



## EVALUACIÓN DE IMPACTO DE POLÍTICAS Y PROGRAMAS A TRAVÉS DE MICROSIMULACIÓN

SONIA ESPERANZA MONROY VARELA

Universidad Nacional de Colombia, Departamento de Ingeniería de Sistemas e Industrial, Colombia  
[semonroy@unal.edu.co](mailto:semonroy@unal.edu.co)

HERNANDO DIAZ MORALES

Universidad Nacional de Colombia, Departamento de Ingeniería Eléctrica y Electrónica, Colombia  
[hdiazmo@unal.edu.co](mailto:hdiazmo@unal.edu.co)

FABIAN HUMBERTO GÓMEZ NIVIA

Universidad Nacional de Colombia, Facultad de Ciencias Económicas, Colombia  
[fhgomezn@unal.edu.co](mailto:fhgomezn@unal.edu.co)

WILLIAM DAVID CAMACHO ARDILA

Universidad Nacional de Colombia, Estudiante de Ingeniería Industrial, Colombia  
[wdcamachoa@unal.edu.co](mailto:wdcamachoa@unal.edu.co)

### RESUMEN

En este artículo se presenta un modelo desarrollado para realizar evaluación de impacto *ex ante* de un programa de transferencias monetarias condicionadas. La evaluación de los impactos de un programa de inversión es fundamental para el diseño de políticas públicas: sus resultados permiten decidir si el programa debe eliminarse, modificarse o continuar tal como fue diseñado. El modelo de evaluación se basa en microsimulaciones que tienen como objetivo predecir las decisiones que tomarían los individuos a partir de sus características y las de su entorno. De esta manera, se puede establecer el comportamiento esperado y predecir los efectos del programa.

El modelo fue utilizado para evaluar los impactos de un programa de inversión social: el programa *Bolsa Familia* de Brasil, el cual fue creado para reducir los índices de pobreza y fomentar la inserción escolar. Con base en las características de los individuos se estimaron las probabilidades de que la familia escoja cada una de las opciones disponibles: Enviar el niño al colegio, ponerlo a trabajar o combinar las dos actividades. A partir de estos datos, usando una simulación probabilística, se estiman los valores esperados de los porcentajes de niños que asisten al colegio. Los resultados, evaluados para el programa *Bolsa Familia* en el año 2011, se comparan con los del anterior Programa *Bolsa Escola*, obtenidos para el año 2003. Se observa que, aunque ambos programas han tenido un impacto notable, los resultados de asistencia escolar por edades han venido mejorando y el programa debería considerarse un éxito y un ejemplo a seguir.

La utilidad del modelo desarrollado yace en la capacidad de predicción de la decisión que tomarían diversos individuos y a partir de ello formular políticas públicas con programas sociales más efectivos. La intención del modelo es que pueda ser replicado en diferentes programas de inversión social y con esto evaluar políticas públicas de manera *ex ante*.



## PALABRAS CLAVES

Evaluación de impacto *Ex ante*. Modelo de microsimulación. Transferencias condicionadas de dinero.

### 1. INTRODUCCIÓN

En la actualidad existe un consenso sobre la importancia de evaluar los impactos de las políticas públicas, dadas las grandes inversiones realizadas y la importancia de la replicabilidad de las mismas. Ante la limitación de recursos de inversión, es fundamental que las políticas escogidas y los programas diseñados sean los que produzcan los impactos deseados en la población.

La evaluación de impacto difiere de la evaluación de resultados de los proyectos o programas. En efecto, para evaluar los resultados, se requiere un análisis cuidadoso de los insumos, actividades y resultados del proyecto o programa con el fin de determinar qué fue lo que produjo como fruto de la financiación recibida. Generalmente se refiere a los efectos de corto plazo de una decisión de inversión o financiación.

La evaluación de impacto, por su parte, analiza “los efectos de largo plazo, tanto positivos como negativos, causados por la intervención de políticas, programas o proyectos”<sup>1</sup>. En general, considera la incidencia del programa de intervención en todos los aspectos sociales, económicos ambientales, políticos, etc., tanto en el mediano como en el largo plazo.

Por lo general, las evaluaciones de impacto se hacen *ex post*, es decir después de que el programa se ha llevado a cabo. Pero, recientemente, se ha hecho patente la necesidad de poder hacer evaluaciones de los efectos de una política económica antes de su implementación (evaluaciones *ex ante*), con el fin de poder seleccionar las políticas que producen el mayor efecto posible a un costo determinado, o que garantizan un efecto mientras se minimiza el costo. En realidad, se hace deseable poder predecir distintos efectos sobre los agentes del sistema económico antes de emprender programas o políticas cuyos efectos globales reales no se conocen con exactitud.

La evaluación de impacto *ex ante* pretende predecir el efecto de la financiación de planes, programas y/o proyectos o de políticas de inversión sobre la sociedad. La predicción se basa en la simulación de los efectos de la financiación de los programas a través de modelos de equilibrio general o de modelos de simulación tanto determinística como estocástica.

Con este propósito, se desarrollaron modelos que pretendían predecir el resultado de las decisiones económicas con base en modelos del comportamiento de los individuos que toman decisiones buscando maximizar su bienestar. A partir de estos modelos se desarrollaron métodos de evaluación de impacto *ex ante*. El modelo de comportamiento permitía predecir el resultado de una política y, a partir de esa predicción, se podía efectuar el proceso de optimización para seleccionar la política óptima en cada situación.

Recientemente, han sido considerados otros métodos de evaluación *ex ante*, basados en modelos de simulación, los cuales permiten considerar en forma mucho más detallada las heterogeneidades siempre presentes en el sistema económico. Tanto la variabilidad entre

---

<sup>1</sup> Definición de la OECD (2002). Glosario de términos principales en materia de evaluación de gestión basada en los resultados.



diferentes individuos que componen una población como las fluctuaciones debidas a condiciones variables por situaciones geográficas o sociales pueden ser incorporadas en los modelos de simulación. Actualmente, los modelos de evaluación de impacto basados en simulación, son de dos tipos: microsimulaciones y modelos basados en agentes.

## **2. MARCO TEÓRICO – ESTADO DEL ARTE**

### **2.1 Modelos de microsimulación**

La microsimulación ha sido definida como “una técnica usada para modelar eventos complejos de la vida real por medio de la simulación de las acciones y los efectos de los cambios de política sobre las unidades individuales (micro unidades) que constituyen el sistema donde los eventos ocurren” (Harding 2007). Este tipo de modelos ha sido utilizado en economía en forma extensa. Las aplicaciones incluyen dinámica de poblaciones para evaluar efectos de políticas de empleo o de regímenes pensionales. También hay aplicaciones en el planeamiento urbano o económico general.

En los modelos de microsimulación la población objeto de estudio se modela por una población sintética con características similares a la población real. Esto significa que es necesario disponer de grandes cantidades de información, dependiendo del estudio. También se incluyen modelos de comportamiento para la toma de decisiones, tal como en los modelos econométricos. Sin embargo, en las microsimulaciones, el modelo de comportamiento se utiliza a nivel individual y existe un modelo para cada individuo que compone la población.

El desarrollo de modelos de microsimulación involucra una serie de problemas, si se van a usar los datos para hacer predicciones económicas significativas. Estos problemas fueron expuestos en forma extensa por Harding (2007). Li, y O’Donoghue (2012), estudiaron algunos de los problemas más importantes. Estos autores analizan el efecto de la información disponible (data set) para crear las poblaciones sintéticas que componen el modelo. Puesto que se requieren numerosos datos tanto demográficos como de empleo, educación, movilidad, etc., es fundamental generar una base de datos que permita describir todos los aspectos relevantes de cada individuo para la simulación. A menudo los datos disponibles no cubren el mismo período de tiempo y se hace necesario completar las series mediante algún mecanismo que conserve la integridad de la información. Li (2011) desarrolló un algoritmo denominado *Simulación en reversa* (“*back simulation*”) que permite reconstruir la información necesaria para tener un modelo consistente con la respuesta histórica del sistema que se está simulando.

El segundo problema afrontado por Li y O’Donoghue (2012) fue el del alineamiento del modelo. Esto consiste en calibrar los parámetros del modelo de microsimulación para que sus resultados agregados a nivel de población o subpoblaciones coincidan con las proyecciones macroeconómicas. Este alineamiento puede hacerse para instantes discretos del tiempo o buscar que se mantenga a lo largo de la simulación. Li (2011) desarrolló también algoritmos para el alineamiento, mediante la técnica de “back simulation”.

Un modelo desarrollado por Li y O’Donoghue (2012) se utilizó para hacer un análisis de impacto de diferentes incentivos monetarios comprendidos en un plan de beneficios tributarios para trabajadores adultos mayores en Irlanda. También construyeron un modelo para predecir las decisiones de jubilación de trabajadores con base en la situación de su hogar. Se considera que la



decisión de retiro de los empleados no depende solamente de su situación personal sino que se ve influenciada por las condiciones familiares, incluyendo los ingresos de todos los miembros de la familia que viven juntos. Usando el modelo se analiza un plan de reforma pensional para cambiar la edad mínima a la cual los mayores tienen derecho a una pensión del Estado.

Canova, Piccoli y Spadaro (2014) utilizaron un modelo combinado micro-macro simulación para evaluar el impacto de la introducción de un programa de bienestar del Estado francés, el Revenu de Solidarité Active (RSA), para remplazar un programa anterior, basado en valores mínimos de indicadores socio-económicos. A partir de la simulación se cuantifican los efectos del programa sobre los aspectos económicos a nivel de hogar y se evalúan indicadores micro y macroeconómicos. Las simulaciones mostraron que aunque el gasto público aumentaba con el nuevo programa, los efectos adicionales esperados: reducción de desempleo voluntario, aumento en la oferta de mano de obra y en el consumo, además de reducciones en pobreza y desigualdad, justifican ampliamente el programa.

De forma similar, Zucchelli, E., Jones, A. M., & Rice, N. (2012) desarrollaron un modelo de microsimulación para evaluar los efectos de políticas alternativas de salud pública y de atención médica. En este modelo se incluyen submodelos específicos para enfermedades determinadas y modelos de riesgo.

Leite, P., A. Narayan, and E. Skoufias, (2011), llevaron a cabo una comparación entre los impactos, estimados por medio de modelos de microsimulación, de programas de transferencias condicionales de dinero con los valores realmente medidos (ex post) en experimentos aleatorizados realizados en México (programa PROGRESA) y Ecuador. El resultado muestra que los valores predichos ex ante mediante la microsimulación aproximan con precisión razonable los obtenidos ex post mediante métodos experimentales y cuasi experimentales.

En un trabajo similar al anterior, Bourguignon *et al* (2003) evaluaron el impacto de un programa de transferencias condicionadas de dinero en Brasil: el programa Bolsa Escola. En dicho proyecto se otorgan subsidios en efectivo a familias cuyos hijos, entre 5 y 15 años, asisten regularmente a la escuela. Los autores desarrollaron un modelo de función de utilidad usado para simular la toma de decisiones en el seno de una familia, basado en la información sobre los ingresos familiares, las características de la familia y las del niño. Este artículo ha servido como punto de partida para el presente trabajo.

## 2.2 Modelos basados en agentes

Todos los modelos económicos, ya sean de equilibrio general global o de microsimulación, se basan en una serie de suposiciones acerca del proceso de toma de decisiones por parte de los individuos que intervienen en la economía. Es necesario hacer algunas suposiciones sobre las características del mercado para garantizar que las funciones de utilidad sobre las cuales se basan los cálculos representen lo que realmente consideran los individuos caen al momento de tomar una decisión. En otras palabras, se supone que los individuos que componen un sistema económico toman decisiones para maximizar su beneficio, en el contexto de unas condiciones de mercado perfecto. Sin embargo, cuando existen distorsiones que hacen inválidas estas condiciones, es mucho más difícil evaluar el efecto de programas o políticas. También se encuentran dificultades para analizar y aun para simular los comportamientos adaptativos de los



individuos a medida que aprenden a reconocer las condiciones del mercado y adaptan sus comportamientos con base en percepciones de efectos directos e indirectos.

Existe un enfoque alternativo, originado en forma independiente en varias áreas del saber (ecología, inteligencia artificial, redes de computadores, etc.), en el cual se simulan los comportamientos de los individuos en forma independiente. Se trata del llamado enfoque de Modelos Basados en Agentes o ABM por su sigla en inglés (“Agent Based Models”). Estos son modelos de simulación, intensivos en computación, en los cuales se simulan comportamientos individuales de una población determinada a través de entidades independientes (agentes) representadas en la simulación por medio de reglas o funciones de decisión. Esto permite una mayor flexibilidad ya que los agentes pueden asumir cualquier tipo de comportamiento, no necesariamente determinado por condiciones económicas. Además, los agentes pueden comunicarse a través de mecanismos previstos en el ambiente donde se desarrolla la dinámica. Y es posible prever la existencia de subsistemas con funciones de coordinación.

Recientemente, se han empleado modelos basados en agentes para hacer evaluaciones de impacto de programas sociales o agrícolas, especialmente para modelar las políticas de intervención en mercados agrícolas. Sin embargo, los modelos basados en agentes han sido usados mucho menos frecuentemente que otras alternativas.

Filipski y Taylor (2012), por ejemplo, utilizaron un modelo basado en agentes para estudiar el impacto de políticas de subsidios a los ingresos de población rural en Ghana y Malawi. Se estudiaron los impactos de pagos directos a los pobladores y de mecanismos alternos de subsidios en insumos, producción o en forma de políticas de bienestar.

En la Comunidad Europea, C. Sahrbacher et al (2014) desarrollaron el modelo AgriPolis para evaluar los cambios estructurales en la agricultura de una región que se pueden presentar como resultado de políticas agrícolas. Cada granja decide qué producir y en qué cantidad. También decide si vale la pena arrendar terrenos para producción adicional o si debe comprar o vender parte de su terreno. El modelo ha sido utilizado para evaluar posibles impactos de políticas en 22 regiones de 11 países de la Unión Europea.

Graupner (2011) desarrolló un modelo de competencia espacial para usar en agricultura, el “Spatial Agent-based Competition Model”. Y lo utilizó para evaluar el impacto de estructuras de mercado en la producción de leche en Alemania.

Aunque no está directamente asociada con la evaluación de decisiones de financiación, es notorio el trabajo de Campos et al (2013), en el cual utilizaron un modelo ABM para analizar la dinámica de la evolución de las redes de colaboración en el trabajo investigativo. En este trabajo se compararon diferentes estrategias de colaboración entre firmas para seleccionar una que optimiza los resultados de la colaboración.

### 3. METODOLOGÍA

El modelo desarrollado se puede utilizar para evaluar el impacto de cualquier programa de transferencias condicionadas en forma *ex ante*. Se usará como ilustración la evaluación de impacto del programa Bolsa Familia de Brasil.



### **3.1 Programa Bolsa Familia de Brasil**

El programa Bolsa Familia consiste en otorgar un subsidio mensual, en efectivo, por cada niño con edad entre seis y quince años, condicionado a que la familia garantice que el niño asista al colegio en forma regular.

El programa Bolsa Familia extendió su cobertura rápidamente y llegó a ser el número uno del mundo con cerca de 11,1 millones de familias (un cuarto de la población) en el 2007. Surgió en el año 2003 de la integración de programas anteriores como: Bolsa Escola (1997), Bolsa Alimentación (2001), Auxilio Gas (2002) y Tarjeta de Alimentación (2003). El programa busca reducir la pobreza y la desigualdad en el país, fomentando la asistencia escolar, la nutrición y la salud. Brinda ayuda financiera de entre R\$20 – R\$182 mensuales, a familias con un ingreso mensual per cápita de R\$137. El programa cuesta en total R\$12,3 mil millones que representan el 0,4% del PIB.

Es importante tener en cuenta que el gobierno de Brasil establece que la renta mensual per capita para familias en estado de pobreza es entre R\$ 77,01 y R\$145, mientras que las familias con ingresos mensuales per capita por debajo de R\$77 se consideran en estado de extrema pobreza (Ministerio de Desarrollo Social y la Lucha contra el Hambre).

#### *3.1.1 Población beneficiada*

Los beneficios son transferidos directamente por el gobierno federal a las familias. En su mayoría se pagan mensualmente a las mujeres de la familia, buscando aumentar su autonomía, capacidad de negociación y autoestima. Se han establecido beneficios variables y básicos que se suministran a familias en situación de pobreza y extrema pobreza, tomando en cuenta la situación de las madres de familia o la edad de los hijos pertenecientes a la misma. Los beneficios que se otorgan son de varios tipos entre los cuales el relevante para el propósito de nuestro estudio es el siguiente:

Beneficio variable que se concede a las familias con niños o adolescentes entre los cero y los quince años, se otorga un máximo de 3 estímulos de R\$35 mensuales por niño.

#### *3.1.2 Condiciones a cumplir*

Para que los estímulos financieros sean otorgados a las familias, éstas deben cumplir ciertas condiciones de ingresos a derechos sociales básicos como la educación, la salud y la nutrición.

1. Salud: Las familias deben cumplir con el compromiso de atender el carné de vacunación y el crecimiento y desarrollo de los niños menores a 7 años. Las mujeres entre los 14 y 44 años deben hacer el acompañamiento familiar y si son madres embarazadas o gestantes deben realizar los cursos pre-natales y el acompañamiento en la salud de su bebe.
2. Educación: Todos los niños entre los seis y los quince años deben estar matriculados y mantener una asistencia mensual mínima de 85%. Para los adolescentes de 16 a 17 años, la asistencia exigida es de mínimo el 75%.

En Brasil existe un acompañamiento del Ministerio de Desarrollo Social articulado con los ministerios de salud y de educación, y de los gestores municipales de Bolsa Familia a las familias que incumplan alguna de estas condiciones. Si la familia incide en el incumplimiento de las condiciones anteriormente mencionadas recibe un aviso sin sanción, al recibir el segundo aviso,



el beneficio se bloquea por 30 días y se desbloqueará al momento de verificar el cumplimiento nuevamente. El tercer aviso produce un bloqueo por 60 días y el último aviso genera la cancelación y remoción de la familia del programa Bolsa Familia.

#### 4. ESTIMACIONES

##### 4.1 Generación del ingreso por estado y estimación de salarios para niños entre los 10 y los 17 años

En el modelo es necesario tener en cuenta que los ingresos familiares pueden variar ampliamente de un estado a otro. Por lo tanto, es necesario considerar una base de ingresos que registre estas disparidades. Para este efecto se tomó como valor representativo del nivel de ingresos por estado la mediana de los ingresos de los habitantes de ese Estado. Una vez calculado el logaritmo de las medianas de los ingresos, se le asigna a cada habitante ese valor de referencia. Los cálculos se hicieron en Stata.

Tomando como eje de estudio la base de datos proporcionada por el Instituto Brasileiro de Geografía y Estadística (IBGE) en su encuesta anual titulada “Pesquisa Nacional por Amostra de Domicílios” (PNAD) del año 2011, se encuentran datos relacionados a los ingresos percibidos por las personas mayores a 10 años de los 27 estados que componen el país. Uno de estos datos es el ingreso recibido del empleo principal en reales, el cual tomamos para hacer análisis estadísticos. Sin embargo, como se ilustra en la Tabla 1, los datos registrados dentro de la encuesta para la población objetivo del estudio constituyen el 7.7% de la muestra, dejando por fuera una vasta proporción de los niños entre los 10 y los 17 años. Es necesario para nuestro estudio de decisiones ex ante captar la mayor parte de la población, por lo cual se elabora un modelo tipo Becker-Mincer de capital humano para predecir cómo se comportan los salarios de los niños en su totalidad teniendo como línea base los salarios de los niños que efectivamente devengan.

Niños entre los 10 y los 17 años		
Valores no registrados	45375	87.49%
Valores registrados	4016	7.74%
Valores registrados cero	2475	4.77%
<b>Total</b>	<b>51866</b>	<b>100%</b>

Tabla 1. Fuente PNAD 2011, cálculos de los autores.

Los modelos de Capital Humano de Becker o Mincer tienen como eje esencial las características educacionales del individuo, pues su hipótesis teórica es que al forjar un mayor espectro de Capital Humano dentro de una Nación, los incrementos en los ingresos nacionales y de producción vendrán concatenados. Con los datos que provee el PNAD del año 2011 y apoyados en Bourguignon (2003) establecimos la siguiente ecuación:

$$\ln y_i = \beta X_i + \alpha D_i + \mu_i$$

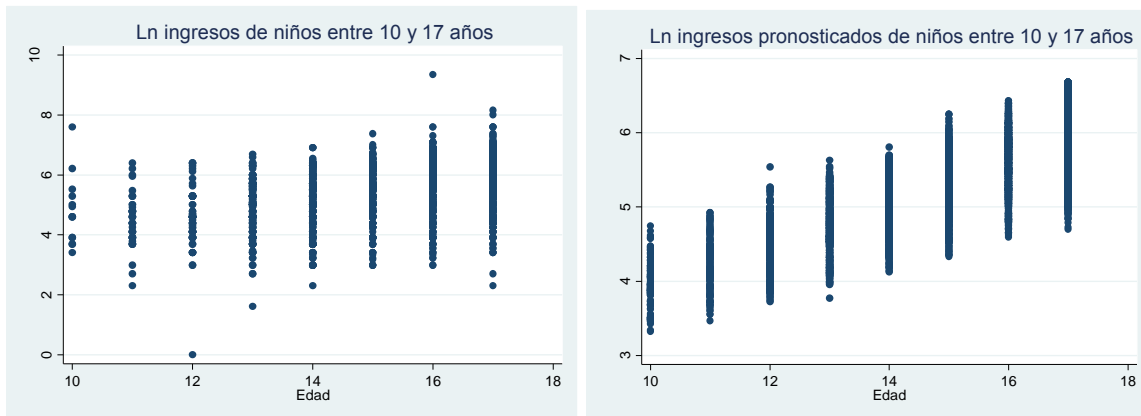
En donde se toma el logaritmo de los salarios observados  $y_i$ , expresado en función de un vector  $X_i$  de características propias del niño  $i$ , tales como la edad, la escolaridad, el sexo entre otras; una variable de decisión  $D_i$  que indica si aquel niño que devenga, y que por supuesto trabaja, va o no al colegio ( $D = 1$ , asistencia escolar). Para realizar la predicción de los salarios faltantes se

realizó una regresión lineal por Mínimos Cuadrados Ordinarios para estimar los parámetros: el vector  $\beta$  y el coeficiente  $\alpha$ . Es cierto que ni la robustez ni el comportamiento de los errores es el esperado, pues presenta algunos problemas de especificación. Cabe anotar que tanto el comportamiento como la estadística descriptiva de los ingresos estimados se asemejan al comportamiento de los ingresos reales.

	Observaciones	Media	Desv. Estándar	Mínimo	Máximo
<b>Ln ingresos pronosticados</b>	6196	5.419542	0.6392875	3.321245	6.681562
<b>Ln ingresos</b>	4016	5.662594	0.83506	0	9.344696

Tabla 2. Fuente PNAD 2011, cálculos de los autores.

En la tabla 2 se puede observar que las medias de ambos ingresos distan muy poco. Igualmente, las brechas de mínimos y máximos se ubican en zonas similares por lo que podemos decir que sus características estadísticas son similares. Las figuras 1a y 1b ilustran la distribución de los datos de los ingresos en función de las edades de los niños. Aunque leve, es evidente el comportamiento creciente de los ingresos observados, lo que estaría de acuerdo con la teoría económica: a mayor edad y mayor nivel educativo los salarios son mayores para la población económicamente activa. La distribución de los ingresos pronosticados también presenta un comportamiento creciente, y aunque su magnitud es claramente mayor que los observados, la tendencia es lo que centra nuestra atención.



Figuras 1a y 1b. Ingresos de los niños entre 10 y 17 años.  
Fuente PNAD 2001, cálculos de los autores.

La regresión por Mínimos Cuadrados Ordinarios se llevó a cabo en el paquete estadístico Stata: Data Analysis and Statistical Software, tomando únicamente los datos de los niños entre los 10 y los 17 años, que trabajaban en la semana referenciada en la encuesta de la PNAD del 2011 y que tuvieran dato referenciado de ingreso por trabajo principal. La regresión arrojó datos concordantes con lo esperado: Un R2 bajo, debido a la baja especificidad del modelo (variables explicativas faltantes) y con la distribución de los residuos un poco distante de la normalidad. Sin embargo, los coeficientes concuerdan con lo teorizado por Bourguignon y los p-valores muestran que las variables son estadísticamente significativas.



La tabla 3 muestra los resultados de la regresión lineal. El coeficiente  $\alpha$  de la asistencia a la escuela D tiene un valor negativo pues es el costo de oportunidad que se tiene por estudiar; es decir, puede verse como el tiempo que se “pierde” estudiando y no devengando dinero. Un año más de escolaridad refleja un aumento en el ingreso logarítmico en promedio de 0.031. Las variables referentes a las características físicas de los niños son positivas, lo que indica que un niño hombre, blanco y que viva en una zona urbana tiende a ganar más que los otros. La variable que indica el logaritmo de la mediana del ingreso estatal es positiva y tiene un coeficiente cercano a uno, lo que significa que una variación en el mismo tendrá como efecto un aumento porcentual de casi la misma magnitud de la variación.

La constante negativa de 4.1028 indica que siendo todas las variables iguales a cero, existe un salario mínimo devengado por cualquier niño igual a la exponencial del número en cuestión, el cual es cercano a cero. Esto cobra lógica desde el punto de vista en que es un ingreso mayor a cero, sin ser un salario mínimo; es un ingreso natural que existe en Brasil según el modelo.

	<b>Coefficiente</b>	<b>Error estándar</b>	<b>P &gt;  t </b>
<b>Observaciones</b>	4016		
<b>R<sup>2</sup></b>	0.3741		
<b>Asistencia a la escuela (D)</b>	-0.2905686	0.0258359	0.000
<b>Escolaridad (años)</b>	0.031071	0.0048985	0.000
<b>Edad</b>	0.1996107	0.0088428	0.000
<b>Masculino</b>	0.2212013	0.022181	0.000
<b>Blanco</b>	0.050121	0.0225968	0.027
<b>Urbano</b>	0.2120087	0.0306822	0.000
<b>Ln mediana ingresos estatales</b>	0.9460122	0.042918	0.000
<b>Constante</b>	-4.102.824	0.3015208	0.000

Tabla 3. Fuente PNAD 2011, cálculos de los autores

## 4.2 Modelo de Utilidad

Se supone que en este modelo las familias toman decisiones de una manera racional, buscando maximizar su utilidad. Las decisiones posibles, de acuerdo con el modelo desarrollado originalmente por Bourguignon et al (2013) para evaluar el programa Bolsa Escola, predecesor del programa actual están caracterizadas por una variable discreta  $X_i$  que describe la decisión adoptada por la familia del niño  $i$ . Los valores posibles de esta variable son:

$$X_i = \begin{cases} 0 & \text{Niño trabaja y no estudia} \\ 1 & \text{Niño trabaja y estudia} \\ 2 & \text{Niño solo estudia} \end{cases}$$



Se supone que cada familia tiene una noción de la utilidad derivada de sus condiciones económicas, la cual puede ser cuantificada por medio de una función de utilidad, dependiente de sus ingresos derivados del trabajo  $Y_i$  más los posibles ingresos recibidos en forma de subsidio, provenientes del programa de transferencias, .

$$\begin{aligned} U_i(j) &= S_j(A_i, X_i, H_i; Y_{-i} + y_{ij}) \\ &= S_j(Z_i; Y_{-i} + y_{ij}) \end{aligned}$$

La cual puede ser linealizada para obtener una expresión de la forma:

$$U_i(j) = Z_i\gamma_i + (Y_{-i} + y_{ij})\alpha_i$$

Donde  $Z_i$  es un vector combinado de variables que caracterizan al niño, su familia y su situación socio-económica, con excepción de los ingresos monetarios. Los coeficientes  $\gamma_i$  y  $\alpha_i$  deben ser estimados a partir de los datos.

#### 4.3 Modelo de regresión logística multinomial para estimar la probabilidades de las distintas decisiones por familia

A partir del modelo de utilidad, es posible estimar las probabilidades de que la persona que toma las decisiones en la familia escoja cada una de las tres opciones disponibles para la variable  $X_i$ .

La forma usual para hacer este tipo de modelo es mediante una regresión logística, en la cual se utiliza una función lineal para aproximar la expresión  $\log(p_i/p_k)$ , donde  $p_i$  es la probabilidad de que la decisión sea  $i$  y se mide con respecto a una decisión base  $k$ , usada como referencia. La expresión  $\log(p_i/p_k)$  se denomina “logit”.

La regresión logística, entonces, usa una expresión de la forma

$$\log\left(\frac{p_i}{p_k}\right) = Z \delta_i + (Y + y)\beta_i \quad i \neq k$$

para estimar los logits de las probabilidades. A partir de la ecuación anterior, se pueden calcular las probabilidades:

$$\begin{aligned} \frac{p_0}{p_2} &= \exp [Z \delta_0 + (Y + y)\beta_0] \\ \frac{p_1}{p_2} &= \exp [Z \delta_1 + (Y + y)\beta_1] \end{aligned}$$

donde se ha tomado como base para los cálculos la decisión 2 (es decir,  $k = 2$ ). Las ecuaciones anteriores junto con la relación

$$p_0 + p_1 + p_2 = 1$$

permiten calcular las probabilidades.

Con el fin de estimar las probabilidades de que una familia, con unas condiciones económicas y familiares determinadas tome una decisión dada, se realizó una regresión logística multinomial en el paquete econométrico Stata tomando 51.815 observaciones de niños y niñas entre los diez y los diecisiete años. Los resultados se presentan en la Tabla 4 en donde se puede observar que la variable estimada de ingresos anteriormente expuesta tiene un efecto marginal positivo cuando la decisión tomada fue trabajar únicamente y un efecto contrario cuando además de trabajar atiende al colegio, lo que tiene sentido pues si sólo trabaja es más probable que genere ingresos más altos a comparación de si distribuye su tiempo entre el estudio y el trabajo.

Un resultado curioso es el del sexo: un niño, a comparación de una niña, tiene menor posibilidad de estar trabajando únicamente y mayor de estar trabajando y estudiando, lo que abre el campo para el estudio de la distribución social del trabajo en Brasil. Siguiendo con las características de los niños, se muestra que es menos probable que un niño blanco o que viva en una zona urbana trabaje únicamente. Las variables correspondientes a la educación, tanto del niño como de los padres, se comportan a lo esperado. Es menos probable que un niño con mayores niveles de educación trabaje sin asistir al colegio y de igual manera con niños cuyos padres tienen niveles altos de educación. Como puede verse, el efecto marginal que tiene la variable que indica la máxima educación del pariente del niño es menor cuando trabaja únicamente (-0.166) que cuando trabaja y atiende al colegio en simultáneo (-0.095). La composición familiar es la cantidad de individuos que tiene una familia, lo que indica la regresión logarítmica multinomial es que hay una menor probabilidad que el niño trabaje únicamente cuando hace parte de una familia numerosa y el efecto cambia a positivo si distribuye su tiempo entre el estudio y el trabajo.

	<b>Trabaja únicamente (0)</b>		<b>Trabaja y atiende al colegio (1)</b>	
	Efecto Marginal	P valor	Efecto Marginal	P valor
<b>y_estimado</b>	9.138	0.000	-0.250	0.000
<b>Orden</b>	0.169	0.000	-0.003	0.882
<b>Masculino</b>	-2.066	0.000	0.832	0.000
<b>Edad</b>	-0.851	0.000	0.450	0.000
<b>comp_fam</b>	-0.118	0.000	0.076	0.000
<b>Blanco</b>	-0.984	0.000	0.046	0.177
<b>Urbano</b>	-2.787	0.000	-1.088	0.000
<b>Escolaridad</b>	-0.557	0.000	0.092	0.000
<b>max_edu_pariente</b>	-0.166	0.000	-0.095	0.000
<b>_cons</b>	-27.225	0.000	-6.996	0.000

Tabla 4. Fuente PNAD 2011. Cálculos de los autores.

## 5. MICROSIMULACIÓN

El proceso de microsimulación se realizó mediante el siguiente procedimiento:

1. Se generó una población sintética a partir de la base de datos original tomada de la encuesta de PNAD. Esta población tiene características similares a la población real del país.
2. Para cada individuo de la población sintética se evaluó la función de utilidad para cada una de las decisiones posibles, a partir de sus características individuales y familiares.
3. Se calcularon las probabilidades de que la familia tomara una de las decisiones posibles: niño trabaja y no estudia, niño trabaja y estudia, niño solo estudia. Se utilizó el modelo

logístico multinomial para calcular estas probabilidades y se evaluó la función de distribución acumulada.

4. Se realizaron las simulaciones mediante la generación de números aleatorios que satisfacen la distribución de probabilidad encontrada en el paso 3. Se llevaron a cabo 50 replicaciones de la simulación, se tabularon y se analizaron los resultados.

## 6. RESULTADOS

Un primer resultado contundente es que el 97% de los niños entre 10 y 15 años van a estudiar, lo cual implica que el programa podría considerarse exitoso dado que uno de sus objetivos primordiales es lograr la asistencia de los niños al colegio. Este resultado es aun mas promisorio si se compara con resultados anteriores de Bourguignon (2003), en el cual para la época de 1999 el total de niños que asistían al colegio era de 93.9%. La cifra de niños asisten a la escuela disminuye hasta 92.77% cuando se considera el grupo total de niños hasta la edad de 17 años.

Las simulaciones muestran una marcada tendencia a disminuir el porcentaje de estudiantes que asisten al colegio a medida que aumenta la edad del joven. Mientras que para edades inferiores a 13 años el porcentaje de niños que solamente asisten al colegio es superior a 90%, este numero se reduce marcadamente a partir de los 15 años hasta llegar a 55% a la edad de 17 años, tal como se muestra en el Figura No. 2.

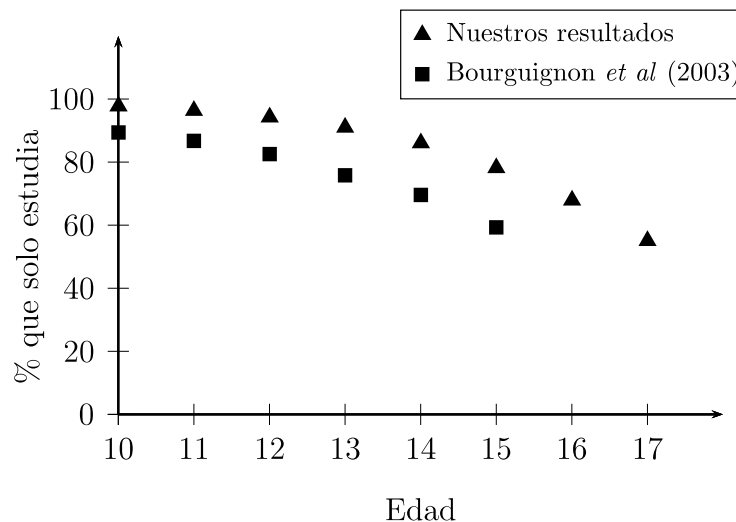


Figura 2. Fracción de los jóvenes que asisten al colegio y no trabajan, por edades.  
Fuente: (Bourguignon *et al* 2003) y cálculos de los autores.

En la figura 2 se presentan los resultados del porcentaje de jóvenes que no trabajan y solamente asisten a la escuela, por edades. Para hacer una comparación, se presentan en la misma también los porcentajes calculados para el año 2009 por Borguignon (Bourguignon *et al* 2003). En dicha figura, los triángulos representan los valores estimados usando nuestro modelo para el período 2011 mientras que los cuadrados son los resultados de las microsimulaciones efectuadas por Bourguignon. Nuestros resultados incluyen las edades de 16 a 17 años, las cuales no estaban



incluidas en el programa Bolsa Escola original y fueron incorporadas al nuevo programa Bolsa Familia.

Aunque no es posible establecer en este momento si las condiciones bajo las cuales se desarrolló nuestro modelo son compatibles con las del modelo de Bourguignon, una comparación simple de los porcentajes de jóvenes que han tomado la decisión de asistir a la escuela en vez de trabajar muestra que las cifras han crecido significativamente para todas las edades. De hecho, el porcentaje actual de jóvenes de 15 años que asisten a la escuela sin trabajar es superior al porcentaje de niños de 13 años del período anterior, que se encontraban en la misma situación.

Estos resultados parecen indicar que, al menos desde el punto de vista de desestimular la deserción escolar, los programas Bolsa Escola y Bolsa Familia han tenido un impacto notable y deberían considerarse un éxito y un ejemplo a seguir. Aun así, sigue siendo preocupante que casi un cuarto de los jóvenes de 17 años (24.83%) está dedicado solo a trabajar.

La tabla 5 muestra los resultados globales por edades para la población estudiada. Es clara la asociación entre una mayor edad y la probabilidad de desertar de la escuela. Para tratar de cuantificar esta tendencia, se calculó la estadística  $\gamma$ , la cual trata de dar valores numéricos a esta relación cualitativa<sup>2</sup>. Para el caso de asociación entre la edad y la importancia relativa de la educación con respecto a los ingresos, el valor estimado es  $\gamma = -0.6053$ . Esto demuestra una asociación fuerte (negativa) entre las dos variables, aunque la relación es claramente no lineal, como lo muestra la figura 2.

Selección	Edad								Total
	10	11	12	13	14	15	16	17	
Sólo trabaja	0.33	0.57	1.26	2.32	4.39	8.81	14.69	24.83	7.23
Trabaja y estudia	1.94	3.06	4.5	6.67	9.55	13.01	17.37	20.14	9.64
Sólo estudia	97.73	96.36	94.24	91.01	86.06	78.18	67.94	55.03	83.14

Tabla 5. Condición de los niños con respecto a la escuela, por edades.  
Fuente: PNAD 2011, cálculos de los autores

La relación entre la edad y la probabilidad de asistir a clases o trabajar se puede ver más claramente en forma gráfica. La figura 3 presenta esta relación. Se ve claramente la asociación positiva fuerte entre la edad y la probabilidad de deserción escolar.

<sup>2</sup> El estadístico  $\gamma$  es una medida de asociación entre variables con datos ordinales y mide el número de concordancias  $C$  y discordancias  $D$  que aparecen al comparar las puntuaciones asignadas a las dos variables. Se define como:  $\gamma = (C - D)/(C + D)$ . Ver Agresti (2007)

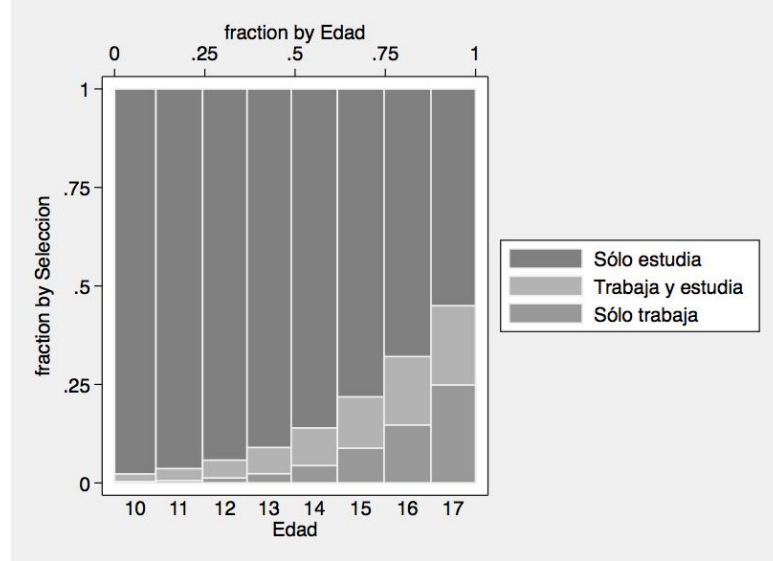


Figura 3. Fracción de niños en condición dada, respecto a la asistencia a la escuela.  
 Fuente: Cálculos de los autores con base en información PNAD 2011.

Adicionalmente, estamos interesados en mirar la influencia de otras variables sobre las decisiones que toma una familia acerca de la educación de sus hijos. En primer lugar, quisimos estudiar si había diferencias entre las decisiones que tomaban las familias residentes en zonas urbanas y las que tomaban familias que habitan en áreas rurales. Se encontró que efectivamente existen diferencias y se halló un valor de  $\gamma = 0.3652$ . La relación entre estas variables se puede apreciar mejor en la figura 4.

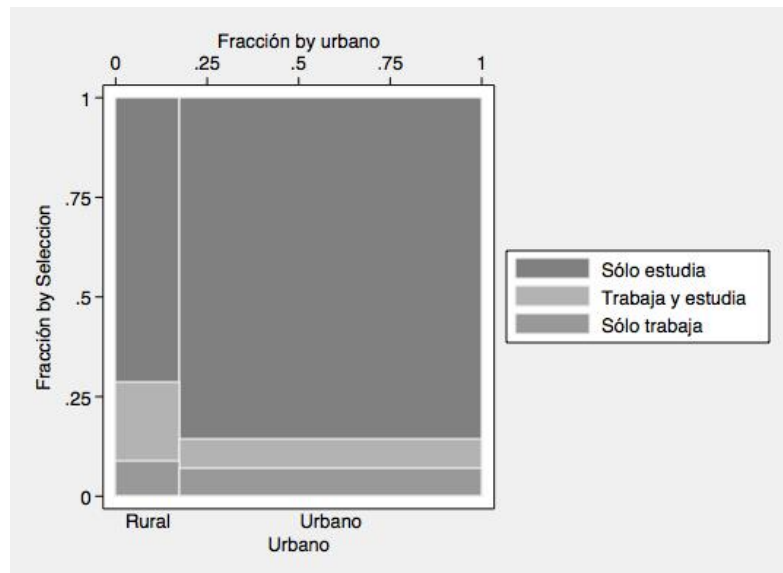


Figura 4. Decisiones sobre escuela, por área urbana o rural.  
 Fuente: Cálculos de los autores con base en información PNAD 2011.

Aunque no hay una asociación tan fuerte como con la edad, aun se puede considerar que existe una influencia considerable del ámbito donde habita la familia sobre la asistencia a la escuela. Es notable, por ejemplo, que para la población rural, el porcentaje de jóvenes que sólo estudian es de 71.5%; en zonas urbanas, ese valor es de 85.64%. Sería interesante tratar de evaluar si las causas

son estructurales tales como mayor distancia a la escuela o mayor contribución relativa a las finanzas familiares o si se trata de un asunto de percepción o de educación de los padres.

Se puede notar que la fracción de niños que solo trabajan no es muy diferente entre los dos grupos. Pero la diferencia fundamental se da entre los que estudian y trabajan simultáneamente. En este caso, la fracción para áreas rurales es 2.7 veces la fracción correspondiente a las áreas urbanas. Un análisis por edades muestra que la diferencia se hace mucho mayor a medida que la edad aumenta.

Es interesante estudiar el efecto potencial que tiene el nivel educativo de los padres sobre la disponibilidad a permitir que el niño trabaje. Sin embargo, la base de datos no permite determinar cuál es el nivel educativo del jefe de hogar. En su lugar, usamos el dato de la persona del grupo familiar con más alta escolaridad. Mediante el análisis se puede notar que tener unos padres más educados si implica una menor probabilidad de permitir el trabajo de su hijo. Esta diferencia no es tan acentuada entre los que han cursado solo primaria y los que cursaron algunos años de la secundaria. Donde se observa una marcada diferencia de prioridades es entre los que terminaron la secundaria comparados con los de menor escolaridad. Las fracciones en cada situación se muestran en la figura 5.

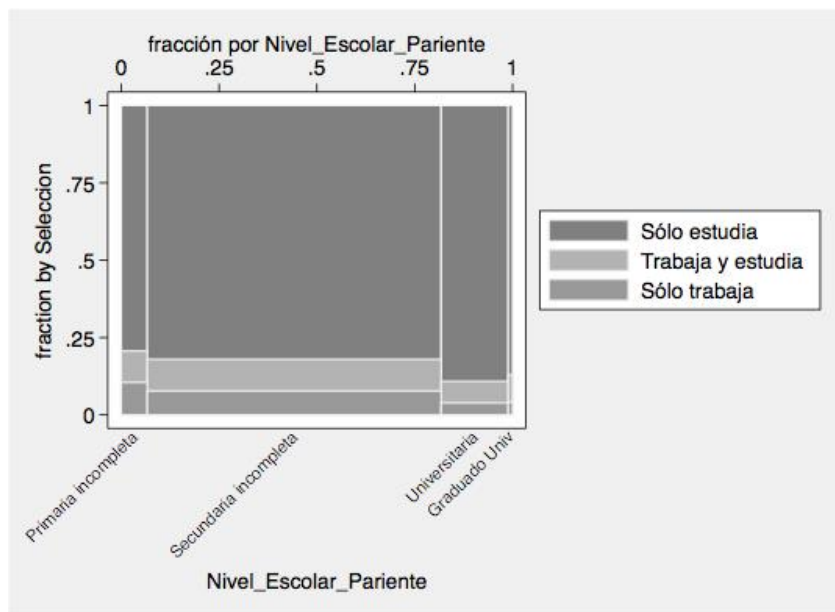


Figura 5. Decisiones sobre asistencia a la escuela, por nivel educativo de los padres.  
Fuente: Cálculos de los autores con base en información PNAD 2011.

A continuación se hizo una comparación de los resultados por género del niño. Este análisis muestra que la proporción de niños y niñas que no asisten a la escuela sino que solamente trabajan es aproximadamente igual. Y esto es válido para todos los grupos de edad. Sin embargo, la fracción de niños que estudian y trabajan al mismo tiempo si es notablemente superior al de las niñas en la misma situación. Un análisis más detallado, por edad, muestra que la proporción es prácticamente constante (la fracción de niños que trabajan y estudian es cercana al doble de la correspondiente a las niñas), para todas las edades.



Figura 6. Decisiones sobre asistencia a la escuela, por género del niño.  
 Fuente: Cálculos de los autores con base en información PNAD 2011.

Por último, estudiamos el efecto de la edad del pariente mayor sobre las decisiones de las familias. En este caso se encontró una asociación bastante débil entre la edad del pariente mayor y el estado de los jóvenes con respecto a la escuela a nivel global. El valor de  $\gamma$  calculado es de  $\gamma = -0.0605$ . No obstante, para el grupo de adultos responsables más jóvenes (edades menores que 30 años) si se puede observar una mayor disposición a permitir que los jóvenes trabajen en vez de estudiar, en comparación con los otros grupos. Esto afecta poco a la población en general, porque este grupo es bastante pequeño, representando solamente el 4% de las familias. Sin embargo, cuando se excluye este grupo, el valor de  $\gamma$  calculado aumenta a  $-0.1064$ , que aunque sigue siendo débil, es bastante mayor que el anterior.

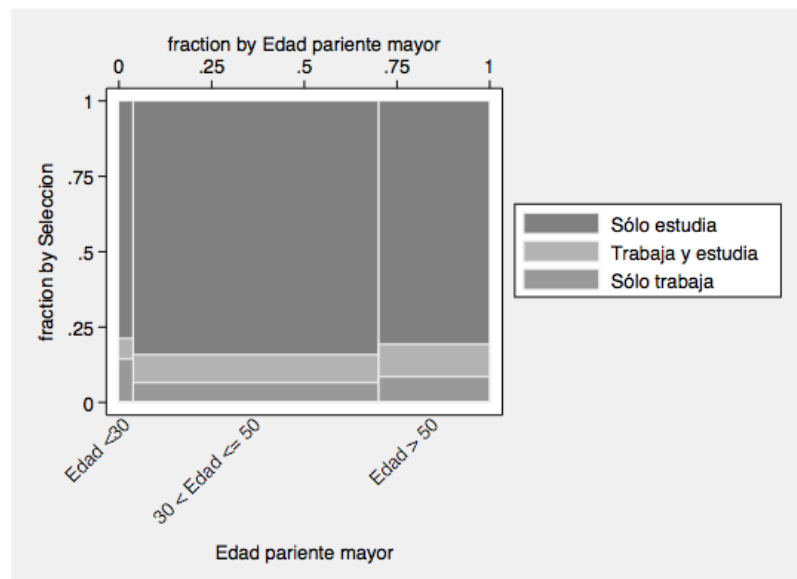


Figura 7. Decisiones sobre asistencia a la escuela, por edad de los padres.  
 Fuente: Cálculos de los autores con base en información PNAD 2011.





## 7. CONCLUSIÓN

En este artículo se describió un método basado en microsimulación para evaluar el impacto de las inversiones de programas sociales del Estado. Se demostró que es posible realizar la evaluación de impacto utilizando un modelo de microsimulación calibrado mediante los datos reales del programa. La existencia de un modelo de este tipo permite hacer evaluaciones *ex ante* con el fin de predecir los resultados e impactos de un programa de inversión. De esta manera, se posibilita la consideración de diferentes alternativas de programa con lo cual los tomadores de decisiones tienen una herramienta para el diseño óptimo de políticas públicas.

El modelo de microsimulación se aplicó a la evaluación del impacto *ex ante* del Programa Bolsa Familia de Brasil. Se utilizaron datos de la encuesta PNAD correspondientes al año 2011. Esta base de datos, aunque de gran amplitud y gran ayuda en la elaboración del presente trabajo, se quedaba corta en estimaciones puntuales como la de los ingresos de los niños para generar los efectos marginales del modelo logarítmico multinomial.

A pesar de que la estimación de los ingresos de los niños se simplificó, la presencia de errores es innegable y la especificación del modelo puede ser mejorada. Es posible incluir variables como las horas trabajadas o generar regresiones de datos panel con datos de años anteriores al año tratado, o generar regresiones más complejas que capten los niveles estatales.

Los supuestos que incluyen modelos económicos tienen las debilidades del comportamiento racional de los agentes. La toma de decisiones en este caso se supone racional pues el individuo debe promover la maximización de su utilidad y la de su núcleo familiar; sin embargo, en ocasiones esto es contrario a la abstracción de la realidad que se hace y se toman decisiones que van en contravía de lo planeado. Este tipo de consideraciones se pueden incluir en un modelo basado en agentes que actualmente está en desarrollo.

Los resultados de esta evaluación muestran que el Programa Bolsa Familia de Brasil ha sido exitoso desde el punto de vista de fomentar la asistencia de los niños al colegio. Por lo menos esto es lo que se puede concluir comparando los resultados de este trabajo con los de Bourguignon (Bourguignon *et al* 2003).

Este tipo de modelo se puede aplicar para la evaluación de impacto de diferentes programas de inversión social siempre y cuando la información requerida esté disponible. En este momento está en desarrollo su adaptación para realizar la evaluación de impacto *ex ante* de la inversión en programas de investigación.



## REFERENCIAS

- Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. Wiley.
- Bourguignon, F. and Ferreira, F. H. G., (2003). Ex Ante Evaluation of Policy Reforms Using Behavioral Models. In F. Bourguignon and L. A. P. da Silva, ed . 2003. *The impact of economic policies on poverty and income distribution: evaluation techniques and tools*. New York: World Bank and Oxford University Press, Ch. 6.
- Campos, Pedro, Brazdil, Pavel and Mota, Isabel, (2013), “Comparing Strategies of Collaborative Networks for R&D: An Agent-Based Study.” *Computational Economics*. 42, issue 1, p. 1-22, <http://EconPapers.repec.org/RePEc:kap:compec:v:42:y:2013:i:1:p:1-22>. [05.06.2015]
- Canova Luciano & Piccoli Luca & Spadaro Amedeo, (2014). “An ex ante evaluation of the Revenu de Solidarité Active by micro-macro simulation techniques,” DEA Working Papers 67, Universitat de les Illes Balears, Departament d'Economía Aplicada.
- Filipski Mateusz & Taylor J. Edward (2012). A simulation impact evaluation of rural income transfers in Malawi and Ghana. *Journal of Development Effectiveness*. Volume 4, Issue 1
- Graupner, Marten (2011). “The Spatial Agent-based Competition Model (SpAbCoM).” Discussion paper. Leibniz Institute of Agricultural Development in Central and Eastern Europe, No. 135.
- Harding, A. (2007). “Challenges and Opportunities of Dynamic Microsimulation modelling.” 1st General Conference of the International Microsimulation Association. Online Conference Paper - CP 107. Vienna.
- Holbrook. J. B. and Frodeman, F. (2011). “Peer review and the ex ante assessment of societal impacts”. *Research Evaluation*, 20(3), September 2011, pages 239–246 DOI: 10.3152/095820211X12941371876788.
- Leite, P., A. Narayan, and E. Skoufias, (2011). How do Ex ante Simulations Compare with Ex Post Evaluations? Evidence from the Impact of Conditional Cash Transfer Programs, <http://elibrary.worldbank.org/doi/pdf/10.1596/1813-9450-5705> , URL <http://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-5705> [2.04.2014]
- Li Jinjing. (2011), *Dynamic Microsimulation for Public Policy Analysis*. Dissertation published by Boekenplan, Maastricht, the Netherlands.
- Li Jinjing and O'Donoghue Cathal. (2012), *A methodological survey of dynamic microsimulation models*. Maastricht Economic and social Research institute on Innovation and Technology United Nations University. UNU-MERIT Working Papers
- Li Jinjing and O'Donoghue Cathal, Sahrbacher C, Sahrbacher A, Kellermann K, Happe K, Balmann A, Brady M, Schnicke H, Ostermeyer A, Schönau F. (2012), *ODD-protocol of AgriPoliS*. Technical report. IAMO. Halle, Germany. [http://www.agripolis.de/documentation/ODD\\_AgriPoliS.pdf](http://www.agripolis.de/documentation/ODD_AgriPoliS.pdf). [03.04.2014]
- Sahrbacher A., Sahrbacher C., Ostermeyer A. (2014), *How financial constraints distort farms' investment behaviour and regional development: a comparative analysis of four European regions*. Leibniz-Institute of Agricultural Development in Transition Economies (IAMO) Department “Structural Development of Farms and Rural Areas”, Halle/Saale, Germany.
- Zucchelli, E, Jones, AM & Rice, N (2012), “The evaluation of health policies through dynamic microsimulation methods.” *International Journal of Microsimulation*, vol 5, no. 1, pp. 2-20.



## APÉNDICE:

### Descripción de variables utilizadas en el modelo

**Orden:** indica el orden de nacimiento en la familia, siendo “1” el valor del individuo con más edad.

**Masculino:** Toma el valor “0” si el género del individuo es femenino ó “1” si el género es masculino.

**Edad:** Indica la edad en años del individuo.

**Comp\_fam:** toma el valor del número de integrantes de la familia.

**Blanco:** Esta variable toma dos valores “0” en caso de que la raza del individuo sea diferente a la blanca y “1” si la raza es blanca. Se creó sintéticamente a partir de los datos de la encuesta, allí aparecen las razas: blanco, negro, amarillo, marrón e indio, que se resumieron en dos tipos, únicamente para la regresión.

**Urbano:** Toma el valor “0” en caso de que el área censada sea urbana y “1” en caso de que sea rural. Se creó sintéticamente a partir de los datos de la encuesta, allí aparecen opciones como: Urbana – aislado, Área urbana, Rural - aislados núcleo rural de partículas, Urbana, Rural - partículas de extensión urbana rural, entre otras, que fueron resumidas en urbano o rural para su fácil manejo.

**Escuela\_dummy:** Puede tomar dos valores posibles “0” o “1”, “0” en caso de que el niño solo trabaje y “1” en caso de que el niño trabaje y estudie.

**Escolaridad:** Hace referencia a los años estudiados, tomando el valor “1” cuando es menos de 1 año, “2” cuando es un año, así sucesivamente hasta llegar a “16” que son 15 años o más de estudio.

**Edad\_esc\_2:** Toma la edad del individuo menos su escolaridad elevando el resultado al cuadrado.

**Trabaja\_todos:** Esta variable hace referencia al trabajo de los individuos, tomará el valor “1” en caso de que trabaje, y, el valor “0” en el caso de que no trabaje

**Ln\_mediana\_estatal:** Se tomó la mediana por estado (27 estados en total) y se le aplicó logaritmo natural a cada uno de los 27 resultados, el resultado obtenido por estado se le colocó a cada uno de los individuos habitantes del estado correspondiente.

**Max\_edu\_pariente:** toma el valor de los años de educación del pariente con más estudio en la familia.

**Edad\_pariente\_mayor:** Toma el valor de la edad del pariente con más años en la familia. Se creó sintéticamente a partir de los datos de la encuesta.

**Below5:** toma el valor de la cantidad de niños menores a 5 años que se encuentran en la familia.

**Dummy\_principal:** Puede tomar tres valores posibles: “0”, “1” o “2”, “0” en caso de que únicamente trabaje, “1” trabaja y va a la escuela, o “2” únicamente va a la escuela.

**Ingresos\_pred:** Después de realizar la regresión del logaritmo natural de los salarios del trabajo principal (obtenidos de la encuesta) tomando como variables independientes: ln\_mediana\_estatal, escuela\_dummy, urbano, masculino, escolaridad, edad. Se tomaron los valores de los coeficientes para realizar la predicción de los salarios para los niños.

**Ingresos\_pred\_mon:** es el valor monetario de los ingresos predichos, puesto que se encontraba expresado en términos del logaritmo natural se obtuvo mediante la aplicación de la función exponencial.

**Ingresos\_pred\_estatales:** Posterior a realizar la regresión de los ingresos de los individuos se observó que existe una correlación entre dichos ingresos y el logaritmo de la mediana estatal,



debido a lo anterior se realizaron regresiones por estado (en total 27 regresiones), quedando como una constante el logaritmo de la mediana estatal en cada regresión. Con los valores de los coeficientes de las variables independientes se realizó la predicción estatal de los ingresos.

**Ingresos\_pred\_estatales\_mon:** es el valor monetario de los ingresos estatales predichos, puesto que se encontraba expresado en términos del logaritmo natural se obtuvo mediante la aplicación de la función exponencial.