de los niños (Qian [2006] y Angrist *et al.* [2010]). Debido a que existen bienes que pueden ser fácilmente transmitidos o compartidos entre los niños, los padres pueden preferir invertir en este tipo de bienes en vez de reducir la inversión en la calidad educativa de los hijos. Un ejemplo de estos bienes son las habitaciones, vestimenta, juguetes, materiales educativos y otro tipo de bienes que pueden ser denominados como “bienes públicos” de la familia. Las regresiones reportadas en este estudio no sustentan la validez de dicha hipótesis, al menos para la probabilidad de que los hijos compartan habitaciones. No obstante, pueden existir efectos distintos para el caso de otros bienes fácilmente transferibles entre los hijos.

Otra posible explicación puede ser la sustitución de un mayor número de hijos por una reducción en el consumo de “bienes no necesarios” en la familia. Debido a que la ENDES no contiene información acerca del gasto per cápita en bienes de consumo en el hogar, ninguna de las especificaciones utilizadas en el análisis empírico mantiene constante el gasto en consumo. En consecuencia, los coeficientes encontrados podrían estar capturando efectos asociados a esta variable.

Una explicación adicional asociada a la reasignación de recursos al interior de los hogares tiene que ver con los niveles de ahorro y acumulación de activos. Un caso interesante es el del Programa de Planificación Familiar y Salud de la Madre y el Niño implementado a partir de 1966 en Matlab, Bangladesh²⁴. La intervención del Programa

24 Ver Schultz [2007].

fue aleatoria entre la población de mujeres en edad fértil y consistía en charlas informativas acerca de métodos de control de la natalidad y provisión de una variedad de métodos e indicaciones de uso sin costo alguno, lo cual se presenta como una oportunidad singular de probar la formulación de la interacción entre la cantidad y calidad de los hijos. Controversialmente, Sinha [2005] no encuentra efectos significativos de la reducción del número de hijos en la familia, a causa de la intervención del Programa, en la tasa de matrícula en la escuela de niños y niñas ni tampoco en la reducción del trabajo infantil. Un hallazgo interesante es que la reducción de la fertilidad de las mujeres tiene un efecto positivo en la acumulación de activos familiares, lo cual sugiere que los hogares pueden compensar el efecto de un mayor número de hijos reduciendo el ahorro o los bienes durables de la familia.

De otro lado, independientemente de sacrificar los niveles actuales de consumo o ahorro de la familia, los padres pueden ajustar el mayor tamaño familiar en márgenes distintos a la inversión en la educación de los hijos. Un ejemplo específico es que los padres pueden consumir menos ocio (asumiendo que el ocio es un bien normal para los individuos). Así, los padres estarían dispuestos a sustituir un mayor número de hijos por jornadas más prolongadas de trabajo o sacrificando sus vacaciones con el objetivo de obtener un mayor ingreso para el hogar.

En referencia a la inversión en educación, se pueden proveer tres argumentos adicionales. Primero, es factible pensar que los padres están dispuestos a sacrificar la calidad educativa de los niños sin privarlos de la asistencia a la escuela. De manera particular, las familias podrían dejar de matricular a sus hijos en escuelas privadas y matricularlos en escuelas públicas, donde el costo de la educación es menor. Al respecto, Cáceres-Delpiano [2006] encuentra que en Estados Unidos los hijos primogénitos de familias que reportan haber tenido nacimientos múltiples en el segundo o tercer parto de la madre tienen una menor probabilidad de asistencia a escuelas privadas. Lamentablemente, la base de datos utilizada en este estudio no contiene información acerca del tipo de escuela a la que asisten los niños y por ello resulta imposible validar dicha hipótesis para el caso peruano.

En segundo lugar, una hipótesis plausible es que los padres pueden asignar más recursos financieros a la educación de los hijos mayores pero dedicar otro tipo de recursos de la familia —como el tiempo— al acompañamiento del proceso de aprendizaje de los hijos menores.

Tercero, los programas de asistencia social pueden haber influido también en los resultados encontrados. De manera específica, puede pensarse en los ingresos por transferencias condicionadas en los hogares. Este tipo de ingresos podría atenuar e incluso compensar en mayor medida los costos asociados al aumento del tamaño familiar. Sin embargo, tal hipótesis no es válida en el contexto peruano debido a que el programa de transferencias condicionadas con mayor cobertura en el país no toma en cuenta la estructura familiar para elegir a los hogares beneficiarios. Por este motivo, los resultados hallados en la estimación por MC2E no estarían capturando ningún efecto positivo asociado a la probabilidad de que las familias sean elegidas como beneficiarias de programas de transferencias condicionadas como consecuencia del mayor número de hijos.

Un argumento adicional, y que no ha sido antes mencionado por la literatura, tiene que ver con las inversiones que realizan los padres en función de las preferencias por la “seguridad” de la familia. Bajo esta hipótesis, los padres asignarán más recursos a los hijos mayores por tres motivos concretos. Primero: debido a que los hijos mayores han gozado en algún período de la totalidad de los recursos de la familia, los padres han invertido más recursos en ellos en comparación con los hijos menores, de modo que resultaría más costoso para los padres (en términos del bienestar de la familia) dejar de invertir en los hijos mayores, quienes, de alguna manera, tienen un mayor nivel de capital humano formado en comparación con los demás hijos. Segundo: los padres esperan que los retornos a la inversión en educación de los hijos mayores se den en un período más próximo en comparación con los hijos nacidos en un orden mayor, dado que, por definición, los hijos mayores se encuentran en años escolares más avanzados en comparación con el resto de sus hermanos. Finalmente, los hijos mayores cumplen un rol distinto en la familia, pues son

quienes toman las responsabilidades de los padres en caso de ausencia temporal o permanente de alguno de ellos (o ambos). Así, los padres preferirán invertir más en los hijos mayores, de tal manera que puedan desempeñar mejor las responsabilidades delegadas que conciernen al cuidado y sustento de los hijos menores.

Dados los resultados obtenidos al restringir por distancia entre nacimientos contiguos y edad de los niños, resulta factible pensar que los padres tienen preferencias por una mayor inversión en educación de los hijos de edades más avanzadas. Este resultado hace que la hipótesis sobre “seguridad” de las familias sea válida en un contexto de mayor restricción de los recursos asignados a la prole como consecuencia de un aumento del tamaño de la familia.

10. CONCLUSIONES

El presente estudio analiza el efecto del número de hijos en la asistencia a la escuela y en el retraso escolar de los hijos mayores en las familias peruanas. Investigar el efecto del número de hijos en los resultados educativos de los niños al interior de las familias es complicado debido a que el tamaño familiar es endógenamente determinado por los padres. A modo de identificar correctamente el efecto de interés, la estrategia de identificación seguida en el presente documento se basa en el uso de variables instrumentales, específicamente los eventos de nacimientos múltiples y las preferencias por balance de sexo de los hijos en las familias, que varían exógenamente la fertilidad en las familias. El uso de los instrumentos descritos debe su validez a la determinación de carácter natural de ambos eventos (nacimientos múltiples y sexo de los hijos).

El análisis empírico se inicia con la validación de los instrumentos propuestos para la identificación del efecto causal del número de hijos en la asistencia a la escuela y el retraso escolar de los hijos mayores de las familias. Para el caso de los nacimientos múltiples, se encuentra que dicho instrumento no está relacionado con el nivel socioeconómico, el uso de métodos contraceptivos ni los antecedentes de fecundidad de las mujeres. Por otro lado, el sexo de los hijos no guarda correlación con la probabilidad de aborto o mortalidad infantil incluso al partir la muestra de acuerdo al ámbito geográfico de nacimiento y residencia de la mujer. Este hallazgo es consistente con la hipótesis de la no existencia de abortos selectivos de acuerdo al sexo de los niños en las familias peruanas. Todos los resultados previamente descritos validan el uso de los nacimientos múltiples y sexo de los dos primeros hijos como potenciales determinantes del tamaño familiar.

Al efectuar las pruebas de instrumentos débiles en la primera etapa de las regresiones se encuentra que los nacimientos múltiples logran explicar adecuadamente la variabilidad del número de hijos en las familias; sin embargo, no sucede lo mismo con el indicador de mismo sexo de los dos primeros hijos. Por este motivo, todo el análisis incluido en este documento se basa en el uso de eventos de nacimientos múltiples como instrumento de identificación.

Con respecto a los resultados educativos de los niños, los hallazgos no sustentan la formulación de la existencia de una interacción entre la cantidad y la calidad de los niños. De manera específica, no se encuentra una relación causal entre el número de hijos y la probabilidad de retraso escolar (margen extensivo) y el año o grado normativo para la edad de los niños entre 6 y 14 años de edad. De manera interesante, los resultados obtenidos en el análisis empírico muestran una relación positiva entre el número de hijos y la asistencia a la escuela de los hijos mayores. Así, se encuentra que un hijo adicional en familias afectadas por eventos de nacimientos múltiples aumenta entre 3,5 y 4,5 puntos porcentuales (dependiendo del tamaño familiar inicial) la probabilidad de asistencia a la escuela de los hijos mayores de las familias. Estos resultados se mantienen robustos aun al añadir nuevos controles, particularmente el orden de nacimiento de los niños y la distancia temporal entre nacimientos contiguos y, al efectuar estimaciones para distintos subgrupos poblacionales. En esta línea se encuentra también que el efecto del tamaño familiar es mayor en los hijos de las familias residentes en zonas rurales del país, en los varones y también en aquellos niños para los cuales se reporta que ambos padres se encuentran presentes en el hogar. Por el contrario, al efectuar estimaciones en niños menores de 10 años de edad y con una distancia temporal menor de dos años entre su nacimiento y el nacimiento del hermano que le sucede en orden, se encuentra que el efecto del tamaño familiar en la asistencia a la escuela, aunque continúa siendo estadísticamente significativo, tiende a cero.

Otro hallazgo interesante tiene que ver con el sesgo a priori del instrumento. Al respecto, se encuentra que mientras menor sea la distancia entre los nacimientos del segundo y tercer hijo de las

familias, la probabilidad de que los hijos primogénitos asistan a la escuela se ve reducida. Extrapolando estos resultados para el caso de los nacimientos múltiples, donde la distancia entre los nacimientos de los niños es la mínima (igual a cero), este instrumento está sesgado a encontrar efectos negativos en los resultados educativos de los niños por sí mismos, lo cual implica que resultaría factible considerar que los resultados encontrados en este estudio son conservadores. Los mismos resultados son encontrados por Black *et al.* [2005].

Al incluir efectos de orden de nacimiento en las regresiones, se encuentra que la inversión en educación de los hijos es menor a medida que el orden de nacimiento de los niños va aumentando. Este resultado cuestiona el supuesto de simetría en la inversión en la calidad por cada hijo al interior de las familias.

Existen dos potenciales limitaciones en el presente estudio. La primera limitación potencial tiene que ver con la totalidad de la asistencia a la escuela en el nivel básico primario. En esta línea, un estudio más interesante podría condicionar la investigación a la probabilidad de continuar estudios básicos secundarios o la probabilidad de matrícula y asistencia de los niños a la secundaria. La segunda limitación potencial tiene que ver con la inferencia. Esto debido a que el análisis se realiza únicamente para los hijos mayores de la familia y no toma en cuenta la situación educativa de los hijos menores.

En el escenario de que la educación de los niños, tanto en inversión como en el progreso escolar, no se vea afectada por la restricción financiera causada por un aumento del tamaño de la familia, una pregunta que queda por responder es qué márgenes alternativos a la inversión en calidad de los hijos pueden ajustar los padres con el objetivo de no perjudicar la educación de su progenie. Una investigación más profunda de esta pregunta podría traer nueva evidencia y conclusiones en la literatura concerniente a la cantidad y la calidad de los niños. Concretamente, es necesario investigar otro tipo de resultados como, por ejemplo, indicadores de salud y morbilidad de los niños, inmunización (vacunas), trabajo infantil (actividades remuneradas y tareas asignadas al interior de los hogares), inversión

en vestimenta, libros y materiales educativos, tipo de escuela a la que asisten los niños (pública o privada), tiempo que los padres asignan a sus hijos, demanda de ocio por parte de los padres y acumulación de activos durables en el hogar, entre otro tipo de indicadores de interés.

A modo de conclusión, los resultados encontrados sugieren que los padres están dispuestos a invertir más en la educación de sus hijos mayores ante variaciones exógenas en el tamaño de la familia. Una posible hipótesis para ello es que los padres prefieren invertir en la “seguridad” de la familia dado que los hijos mayores cumplen un rol diferente, pues son quienes adquieren mayores responsabilidades en comparación con los demás hijos. Otra posible explicación es que los padres pueden reasignar mejor los recursos de la familia cuando los niños son mayores, ya que obtienen ganancias por experiencia en la crianza de los hijos, lo cual no es considerado por los modelos teóricos. Hasta este punto, es posible concluir, al menos, que la interacción entre cantidad y calidad puede no ser monótona.

REFERENCIAS BIBLIOGRÁFICAS

- Agüero, J. M. y Marks, M. S. (2008). *A new approach to estimate the effect of family size on child development*. Recuperado el 19 de setiembre de 2011, de <http://faculty.ucr.edu/~jorgea/research/qq.pdf>
- Angrist, J. D. y Evans, W. N. (1998). Children and their parent's labor supply: evidence from exogenous variation in family size. *American Economic Review*, 88 (3), 450-477.
- Angrist, J. D., Lavy, V. y Schloseer, A. (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics*, 28 (4), 773-823.
- Becker, G. S. y Lewis, G. H. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy*, 81 (2), S279-288.
- Becker, G. S. y Tomes, N. (1976). Child endowments and the quantity and quality of children", *Journal of Political Economy*, 84 (4), S143-162.
- Behrman, J. R., Pollak, R. A. y Taubman, P. (1982). Parental preferences and provision for progeny. *Journal of Political Economy*, 90 (1), 52-73.
- Ben-Porath, Y. (1973). Economics analysis of fertility in Israel. *Journal of Political Economy*, 81 (2), S202-233.
- Ben-Porath, Y. y Welch, F. (1976). Do sex preferences really matter? *Quarterly Journal of Economics*, 90 (2), 285-307.

- Birdsall, N. (1991). Birth order effects and time allocation. En Schultz, T.P (Ed.). *Research in population economics: Vol. 7. A research annual* (pp. 191-213). London: JAI Press.
- Blake, J. (1981). Family size and the quality of children. *Demography*, 18 (4), 421-442.
- Black, S. E., Devereux, P. J. y Salvanes, K. G. (2005). The more the merrier? The effect of family size and birth order on children's education. *Quarterly Journal of Economics*, 120 (2), 669-700.
- Butcher, K. F. y Case, A. (1994). The effect of sibling sex composition on women's education and earnings. *Quarterly Journal of Economics*, 109 (3), 531-563.
- Cáceres-Delpiano, J. (2006). The impacts of family size on investments in child quality. *Journal of Human Resources*, 41 (4), 738-754.
- Conley, D. (2000). Sibship sex composition: effects on educational attainment. *Social Science Research*, 29 (3), 441-457.
- Conley, D. y Glauber, R. (2006). Parental educational investment and children's academic risk. *Journal of Human Resources*, 41 (4), 722-737.
- Dahl, G. B. y Moretti, E. (2004). *The demand for sons: evidence from divorce, fertility and shotgun marriage*. Cambridge, MA: NBER. Working paper, W10281. Recuperado el 12 de setiembre de 2011, de http://www.nber.org/papers/w10281.pdf?new_window=1
- Dammert, A. C. (2010). Siblings, child labor, and schooling in Nicaragua and Guatemala. *Journal of Population Economics*, 23 (1), 199-224.
- Deaton, A. (2009). *Instruments of development: randomization in the tropics, and the search for the elusive keys to economic development*. Cambridge, MA: NBER. Working paper, W14690. Recuperado el 12 de setiembre de 2011, de http://www.nber.org/papers/w14690.pdf?new_window=1

- DeTray, D. N. (1974). Child quality and the demand for children. *Journal of Political Economy*, 81 (2), 70-95.
- Ejrnæs, M. y Pörtner, C. C. (2004). Birth order and the intrahousehold allocation of time and education. *Review of Economics and Statistics*, 86 (4), 1008-1019.
- Ginther, D. K. y Pollak, R. A. (2000). *Does family structure affects children outcomes?* St. Louis, CO: Department of Economics. Washington University.
- Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy*, 100 (1), 84-117.
- Hauser, R. M. y Kuo, HH. D. (1998). Does the gender composition of sibships affect women's educational attainment? *Journal of Human Resources*, 33 (3), 644-657.
- Hauser, R. M. y Sewell, W. H. (1985). Birth order and educational attainment in full sibships. *American Educational Research Journal*, 22 (1), 1-23.
- Haverman, R. y Wolfe, B. (1995). Determinants of children's attainment: review of methods and findings. *Journal of Economics Literature*, 33 (4), 1829-1878.
- Heckman, J. J. (1997). Instrumental variables: a study of implicit behavioral assumptions using in making program evaluations. *Journal of Human Resources*, 32 (3), 441-462.
- Heckman, J. J. y Urzúa, S. (2009). *Comparing IV with structural models: what simple IV can and cannot identify*. Cambridge, MA: NBER. Working paper, W14706. Recuperado el 12 de diciembre de 2011, de http://www.nber.org/papers/w14706.pdf?new_window=1
- Iacovou, M. (2001). *Family Composition and Children's Educational Outcomes*. Colchester: University of Essex. Working Papers of the Institute for Social and Economic Research, 2001-12.

- Recuperado el 12 de setiembre de 2011, de <http://www.iser.essex.ac.uk/publications/working-papers/iser/2001-12.pdf>
- Iacovou, M. (2001). *Fertility and female labor supply*. Colchester: University of Essex. Working Papers of the Institute for Social and Economic Research, 2001-19. Recuperado el 12 de setiembre de 2011, de <http://www.iser.essex.ac.uk/publications/working-papers/iser/2001-19.pdf>
- Imbens, G. W. (2010). Better LATE than nothing: some comments on Deaton (2009) and Heckman and Urzua (2009). *Journal of Economic Literature*, 48 (2), 399-423.
- Imbens, G. W. y Angrist, J. A. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 61 (2), 467-476.
- Jaramillo, M. y Montalva, V. (2010). *Fertility decline and women's work in Peru* (Informe de consultoría para el Banco Mundial). Lima: GRADE.
- Leamer, E. E. (2010). Tantalus on the road to asymptopia. *Journal of Economic Perspectives*, 24 (2), 31-46.
- Lee, J. (2008). Sibling size and investments in children's education: an asian instrument. *Journal of Population Economics*, 21 (4), 855-875.
- Levy, V. (1985). Cropping pattern, mechanization, child labor, and fertility behavior in a farming economy: rural Egypt. *Economic Development and Cultural Change*, 33 (4), 777-791.
- Li, H., Zhang, J. y Zhung, Y. (2007). *The quantity-quality tradeoff of children in a developing country: identification using chinese twins*. Bonn: IZA. Discussion papers, 3012. Recuperado el 12 de setiembre de 2011, de <ftp://repec.iza.org/RePEc/Discussionpaper/dp3012.pdf>
- Lundberg, S. (2005). *Sons, daughters, and parental behaviour*. Bonn: IZA. Discussion papers, 2103. Recuperado el 12 de setiembre de 2011, de <ftp://repec.iza.org/RePEc/Discussionpaper/dp2103.pdf>

- Manski, Ch., Sandefur, G., McLanahan, M. y Powers, D. (1992). Alternative estimates of the effect of family structure during adolescence on high school graduation. *Journal of the American Statistical Association*, 87 (1), 25-37.
- Patrinos, H. A. y Psacharopoulos, G. (1997). Family size, schooling and child labor in Peru – an empirical analysis. *Journal of Population Economics*, 10 (4), 387-405.
- Powell, B. y Steelman, L. C. (1990). Beyond sibship size: sibling density, sex composition and educational outcomes. *Social Forces*, 69 (1), 181-206.
- Qian, N. (2006). *Quantity-quality and the one child policy: the positive effect of family size on school enrollment in China*. Mimeo, Department of Economics, Brown University.
- Ronsenzweig, M. R. y Evenson, R. E. (1977). Fertility, schooling and the economic contribution of children in rural India. *Econometrica*, 45 (5), 1065-1079.
- Rosenzweig, M. R. y Schultz, T. P. (1985). The demand for and supply of births: fertility and its life cycle consequences. *American Economic Review*, 75 (5), 992-1015.
- Rosenzweig, M. R. y Wolpin, K. I. (1980). Testing the quantity-quality model of fertility: the use of twins as a natural experiment. *Econometrica*, 48 (1), 227-240.
- Rosenzweig, M. R. y Wolpin, K. I. (2000). Natural ‘Natural experiments’ in economics. *Journal of Economic Literature*, 38 (4), 827-874.
- Rosenzweig, M. R. y Zhang, J. (2009). *Do population control policies induce more human capital investment? Twins, birthweight, and China’s ‘One child’ policy*. New Haven, CT: Economic Growth Center. Yale University. Discussion paper, 933. Recuperado el 12 de setiembre de 2011, de http://www.econ.yale.edu/growth_pdf/cdp933.pdf

- Sinha, N. (2005). Fertility, child work, and schooling consequences of family planning programs. *Economic Development and Cultural Change*, 54 (1), 97-128.
- Schultz, T. P. (1997). The demand for children in low income countries. En M. Rosenzweig y O. Stark (eds.). *Handbook of Population and Family Economics*: Vol. 1A (pp. 349-430). North-Holland: Elsevier.
- Schultz, T. Paul (2007). *Population policies, fertility, women's human capital, and child quality*. Bonn: IZA. Discussion papers, 2815. Recuperado el 12 de setiembre de 2011, de <ftp://repec.iza.org/RePEc/Discussionpaper/dp2815.pdf>
- Van Eijck, K. y De Graff, P.M. (1995). The effects of family structure on the educational attainment of siblings in Hungary. *European Sociological Review*, 11 (3), 273-292.
- Zajonc, R.B. (1976). Family configuration and intelligence. *Science*, 192 (4236), 227-236.
- Zajonc, R. B. y Markus, G. B. (1975). Birth order and intellectual development. *Psychological Review*, 82 (1), 74-88.
- Zajonc, R. B., Markus, H. y Markus, G. B. (1979). The birth order puzzle. *Journal of Personality and Social Psychology*, 37 (8), 1325-1341.
- Zajonc, R. B. y Mullaly, P. R. (1997). Birth order: reconciling conflicting effects. *American Psychologist*, 52 (7), 685-699.

TABLAS

Tabla 1
Validez de instrumentos: composición de la muestra

	Año de encuesta				
	1996	2000	2004-2008	2009	Total
Número de mujeres entre 15 y 49 años	32.646	30.718	44.439	25.460	133.263
Mujeres entrevistadas	28.951	27.843	41.648	24.212	122.654
Mujeres con hijos ^a	19.835	18.931	28.613	16.887	84.266
Todos los hijos vivos	14.958	14.656	23.822	14.481	67.917
2 a más hijos ^b	10.827	10.442	16.796	10.146	48.211
3 a más hijos ^c	7.067	6.562	10.120	5.956	29.705
4 a más hijos ^d	4.210	3.871	5.548	3.263	16.892

Fuente: ENDES 1996, 2000, 2004-2008 y 2009.

a/ Muestra "completa de mujeres". Esta muestra está compuesta por mujeres entrevistadas entre 15 y 49 años de edad que reportan haber tenido por lo menos 1 hijo nacido vivo.

b/ Muestra compuesta por mujeres entrevistadas de 15 a 49 años de edad que reportaron tener por lo menos 2 hijos nacidos vivos (2 partos). Se excluye de esta muestra a las mujeres que reportaron tener al menos un hijo que murió después de su nacimiento.

c/ Muestra compuesta por mujeres entrevistadas de 15 a 49 años de edad que reportaron tener por lo menos 3 hijos nacidos vivos (3 partos). Se excluye de esta muestra a las mujeres que reportaron tener al menos un hijo que murió después de su nacimiento.

d/ Muestra compuesta por mujeres entrevistadas de 15 a 49 años de edad que reportaron tener por lo menos 4 hijos nacidos vivos (4 partos). Se excluye de esta muestra a las mujeres que reportaron tener al menos un hijo que murió después de su nacimiento.

Tabla 2**Tamaño familiar y resultados educativos: composición de la muestra**

	Año de encuesta				
	1996	2000	2004-2008	2009	Total
Número de niños de 6 a 14 años	33.179	30.217	39.203	20.904	123.503
Miembro del hogar	32.495	29.735	38.529	20.652	121.411
Madre viva	31.488	29.028	37.819	20.326	118.661
Madre presente en el hogar	28.491	26.144	34.236	18.342	107.213
Madre entrevistada	24.175	22.538	30.226	16.393	93.332
Todos los hermanos (hijos) vivos	15.946	15.545	23.500	13.365	68.356
2 a más hermanos (hijos) en la familia	15.254	14.752	21.804	12.310	64.120
2 a más hijos ^a	4.033	3.924	6.109	3.662	17.728
3 a más hijos ^b	3.730	3.311	4.326	2.397	13.764
4 a más hijos ^c	2.018	1.748	1.851	896	6.513

Fuente: ENDES 1996, 2000, 2004-2008 y 2009.

a/ Muestra compuesta por el hijo mayor de cada hogar y que estuvo presente en el hogar al momento de la entrevista. Se excluye de la muestra a los niños que son gemelos o mellizos, los niños para los cuales el nacimiento siguiente es de trillizos o cuatrillizos y los niños que no presentan información completa en el módulo de educación de la encuesta.

b/ Muestra compuesta por los dos primeros hijos de cada hogar y que estuvieron presentes en el hogar al momento de la entrevista. Se excluye de la muestra a los niños que son gemelos o mellizos, los niños para los cuales el nacimiento siguiente es de trillizos o cuatrillizos, los niños cuyos hermanos no están presentes en el hogar y los niños que no presentan información completa en el módulo de educación de la encuesta.

c/ Muestra compuesta por los tres primeros hijos de cada hogar y que estuvieron presentes en el hogar al momento de la entrevista. Se excluye de la muestra a los niños que son gemelos o mellizos, los niños para los cuales el nacimiento siguiente es de trillizos o cuatrillizos, los niños cuyos hermanos no están presentes en el hogar y los niños que no presentan información completa en el módulo de educación de la encuesta.

Tabla 3
Muestra de mujeres entrevistadas (15 a 49 años):
estadísticas descriptivas

Variable	Muestra completa	2 a más hijos	3 a más hijos	4 a más hijos
Antecedentes				
Edad	33,73 (8,49)	35,02 (7,54)	36,83 (7,00)	38,30 (6,41)
Lugar de nacimiento (igual a 1 si es en ciudad)	0,64 (0,48)	0,65 (0,48)	0,60 (0,49)	0,54 (0,50)
Lugar de residencia actual (igual a 1 si es urbano)	0,60 (0,49)	0,62 (0,49)	0,43 (0,50)	0,50 (0,50)
Nivel educativo alcanzado: sin educación	0,07 (0,26)	0,06 (0,24)	0,08 (0,27)	0,11 (0,32)
Nivel educativo alcanzado: básico primario	0,38 (0,49)	0,40 (0,49)	0,47 (0,50)	0,55 (0,50)
Nivel educativo alcanzado: básico secundario	0,36 (0,48)	0,37 (0,48)	0,33 (0,47)	0,27 (0,44)
Nivel educativo alcanzado: superior	0,19 (0,39)	0,18 (0,39)	0,12 (0,32)	0,06 (0,24)
Fertilidad				
Número de hijos	3,29 (2,31)	3,38 (1,64)	4,24 (1,56)	5,18 (1,49)
Edad al primer parto/nacimiento	20,87 (4,23)	20,33 (3,92)	19,65 (3,43)	19,12 (3,08)
Tuvo algún embarazo antes de los 19 años	0,38 (0,48)	0,37 (0,48)	0,42 (0,49)	0,47 (0,50)
Número de hijos que han fallecido	0,29 (0,72)	- -	- -	- -
Número de hijos que han fallecido (menores a 1 año)	0,20 (0,57)	- -	- -	- -
Tuvo algún aborto	0,20 (0,40)	0,22 (0,41)	0,22 (0,42)	0,22 (0,42)
Porcentaje de hijos varones	0,51 (0,34)	0,51 (0,29)	0,51 (0,26)	0,51 (0,23)
Madre tuvo nacimientos múltiples	0,05 (0,22)	0,05 (0,22)	0,05 (0,22)	0,05 (0,21)
Edad al <i>k</i> -ésimo parto/nacimiento	- -	23,84 (4,71)	26,25 (4,80)	28,19 (4,55)

Variable	Muestra completa	2 a más hijos	3 a más hijos	4 a más hijos
Antecedentes				
Nacimiento múltiple (<i>k-ésimo</i> parto/nacimiento) ^a	0,02 (0,15)	0,01 (0,09)	0,01 (0,10)	0,01 (0,11)
Métodos anticonceptivos hormonales	0,78 (0,41)	0,82 (0,38)	0,82 (0,39)	0,80 (0,40)
Dos primeros hijos tienen el mismo sexo	-	0,49 (0,50)	0,49 (0,50)	0,50 (0,50)
Quintil de riqueza				
Quintil inferior	0,20 (0,40)	0,19 (0,39)	0,22 (0,41)	0,27 (0,45)
Segundo quintil	0,24 (0,43)	0,23 (0,42)	0,26 (0,44)	0,29 (0,45)
Tercer quintil	0,23 (0,42)	0,23 (0,42)	0,23 (0,42)	0,22 (0,42)
Cuarto quintil	0,18 (0,39)	0,19 (0,40)	0,17 (0,37)	0,14 (0,34)
Quintil superior	0,14 (0,35)	0,16 (0,36)	0,13 (0,33)	0,08 (0,27)
Observaciones	84.266	48.211	29.705	16.892

Desviaciones estándar entre paréntesis. La composición de las muestras se detalla en la tabla 1.

a/ Indica si la mujer tuvo algún nacimiento múltiple en general para la muestra completa de mujeres.

Tabla 4
Muestra de niños (6 a 14 años): estadísticas descriptivas

Variable	2 a más hijos	3 a más hijos	4 a más hijos
Características del niño			
Edad	9,99 (2,53)	10,13 (2,35)	10,17 (2,32)
Sexo (igual a 1 si es hombre)	0,51 (0,50)	0,51 (0,50)	0,51 (0,50)
Orden de nacimiento: hijo mayor	1,00 (0,00)	0,50 (0,50)	0,33 (0,47)
Orden de nacimiento: segundo hijo	- -	0,50 (0,50)	0,33 (0,47)
Orden de nacimiento: tercer hijo	- -	- -	0,33 (0,47)
Intervalo de tiempo con nacimiento siguiente: 1 año o menos	0,18 (0,39)	0,30 (0,46)	0,36 (0,48)
Intervalo de tiempo con nacimiento siguiente: 2 a 3 años	0,42 (0,49)	0,58 (0,49)	0,60 (0,49)
Intervalo de tiempo con nacimiento siguiente: 4 años o más	0,40 (0,49)	0,12 (0,33)	0,04 (0,20)
Asiste a la escuela	0,96 (0,19)	0,96 (0,19)	0,95 (0,21)
Retraso escolar	0,28 (0,45)	0,39 (0,49)	0,50 (0,50)
Año normativo para la edad	-0,47 (0,96)	-0,68 (1,10)	-0,92 (1,24)
Características del hogar			
Sexo del jefe de hogar (igual a 1 si es hombre)	0,87 (0,33)	0,90 (0,30)	0,92 (0,27)
Edad del jefe de hogar	39,37 (11,44)	38,48 (9,67)	38,10 (7,93)
Lugar de residencia (igual a 1 si es urbano)	0,63 (0,48)	0,54 (0,50)	0,40 (0,49)
Quintil de riqueza: quintil inferior	0,18 (0,38)	0,26 (0,44)	0,36 (0,46)
Quintil de riqueza: segundo quintil	0,23 (0,42)	0,27 (0,45)	0,31 (0,48)
Quintil de riqueza: tercer quintil	0,25 (0,43)	0,23 (0,42)	0,20 (0,40) ▶

Variable	2 a más hijos	3 a más hijos	4 a más hijos
Características del niño			
Quintil de riqueza: cuarto quintil	0,20 (0,40)	0,14 (0,35)	0,09 (0,28)
Quintil de riqueza: quintil superior	0,14 (0,35)	0,09 (0,29)	0,04 (0,19)
Características de la madre			
Número de hijos	2,84 (1,08)	3,74 (0,96)	4,72 (0,88)
Edad al primer parto/nacimiento	20,89 (4,22)	20,09 (3,67)	19,55 (3,26)
Edad al <i>k</i> -ésimo parto/nacimiento	24,62 (4,92)	26,44 (4,57)	27,59 (3,93)
Nacimiento múltiple (<i>k</i> -ésimo parto/nacimiento)	0,01 (0,08)	0,01 (0,07)	0,00 (0,05)
Dos primeros hijos del mismo sexo	0,49 (0,50)	- -	- -
Años de educación	8,50 (4,41)	7,05 (4,27)	5,53 (3,71)
Observaciones	17.728	13.764	6.513

Desviaciones estándar entre paréntesis. La composición de las muestras se detalla en la tabla 2.

Tabla 5
Efecto de nacimientos múltiples e indicador de mismo sexo de los
dos hermanos mayores en el tamaño familiar

Muestra	Nacimiento múltiple	Mismo sexo	Observaciones
2 a más hijos	0,828*** (0,063) [0,337]	0,039*** (0,013) [0,334]	17.728
3 a más hijos	0,847*** (0,108) [0,339]	- - -	13.764
4 a más hijos	0,781*** (0,194) [0,235]	- - -	6.513

Errores estándar agrupados a nivel de la madre entre paréntesis. R2 entre corchetes.

El vector de variables de control en la ecuación estructural incluye la edad y el sexo del niño, edad de la madre al primer hijo, edad de la madre en el *k-ésimo* parto, años de educación de la madre, edad del jefe de hogar, sexo del jefe de hogar, dicotómicas por quintil de riqueza del hogar, indicador de ámbito geográfico, dicotómicas por departamento de residencia y dicotómicas por año de la encuesta. La composición de la muestra utilizada se detalla en la tabla 2.

Tabla 6
MCO y estimados de Wald: efectos del tamaño familiar en resultados educativos

Resultados educativos	2 a más hijos		3 a más hijos		4 a más hijos	
	MCO	Wald	MCO	Wald	MCO	Wald
		2do. parto fue múltiple		3er. parto fue múltiple		4to. parto fue múltiple
Asistencia a la escuela	-0,010*** (0,002) [0,003]	0,074*** (0,010) [0,030]	-0,010 (0,078) [0,003]	0,066*** (0,012) [0,025]	-0,018*** (0,004) [0,005]	0,021 (0,038) [0,028]
Retraso escolar	0,119*** (0,003) [0,082]	-0,128 (0,079) [0,020]	0,024 (0,180) [0,030]	-0,227* (0,121) [0,021]	0,129*** (0,008) [0,052]	-0,208 (0,383) [0,020]
Año o grado normativo	-0,287*** (0,009) [0,105]	0,243 (0,162) [0,120]	-0,354 (0,370) [0,100]	0,496** (0,240) [0,100]	-0,349*** (0,026) [0,062]	0,104 (0,657) [0,100]
Observaciones	17.728	17.728	17.728	13.764	6.513	6.513

Errores estándar agrupados a nivel de la madre entre paréntesis. R2 entre corchetes.

No se incluyen variables de control en ninguna de las especificaciones. La composición de las muestras se detalla en la tabla 2.

Tabla 7
MCO y MC2E: efecto del tamaño familiar
en la probabilidad de asistencia a la escuela

Muestra	MCO	MC2E		Observaciones
		Nacimiento múltiple	Mismo sexo	
2 a más hijos	-0,004** (0,002) [0,032]	0,036*** (0,005) [0,031]	-0,004 (0,072) [0,032]	17.728
3 a más hijos	-0,007*** (0,003) [0,025]	0,046*** (0,010) [0,024]	- - -	13.764
4 a más hijos	-0,014*** (0,005) [0,035]	0,008 (0,012) [0,028]	- - -	6.513

Errores estándar agrupados a nivel de la madre entre paréntesis. R2 entre corchetes.

El vector de variables de control incluye la edad y el sexo del niño, edad de la madre al primer hijo, edad de la madre en el *k*-ésimo parto, años de educación de la madre, edad del jefe de hogar, sexo del jefe de hogar, dicotómicas por quintil de riqueza del hogar, indicador de ámbito geográfico, dicotómicas por departamento de residencia y dicotómicas por año de la encuesta. La composición de la muestra utilizada se detalla en la tabla 2.

Tabla 8
MCO y MC2E: efecto del tamaño familiar
en la probabilidad de retraso escolar

Muestra	MCO	MC2E		Observaciones
		Nacimiento múltiple	Mismo sexo	
2 a más hijos	0,041*** (0,004) [0,245]	0,002 (0,041) [0,240]	0,030 (0,149) [0,245]	17.728
3 a más hijos	0,055*** (0,005) [0,250]	-0,058 (0,057) [0,213]	- - -	13.764
4 a más hijos	0,060*** (0,009) [0,209]	-0,223 (0,183) [0,200]	- - -	6.513

Errores estándar agrupados a nivel de la madre entre paréntesis. R2 entre corchetes.

El vector de variables de control incluye la edad y el sexo del niño, edad de la madre al primer hijo, edad de la madre en el *k*-ésimo parto, años de educación de la madre, edad del jefe de hogar, sexo del jefe de hogar, dicotómicas por quintil de riqueza del hogar, indicador de ámbito geográfico, dicotómicas por departamento de residencia y dicotómicas por año de la encuesta. La composición de la muestra utilizada se detalla en la tabla 2.

Tabla 9
MCO y MC2E: efecto del tamaño familiar en la edad normativa para el año/grado de educación básica

Muestra	MCO	MC2E		Observaciones
		Nacimiento múltiple	Mismo sexo	
2 a más hijos	-0,115*** (0,009) [0,267]	-0,034 (0,087) [0,261]	-0,337 (0,321) [0,225]	17.728
3 a más hijos	-0,157*** (0,013) [0,295]	0,106 (0,113) [0,255]	- - -	13.764
4 a más hijos	-0,150*** (0,023) [0,300]	0,156 (0,296) [0,146]	- - -	6.513

Errores estándar agrupados a nivel de la madre entre paréntesis. R2 entre corchetes.

El vector de variables de control incluye la edad y el sexo del niño, edad de la madre al primer hijo, edad de la madre en el *k-ésimo* parto, años de educación de la madre, edad del jefe de hogar, sexo del jefe de hogar, dicotómicas por quintil de riqueza del hogar, indicador de ámbito geográfico, dicotómicas por departamento de residencia y dicotómicas por año de la encuesta. La composición de la muestra utilizada se detalla en la tabla 2.

Tabla 10

Efecto de nacimientos múltiples e indicador de mismo sexo de los dos hermanos mayores en el tamaño familiar controlando por orden de nacimiento y distancia temporal

Muestra	Nacimiento múltiple (1)	Nacimiento múltiple (2)	Mismo sexo (1)	Mismo sexo (2)	Observaciones
2 a más hijos	0,857*** (0,059) [0,359]	0,857*** (0,059) [0,359]	0,040*** (0,013) [0,355]	0,040*** (0,013) [0,355]	17.728
3 a más hijos	0,873*** (0,112) [0,339]	0,812*** (0,104) [0,322]	- - -	- - -	13.764
4 a más hijos	0,726*** (0,179) [0,261]	0,764*** (0,196) [0,251]	- - -	- - -	6.513

Errores estándar agrupados a nivel de la madre entre paréntesis. R2 entre corchetes.

El vector de variables de control en la ecuación estructural incluye la edad y el sexo del niño, indicadores de orden de nacimiento del niño (para las muestras de 3 a más hijos y 4 a más hijos), dicotómicas de distancia temporal con respecto al nacimiento subsiguiente (entre 2 y 4 años y 4 a más años), edad de la madre al primer hijo, edad de la madre en el *k*-ésimo parto, años de educación de la madre, edad del jefe de hogar, sexo del jefe de hogar, dicotómicas por quintil de riqueza del hogar, indicador de ámbito geográfico, dicotómicas por departamento de residencia y dicotómicas por año de la encuesta. La segunda y cuarta columnas incluyen dicotómicas por orden de nacimiento para las muestras de 3 a más hijos y 4 a más hijos. La tercera y quinta columnas incluyen el indicador de orden de nacimiento relativo para las muestras de 3 a más hijos y 4 a más hijos. La composición de la muestra utilizada se detalla en la tabla 2.

Tabla 11
MCO y MC2E: efecto del tamaño familiar, orden de nacimiento y
distancia temporal en la probabilidad de asistencia a la escuela

Muestra	MCO		MC2E			Observaciones	
	(1)	(2)	Nac. múltiple (1)	Nac. múltiple (2)	Mismo sexo (1)	Mismo sexo (2)	
2 a más hijos							
Número de hijos	-0,004** (0,002)	-0,004** (0,002)	0,035*** (0,005)	0,035*** (0,005)	-0,004 (0,071)	-0,004 (0,071)	17.728
Distancia: 2-4 años	-0,007* (0,004)	-0,007* (0,004)	0,007 (0,004)	0,007 (0,004)	-0,007 (0,026)	-0,007 (0,026)	
Distancia: 4 a más años	-0,007 (0,005)	-0,007 (0,005)	0,014*** (0,005)	0,014*** (0,005)	-0,006 (0,039)	-0,006 (0,039)	
R2	[0,033]	[0,033]	[0,032]	[0,032]	[0,032]	[0,032]	
3 a más hijos							
Número de hijos	-0,008*** (0,003)	-0,008*** (0,003)	0,044*** (0,010)	0,047*** (0,011)	-	-	13.764
Orden: 2do. hijo	-0,003 (0,003)	-	-0,017*** (0,004)	-	-	-	
Orden relativo	-	-0,020*** (0,007)	-	0,010 (0,009)	-	-	
Distancia: 2-4 años	-0,004 (0,004)	-0,004 (0,004)	0,005 (0,004)	0,006 (0,004)	-	-	
Distancia: 4 a más años	-0,008 (0,007)	-0,008 (0,007)	0,014* (0,008)	0,019** (0,009)	-	-	
R2	[0,025]	[0,025]	[0,024]	[0,024]	-	-	
4 a más hijos							
Número de hijos	-0,014*** (0,005)	-0,015*** (0,005)	0,011 (0,013)	0,008 (0,012)	-	-	6.513
Orden: 2do. hijo	-0,032*** (0,006)	-	-0,025*** (0,007)	-	-	-	
Orden: 3er. hijo	-0,004 (0,009)	-	-0,029*** (0,008)	-	-	-	
Orden relativo	-	-0,035** (0,016)	-	-0,019 (0,019)	-	-	
Distancia: 2-4 años	0,002 (0,006)	0,001 (0,006)	0,002	0,002 (0,006)	-	-	
Distancia: 4 a más años	0,001 (0,015)	-0,011 (0,015)		0,002 (0,015)	-0,010 (0,015)	-	
R2	[0,040]	[0,035]		[0,033]	[0,029]	-	

- Errores estándar agrupados a nivel de la madre entre paréntesis. R2 entre corchetes.

El vector de variables de control en la ecuación estructural incluye la edad y el sexo del niño, edad de la madre al primer hijo, edad de la madre en el *k-ésimo* parto, años de educación de la madre, edad del jefe de hogar, sexo del jefe de hogar, dicotómicas por quintil de riqueza del hogar, indicador de ámbito geográfico, dicotómicas por departamento de residencia y dicotómicas por año de la encuesta. La segunda, cuarta y sexta columnas incluyen dicotómicas por orden de nacimiento para las muestras de 3 a más hijos y 4 a más hijos. La tercera, quinta y séptima columnas incluyen el indicador de orden de nacimiento relativo para las muestras de 3 a más hijos y 4 a más hijos. La composición de la muestra utilizada se detalla en la tabla 2.

Tabla 12
MCO y MC2E: efecto del tamaño familiar, orden de nacimiento y
distancia temporal en la probabilidad de retraso escolar

Muestra	MCO		MC2E				Observaciones
	(1)	(2)	Nac. múltiple (1)	Nac. múltiple (2)	Mismo sexo (1)	Mismo sexo (2)	
2 a más hijos							
Número de hijos	0,047*** (0,004)	0,047*** (0,004)	-0,002 (0,039)	-0,002 (0,039)	0,027 (0,148)	0,027 (0,148)	17.728
Distancia: 2-4 años	0,081*** (0,009)	0,081*** (0,009)	0,064*** (0,016)	0,064*** (0,016)	0,074 (0,053)	0,074 (0,053)	
Distancia: 4 a más años	0,105*** (0,011)	0,105*** (0,011)	0,078*** (0,024)	0,078*** (0,024)	0,094 (0,081)	0,094 (0,081)	
R2	[0,245]	[0,245]	[0,240]	[0,240]	[0,245]	[0,245]	
3 a más hijos							
Número de hijos	0,054*** (0,005)	0,059*** (0,005)	-0,050 (0,054)	-0,055 (0,060)	-	-	13.764
Orden: 2do. hijo	0,033*** (0,008)	-	0,061*** (0,017)	-	-	-	
Orden relativo	-	0,080*** (0,019)	-	0,018 (0,038)	-	-	
Distancia: 2-3 años	0,031*** (0,009)	0,030*** (0,009)	0,014 (0,013)	0,011 (0,014)	-	-	
Distancia: 4 a más años	0,041*** (0,015)	0,041*** (0,015)	-0,004 (0,028)	-0,015 (0,034)	-	-	
R2	[0,252]	[0,252]	[0,221]	[0,215]	-	-	
4 a más hijos							
Número de hijos	0,050*** (0,009)	0,067*** (0,009)	-0,257 (0,201)	-0,225 (0,189)	-	-	6.513
Orden: 2do. hijo	0,142*** (0,015)	-	0,228*** (0,058)	-	-	-	
Orden: 3er. hijo	0,169*** (0,024)	-	0,349*** (0,121)	-	-	-	
Orden relativo	-	0,277*** (0,041)	-	0,072 (0,139)	-	-	
Distancia: 2-3 años	0,001 (0,013)	-0,001 (0,013)	-0,005 (0,015)	-0,013 (0,016)	-	-	
Distancia: 4 a más años	0,025 (0,029)	0,003 (0,029)	0,019 (0,032)	-0,016 (0,033)	-	-	
R2	[0,219]	[0,216]	[0,05]	[0,018]	-	-	

- ▶ Errores estándar agrupados a nivel de la madre entre paréntesis. R2 entre corchetes.
- El vector de variables de control en la ecuación estructural incluye la edad y el sexo del niño, edad de la madre al primer hijo, edad de la madre en el *k-ésimo* parto, años de educación de la madre, edad del jefe de hogar, sexo del jefe de hogar, dicotómicas por quintil de riqueza del hogar, indicador de ámbito geográfico, dicotómicas por departamento de residencia y dicotómicas por año de la encuesta. La segunda, cuarta y sexta columnas incluyen dicotómicas por orden de nacimiento para las muestras de 3 a más hijos y 4 a más hijos. La tercera, quinta y séptima columnas incluyen el indicador de orden de nacimiento relativo para las muestras de 3 a más hijos y 4 a más hijos. La composición de la muestra utilizada se detalla en la tabla 2.

Tabla 13
MCO y MC2E: efecto del tamaño familiar, orden de nacimiento
y distancia temporal en la edad normativa para el año/grado
de educación básica

Muestra	MCO		MC2E			Observaciones	
	(1)	(2)	Nac. múltiple (1)	Nac. múltiple (2)	Mismo sexo (1)	Mismo sexo (2)	
2 a más hijos							
Número de hijos	-0,129*** (0,009)	-0,129*** (0,009)	-0,024 (0,084)	-0,024 (0,084)	-0,332 (0,317)	-0,332 (0,317)	17.728
Distancia: 2-4 años	-0,158*** (0,019)	-0,158*** (0,019)	-0,121*** (0,035)	-0,121*** (0,035)	-0,231** (0,114)	-0,231** (0,114)	
Distancia: 4 a más años	-0,226*** (0,022)	-0,226*** (0,022)	-0,169*** (0,051)	-0,169*** (0,051)	-0,336* (0,174)	-0,336* (0,174)	
R2	[0,267]	[0,267]	[0,261]	[0,261]	[0,225]	[0,225]	
3 a más hijos							
Número de hijos	-0,157*** (0,013)	-0,169*** (0,013)	0,091 (0,109)	0,089 (0,119)	-	-	13.764
Orden: 2do. hijo	-0,073*** (0,017)	-	-0,141*** (0,034)	-	-	-	
Orden relativo	-	-0,289*** (0,037)	-	-0,148* (0,076)	-	-	
Distancia: 2-4 años	-0,046** (0,020)	-0,045** (0,020)	-0,005 (0,027)	-0,001 (0,029)	-	-	
Distancia: 4 a más años	-0,123*** (0,035)	-0,126*** (0,035)	-0,015 (0,059)	0,001 (0,069)	-	-	
R2	[0,297]	[0,298]	[0,262]	[0,261]	-	-	
4 a más hijos							
Número de hijos	-0,130*** (0,023)	-0,169*** (0,023)	0,541 (0,331)	0,468 (0,307)	-	-	6.513
Orden: 2do. hijo	-0,225*** (0,036)	-	-0,412*** (0,101)	-	-	-	
Orden: 3er. hijo	-0,335*** (0,053)	-	-0,727*** (0,205)	-	-	-	
Orden relativo	-	-0,726*** (0,090)	-	-0,281 (0,234)	-	-	
Distancia: 2-4 años	-0,029 (0,029)	-0,029 (0,029)	-0,016 (0,034)	-0,003 (0,036)	-	-	
Distancia: 4 a más años	-0,153* (0,079)	-0,133* (0,078)	-0,141* (0,084)	-0,092 (0,085)	-	-	
R2	[0,305]	[0,308]	[0,139]	[0,154]	-	-	

- Errores estándar agrupados a nivel de la madre entre paréntesis. R2 entre corchetes.

El vector de variables de control en la ecuación estructural incluye la edad y el sexo del niño, edad de la madre al primer hijo, edad de la madre en el *k-ésimo* parto, años de educación de la madre, edad del jefe de hogar, sexo del jefe de hogar, dicotómicas por quintil de riqueza del hogar, indicador de ámbito geográfico, dicotómicas por departamento de residencia y dicotómicas por año de la encuesta. La segunda, cuarta y sexta columnas incluyen dicotómicas por orden de nacimiento para las muestras de 3 a más hijos y 4 a más hijos. La tercera, quinta y séptima columnas incluyen el indicador de orden de nacimiento relativo para las muestras de 3 a más hijos y 4 a más hijos. La composición de la muestra utilizada se detalla en la tabla 2.

Tabla 14
MC2E: efecto del tamaño familiar en la probabilidad de asistencia a la escuela

Muestra	Ámbito geográfico		Sexo		Composición del hogar		Distancia temporal entre nacimientos	
	Urbano	Rural	Niños	Niñas	Padre presente	Padre ausente	Menos de 2 años	Menos de 2 años y edad menor o igual a 10 años
2 a más hijos	0,037*** (0,005) [0,026] {11.197}	0,046*** (0,016) [0,031] {6.531}	0,050*** (0,008) [0,042] {9.107}	0,034*** (0,007) [0,029] {8.621}	0,042*** (0,006) [0,033] {14.158}	0,038*** (0,012) [0,033] {3.570}	0,049** (0,020) [0,040] {3.233}	0,027** (0,014) [0,036] {1.757}
	0,038*** (0,007) [0,024] {7.388}	0,068*** (0,020) [0,033] {6.376}	0,050*** (0,010) [0,039] {6.960}	0,038*** (0,011) [0,029] {6.804}	0,042*** (0,009) [0,028] {11.743}	0,035*** (0,019) [0,026] {2.021}	0,031*** (0,007) [0,031] {4.132}	0,021** (0,018) [0,033] {2.258}

Errores estándar agrupados a nivel de la madre entre paréntesis. R2 entre corchetes. Número de observaciones entre llaves.

El instrumento utilizado para la identificación corresponde al indicador de nacimientos múltiples en el segundo (muestra de 2 a más hijos) y tercer (muestra de 3 a más hijos) parto de la madre. El vector de variables de control incluye la edad y el sexo del niño (para las columnas 2-3 y 6-9), dicotómicas por orden de nacimiento (para la muestra de 3 a más hijos), dicotómicas por distancia temporal entre nacimientos (para las columnas 2-7), edad de la madre al primer hijo, edad de la madre en el *k-ésimo* parto, años de educación de la madre, edad del jefe de hogar, sexo del jefe de hogar, dicotómicas por quintil de riqueza del hogar, indicador de ámbito geográfico (para las columnas 4-9), dicotómicas por departamento de residencia y dicotómicas por año de la encuesta. La composición de la muestra utilizada se detalla en la tabla 2.

Tabla 15
MCO y MC2E: efecto del tamaño familiar en la probabilidad de
compartir habitación y en el estado civil de la madre

Muestra	Compartir habitación		Madre divorciada		Observaciones
	MCO	MC2E	MCO	MC2E	
2 a más hijos	0,041*** (0,002) [0,094]	0,064** (0,029) [0,090]	-0,030*** (0,002) [0,017]	-0,038 (0,028) [0,016]	17.728
3 a más hijos	0,011*** (0,002) [0,034]	0,011 (0,031) [0,034]	-0,021*** (0,003) [0,014]	-0,049 (0,032) [0,006]	13.764
4 a más hijos	0,001 (0,001) [0,025]	0,011 (0,009) [0,013]	-0,012** (0,005) [0,029]	-0,072 (0,051) [0,029]	6.513

Errores estándar agrupados a nivel de la madre entre paréntesis. R2 entre corchetes.

El instrumento utilizado para la identificación corresponde al indicador de nacimientos múltiples en el segundo (muestra de 2 a más hijos), tercer (muestra de 3 a más hijos) y cuarto (muestra de 4 a más hijos) parto de la madre. El vector de variables de control incluye la edad y el sexo del niño, indicadores de orden de nacimiento del niño (para las muestras de 3 a más hijos y 4 a más hijos), dicotómicas de distancia temporal con respecto al nacimiento subsiguiente (entre 2 y 4 años y 4 a más años), edad de la madre al primer hijo, edad de la madre en el *k-ésimo* parto, años de educación de la madre, edad del jefe de hogar, sexo del jefe de hogar, dicotómicas por quintil de riqueza del hogar, indicador de ámbito geográfico, dicotómicas por departamento de residencia y dicotómicas por año de la encuesta. La composición de la muestra utilizada se detalla en la tabla 2.

ANEXO A: VALIDEZ DE LOS INSTRUMENTOS

Tabla A.1
Nacimientos múltiples y sexo de los dos primeros hijos: Diferencia de promedios muestrales

Variables	Nacimientos múltiples										Sexo de los dos primeros hijos	
	2 a más hijos		3 a más hijos		4 a más hijos		4 a más hijos		Sing.		Múlt.	
	Sing.	Múlt.	Est. t	Sing.	Múlt.	Est. t	Sing.	Múlt.	Est. t	Sing.	Múlt.	Est. t
Características del niño												
Edad	10,00	9,76	1,02	10,13	9,69	1,57	10,16	11,27	-1,84*	10,01	9,99	0,51
Sexo (igual a 1 si es hombre)	0,87	0,82	1,73*	0,90	0,92	-0,39	0,92	0,80	1,76*	0,87	0,88	-1,12
Distancia con nacimiento siguiente:												
1 año o menos	0,18	0,11	1,97**	0,30	0,32	-0,36	0,36	0,60	-1,93**	0,18	0,18	0,59
Distancia con nacimiento siguiente:												
2 a 3 años	0,42	0,22	4,52***	0,58	0,54	0,61	0,60	0,40	1,56	0,42	0,42	-0,17
Distancia con nacimiento siguiente:												
4 años o más	0,39	0,67	-6,12***	0,12	0,14	-0,41	0,04	0,00	0,80	0,39	0,40	-0,29
Asiste a la escuela	0,94	0,98	-2,14**	0,95	0,99	-1,66*	0,95	0,96	-0,87	0,96	0,96	0,13
Retraso escolar	0,28	0,22	1,57	0,39	0,36	1,19	0,58	0,48	0,91	0,28	0,28	-0,13
Año o grado normativo para la edad	-0,47	-0,34	-1,40	-0,68	-0,60	-1,13	-0,92	-0,87	-1,05	-0,46	-0,47	0,91
Características del hogar												
Sexo del jefe de hogar (igual a 1 si es hombre)	0,87	0,82	1,73*	0,90	0,92	-0,39	0,92	0,80	1,76*	0,87	0,88	-1,12
Edad del jefe de hogar	39,37	39,35	0,02	38,47	39,75	-1,12	38,08	45,60	-3,67***	39,39	39,35	0,20
Lugar de residencia (igual a 1 si es urbano)	0,63	0,70	-1,49	0,54	0,63	-1,51	0,40	0,80	-3,14***	0,64	0,63	1,06

Variables	Nacimientos múltiples						Sexo de los dos primeros hijos								
	2 a más hijos		3 a más hijos		4 a más hijos		4 a más hijos		Sing.		Múlt.		Est. t		
	Sing.	Múlt.	Est. t	Sing.	Múlt.	Est. t	Sing.	Múlt.	Est. t	Sing.	Múlt.	Est. t	Sing.	Múlt.	Est. t
Quintil de riqueza: quintil inferior	0,18	0,07	3,02***	0,26	0,21	0,98	0,37	0,00	2,94***	0,17	0,18	-0,85	0,17	0,18	-0,85
Quintil de riqueza: segundo quintil	0,23	0,19	1,09	0,27	0,17	2,05**	0,31	0,20	0,93	0,24	0,23	1,03	0,24	0,23	1,03
Quintil de riqueza: tercer quintil	0,25	0,29	-1,03	0,23	0,29	-1,28	0,20	0,40	-1,97**	0,25	0,25	-0,39	0,25	0,25	-0,39
Quintil de riqueza: cuarto quintil	0,20	0,26	-1,67	0,14	0,17	-0,53	0,09	0,40	-4,30***	0,20	0,20	-0,18	0,20	0,20	-0,18
Quintil de riqueza: quintil superior	0,14	0,19	-1,41	0,09	0,17	-2,11**	0,04	0,00	0,79	0,14	0,14	0,36	0,14	0,14	0,36
Características de la madre															
Número de hijos	2,84	3,35	-5,09***	3,73	4,29	-4,94***	4,72	5,20	-2,10**	2,83	2,86	-2,27**	2,83	2,86	-2,27**
Edad al primer parto/nacimiento	20,89	21,41	-1,33	20,08	22,42	-5,4***	19,55	19,60	-0,06	20,92	20,87	0,82	20,92	20,87	0,82
Edad al <i>k-ésimo</i> parto/nacimiento	24,61	26,20	-3,46***	26,42	29,72	-6,13***	27,59	29,00	-1,39	24,63	24,61	0,28	24,63	24,61	0,28
Años de educación	8,49	9,46	-2,35**	7,05	7,40	-0,71	5,52	6,60	-1,12	8,52	8,48	0,55	8,52	8,48	0,55

La composición de las muestras se detalla en la tabla 2.

Tabla A.2
MCO: validez de instrumentos

Variable dependiente	Variable(s) explicativa(s) de interés	Controles adicionales	Estadístico Prueba F	Prob > F
Nacimiento múltiple en el segundo parto	- Dicotómicas por nivel educativo (primaria, secundaria, superior)	- Edad de la mujer al primer parto	F(7, 48198) = 0,81	0,583
	- Dicotómicas por quintil de riqueza (segundo quintil, tercer quintil, cuarto quintil, quintil superior)	- Edad de la mujer al segundo parto - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)		
Nacimiento múltiple en el segundo parto	- Dicotómica igual a 1 si la madre tuvo gemelos o mellizos	- Edad de la mujer al primer parto	F(1, 48204) = 1,59	0,111
		- Edad de la mujer al segundo parto - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)		
Nacimiento múltiple en el segundo parto	- Dicotómica igual a 1 si la mujer ha usado métodos anticonceptivos hormonales	- Edad de la mujer al primer parto - Edad de la mujer al segundo parto - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 48204) = 1,29	0,197

La muestra utilizada corresponde a la de mujeres con "2 a más hijos". La composición de la muestra se detalla en la tabla 1.

Tabla A.3
MCO: validez de instrumentos

Variable dependiente	Variable(s) explicativa(s) de interés	Controles adicionales	Estadístico Prueba F	Prob > F
Nacimiento múltiple en el tercer parto	-Dicotómicas por nivel educativo (primaria, secundaria, superior) - Dicotómicas por quintil de riqueza (segundo quintil, tercer quintil, cuarto quintil, quintil superior)	- Edad de la mujer al primer parto - Edad de la mujer al tercer parto - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(7, 29692) = 1,39	0,203
Nacimiento múltiple en el tercer parto	- Dicotómica igual a 1 si la madre tuvo gemelos o mellizos	- Edad de la mujer al primer parto - Edad de la mujer al tercer parto - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 29698) = 0,88	0,347
Nacimiento múltiple en el tercer parto	- Dicotómica igual a 1 si la mujer ha usado métodos anticonceptivos hormonales	- Edad de la mujer al primer parto - Edad de la mujer al segundo parto - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 29698) = 2,55	0,111

La muestra utilizada corresponde a la de mujeres con "3 a más hijos". La composición de la muestra se detalla en la tabla 1.

Tabla A.4
MCO: validez de instrumentos

Variable dependiente	Variable(s) explicativa(s) de interés	Controles adicionales	Estadístico Prueba F	Prob > F
Nacimiento múltiple en el cuarto parto	- Dicotómicas por nivel educativo (primaria, secundaria, superior) - Dicotómicas por quintil de riqueza (segundo quintil, tercer quintil, cuarto quintil, quintil superior)	- Edad de la mujer al primer parto - Edad de la mujer al cuarto parto - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(7, 16879) = 1,10	0,358
Nacimiento múltiple en el cuarto parto	- Dicotómica igual a 1 si la madre tuvo gemelos o mellizos	- Edad de la mujer al primer parto - Edad de la mujer al cuarto parto - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 16885) = 0,45	0,502
Nacimiento múltiple en el cuarto parto	- Dicotómica igual a 1 si la mujer ha usado métodos anticonceptivo hormonales	- Edad de la mujer al primer parto - Edad de la mujer al segundo parto - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 16885) = 1,33	0,248

La muestra utilizada corresponde a la de mujeres con "4 a más hijos". La composición de la muestra se detalla en la tabla 1.

Tabla A.5
MCO: validez de instrumentos

Variable dependiente	Muestra	Variable(s) explicativa(s) de interés	Controles adicionales	Estadístico Prueba F	Prob > F
Asiste a la escuela	-	- Inversa de la distancia temporal entre los dos nacimientos que siguen	- Edad - Sexo - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 13684) = 3,82	0,051
Retraso escolar	-	- Inversa de la distancia temporal entre los dos nacimientos que siguen	- Edad - Sexo - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 13684) = 3,71	0,058
Edad normativa para el año/grado de educación básica	-	- Inversa de la distancia temporal entre los dos nacimientos que siguen	- Edad - Sexo - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 13684) = 2,71	0,100
Asiste a la escuela	2 a más hijos	- Nacimiento múltiple (2do. parto)	- Edad - Sexo - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 17718) = 0,03	0,873
	3 a más hijos	- Nacimiento múltiple (3er. parto)	- Número de hijos	F(1, 13754) = 0,02	0,899
	4 a más hijos	- Nacimiento múltiple (4to. parto)	- Edad de la madre al primer nacimiento - Edad de la madre al <i>k-ésimo</i> parto - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 6503) = 0,69	0,408



Variable dependiente	Muestra	Variable(s) explicativa(s) de interés	Controles adicionales	Estadístico Prueba F	Prob > F
Retraso escolar	2 a más hijos	- Nacimiento múltiple (2do. parto)	- Edad - Sexo	F(1, 17718) = 0,02	0,892
	3 a más hijos	- Nacimiento múltiple (3er. parto)	- Número de hijos - Edad de la madre al primer nacimiento	F(1, 13754) = 2,02	0,155
	4 a más hijos	- Nacimiento múltiple (4ro. parto)	- Edad de la madre al <i>k-ésimo</i> parto - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 6503) = 0,62	0,433
Edad normativa para el año/grado de educación básica	2 a más hijos	- Nacimiento múltiple (2do. parto)	- Edad - Sexo	F(1, 17719) = 0,54	0,461
	3 a más hijos	- Nacimiento múltiple (3er. parto)	- Número de hijos - Edad de la madre al primer nacimiento	F(1, 13754) = 1,01	0,315
	4 a más hijos	- Nacimiento múltiple (4ro. parto)	- Edad de la madre al <i>k-ésimo</i> parto - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 6503) = 0,76	0,382

Tabla A.6
MCO: validez de instrumentos

Variable dependiente	Variable(s) explicativa(s) de interés	Controles adicionales	Estadístico Prueba F	Prob > F
Aborto (urbano)	- Razón de hijos varones respecto del total de hijos nacidos vivos	- Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 50450) = 0,41	0,521
Aborto (rural)	- Razón de hijos varones respecto del total de hijos nacidos vivos	- Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 33806) = 0,05	0,815
Aborto (urbano)	- Interacción entre el lugar de nacimiento de la madre (igual a 1 si es ciudad) y la razón de hijos varones respecto del total de hijos nacidos vivos	- Dicotómica igual a 1 si la mujer ha nacido en ciudad - Porcentaje de hijos varones del total de hijos nacidos vivos - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 50449) = 0,40	0,526
Aborto (rural)	- Interacción entre el lugar de nacimiento de la madre (igual a 1 si es ciudad) y la razón de hijos varones respecto del total de hijos nacidos vivos	- Dicotómica igual a 1 si la mujer ha nacido en ciudad - Porcentaje de hijos varones del total de hijos nacidos vivos - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 33805) = 0,07	0,792
Mortalidad infantil (urbano)	- Razón de hijos varones respecto del total de hijos nacidos vivos	- Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 50449) = 0,31	0,572
Mortalidad infantil (rural)	- Razón de hijos varones respecto del total de hijos nacidos vivos	- Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 33806) = 0,64	0,423



Variable dependiente	Variable(s) explicativa(s) de interés	Controles adicionales	Estadístico Prueba F	Prob > F
Mortalidad infantil (urbano)	- Interacción entre el lugar de nacimiento de la madre (igual a 1 si es ciudad) y la razón de hijos varones respecto del total de hijos nacidos vivos	- Dicotómica igual a 1 si la mujer ha nacido en ciudad - Porcentaje de hijos varones del total de hijos nacidos vivos - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 50449) = 0,40	0,526
Mortalidad infantil (rural)	- Interacción entre el lugar de nacimiento de la madre (igual a 1 si es ciudad) y la razón de hijos varones respecto del total de hijos nacidos vivos	- Dicotómica igual a 1 si la mujer ha nacido en ciudad - Porcentaje de hijos varones del total de hijos nacidos vivos - Dicotómicas por año de encuesta (2000, 2004-2008, 2009)	F(1, 33805) = 0,75	0,388

La muestra utilizada corresponde a la "muestra completa de mujeres". La composición de la muestra se detalla en la tabla 1.

Número de hermanos, orden de nacimiento
y resultados educativos en la niñez:
evidencia en Perú.

Se terminó de editar en el
mes de diciembre de 2011

