
V. Otros Temas Económicos

Una Nota sobre la Concentración Bancaria en la República Dominicana

Por: *Fadua Camacho y Oscar Iván Pascual¹*

Año 2013, Vol. VII, No. 1

I. Introducción

En años recientes, el sistema financiero mundial ha vivido una marcada tendencia hacia la concentración a través de la desaparición de instituciones, adquisiciones, expansiones y absorciones por parte de grandes bancos multinacionales (Beck, 2008). A su vez, la importancia del papel que juegan las instituciones bancarias en el sistema económico de cualquier nación ha ido incrementándose, particularmente para países en desarrollo (Krueger, 2006).

En los últimos diez años, el volumen de activos del sistema financiero de la República Dominicana se ha incrementado sin estar acompañado de un incremento significativo del número de instituciones participantes (pasando de un promedio de total de activos de RD\$23,529.74 millones por cada entidad bancaria en 2003 a un promedio de RD\$55,420.36 millones en 2012). Un número importante de investigaciones recientes han demostrado que las condiciones del mercado financiero son importantes en los canales de transmisión de la política monetaria Kamin, Turner y Van Oudiz (1998), por tanto, es de gran interés evaluar la evolución de la estructura del sistema financiero en la República Dominicana, particularmente la concentración bancaria.

El presente análisis pretende medir la concentración bancaria a partir de dos metodologías: los índices de medida de concentración bancaria que utilizan la suma de las principales variables de las instituciones de mayor volumen y el índice de Herfindahl-Hirschman que se elabora a partir de las cuotas mercado o tamaño relativo de cada una de las instituciones. Estos indicadores de concentración bancaria se construirán para los activos, préstamos y depósitos de los bancos múltiples del sistema financiero dominicano.

II. Revisión de literatura

Existen teorías divergentes en cuanto a las consecuencias de la concentración bancaria; Por un lado, el enfoque Estructura-Conducta-Desempeño (ECD) sostiene que una estructura bancaria de escasas y grandes instituciones impulsa un comportamiento oligopólico por parte de los oferentes del mercado. Por otro lado, la Nueva Teoría Empírica de la Organización Industrial

¹ Subdirección de Estadísticas Monetarias y Financieras. Los autores agradecen la colaboración de Pilar Mateo Mejía. Para preguntas y comentarios escribir a f.camacho@bancentral.gov.do y o.pascual@bancentral.gov.do.

(TEOI), sostiene que la estructura no determina la competitividad del mercado y que lo relevante para la eficiencia es si existen o no barreras de entrada y/o salida (Asobancaria, 2008).

En el Enfoque Estructura-Conducta-Desempeño, el entorno oligopólico y poder de mercado de las instituciones financieras existentes les permitiría mayores tasas de interés activas, lo que afectaría el crédito e inversión y, a su vez, el desarrollo económico. Estas mayores tasas de interés implicarían un encarecimiento del crédito (Cerqueiro, 2008) y la discriminación en contra de las inversiones menos rentables que podría conducir a una concentración industrial.

La Nueva Teoría Empírica de la Organización Industrial, por el contrario, sostiene que los mercados concentrados pueden ser competitivos si las barreras de entrada y salida son mínimas y por consiguiente las instituciones están presionadas a competir. Este modelo incorpora asimetría de información, y argumenta que una alta concentración bancaria y la capacidad de cargar a la cartera de clientes con tasas de interés más elevadas podrían compensar el costo de otorgar préstamos de menor calidad, otorgándose estos últimos a menores tasas de interés que bajo un sistema de baja concentración (Cifuentes, 2003).

Esto implicaría un aumento del crédito disponible y una mayor bancarización de la economía. Ciertos autores argumentan que esto posiblemente conduciría a mayor inestabilidad del sector financiero, sin embargo, Davis (2007) destaca que no existe evidencia sobre esto. Por tanto, el efecto neto sobre la disponibilidad de crédito de un sistema financiero altamente concentrado no es claro al evaluar la literatura sobre el tema.

Otras consecuencias de la concentración bancaria corresponden a la probabilidad de fallos y crisis en una institución y el riesgo de la estabilidad del sistema financiero. En estos aspectos, el debate está impulsado en primera instancia por la percepción de que existe un trade-off entre eficiencia y estabilidad en el sistema bancario. La estabilidad se explica por un reducido número de instituciones con mayor diversidad de instrumentos financieros y mejores sistemas de manejo de riesgo; además, una menor cantidad de instituciones implican mayor facilidad de monitoreo por parte de las autoridades (Davis, 2007).

El argumento contrario señala que pocas y grandes instituciones financieras podrían ser muy grandes como para dejarlas fracasar (“too big to fail”), como ha sucedido en diversos países con la crisis financiera internacional. Esto podría suponer la toma de riesgos excesiva por parte de los bancos y altos costos en el caso de una crisis en uno de ellos.

III. Midiendo la concentración bancaria en República Dominicana

Los indicadores frecuentemente utilizados para medir la concentración bancaria son los índices C3, C4, C5 y el Herfindahl-Hirschman (HHI). Los índices C3, C4 y C5 son calculados como la suma de las principales variables de las 3, 4 y 5 instituciones bancarias de mayor volumen como proporción del total. El HHI se calcula a partir de la siguiente expresión:

$$(1) \quad HHI = \sum_i s_i^2$$

, donde s_i es la cuota de mercado de cada institución, elevándolas al cuadrado para otorgar mayor ponderación a las instituciones más grandes.

Este índice toma valores entre 0 y 10,000; el caso de competencia perfecta se da cuando existen un número elevado de instituciones y cada una de estas posee una proporción mínima que tiene a cero (02), mientras que el caso de monopolio se da cuando existe sólo una institución y por tanto posee el cien por ciento del total (1002).

Según la División Antimonopolio del Ministerio de Justicia de los Estados Unidos que se encarga de promover la competencia económica a través del establecimiento y la aplicación de leyes y principios antimonopolio, un resultado de menos de 1,000 es generalmente considerado como un mercado no concentrado, entre 1,000 y 1,800 es moderadamente concentrado; y mayor de 1,800 se considera altamente concentrado.

Las Tablas 1, 2 y 3 muestran los resultados de los índices C3, C4, C5 y HHI para el total de activos, total de depósitos y total de préstamos de los bancos múltiples que operaron en el sistema financiero dominicano, respectivamente, para el período 2003 - 2012. La columna 2 resume el total y las columnas 3-6 contienen las medidas de concentración para los años indicados. Por último, la columna 7 presenta el número de instituciones registradas por año.

Tabla 1. Concentración de activos

Año	Totales (Millones de RD\$)	C3	C4	C5	HHI	N
		(Porcentajes)				
2003	329416	61.2	71.7	79.8	1530	14
2004	343461	61.6	71.8	79.5	1606	14
2005	335898	71.0	78.6	83.8	2039	13
2006	384708	72.7	79.7	86.5	2112	12
2007	454444	73.7	80.9	87.4	2158	13
2008	505332	76.1	83.4	89.7	2254	12
2009	569347	75.7	82.4	88.8	2195	12
2010	660815	74.9	81.2	87.4	2127	14
2011	761374	74.8	81.2	87.0	2152	15
2012	831305	73.5	79.8	85.7	2447	15

Fuente: Cálculos de los autores.

A pesar de que entre 2003 y 2012 el número de instituciones se ha mantenido estable, pasando de catorce a quince instituciones, el crecimiento promedio del total de activos fue de 11.0% anual. Las medidas de concentración calculadas muestran un incremento de la concentración durante este mismo período. Las tres instituciones con el mayor volumen de activos a 2003

poseían 61.2% de los activos del sistema, mientras que a 2012 concentraban 73.5% de los mismos.

Tabla 2. Concentración de depósitos

Año	Totales (Millones de RD\$)	C3	C4	C5	HHI	N
		(Porcentajes)				
2003	150102	63.1	74.3	80.5	1649	14
2004	162762	64.7	75.1	85.3	1772	14
2005	132719	66.7	75.1	83.1	2239	13
2006	239000	67.1	76.4	85.3	1953	12
2007	294056	69.2	78.6	86.8	2049	13
2008	295602	69.5	79.8	88.7	2099	12
2009	337822	69.6	78.6	87.4	2053	12
2010	410181	70.2	78.3	86.3	2067	14
2011	465927	70.0	77.8	85.6	2078	15
2012	511826	68.3	76.3	83.9	1955	15

Fuente: Cálculos de los autores.

Tabla 3. Concentración de préstamos

Año	Totales (Millones de RD\$)	C3	C4	C5	HHI	N
		(Porcentajes)				
2003	171841	66.0	77.2	84.8	1725	14
2004	160887	67.7	77.5	83.8	1879	14
2005	179037	66.6	75.5	82.3	1975	13
2006	196292	74.4	82.8	89.8	2335	12
2007	246836	75.2	83.8	90.8	2320	13
2008	286522	77.3	85.2	92.1	2414	12
2009	321911	77.1	84.4	91.5	2392	12
2010	368640	75.2	82.6	89.2	2233	14
2011	410789	74.7	82.5	88.1	2163	15
2012	466862	76.6	83.8	89.1	2238	15

Fuente: Cálculos de los autores.

Estas mismas tres instituciones a 2003 contaban con 63.1% de los depósitos y a 2012 con 68.3%. Cerca del 85% de los activos y depósitos en los bancos múltiples dominicanos están concentrados en cinco instituciones (un tercio del total).

Sin embargo, la mayor concentración a partir de las medidas C3, C4 y C5 se observa en los préstamos; a 2012 casi el 90% de éstos estaban concentrados en cinco instituciones.

El cálculo del HHI para el período indica un incremento de la concentración de activos de un HHI de 1,530 (apenas moderadamente concentrado) a 2,447 (dentro del rango considerado de alta concentración). Por otro lado, los depósitos muestran un menor crecimiento de la concentración, con resultados que lo sitúan en un mercado moderadamente concentrado.

El HHI calculado para el total de préstamos en 2011 ubica al sistema bancario dominicano entre los de mayor concentración entre los países de Centroamérica, República Dominicana y Panamá, sólo superado por Nicaragua (2,458), y lejos del promedio de la región (1,621).

Tabla 4. HHI Cartera de créditos centroamérica y la República Dominicana, Diciembre 2011

País	IHH (Puntos)
Honduras	1127
Guatemala	1701
República Dominicana	2138
Nicaragua	2458
El Salvador	1591
Costa Rica	1520
Panamá	811
Promedio	1621

Fuente: SECMCA.

IV. Conclusión

Los resultados señalan una alta concentración de activos y depósitos en la banca múltiple de la República Dominicana. La inquietud sobre la tendencia a la concentración bancaria viene dada principalmente por la preocupación de los efectos que puede tener una estructura oligopólica del sistema financiero sobre el crédito y la economía en general. Sin embargo, dado que la relación entre concentración y competencia es poco clara, no es posible concluir que el sistema financiero dominicano es poco competitivo a partir de los resultados anteriores (Davis, 2007; Fandiño, 2007).

No obstante la alta concentración, medir la eficiencia económica y el riesgo de inestabilidad requeriría de un análisis de las barreras de entrada y salida de las instituciones, basándose en hallazgos de que la estabilidad bancaria es mayor en los países donde existen menores regulaciones que limitan la entrada al mercado bancario y por consiguiente existe mayor competencia.

Referencias

Asobancaria (2008). "Algunas reflexiones sobre la competencia bancaria en Colombia". Semana Económica.

- Beck, T. (2008). "Bank Competition and Financial Stability: Friends or Foes?". World Bank.
- Bernal, R. (2007). "Concentración y Competencia en el Sistema Financiero Crediticio Colombiano en la Última Década". Borradores de Economía, Banco de la República de Colombia.
- Bikker, J. y Haaf, K. (2002). "Measures of Competition and Concentration in the Banking Industry: a Review of the Literature". Central Bank of Neatherlands.
- Cerqueiro, G. (2008). "Bank Concentration, Credit Quality and Loan Rates". CentER – Tilburg University.
- Cifuentes, R. (2003). "Banking concentration: implications for systemic risk and safety net design". Central Bank of Chile Working Papers.
- Davis, K. (2007). "Banking Concentration, Financial Stability and Public Policy". The University of Melbourne.
- Kamin, S., Turner, P., y Van Öt dack, J. (1998). "The transmission mechanism of monetary policy in emerging market economies: an overview".
- Krueger, A. (2006). "Financial Markets and Economic Growth". Speeches, International Monetary Fund.
- Maudos, J. y Pérez, F. (2003). "Competencia versus poder de mercado en la banca española". Ivie y Universitat de València.
- Ministerio de Justicia de los Estados Unidos (2012). "Horizontal Merger Guidelines". www.justice.gov
- Pérez, L. (2009). "Competencia y Estabilidad en el Sistema Bancario: Evidencia para la República Dominicana". Instituto de Economía, Pontificia Universidad Católica de Chile.
- Peria, M.S. (2010). "Measuring Bank Competition: How Should We Do It? All About Finance".
- Rojas-Suárez, L. (2006). "El Acceso a los Servicios Bancarios en América Latina: Identificación de Obstáculos y Recomendaciones". Center for Global Development.
- Secretaría Ejecutiva Consejo Monetario Centroamericano, SECMCA (2012). "Informe del Sistema Bancario de Centroamérica, República Dominicana y Panamá 2011". Consejo Monetario Centroamericano.

Propiedades de Estimadores Con Datos de Panel Dinámicos: Un Experimento Monte Carlo para Distintos Tamaños de Muestras

Por: Alexander Medina Félix¹

Año 2014, Vol. VIII, No. 2

1. Introducción

Los modelos dinámicos con datos de panel han sido retomados como una herramienta de análisis de la literatura económica en el transcurso de las dos últimas décadas con el fin de analizar el crecimiento de largo plazo (ver Mankiw, Romer y Weil, 1992; Levine y Renelt, 1992).

La literatura de la economía laboral también ha estado utilizando estos modelos para el análisis de la relación existente entre salarios y desempleo (ver Blanchflower y Oswald, 1994; Blanchard y Katz, 1997; Galiani, 1999). También en el área financiera se han estado aplicando estas metodologías para los análisis aplicados (Ayuso et al., 2002; Jokipii y Milne, 2007).

Acorde con esta tendencia de un mayor uso de modelos dinámicos con datos de panel, los investigadores latinoamericanos han mostrado en la última década un interés en esta herramienta de análisis para examinar los principales preceptos de la teoría económica en los países de esta región (Le Fort y Budnevich, 2005; Bailey-Tapper, 2011).

Los estudios con datos de panel dinámicos utilizan distintas dimensiones de observaciones temporales (T) y observaciones de corte transversal (N) en los análisis de los datos macroeconómicos.

La magnitud de T y N es fundamental en la selección del método de estimación en los modelos de datos de panel dinámicos debido a que los estimadores tradicionales como *Least Square Dummy Variables (LSDV)* son sesgados e inconsistentes para un T fijo y N tendiendo a infinito; mientras que los estimadores que utilizan procedimientos de Variables Instrumentales (IV) ó Método Generalizado de Momentos (GMM) son consistentes cuando N tiende a infinito.

En este artículo se estudian los métodos de estimación e inferencia en paneles dinámicos siguiendo la línea del estimador de variables binarias (LSDV) y los estimadores desarrollados en los trabajos de Arellano y Bond (1991), Blundell y Bond (2000) y Kiviet (1995) en el contexto en el cual el conjunto de datos macroeconómicos no tienen una dimensión muy pequeña en T y no tienen una dimensión muy grande en N.

El objetivo de este trabajo es evaluar el desempeño de los diferentes estimadores en términos de sesgo, consistencia y eficiencia con distintos tamaños de N y T y con distintos niveles de

¹ Subdirección de Estudios Económicos. Departamento de Programación Monetaria y Estudios Económicos. Para preguntas y comentarios escribir a a.medina@bancentral.gov.do.

persistencia (α) de las series macroeconómicas generadas, en un contexto de datos de panel dinámicos.

El análisis de los estimadores en un contexto de panel dinámico nos permitirá responder algunas preguntas, tales como: ¿Es un valor de $T=40$ lo suficientemente grande como para ignorar el sesgo del estimador LSDV?, ¿Es $N=60$ lo suficientemente grande para que los métodos de IV ó GM arrojen estimadores consistentes?, ¿Cuál estimador es más eficiente para modelar series de datos de panel con persistencia alta? Para responder este tipo de interrogantes, se realizan varios experimentos Monte Carlo.

El artículo se divide en 6 secciones. En la segunda sección se presenta la descripción y propiedades de los distintos estimadores evaluados. La sección III describe el modelo a estimar mediante un experimento Monte Carlo y se definen los indicadores de sesgo y eficiencia a utilizar. En la cuarta sección se presentan los resultados bajo distintos escenarios de N , T y α . En la sección V se lleva a cabo un experimento Monte Carlo para evaluar los estimadores mencionados con un tamaño muestral de $N=13$, $T=6$ y un $\alpha=0.37$. Este tamaño de muestra es el utilizado en un estudio para Latinoamérica realizado por la CEPAL.² En la sección VI se presentan las conclusiones.

II. Estimadores: descripción y propiedades

En este artículo utilizamos un concepto restringido de modelos dinámicos de datos de panel donde los modelos son lineales y la variable dependiente está rezagada sólo un periodo.

El modelo a analizar es el siguiente:

$$(1) \quad y_{jt} = \alpha y_{jt-1} + \beta x_{jt} + c_j + u_{jt}, \quad j = 1, 2, \dots, N; \quad t = 1, 2, \dots, T + 10$$

Donde c_j representa la heterogeneidad no observada y es generada como $c_j \sim N(0,1)$; u_{jt} es el error idiosincrático generado como $u_{jt} \sim N(0,1)$. Las primeras 10 observaciones de series temporales se descartan por lo que el tamaño muestral es NT .

El regresor adicional, x_{jt} , es una variable estrictamente exógena condicionada en c_j , generada de la siguiente manera: $x_{jt} = 0.8x_{jt-1} + v_{jt}$, donde $v_{jt} \sim N(0,0.9)$. La primera observación es $y_{j0} = 0$. Se asume el cumplimiento de la condición de exogeneidad secuencial (Chamberlain, 1992).³

Analizaremos las propiedades de consistencia y eficiencia del estimador α cuando se utilizan como métodos de estimación *Least Square Dummy Variables (LSDV)*, Arellano-Bond (1991), Blundell y Bond (2000) y Kiviet (1995).

² Le Fort, G. y Budnevich, C. (2005)

³ Se dice que las variables explicativas son secuencialmente exógenas condicionadas en el efecto no observable cuando se cumple que: $E(y_{jt} | x_{jt}, x_{jt-1}, \dots, x_{j1}, c_j) = E(y_{jt} | x_{jt}, c_j) = x_{jt}\beta + c_j$. Es decir, después de haber controlado por x_{jt} y c_j , ningún valor presente y pasado afecta el valor esperado de y_{jt} . El cual es un concepto menos restrictivo que la exogeneidad estricta, pues permite relación con los valores futuros.

2.1 Estimador de Least Square Dummy Variables (LSDV)

En un modelo dinámico como el de la ecuación (1) es importante eliminar la heterogeneidad no observada c_j mediante alguna transformación como la que se realiza bajo el modelo de variables binarias (LSDV).⁴

Para estimar el modelo por LSDV se definen N variables binarias, una para cada observación de corte transversal. Luego se estima por POLS (Pool Ordinary Least Squares) una regresión de y_{jt} sobre las variables binarias, y x_{jt} con $t=1,2,\dots,T$; $j=1,2,\dots,N$. Los coeficientes que acompañan a las variables binarias son las estimaciones de los c_j .

El estimador obtenido utilizando esa transformación en un POLS es el siguiente:

$$(2) \quad \hat{\beta}^{LSDV} = [X'(I_N \otimes Q_T)X]^{-1}X'(I_N \otimes Q_T)y$$

Donde $Q_T = I_T - J_T(J_T'J_T)^{-1}J_T'$ es la matriz *time-demeaning* que es una matriz simétrica e idempotente de rango $T-1$; $(J_T'J_T)^{-1}$ representa la matriz de promedios temporal y I_T es la matriz identidad.

La varianza del estimador LSDV viene dada por:

$$(3) \quad \hat{V} = \left(\sum_{j=1}^N \ddot{X}'_j \ddot{X}_j \right)^{-1} \left(\sum_{j=1}^N \ddot{X}'_j \hat{u}_j \hat{u}'_j \ddot{X}_j \right) \left(\sum_{j=1}^N \ddot{X}'_j \ddot{X}_j \right)^{-1}$$

Donde los \hat{u}_j son los residuos estimados usando la metodología de Efectos Fijos (FE) y los puntitos sobre las variables representan variables *time-demeaning*. Una desventaja con el enfoque LSDV es que los estimadores de α son no consistentes bajo T fijo.

2.2 El Estimador de Arellano y Bond (AB)

Para tratar con el problema de variables endógenas del modelo, que causa inconsistencia del estimador LSDV (visto en 2.1) en muestras con dimensión temporal pequeña, necesitaremos un estimador de variables instrumentales. Para obtener una variable instrumental de la variable endógena es más sencillo transformar el modelo usando diferencias finitas que utilizando efectos fijos. En particular Arellano y Bond (1991) utilizan la transformación de diferencias finitas de orden 1. De esta forma, obtenemos el modelo transformado de la ecuación 4.

$$(4) \quad \Delta y_{it} = \rho \Delta y_{it-1} + \Delta u_{it}$$

Donde Δ representa el operador de diferencias finitas.

⁴ En LSDV se ve c_j como parámetros a ser estimados.

Se podría estimar de forma consistente el parámetro ρ a través de un estimador de variables instrumentales eficiente, que utilice como estimador todos los instrumentos posibles, usando el método de los momentos generalizados (GMM) con la siguiente matriz de instrumentos Z_i para cada observación de corte transversal i :

$$Z_i = \begin{bmatrix} y_{i1} & 0 & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & y_{i1} & y_{i2} & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & 0 & \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & \dots & y_{i1} & \dots & y_{iT-2} \end{bmatrix}$$

donde los y_{it} son los instrumentos en cada periodo temporal, para una muestra aleatoria de tamaño NT .⁵

$$Z = \begin{bmatrix} Z_1 \\ \vdots \\ Z_N \end{bmatrix}$$

Z es una matriz de matrices que compila las matrices de instrumentos de las distintas observaciones de corte transversal.

El estimador de Arellano y Bond (AB) de ρ viene dado por:

$$(5) \quad \hat{\rho}^{AB}(W) = [\Delta' y_{t-1} (ZWZ') \Delta y_{t-1}]^{-1} \Delta' y_{t-1} (ZWZ') \Delta y$$

Donde $W = (\sum_{i=1}^N \Delta Z_i' \Delta Z_i)^{-1}$ es la matriz ponderadora de AB de un paso.⁶

La estimación de la varianza del estimador de AB es:

$$(6) \quad \widehat{var}(\hat{\rho}^{AB}) = \hat{\sigma}^2 [\Delta' y_{t-1} (ZWZ') \Delta y_{t-1}]^{-1}$$

Donde $\hat{\sigma}^2$ es una estimación consistente de σ^2 .

2.3 Estimador de Blundell-Bond (BB)

Cuando ρ se aproxima a uno, los instrumentos utilizados por Arellano-Bond se vuelven instrumentos débiles. Por tanto, Blundell y Bond (2000) proponen agregar nuevos instrumentos a los usados por AB a través de un estimador de sistemas (GMM), y esos instrumentos se introducen a la ecuación original en niveles en forma de diferencias finitas rezagadas, por lo que

⁵ Este estimador es más eficiente que el de Anderson y Hsiao (1982) (AH) el cual es un procedimiento que genera estimaciones consistentes y se basa en variables instrumentales de la variable endógena en una ecuación dinámica del tipo del modelo (1). Debido que AH es un estimador que utiliza como instrumento de la variable diferencia finita a la variable dependiente en niveles, rezagada dos periodos, se podría estimar entonces de forma consistente el estimador de α usando como variable instrumental a la variable dependiente rezagada dos periodos (y_{t-2}).

⁶ El estimador de AB se calcula también utilizando una matriz de ponderadores GMM de dos pasos en el que se utiliza los residuos del estimador de 1 paso de AB para calcular una nueva matriz ponderadora.

los nuevos instrumentos no están relacionados con el error en niveles. Sólo se agrega un instrumento de diferencias finitas por período temporal.

La matriz Z_i^+ es la matriz de instrumentos de BB para una observación de corte transversal i .

$$Z_i^+ = \begin{bmatrix} Z_i & 0 & 0 & 0 & \dots & 0 \\ 0 & \Delta y_{i2} & 0 & 0 & \dots & 0 \\ 0 & 0 & \Delta y_{i3} & 0 & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & \dots & \vdots & 0 \\ 0 & 0 & 0 & \dots & 0 & \Delta y_{i(T-1)} \end{bmatrix}$$

Donde Z_i es la matriz de instrumentos de AB.

$$Z^+ = \begin{bmatrix} Z_i^+ \\ \vdots \\ Z_N^+ \end{bmatrix}$$

Z^+ es una matriz de matrices que compila las matrices de instrumentos de las distintas observaciones de corte transversal de Z_i^+ .

El estimador de Blundell y Bond de un paso es el siguiente estimador:⁷

$$(7) \quad \hat{\rho}^{(1)} = \left\{ \left[[y'_{t-1} (I_N \otimes H')] Z^+ \right] W^{(1)} \left[[Z^+ (I_N \otimes H')] \right] \right\}^{-1} \times \left\{ \left[[y'_{t-1} (I_N \otimes H')] Z^+ \right] W^{(1)} \left[[Z^+ (I_N \otimes H')] \right] \right\}$$

Donde $W^{(1)} = [\sum_{i=1}^N Z_i^{+'} H H' Z_i^+]^{-1}$. H es una matriz en bloques de transformación de diferencias finitas y \otimes es el producto de Kronecker. La varianza del estimador de BB es:

$$(8) \quad \widehat{var}(\hat{\rho}^{BB}) = \hat{\sigma}^2 \left\{ \left[[y'_{t-1} (I_N \otimes H')] Z^+ \right] W^{(1)} \left[[Z^+ (I_N \otimes H')] \right] \right\}^{-1}$$

El $\hat{\sigma}^2$ es una estimación consistente de σ^2 .

2.4 Estimador de Kiviet

El estimador de Kiviet es una versión corregida del LSDV, donde se le suma al LSDV una estimación del sesgo de ese estimador. Kiviet (1995) muestra que el sesgo de LSDV puede aproximarse con un error de tamaño consistente.⁸

⁷ El estimador de dos pasos de BB utiliza en la matriz ponderadora los residuos del estimador de 1 paso. En este trabajo sólo se calcula el estimador de un paso.

⁸ Una estimación del sesgo de LSDV vendría dada por el sesgo de Nickel (1981), que es la diferencia entre el estimador de Kiviet y el estimador de efectos fijos. Este sesgo sólo disminuye si $t \rightarrow \infty$. Si $\alpha > 0$ el sesgo es siempre negativo y nunca converge a cero aún si $\alpha = 0$. El sesgo, en este caso, se vuelve más grande si se adicionan regresores exógenos a la ecuación. Y los estimadores de los coeficientes que acompañan a los regresores son también inconsistentes.

El estimador de Kiviet (1995, pp. 64) se calcula de la siguiente forma:

$$(9) \quad \hat{\delta}_{LSDVC} = \hat{\delta}_{FE} - \text{Sesgo}_{Kiviet}$$

El sesgo de Kiviet es calculado una estimación de variables instrumentales de Anderson y Hsiao (1982) (ver Judson y Owen, 1999). La matriz de varianzas y covarianzas de Kiviet es:

$$(10) \quad \text{Var}(\hat{\delta}_{LSDVC}) = \hat{\sigma}_u^2 \left(\sum_{i=1}^N \ddot{Z}_i' \ddot{Z}_i \right)^{-1}$$

El $\hat{\sigma}_u^2$ es una estimación consistente de σ_u^2 .

III. Experimento de Montecarlo

En este artículo realizaremos un experimento de Montecarlo con 1000 simulaciones, con el fin de calcular la media del estimador ($\bar{\alpha}$) y su desvío estándar. Con esos datos de las simulaciones calcularemos el sesgo y la raíz del error cuadrático medio (RECM) usando los estimadores: LSDV, AB, BB y Kiviet.

El RECM es la raíz cuadrada de la suma de dos cantidades no negativas: una es la varianza del estimador y la otra es el cuadrado del sesgo del estimador.

$$(11) \quad RECM = \sqrt{\text{Var}(\hat{\alpha}) + (\alpha - \bar{\alpha})^2}$$

IV. Resultados

4.1 Resultados cuando T es fijo y N aumenta

En la Tabla 1 se presentan los resultados de los cuatro estimadores analizados anteriormente con un $\alpha = 0.5$ un T fijo y N aumentando. En el caso 1 se realiza un experimento Monte Carlo con N=20 y en el caso 2 se usa N=60. En ambos casos se observa que los estimadores con menor sesgo y mayor eficiencia (menor RECM) son los de Kiviet y Blundell-Bond.

El estimador con mayor sesgo es el de LSDV, esto era de esperarse debido a que en modelos de paneles dinámicos como el de la ecuación (1), donde asumimos exogeneidad secuencial de las variables explicativas de la ecuación en el modelo original (no transformado) y un T pequeño, no se puede garantizar consistencia del estimador aun cuando N aumente.

Dada la inconsistencia de los estimadores LSDV en el modelo dinámico (1) bajo el supuesto de exogeneidad secuencial en las variables originales, se recomienda el uso de los estimadores de Arellano-Bond, Blundell-Bond y Kiviet. El estimador de Arellano-Bond muestra un fuerte sesgo cuando N es pequeño, el cual va disipándose a medida que N es mayor.

Tabla 1. Resultados cuando T es fijo y N aumenta

Estimadores	Alfa estimado Promedio	Sesgo	Desviación Estándar	RECM
	<i>Caso 1: $\beta=0, T=10, N=20, a=0.5, \beta=0$</i>			
LSDV	0.3121	-0.1879	0.0732	0.2017
Arellano-Bond	0.3576	-0.1424	0.1237	0.1886
Blundell-Bond	0.4931	-0.0069	0.1264	0.1266
Kiviet	0.4932	-0.0068	0.0939	0.0941
<i>Caso 2: $\beta=0, T=10, N=60, a=0.5, \beta=0$</i>				
LSDV	0.3151	-0.1849	0.0428	0.1898
Arellano-Bond	0.4390	-0.0610	0.0787	0.0996
Blundell-Bond	0.4939	-0.0061	0.0731	0.0734
Kiviet	0.4905	-0.0095	0.0540	0.0548

Fuente: Generada con simulaciones Monte Carlo.

Además de la condición de exogeneidad secuencial de las variables explicativas en niveles, se necesitan, adicionalmente, dos condiciones de estabilidad para que los estimadores surgidos con los métodos de VI y GMM sean estables y sus varianzas sean estacionarias: 1) $\alpha < |1|$ y (2) que la esperanza matemática y la varianza en su estado estacionario coincidan con la media de la primera observación y la varianza de la primera observación, respectivamente.

Si las dos condiciones de estabilidad de los estimadores no se cumplen, los estimadores no son consistentes. Como en este ejercicio asumimos que se cumplen ambos supuestos los estimadores de GMM (Arellano y Bond, Blundell) son consistentes como muestran los resultados.

En Rangel (2012) se obtienen resultados similares para un aumento de $N=30$ a $N=50$, donde se evidencia la consistencia de los estimadores de Arellano y Bond y la inconsistencia de los LSDV en paneles dinámicos cuando T es fijo (pequeño) y N aumenta.

4.2 Resultados cuando N es fijo y T aumenta

En la Tabla 2 se analizan las propiedades de los distintos estimadores cuando la dimensión temporal aumenta de $T=10$ a $T=40$, manteniendo el N fijo en 20 y el $\alpha = 0.5$. En la Tabla 1 vimos que cuando N aumentaba en un contexto de T pequeño el estimador LSDV no variaba mucho, y reflejaba un sesgo significativo que se mantenía luego del aumento de N.

En la Tabla 2 se observa que al aumentar T, aún con un N pequeño, se observa que el estimador LSDV se acerca bastante al parámetro poblacional (0.4583) y asimismo registra una reducción significativa del RECM. Esto suele ocurrir debido a que en términos generales los estimadores LSDV son consistentes al aumentar la dimensión temporal. No obstante, es el estimador de Kiviet el que registra el menor sesgo y su RECM (0.0332) es el mejor en términos relativos.

Tabla 2. Resultados cuando N es fijo y T aumenta

Estimadores	Alfa estimado Promedio	Sesgo	Desviación Estándar	RECM
<i>Caso 3: $\beta=0$, $T=40$, $N=20$, $a=0.5$, $\beta=0$</i>				
LSDV	0.4583	-0.0417	0.0320	0.0526
Arellano-Bond	0.4535	-0.0465	0.0340	0.0576
Blundell-Bond	0.4930	-0.0070	0.0419	0.0425
Kiviet	0.4996	-0.0004	0.0332	0.0332
<i>Caso 1: $\beta=0$, $T=10$, $N=20$, $a=0.5$, $\beta=0$</i>				
LSDV	0.3121	-0.1879	0.0732	0.2017
Arellano-Bond	0.3576	-0.1424	0.1237	0.1886
Blundell-Bond	0.4931	-0.0069	0.1264	0.1266
Kiviet	0.4932	-0.0068	0.0939	0.0941

Fuente: Generada con simulaciones Monte Carlo.

Estos resultados coinciden con los de Rangel (2012) quien hace el experimento para un aumento de $T=10$ a $T=40$, con $N=50$ fijo. El obtiene que el estimador de Kiviet es el de menor sesgo y más eficiencia y que el estimador LSDV es consistente en ese escenario.

4.3 Efectos del cambio de la persistencia: resultados cuando N es pequeño y T grande⁹

En la Tabla 3 se observa un panel con 20 observaciones en la dimensión transversal (N) y 40 en la dimensión temporal (T). En este caso lo que varía es el coeficiente autorregresivo α , que pasa de 0.2 a 0.9 entre cada uno de los casos analizados en la tabla (caso 4 y 5).

Tabla 3. Efectos del cambio de la persistencia
Resultados cuando N es pequeño y T grande*

Estimadores	Alfa estimado Promedio	Sesgo	Desviación Estándar	RECM
<i>Caso 4: $\beta=0$, $T=40$, $N=20$, $\alpha=0.2$, $\beta=0$</i>				
LSDV	0.1676	-0.0324	0.0344	0.0473
Arellano-Bond	0.1656	-0.0344	0.0357	0.0496
Blundell-Bond	0.1968	-0.0032	0.0416	0.0417
Kiviet	0.1999	-0.0001	0.0355	0.0355
<i>Caso 5: $\beta=0$, $T=40$, $N=20$, $\alpha=0.9$, $\beta=0$</i>				
LSDV	0.8507	-0.0493	0.0202	0.0533
Arellano-Bond	0.8354	-0.0646	0.0252	0.0693
Blundell-Bond	0.9582	0.0582	0.0237	0.0628
Kiviet	0.8752	-0.0248	0.0298	0.0388

*Asumimos que N es pequeño en relación a T.

Fuente: Generada con simulaciones Monte Carlo.

⁹ Asumimos N es pequeño en relación a T.

En primer lugar, se advierte que cuando $\alpha = 0.2$ los estimadores de menor sesgo son el de Kiviet y Blundell-Bond, mientras que los de LSDV y Arellano-Bond presentan sesgos altos de alrededor de 20% de su media estimada. Al analizar lo sucedido cuando el parámetro pasa de 0.2 a 0.9 se observa que el estimador de menor sesgo y menor RECM es el de Kiviet.

El sesgo de Nickell (diferencia entre LSDV y Kiviet) baja de 0.0323 a 0.0245, al pasar de $\alpha = 0.2$ a $\alpha = 0.9$, lo cual es esperado debido a que con un T relativamente alto el sesgo de Nickell tiende a ser bajo. Sin embargo, a medida que α aumenta el sesgo tiende a persistir.¹⁰

Un punto importante es que el estimador de Arellano-Bond es el de mayor sesgo y mayor RECM cuando $\alpha = 0.9$, lo cual es consistente con la literatura teórica que especifica que cuando α es cercano a 1 los instrumentos utilizados se vuelven débiles (que consiste en que la matriz de instrumentos a utilizar, a pesar de ser exógenos no tienen relación significativa con la variable endógena que en este caso es la diferencia finita rezagada).

4.4 Efectos del cambio de la persistencia: resultados cuando T es pequeño y N grande¹¹

En la Tabla 4 se analizan los casos 6 (Caso 6: $\beta=0, T=10, N=60, \alpha=0.2$) y caso 7 (Caso 7: $\beta=0, T=10, N=60, \alpha=0.9$) donde el coeficiente de persistencia pasa de $\alpha = 0.2$ a $\alpha = 0.9$, pero a diferencia de la Tabla 3, el T es pequeño y el N grande. En este caso se advierte que el estimador LSDV es altamente sesgado en ambos casos.

**Tabla 4. Efectos del cambio de la persistencia
Resultados cuando T es pequeño y N grande***

Estimadores	Alfa estimado Promedio	Sesgo	Desviación Estándar	RECM
<i>Caso 6: $\beta=0, T=10, N=60, \alpha=0.2, \beta=0$</i>				
LSDV	0.0626	-0.1374	0.043	0.144
Arellano-Bond	0.1697	-0.0303	0.0641	0.0709
Blundell-Bond	0.1977	-0.0023	0.0632	0.0632
Kiviet	0.1979	-0.0021	0.0504	0.0504
<i>Caso 7: $\beta=0, T=10, N=60, \alpha=0.9, \beta=0$</i>				
LSDV	0.6876	-0.2124	0.0382	0.2158
Arellano-Bond	0.7913	-0.1087	0.094	0.1437
Blundell-Bond	0.9924	0.0924	0.0197	0.0945
Kiviet	0.8491	-0.0509	0.0492	0.0708

*Asumimos que T es pequeño en relación a N.
Fuente: Generada con simulaciones Monte Carlo.

¹⁰ Ver Nickell (1981).

¹¹ Asumimos T es pequeño en relación a N.

El estimador de AB sigue mostrando el problema de instrumentos débiles, analizado en la sección anterior. El estimador de Kiviet es el que menor sesgo y menor RECM presenta en el experimento.

Se nota que el sesgo de Nickell se mantiene persistente, e incluso aumentando cuando aumenta el α , reflejando el problema de inconsistencia de los estimadores LSDV (independientemente del nivel de persistencia) cuando la dimensión temporal es pequeña.

V. Caso estudio con tamaño muestral $N=13$, $T=6$

En esta sección llevamos a cabo un experimento Monte Carlo usando como referencia el tamaño de la muestra y el nivel de persistencia usado por Le Fort y Budnevich (2005, pp. 25-31), en su estudio de brecha del PIB para Latinoamérica. Los autores utilizan un modelo de panel dinámico del tipo de la ecuación 1 de este artículo, que incluye la variable dependiente rezagada un período y variables exógenas.

La metodología utilizada por Le Fort y Budnevich (2005) es el método generalizado de momentos de Arellano- Bond en un contexto de datos de panel dinámicos que incluye a trece países de Latinoamérica ($N=13$) que son: Argentina, Bolivia, Brasil, Chile, Colombia, Costa Rica, Ecuador, México, Panamá, Perú, República Dominicana, Uruguay y Venezuela.

La evaluación de la metodología usada en el trabajo de la CEPAL para estimar el panel dinámico (Arellano-Bond) es comparada en esta sección con otros métodos (LSDV, Blundell-Bond y Kiviet), usando un experimento Monte Carlo que parte de los niveles de N , T y α del estudio citado.

Utilizando nueve especificaciones diferentes Le Fort y Budnevich encuentran coeficientes de persistencia α que varían desde 0.3328 a 0.4321 (con un promedio de 0.37) todos significativos al 1% y usando $T=6$ y $N=13$. Estos resultados muestran una robustez en $\hat{\alpha}$ independientemente de cuáles variables exógenas son utilizadas en la estimación.¹² Llevamos a cabo un experimento Monte Carlo con simulaciones de 1000 repeticiones, con $\alpha=0.37$, $N=13$ y $T=6$.

Tabla 5. Experimento Montecarlo con modelo dinámico del Ciclo del PIB con α encontrados por Le Fort y Budnevich (2005) de la CEPAL

Estimadores	Alfa estimado Promedio	Sesgo	Desviación Estándar	RECM
<i>Caso 8: $\beta=0$, $T=6$, $N=13$, $\alpha=0.37$</i>				
LSDV	0.0626	-0.1374	0.043	0.144
Arellano-Bond	0.1697	-0.0303	0.0641	0.0709
Blundell-Bond	0.1977	-0.0023	0.0632	0.0632
Kiviet	0.1979	-0.0021	0.0504	0.0504

Fuente: Generada con simulaciones Monte Carlo.

¹² Ver Tabla 4 de Le Fort y Budnevich (2005).

Los resultados de la Tabla 5 muestran que los estimadores de Arellano-Bond y LSDV son sesgados e ineficientes para el tamaño muestral utilizado en el experimento. Los estimadores Blundell-Bond y Kiviet son los mejores estimadores en este caso.

VI. Conclusiones

En este artículo se estudiaron y compararon las propiedades en muestras finitas de distintos métodos de estimación e inferencia en paneles dinámicos. Los estimadores estudiados fueron el de Least Square Dummy Variable (LSDV), Arellano-Bond, Blundell-Bond y Kiviet con datos muestrales de distintas dimensiones temporales y transversales y varios grados de persistencia del modelo.

Se utilizaron ejercicios tipo Monte Carlo que se usaron en la investigación de las propiedades de los estimadores. Se comparó el desempeño de los estimadores en varios contextos, uno de dimensión temporal (T) pequeña y fija con dimensión transversal (N) aumentando; otro donde el N estaba fijo y T aumentando; el tercero donde se evaluaron los efectos del cambio de la persistencia cuando N es pequeño y T grande y el cuarto se evaluó el efecto del cambio de la persistencia pero, a diferencia del tercer escenario, en el cuarto el T es pequeño y el N grande.

Los resultados evidencian que en el punto 1 donde T es pequeño y fijo y N aumenta, los estimadores con mejor desempeño son los de Blundell-Bond y Kiviet, mientras que el LSDV es sesgado e inconsistente, como sugiere la teoría econométrica. En el contexto del punto 2 cuando N es fijo y T aumenta se observa que el estimador de LSDV es consistente y aumenta su eficiencia. El estimador de Kiviet es, sin embargo, el de menor sesgo y mayor eficiencia.

Cuando se evaluaron las propiedades de los estimadores ante un incremento de la persistencia desde $\alpha = 0.2$ a $\alpha = 0.9$, cuando N es pequeño y T grande se advierte primeramente que cuando el $\alpha = 0.2$ los estimadores de LSDV y Arellano - Bond son los que presentan un mayor sesgo mientras que los de Kiviet y Blundell-Bond son los de menor sesgo. Cuando se pasa de $\alpha = 0.2$ a $\alpha = 0.9$ el estimador de menor sesgo y mayor eficiencia es el Kiviet en un contexto de alta persistencia y dimensión temporal grande y N pequeño.

Bajo el supuesto de T pequeño y N grande se observó que el estimador LSDV es sesgado independientemente del grado de persistencia supuesto del modelo. El estimador de Kiviet es de menor sesgo y menor RECM en este caso.

El estimador de Arellano-Bond resultó sesgado e ineficiente cuando la persistencia es alta, reflejando un posible problema de instrumentos débiles, descrito en la literatura teórica.

En el caso estudio presentado en la sección V donde se llevó a cabo un experimento de Monte Carlo para analizar la propiedad de los estimadores utilizando el tamaño muestral de un documento de la CEPAL, se observa que el estimador de Arellano-Bond utilizado en ese trabajo para calcular la persistencia y los demás coeficientes exógenos de la ecuación, pudiera ser

complementado con los estimadores de Blundell-Bond y de Kiviet, que son estimadores insesgados y consistentes para el tamaño de la muestra.

Bibliografía

Anderson, T.W. y Hsiao, C. (1981). "Estimation of Dynamic models with errors components". *Journal of the American Statistical Association*, 76. pp. 598-606.

Arellano, M. y Bond, S. (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations". *The Review of Economic Studies*, 58. pp. 277-297.

Ayuso, J., Pérez, D. y Saurina, J. (2002). "Are Capital Buffers Pro-Cyclical? Evidence from Spanish Panel Data". Banco de España (Research Working Paper, No.0224).

Bailey-Tapper, S.A. (2011). "Investigating the link between bank capital and economic activity: evidence on Jamaican panel data". *Money Affairs*, Vol. XXIV, No. 2. pp. 163-188.

Blanchard, O. y Katz, L. (1997). "What we know and do not know about the natural rate of unemployment". *Journal of Economic Perspectives*, 11. pp. 51-72.

Blanchflower, D. y Oswald, A. (1996). "The wage curve". The MIT Press, Cambridge, Massachusetts.

Blundell, R. y Bond, S. (2000). "Initial conditions and moment restrictions in dynamic panel data models". *Journal of Econometrics*, 87. pp. 115-143.

Chamberlain, G. (1992). "Comment: Sequential moment restrictions in panel data". *Journal of Business and Economic Statistics*, Vol. 10, No.1. pp. 20-26.

Galiani, S. (1999). "Wage determination in Argentina: An econometric analysis with methodology discussion". ITDT WP 218.

Jokipii, T. y Milne, A. (2007). "The Cyclical Behavior of European Capital Buffers". Swedish Institute for Financial Research (Report).

Judson, R. y Owen, A. (1999). "Estimating dynamic panel data models: A guide for macroeconomists". *Economics Letters*, 79. pp. 145-152.

Kiviet, J. (1995). "On bias, inconsistency, and efficiency of various estimators in dynamic panel data models". *Journal of Econometrics*, 68. pp. 53-78.

Le Fort, G. y Budnevich, C. (2005). "Entendiendo las volatilidades cíclicas de la cuenta de capitales y del PIB: Un estudio de panel para países latinoamericanos". *Documentos de la CEPAL, Serie Macroeconomía del Desarrollo*, 42. pp. 25-31.

Levine, R. y Renelt, D. (1992). "A sensitivity analysis of cross country growth regressions". *American Economic Review*, September 1992. pp. 942-963.

Mankiw, N.G., Romer, D. y Weil, D. (1992). "A contribution to the empirics of economic growth". *Quarterly Journal of Economics*. pp. 407-438.

Nickell, S.J. (1981). "Biases in dynamic models with fixed effects". *Econometrica*, 49. pp. 1417-1426.

Rangel, A.E. (2012). "Superioridad relativa de los estimadores Kiviet y Blundell-Bond (GMM1) en paneles dinámicos. Un experimento Monte Carlo con muestras finitas". *Estudios Gerenciales*, 28. pp. 81-86.

Determinantes de la Deserción y Repitencia Escolar de los Hogares en Condiciones de Pobreza en la República Dominicana: Evidencia de la Encuesta de Evaluación de la Protección Social 2010

Por: José Manuel Lozano¹

Año 2014, Vol. VIII, No. 3

I. Introducción

En este artículo, el término deserción escolar se refiere a los estudiantes que, por alguna razón, dejan de asistir a clase y se mantienen fuera del sistema escolar. Por otro lado, la repitencia escolar para los estudiantes ocurre en aquellos casos donde estos no son promovidos al curso siguiente, luego de finalizado el año escolar. Por lo tanto, el estudiante se ve precisado a repetir el grado que estaba cursando.

La investigación aquí presentada consiste precisamente en analizar el impacto de un conjunto de variables correspondientes a características individuales, aspectos institucionales y aspectos del hogar sobre la deserción y repitencia escolar. Entre las características individuales tenemos variables tales como ayuda en la tarea, asistencia escolar, edad del estudiante, edad de inicio en el nivel básico, género del estudiante, medio de transporte, tiempo dedicado a hacer la tarea, tiempo de transportación y trabajo. Entre los aspectos institucionales, tenemos variables como la obtención de libros, Incentivo a la Asistencia Escolar (ILAE), obtención de útiles y centro educativo. Por último, entre las variables referentes a aspectos del hogar, tenemos a la educación del jefe del hogar, el estado civil del jefe del hogar, el género del jefe del hogar, el zona de residencia del estudiante y su familia y el ingreso familiar mensual.

El objetivo es cuantificar la incidencia de los determinantes del desempeño académico en los niveles básico y medio para hogares en condiciones de pobreza en RD. Específicamente, se busca estimar el impacto de las características individuales de los estudiantes y factores relacionados con los hogares, así como el programa de Incentivo a la Asistencia Escolar (ILAE) en la repitencia y el abandono escolar. Dentro de los determinantes se hará mayor énfasis en aquellos que pueden ser modificados a través de políticas públicas.

La literatura que estudia los determinantes de la deserción y repitencia escolar en la RD es bastante escasa. Este estudio, no obstante, contribuye al tema de los factores causales del fenómeno del desempeño escolar en RD desde el punto de vista econométrico, lo que constituye un aporte novedoso. Entre los beneficiarios de esta investigación están los segmentos poblacionales de bajos ingresos, es decir, aquellos hogares que se encuentran en condiciones de pobreza, dado que la identificación de las principales variables que se encuentran detrás de la deserción y repitencia

¹ División de Consolidación de Estadísticas Monetarias y Financieras, Departamento de Programación Monetaria y Estudios Económicos. Para preguntas y comentarios escribir a j.lozano@bancentral.gov.do.

escolares permitirá a los tomadores de decisión de política en materia educativa diseñar políticas con mayor efectividad.

En el presente estudio se emplean tanto modelos probit como logit (Gujarati, 2004) para relacionar la deserción y repitencia escolar con variables que agrupan características individuales de los estudiantes, aspectos institucionales y aspectos del hogar.

El documento está organizado de la siguiente manera. En la sección 2 que sigue a esta introducción se explica la metodología a utilizar. En la sección 3 se describen los datos y variables a utilizar en la modelización. En la sección 4 se analizan los resultados y la sección 5 presenta las conclusiones y recomendaciones.

II. Metodología empírica

Generalmente, el desempeño escolar se mide con las calificaciones de los estudiantes. No obstante, los datos que disponemos no nos permiten observar directamente esta variable. Por esta razón, se ha optado por utilizar la repitencia como medida de desempeño escolar. Si un estudiante repite un curso es porque sus calificaciones son bajas y en consecuencia, su desempeño ha sido bajo. Se utiliza este tipo de medida debido a que la encuesta no nos da información sobre las calificaciones de los estudiantes, medida de desempeño más utilizada en la literatura internacional. De igual manera, si el estudiante abandona la escuela se presume que no ha tenido un buen desempeño en cuanto a sus calificaciones o que el sistema escolar y social no es capaz de crear las condiciones para que el individuo siga su formación, lo que reflejaría un problema de eficiencia en el sistema educativo.

Para analizar situaciones en las cuales se tienen variables dependientes dicotómicas se utilizan modelos de variable dependiente limitada, específicamente los modelos de probabilidad logit y probit son una buena opción.² Estos dos modelos permiten estimar el impacto de cada uno de los factores determinantes en la probabilidad de que un estudiante repita un grado o abandone la escuela. Se eligieron estos modelos debido a los problemas que presenta la modelización de variable dependiente binaria mediante el Modelo Lineal de Probabilidad (MLP).

En un primer momento, se asume que la probabilidad de que un estudiante repita un grado/abandone la escuela sigue una distribución logística como en la ecuación 1:

$$(1) \quad P_i = \frac{1}{1+e^{-z}}$$

Donde P_i es la probabilidad de que un estudiante repita un grado (abandone la escuela) y $1 - P_i$ es la probabilidad de promover el curso (permanecer en la escuela). La variable Z se describe en la ecuación 2.

$$(2) \quad Z = X\beta + \varepsilon$$

² Para detalles sobre la metodología logit y probit ver Gujarati (2004).

Donde Z es un vector $N \times 1$ que toma valores entre más y menos infinito (en el conjunto de números reales). La variable X es una matriz de variables explicativas o de los factores determinantes del desempeño académico. Se puede demostrar que el logaritmo de la razón entre la probabilidad de repetencia P_i (abandono escolar) y la de promover de curso $1 - P_i$ es igual a Z :

$$(3) \quad L_i = \log \left[\frac{P_i}{1 - P_i} \right] = Z = X\beta + \varepsilon$$

El coeficiente β es un vector que incluye el intercepto y las pendientes, los antilogaritmos de estas últimas se interpretan como el cambio porcentual en la probabilidad de repetir (abandonar la escuela) ante una unidad de incremento del determinante correspondiente.

En el caso del modelo probit, se parte de una función de distribución normal estandarizada, similar a la utilizada en los textos de estadísticas. El procedimiento consiste en lo siguiente. Para cada nivel dado de probabilidad se obtiene un valor de Z y se utiliza este valor como variable dependiente. La probabilidad de Z viene dada por la siguiente ecuación:

$$(4) \quad f(Z_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^Z \frac{-1}{2} Z^2 dZ$$

Por vía de consecuencia,

$$(5) \quad Z = f^{-1}(Z_i)$$

$$(6) \quad Z = X\beta + \varepsilon$$

En el modelo probit, el impacto de un cambio unitario en las variables explicativas no es igual al β correspondiente como en el caso del modelo logit. En este caso, el impacto ha cambiado con el nivel de la variable explicativa, según la siguiente ecuación:

$$(7) \quad \frac{dp}{dx_i} = f(X\beta + \varepsilon) * b_i$$

El coeficiente b_i denota las pendientes o coeficientes correspondientes a las variables explicativas (es decir, excluyendo el intercepto).

III. Datos y variables:

Para la realización de esta investigación, se utilizan datos de la Encuesta de Evaluación de la Protección Social (EEPS) 2010, la cual contiene información sobre 6,791 hogares en condiciones de pobreza a nivel nacional. Los contenidos temáticos de la EEPS abarcan información sobre las características demográficas, educativas, laborales, y de salud de estos hogares, así como sobre la vivienda y condiciones generales de acceso a servicios, bienes del hogar, entre otros temas.

El universo de selección se definió a partir de hogares con información en el Sistema Único de Beneficiarios (SIUBEN), un sistema de información y registro de beneficiarios y potenciales

beneficiarios implementado por el Gabinete Social de la República Dominicana. Este sistema permite la clasificación socioeconómica de los hogares, y su priorización para la participación en programas sociales. Para los cuatro casos a estimar (deserción escolar nivel básico y medio, repitencia nivel básico y medio) se utilizarán modelos logit y probit.

Para la construcción de estas muestras se utilizaron aquellas observaciones en las cuales los estudiantes encuestados presentan respuesta afirmativa o negativa según el caso de estudio (deserción escolar en el nivel básico, deserción escolar en el nivel medio, repitencia escolar en el nivel básico, repitencia escolar en el nivel medio) para después agrupar estas con la variable folio común a las dos bases de datos (individual y hogares), la cual se refiere al número del hogar al que pertenece el o la estudiante en cuestión.

La especificación del modelo a estimar es la siguiente:

$$(8) \quad D_e = \beta_0 - \beta_1 A_t - \beta_2 A_s - \beta_3 C_e + \beta_4 E_d + \beta_5 E_{ib} - \beta_6 E_{jf} - \beta_7 E_{cj} - \beta_8 G_e + \beta_9 G_j \\ + \beta_{10} I_s - \beta_{11} I_e - \beta_{12} L_{im} - \beta_{13} M_t - \beta_{14} O_l - \beta_{15} O_u + \beta_{16} T_{tr} - \beta_{17} T_{ta} + \beta_{18} T_b + u$$

Donde:

D_e = Desempeño escolar

A_t = Ayuda en la tarea

A_s = Asistencia

C_e = Centro educativo

E_d = Edad

E_{ib} = Edad de inicio en el nivel básico

E_{jf} = Nivel de educación del jefe del hogar

E_{cj} = Estado civil del jefe del hogar

G_e = Género del estudiante

G_j = Género del jefe del hogar

I_s = Zona de residencia

I_e = Programa de Incentivo a la Asistencia Escolar (ILAE)

L_{im} = Logaritmo del Ingreso mensual familiar por miembros

M_t = Medio de transporte

O_l = Obtención de libros

O_u = Obtención de útiles

T_{tr} = Tiempo de transportación

T_{ta} = Tiempo para hacer la tarea

T_b = Trabajo

Las variables dependientes Desempeño Escolar (D_e) que está compuesta por deserción y repitencia adopta cuatro formas: Deserción a Nivel Básico (Dnb); Deserción a Nivel Medio (Dnm); Repitencia a Nivel Básico (Rnb); y Repitencia a Nivel Medio (Rnm). Dichas variables se construyeron de la siguiente manera:

- La variable dependiente Dnb es una variable dummy que se refiere a la deserción escolar en el nivel básico, y que toma el valor de 1 cuando el ó la estudiante ha desertado y 0 en el caso contrario.
- La variable dependiente Dnm es una variable dummy que trata sobre la deserción escolar en el nivel medio, toma el valor de 1 cuando el ó la estudiante ha desertado y 0 en el caso contrario.

- La variable dependiente R_{nb} es una variable dummy que trata sobre la repitencia escolar en el nivel básico. Toma el valor de 1 cuando el ó la estudiante ha repetido y 0 en el caso contrario.
- La variable dependiente R_{nm} es una variable dummy que trata sobre la repitencia escolar en el nivel medio. Toma el valor de 1 cuando el ó la estudiante ha repetido y 0 en el caso contrario.

En cuanto a las variables independientes, se ha elaborado la Tabla 1 que se presenta más abajo y que contiene información sobre el símbolo que representa a cada variable, el signo esperado del coeficiente que relaciona cada variable con la variable dependiente y el código de la pregunta(s) realizada(s) en la encuesta que sirve como fuente de datos.

Tabla 1. Variables independientes a utilizar en los modelos

Nombre de la variable	Símbolo	Signo esperado	Código(s) en EEPS
1. Ayuda en la tarea	A_t	-	p3c30
2. Asistencia	A_s	-	p3c27
3. Centro educativo	C_e	-	p3c7
4. Edad del estudiante	E_d	+	p2a3
5. Edad de inicio en el nivel básico	E_{ib}	+	p3c2
6. Educación del jefe del hogar	E_{jf}	-	p2a5, p3a6a
7. Estado civil del jefe del hogar	E_{cj}	-	p2a5, p2a8
8. Género del estudiante	G_e	-	p2a2
9. Género del jefe del hogar	G_j	+	p2a2, p2a5
10. Zona de residencia	I_s	+	id_seccion
11. Log ingreso familiar mensual	L_{im}	-	p6a1a
12. ILAE	I_e	-	p13021
13. Medio de transporte	M_t	-	p3c9c
14. Obtención de libros	O_l	-	p3c13
15. Obtención de útiles	O_u	-	p3c15
16. Tiempo de transportación	T_{tr}	+	p3c8
17. Tiempo dedicado a hacer tarea	T_{ta}	-	p3c29
18. Trabajo	T_b	+	p3c31

Fuente: Elaboración propia basados en la encuesta.

IV. Resultados e interpretación

4.1 Deserción escolar nivel básico

La presente sección realiza un análisis de los resultados de los modelos econométricos estimados con el objetivo de contrastar las hipótesis planteadas con la evidencia empírica. En este apartado, se presenta el efecto de las diferentes variables en la probabilidad de que un individuo abandone la escuela o repita un grado. La Tabla 2 muestra los resultados de las variables independientes para el caso de la deserción escolar en el nivel básico.

Tabla 2. Deserción escolar nivel básico

Variables Independientes	Logit		Probit	
	Coefficiente	Estad-T	Coefficiente	Estad-T
Constante	-1.69*	-1.95	-0.90**	-2.11
<u>Características individuales:</u>				
Ayuda en la tarea (A_t)	0.26**	2.20	0.12**	2.11
Asistencia escolar (A_s)	-0.86***	-10.57	-0.41***	-10.11
Edad del estudiante (E_d)	0.10***	7.37	0.05***	7.24
Edad de inicio en el nivel básico (E_{ib})	0.09*	1.71	0.04	1.63
Género del estudiante (G_e)	-0.19*	-1.86	-0.09*	-1.75
Medio de transporte (M_t)	-0.51***	-3.34	-0.24***	-3.39
Tiempo dedicado a hacer la tarea (T_{ta})	-0.00	-1.04	-0.00	-1.13
Tiempo de transportación (T_{tr})	-0.00	-1.02	-0.00	-0.85
Trabajo (T_b)	0.45***	2.90	0.23***	2.81
<u>Aspectos institucionales:</u>				
Obtención de libros (O_l)	0.57	1.28	0.20	0.99
ILAE (I_e)	-0.11	-1.05	-0.06	-1.18
Obtención de útiles (O_u)	-0.37	-0.80	-0.24	-0.98
Centro educativo (C_e)	0.12	1.28	0.05	1.17
<u>Aspectos del hogar:</u>				
Educación del jefe del hogar (E_{jf})	0.02	0.57	0.01	0.49
Estado civil del jefe del hogar (E_{cj})	0.02	0.14	0.00	0.03
Género del jefe del hogar (G_j)	-0.02	-0.21	-0.02	-0.41
Zona de residencia (I_s)	0.01	0.05	-0.00	-0.01
Log ingreso familiar mensual (L_{im})	-0.14**	-2.37	-0.06**	-2.25

Fuente: Elaboración propia basado en la encuesta. Niveles de significancia: *** = 1%, ** = 5%, * = 10%

A continuación se realiza un análisis detallado sobre los resultados de las estimaciones para verificar si son consecuentes con las hipótesis planteadas al principio del trabajo.

En lo concerniente a la parte de características individuales de los modelos logit y probit de deserción escolar en el nivel básico, se tiene que la variable Ayuda en la tarea (A_t) aparenta estar correlacionada positivamente con la deserción escolar en el nivel básico, por lo cual se rechaza la hipótesis de que los

estudiantes que reciben ayuda en la tarea tienden a registrar mayores probabilidades de permanecer en la escuela.

La variable Asistencia escolar (As) se encuentra correlacionada negativamente con la deserción escolar en el nivel básico, lo cual va acorde con lo esperado en la evidencia convencional, es decir, no se rechaza la hipótesis de que los estudiantes que asisten a la escuela regularmente tendrán menores probabilidades de desertar.

En cuanto a la variable Edad del Estudiante (Ed), se tiene que esta se encuentra correlacionada positivamente con la deserción escolar en el nivel básico, por lo cual no se rechaza la hipótesis de que los estudiantes con mayor edad que la correspondiente al grado que están cursando tendrán mayores probabilidades de desertar en comparación con sus condiscípulos, lo cual expone el problema de la sobre edad que padecen muchos estudiantes en el caso dominicano.

La variable Edad de inicio en el nivel básico (Eib), al menos en el caso del modelo logit, es significativa y se encuentra correlacionada positivamente con la deserción escolar en el nivel básico, es decir, no se rechaza la hipótesis de que los estudiantes con edad de inicio tardía en el nivel básico tendrán mayores probabilidades de desertar la escuela.

El signo y la significancia estadística en la variable Género del estudiante (Ge) indica que el riesgo de desertar disminuye si el estudiante es de género femenino, lo cual concuerda con lo visto en el estudio de Santos (2009) para el caso chileno. En lo concerniente a la variable Medio de transporte (Mt), se tiene que no se rechaza la hipótesis de que los estudiantes con acceso a medio de transporte tendrán menores probabilidades de desertar en comparación a aquellos que tienen que trasladarse a pie a su centro de estudio.

La variable Trabajo (Tb) tiene el símbolo esperado y es estadísticamente significativa, por lo cual no se rechaza la hipótesis de que los estudiantes que trabajan además de estudiar tendrán mayores probabilidades de desertar, lo cual concuerda con lo expuesto por Guzman y Cruz (2009) acerca de la relación negativa entre el trabajo infantil y el desempeño escolar para el caso dominicano. En este estudio se desarrolló un análisis de las causas y consecuencias de la deserción escolar por parte de niños, niñas y adolescentes, además de las políticas actuales dirigidas a combatir este flagelo social.

En lo concerniente a los aspectos del hogar, se tiene que el logaritmo del ingreso familiar mensual (Lim) es la única variable estadísticamente significativa, tiene el signo indicado, lo cual quiere decir que a mayor ingreso mensual en el hogar, menores serán las probabilidades de desertar del estudiante en el nivel básico.

4.2 Deserción escolar nivel medio

En esta sección se analizan los resultados de la Tabla 3, la cual muestra los resultados para el caso de la deserción escolar en el nivel medio:

Tabla 3. Deserción escolar nivel medio

Variable Dependiente: Deserción Escolar Nivel Medio				
Variables Independientes	Logit		Probit	
	Coefficiente	Estad-T	Coefficiente	Estad-T
Constante	0.79	1.01	0.45	0.97
<u>Características individuales:</u>				
Ayuda en la tarea (A_t)	0.06	0.70	0.04	0.72
Asistencia escolar (A_s)	-0.42***	-5.99	-0.25***	-6.01
Edad del estudiante (E_d)	-0.03	-1.49	-0.02	-1.45
Edad de inicio en el nivel básico (E_{ib})	0.07*	1.66	0.04*	1.75
Género del estudiante (G_e)	-0.69***	-8.08	-0.41***	-8.12
Medio de transporte (M_t)	-0.17*	-1.71	-0.10*	-1.73
Tiempo dedicado a hacer la tarea (T_{ta})	-0.00***	-2.83	-0.00***	-2.81
Tiempo de transportación (T_{tr})	-0.01***	-3.49	-0.00***	-3.47
Trabajo (T_b)	0.27**	2.32	0.16**	2.31
<u>Aspectos institucionales:</u>				
Obtención de libros (O_l)	0.54	1.49	0.30	1.50
ILAE (I_e)	0.21**	2.44	0.13**	2.44
Obtención de útiles (O_u)	-0.90**	-2.15	-0.54**	-2.13
Centro educativo (C_e)	-0.26***	-2.89	-0.15***	-2.93
<u>Aspectos del hogar:</u>				
Educación del jefe del hogar (E_{jf})	-0.02	-0.64	-0.01	-0.65
Estado civil del jefe del hogar (E_{cj})	-0.07	-0.64	-0.04	-0.65
Género del jefe del hogar (G_j)	0.11	1.14	0.07	1.12
Zona de residencia (I_s)	0.25***	2.69	0.15***	2.73
Log ingreso familiar mensual (L_{im})	0.08*	1.69	0.05*	1.70

Fuente: Elaboración propia basados en la encuesta. Niveles de significancia: *** = 1%, ** = 5%, * = 10%

En lo referente a la deserción escolar en el nivel medio, se tiene en la parte de características individuales que la variable Asistencia (As) tiene el símbolo esperado y es estadísticamente significativa, por lo cual no se rechaza la hipótesis de que a mayor asistencia del estudiante a las clases, menores serán las probabilidades de este de desertar la escuela en el nivel medio. En el caso de la Edad de inicio en el nivel básico (Eib), no se rechaza la hipótesis de que mientras mayor sea la edad en la cual el estudiante se inscriba en el nivel básico, mayores serán sus probabilidades de desertar la escuela en el nivel medio.

En cuanto a Género del estudiante (Ge), la evidencia sugiere que si el género del estudiante es femenino, menores serán las probabilidades de desertar de la escuela, lo cual podría estar relacionado con una mayor inserción de los estudiantes de género masculino en el mercado laboral a partir de la adolescencia, lo cual se refleja en la variable Trabajo (Tb) del estudiante en estos modelos, que tiene el símbolo esperado y es estadísticamente significativa, por lo cual los estudiantes que trabajan para ayudar en los gastos del hogar tendrían mayores probabilidades de desertar en el nivel medio, es decir, no se rechaza la hipótesis de que los estudiantes con trabajo se verán con mayores

probabilidades de desertar de la escuela en el nivel medio en comparación con sus compañeros que no trabajan.

La variable Medio de transporte (Mt) tiene el símbolo esperado y es estadísticamente significativa, indicando que no se rechaza la hipótesis de que los estudiantes que poseen medio de transporte tendrán menores probabilidades de desertar la escuela en el nivel medio, ya que la posesión de vehículo le permitirá prevenir la pérdida de clases por motivos de tardanza. Es necesario señalar la coincidencia entre los resultados de estas variables individuales con los que se presentan en el caso del nivel básico, especialmente en lo referente a la asistencia escolar y al trabajo, mostrando en esto último concordancia con el estudio de Guzmán y Cruz (2009) descrito anteriormente.

En la variable Tiempo dedicado a hacer la tarea (Tta), se tiene que los estudiantes que dedican más tiempo a los estudios son aquellos que menores probabilidades tendrán de desertar de la escuela. En lo concerniente a Tiempo de transportación (Ttr), no obstante ser estadísticamente significativa, no muestra el signo esperado.

En los aspectos institucionales, se tiene que la variable Obtención de útiles (Ou) muestra el símbolo esperado y es estadísticamente significativa, por lo cual no se rechaza la hipótesis de que los estudiantes que obtienen útiles tendrán menores probabilidades de desertar de la escuela en el nivel medio. La variable ILAE (Ie), no obstante ser estadísticamente significativa, no posee el símbolo esperado, lo cual aparenta indicar que los estudiantes dentro de este programa están en riesgo de desertar la escuela en el nivel medio a pesar del mismo.

La variable Centro educativo (Ce) indica que los estudiantes que asisten a escuelas privadas tendrán menores probabilidades de desertar de la escuela en el nivel medio que aquellos que asisten a escuelas públicas, lo cual concuerda con lo que se ve convencionalmente en la realidad actual dominicana.

En lo referente a las variables que tratan sobre el hogar del estudiante se tiene que la variable Zona de residencia (Is) indica que no se rechaza la hipótesis de que los estudiantes que viven en zonas rurales tendrán mayores probabilidades de desertar la escuela en el nivel medio. Mientras que la variable Logaritmo del ingreso familiar mensual (Lim) opera en sentido inverso a lo visto en el caso del nivel básico, es decir, los estudiantes con mayor ingreso en sus hogares aparentan tener mayores probabilidades de desertar de la escuela en el nivel medio, lo cual puede estar relacionado con la inserción en el mundo laboral por parte de los estudiantes en esta etapa de sus vidas, sobre todo si pertenecen al género masculino.

4.3 Repitencia escolar nivel básico

La Tabla 4 que se presenta más abajo detalla los resultados para la repitencia escolar de los estudiantes del nivel básico.

Tabla 4. Repitencia escolar nivel básico

Variable Dependiente: Repitencia Escolar Nivel Básico				
Variables Independientes	Logit		Probit	
	Coficiente	Estad-T	Coficiente	Estad-T
Constante	-7.38***	-4.08	-3.55***	-4.01
<u>Características individuales:</u>				
Ayuda en la tarea (A_t)	-0.2	-0.84	-0.11	-0.94
Asistencia escolar (A_s)	-1.01***	-5.95	-0.45***	-5.44
Edad del estudiante (E_d)	0.24***	5.05	0.11***	4.88
Edad de inicio en el nivel básico (E_{ib})	0.19*	1.79	0.10*	1.88
Género del estudiante (G_e)	-0.09	-0.39	-0.07	-0.6
Medio de transporte (M_t)	-0.53*	-1.89	-0.24*	-1.84
Tiempo dedicado a hacer la tarea (T_{ta})	-0.00	-1.23	-0.00	-1.24
Tiempo de transportación (T_{tr})	-0.00	-0.67	-0.00	-0.65
Trabajo (T_b)	1.12***	4.45	0.55***	4.45
<u>Aspectos institucionales:</u>				
Obtención de libros (O_l)	1.18	1.45	0.52	1.38
ILAE (I_e)	-0.18	-0.67	-0.09	-0.76
Obtención de útiles (O_u)	-0.68	-0.9	-0.35	-0.84
Centro educativo (C_e)	-0.17	-0.83	-0.09	-0.89
<u>Aspectos del hogar:</u>				
Educación del jefe del hogar (E_{jf})	0.04	0.45	0.01	0.33
Estado civil del jefe del hogar (E_{cj})	-0.36	-1.4	-0.19	-1.62
Género del jefe del hogar (G_j)	0.63***	2.59	0.28**	2.49
Zona de residencia (I_s)	0.28	1.11	0.16	1.29
Log ingreso familiar mensual (L_{im})	0.12	0.95	0.05	0.83

Fuente: Elaboración propia basados en la encuesta. Niveles de significancia: *** = 1%, ** = 5%, * = 10%

Al igual que en caso de deserción la razón del logaritmo de verosimilitud permite rechazar la hipótesis de no significancia estadística del conjunto de variables explicativas. Razón por la cual se puede afirmar que el conjunto de variables explicativas tienen incidencia en la repitencia. Los efectos reparados serán analizados a continuación.

En lo concerniente a las características individuales para este caso, se tiene que solamente las variables Asistencia escolar (A_s), Edad del estudiante (E_d), Edad de inicio en el nivel básico (E_{ib}), Medio de transporte (M_t) y Trabajo (T_b) son estadísticamente significativas. En el caso de Asistencia (A_s), esto indicaría que no se rechaza la hipótesis de que a mayor nivel de asistencia de los estudiantes, menores serán sus probabilidades de repetir la escuela en el nivel básico.

En cuanto a la variable Edad del estudiante (E_d), no se rechaza la hipótesis de que a mayor edad de estos mayores serán sus probabilidades de repetir. La significancia de la variable Edad de inicio en el nivel básico (E_{ib}) indica que los estudiantes que se inscriben en el nivel básico a una edad avanzada en relación con la edad oficial de inicio de este ciclo educativo tendrán mayores probabilidades de

repetir la escuela que aquellos de sus compañeros que se inscribieron en la edad indicada o más temprano.

En la variable Trabajo (T_b), se tiene que los estudiantes que trabajan para ayudar a sostener los gastos del hogar tendrán mayores probabilidades de repetir la escuela en el nivel básico.

En cuanto a los aspectos del hogar, se tiene que la variable Género del jefe del hogar (G_j) sería la única que sería estadísticamente significativa, y su símbolo indicaría que aquellos estudiantes provenientes de hogares en los cuales el jefe del hogar es mujer tienen mayores probabilidades de repetir la escuela en el nivel básico.

4.4 Repitencia escolar nivel medio

A continuación, la Tabla 5 presenta los resultados correspondientes a la repitencia escolar para los estudiantes del nivel medio.

Tabla 5. Repitencia escolar nivel medio

Variable Dependiente: Repitencia Escolar Nivel Medio				
Variables Independientes	Logit		Probit	
	Coefficiente	Estad-T	Coefficiente	Estad-T
Constante	-2.57	-1.31	-1.63	-1.62
<u>Características individuales:</u>				
Ayuda en la tarea (A_t)	-0.16	-0.87	-0.09	-0.97
Asistencia escolar (A_s)	-0.89***	-6.28	-0.43***	-5.96
Edad del estudiante (E_d)	0.05	1.19	0.03	1.41
Edad de inicio en el nivel básico (E_{ib})	-0.06	-0.59	-0.03	-0.58
Género del estudiante (G_e)	-0.71***	-3.69	-0.34***	-3.58
Medio de transporte (M_t)	-0.30	-1.38	-0.17	-1.52
Tiempo dedicado a hacer la tarea (T_{ta})	-0.01***	-2.73	-0.00***	-2.77
Tiempo de transportación (T_{tr})	-0.01*	-1.65	-0.00*	-1.66
Trabajo (T_b)	0.29	1.24	0.16	1.29
<u>Aspectos institucionales:</u>				
Obtención de libros (O_l)	2.06*	1.96	1.09**	2.11
ILAE (I_e)	-0.37*	-1.79	-0.19*	-1.83
Obtención de útiles (O_u)	0.65	0.58	0.42	0.69
Centro educativo (C_e)	0.05	0.29	0.02	0.28
<u>Aspectos del hogar:</u>				
Educación del jefe del hogar (E_{jf})	-0.01	-0.10	-0.01	-0.34
Estado civil del jefe del hogar (E_{cj})	0.04	0.15	0.02	0.13
Género del jefe del hogar (G_j)	0.23	1.06	0.13	1.16
Zona de residencia (I_s)	-0.16	-0.72	-0.07	-0.60
Log ingreso familiar mensual (L_{im})	0.03	0.25	0.00	0.07

Fuente: Elaboración propia basados en la encuesta. Niveles de significancia: *** = 1%, ** = 5%, * = 10%.

Para el caso de la repitencia escolar en el nivel medio se tiene en la parte de las características individuales que las variables Asistencia escolar (As), Género del estudiante (Ge), Tiempo dedicado a hacer la tarea (Tta) y Tiempo de transportación (Ttr) son las únicas variables estadísticamente significativas.

En los aspectos institucionales, se tiene que, si bien la variable Obtención de libros (Ol) fue estadísticamente significativa, esta no muestra el símbolo esperado por la evidencia convencional, por lo cual este resultado debe ser tomado con cautela, ya que indicaría que los estudiantes están repitiendo la escuela en el nivel medio a pesar de estos contar los libros requeridos en el grado que están cursando. La variable ILAE (Ie), mostró el signo esperado y es estadísticamente significativa, lo cual aparenta indicar que no se rechaza la hipótesis de que los estudiantes beneficiarios de este programa muestran menores probabilidades de repetir la escuela para el nivel medio en comparación con aquellos que no son beneficiarios del mismo.

V. Conclusiones

Entre las principales conclusiones de la investigación se encuentran:

- El trabajo infantil en niños, niñas y adolescentes impacta negativamente en la mayoría de los casos, con excepción de la repitencia escolar en el nivel medio. Se confirma relación negativa entre aprendizaje y trabajo infantil expuesta en la literatura a nivel nacional e internacional.
- En el caso de la deserción escolar en el nivel básico, ingreso familiar tiene el símbolo y significancia esperados, por lo cual los estudiantes provenientes de familias con mayor nivel de ingreso tienen menor probabilidad de desertar de la escuela que sus pares de menores ingresos.
- Persiste desigualdad urbana/rural para el caso de la deserción escolar en el nivel medio, en el sentido de que los estudiantes que viven en zonas rurales tendrán mayores probabilidades de desertar.
- La obtención de útiles aparenta ser importante para prevenir la deserción escolar en el nivel medio. Asimismo, los estudiantes de centros educativos privados tienen mayores probabilidades de aprobar la escuela en comparación con sus pares públicos.
- En la repitencia escolar en el nivel básico, el género del jefe del hogar muestra los resultados esperados, por lo cual los estudiantes cuyo jefe de hogar es femenino tienen mayores probabilidades de repetir la escuela, lo cual pone de manifiesto uno de los tantos problemas que afrontan los hogares monoparentales (usualmente el jefe del hogar siendo en este caso la madre soltera del o la estudiante en cuestión).
- El género del estudiante no es significativo en el caso de la repitencia escolar para el nivel básico, por lo que, a diferencia de lo ocurrido en los demás casos, no hay diferencias en la probabilidad de repitencia entre niños y niñas para este nivel educativo.

- A diferencia de lo que sucede en la repitencia escolar en el nivel básico, para el nivel medio los aspectos institucionales aparentan tener importancia. El programa ILAE muestra el signo y significancia esperadas, por lo cual los estudiantes que disfrutaron del programa muestran menores probabilidades de repetir la escuela que aquellos que no son beneficiarios. No obstante, la obtención de libros tiene signo contrario al esperado, por lo cual la obtención de estos no impide que los estudiantes repitan la escuela en el nivel medio.

A nivel general, las características individuales prevalecieron en todos los casos estudiados, siendo notables en los casos de estudio para el nivel básico, mientras que las institucionales solamente mostraron importancia para el análisis de la deserción y repitencia escolares del nivel medio.

Referencias

Gujarati, D.N. (2004). "Econometría". Cuarta Edición. McGraw-Hill.

Guzmán, R. y Cruz, C. (2009). "Niños, Niñas y Adolescentes fuera del sistema educativo". Proyecto Un Foro para la Mejora de la Educación Dominicana, Santiago, República Dominicana.

Lizardo, M.M. y Guzmán, R.M. (1999). "Niveles de Escolaridad y sus Factores Determinantes: Una Cuantificación Econométrica". Revista Ciencia y Sociedad Vol. XXIV, Numero 2.

Maradona, G. y Calderón, M.I. (2004). "Una aplicación del enfoque de la función de producción en educación". Revista de Economía y Estadística Vol. XLII, Instituto de Economía y Finanzas - Facultad de Ciencias Económicas, Universidad Nacional de Córdoba - Argentina.

Santos, H. (2009). "Dinámica de la Deserción Escolar en Chile". Documento de Trabajo CPCE N° 3.

Anexos

El Programa Solidaridad

El Programa Solidaridad de la República Dominicana coordina un conjunto de estrategias de transferencias condicionadas a los hogares, como parte de la política de desarrollo social dominicana. El objetivo del programa es "contribuir a romper la transmisión intergeneracional de las causas que generan o arraigan la pobreza, mejorando la inversión que los hogares pobres realizan en educación, salud y alimentación, potenciando la acumulación del capital humano de la descendencia".

Solidaridad comenzó a operar en 2004, y ha seguido un proceso de crecimiento tanto cuantitativo como cualitativo importante. Por un lado, el mayor presupuesto ha permitido ampliar de forma significativa la cobertura del Programa, a la vez que se han incrementado el tipo de transferencias que se entregan, y que actualmente comprenden las transferencias i. Comer es Primero, ii. Incentivo a la Asistencia Escolar (ILAE), y iii. el Bono Gas.

Como parte de la evaluación del Programa, se implementó en 2010 la Encuesta de Evaluación de la Protección Social (EEPS 2010), para recolectar información sobre hogares dominicanos en condiciones de pobreza.

Fuente: Ficha técnica de la Encuesta de Evaluación de la Protección Social 2010, Banco Interamericano de Desarrollo.