

# **Un análisis de las diferencias rurales y urbanas en el rendimiento educativo de los estudiantes colombianos a partir de los microdatos de PISA \***

**Raul Ramos<sup>1</sup>, Juan Carlos Duque<sup>2</sup>, Sandra Nieto<sup>1</sup>**

<sup>1</sup>AQR-IREA, Universitat de Barcelona

<sup>2</sup> Research in Spatial Economics (RiSE-group), EAFIT University.

**Resumen:** Aunque hay cientos de trabajos que utilizan los microdatos de PISA para analizar los determinantes de los resultados educativos, sólo unos pocos de ellos han considerado la importancia de la geografía. En este artículo, nos centramos en el análisis de las diferencias en los resultados educativos entre los alumnos de las escuelas rurales y urbanas. Para ello, utilizamos los microdatos de las olas de 2006 y 2009 de PISA para Colombia. El caso colombiano es particularmente interesante desde esta perspectiva debido a los cambios estructurales sufridos por este país durante los últimos años, tanto en términos de estabilidad política como por las reformas educativas realizadas. El análisis descriptivo de los datos muestra que los resultados educativos de los estudiantes urbanos son más altos que los rurales. Sin embargo, y con el fin de identificar los factores que explican este diferencial, utilizamos la descomposición de Oaxaca-Blinder y, a continuación, se pretende explotar la variación temporal de los datos utilizando la metodología propuesta por Juhn-Murphy-Pierce. Los resultados muestran que la mayor parte del diferencial está relacionado con las características de la familia y no tanto con las características de la escuela. Desde una perspectiva de política educativa, la evidencia obtenida refuerza la idoneidad de todas aquellas medidas dirigidas a mejorar las condiciones de la familia y no tanto a la discriminación positiva de las escuelas rurales como medio para mejorar el rendimiento educativo.

**Palabras clave:** resultados educativos, diferencias entre zonas urbanas y rurales, métodos de descomposición.

---

\* Sandra Nieto y Raúl Ramos agradecen el apoyo recibido del Ministerio de Ciencia e Innovación de España a través del proyecto ECO2010-16006.

## **Decomposing rural-urban differences in student achievement in Colombian using PISA microdata**

Although there are hundreds of works using PISA microdata to analyse the determinants of educational outcomes, only a few of them have considered the relevance of geography. In this paper, we focus on the analysis of differences in educational outcomes between students in rural and urban schools. We use microdata from the 2006 and 2009 PISA waves for Colombia. The Colombian case is particularly interesting from this perspective due to the structural changes suffered by this country during the last years, both in terms of political stability and educational reforms. The descriptive analysis of the data shows that the educational outcomes of urban students are higher than the rural ones. In order to identify the factors behind this differential, we use the Oaxaca-Blinder decomposition and, next, we exploit the time variation of the data using the methodology proposed by Juhn-Murphy-Pierce. The results show that most part of the differential is related to family characteristics more than to school characteristics. From a policy perspective, the obtained evidence support actions addressed to improve family conditions and not so much to positive discrimination of rural schools.

Keywords: educational outcomes, urban-rural differences, decomposition methods.

JEL codes: J24, I25, R58

## 1 Introducción

Una de las ramas de la economía de la educación que ha despertado un mayor interés de los investigadores en las últimas décadas ha sido el análisis de los factores que influyen sobre el rendimiento educativo de los estudiantes. La mayor disponibilidad de información estadística ha facilitado el análisis de este tema en un mayor número de países y, en concreto, para un amplio número de países en vías de desarrollo. Un aspecto importante en este contexto es el análisis de las posibles diferencias existentes entre zonas rurales y urbanas y de sus factores explicativos.

En los distintos estudios realizados se han identificado numerosos factores que determinan el rendimiento educativo de los estudiantes y que, por su naturaleza, pueden agruparse en tres grupos. En el primer grupo se encuentran las características individuales, en las que destacan las variables relativas a la nacionalidad y al idioma principal del alumno. En este sentido, se encuentra que los inmigrantes tienen un rendimiento educativo inferior a los nativos (Meunier 2011, Chiswick y DebBurman 2004) y se argumenta que este efecto está relacionado con los diferentes entornos familiares de cada uno de los colectivos analizados (Ammermueller, 2007a y Entorf y Lauk, 2008). Respecto al idioma, existe evidencia de que los inmigrantes mejoran sus resultados académicos cuando hablan el idioma oficial en su entorno familiar (Entorf y Minoiu, 2005).

Un segundo grupo de variables hace referencia a los antecedentes familiares. Coleman et al. (1966) fue uno de los precursores en mostrar la incidencia de variables relativas a la familia en el rendimiento educativo de los alumnos. Algunos trabajos como por ejemplo Haveman y Wolfe (1995) y Feinstein y Symons (1999) explican que este tipo de variables son las que tienen una mayor influencia sobre el rendimiento educativo de los alumnos. Se encuentra que los estudiantes cuyos progenitores disponen de un nivel de estudios elevado, obtienen un rendimiento superior que los estudiantes con padres con un menor nivel educativo (Häkkinen et al. 2003, Woßmann 2003). Además, el nivel socio-económico familiar también guarda relación con el rendimiento educativo del estudiante, siendo éste superior cuanto mayor es el nivel social y económico de los progenitores. La transmisión genética de habilidades cognitivas es uno de los argumentos más utilizados para explicar el mayor rendimiento de los estudiantes cuyos progenitores tienen un nivel educativo elevado. Por otro lado, la dotación de un buen entorno cultural y un entorno familiar estable también contribuye a un mayor rendimiento escolar de los jóvenes. De hecho, suele existir una cierta correlación positiva entre el nivel educativo de los padres y el nivel socio-económico y cultural de la familia.

Por último, el tercer grupo de variables está relacionado con distintas características de la escuela a la que asisten los alumnos como, por ejemplo, su localización urbana o rural, el tipo de escuela –pública o privada-, el ratio alumno-profesor, su tamaño o distintas características de los compañeros de clase (los conocidos como *peer effects*).

Los resultados suelen coincidir en la influencia de las características individuales y del entorno familiar sobre el rendimiento educativo de los estudiantes. Este consenso no es tan amplio en los trabajos que analizan la influencia de las variables relativas a las escuelas a las que asisten los alumnos. Estudios realizados por Heyneman y Loxley (1983), Harbison y Hanushek (1992), Fuller y Clarke (1994), Gamoran y Long (2006) y Behrman (2010) encontraron que las características de la escuela tienen un importante efecto en el rendimiento académico en países en vía de desarrollo. Estudios como Coleman y Hoffer (1987), Hanushek (1986), Stevans y Sessions (2000) o, Vandenberghe y Robin (2004) y Opdenakker y Van Damme (2006) entre otros, encuentran que los alumnos obtienen un rendimiento superior en escuelas privadas que públicas. No obstante, otros estudios como, por ejemplo, Noell (1982), Sander (1996), Fertig (2003), Somers et al (2004) o Smith y Naylor (2005), muestran un efecto nulo de la titularidad de la escuela sobre el rendimiento de los alumnos. Por otro lado, el efecto del tamaño de la escuela sobre el rendimiento de los alumnos tampoco está claro. Mientras que Barnett et al. (2002) encuentran una relación positiva entre el tamaño de la escuela y el rendimiento educativo de los estudiantes, Hanushek y Luque (2003) no observan un efecto significativo de dicha variable en gran parte de los países analizados. Tampoco son concluyentes los resultados relativos a la variable número de alumnos por profesor. Arum (2000) y Krueger (2003) muestra que los alumnos obtienen un mejor rendimiento en clases reducidas, mientras que Hanushek (2003) y Rivkin et al. (2005) no encuentran un efecto estadísticamente significativo de dicha variables sobre el rendimiento educativo de los alumnos. En cambio, la mayoría de trabajos, como Coleman et al. (1966), Henderson et al. (1978), Caldas y Bankston (1997), Lee et al (1997), Feinstein y Symons (1999) y Hanushek et al. (2003) entre otros, sí coinciden en la importancia de las características de los compañeros sobre el rendimiento educativo de los estudiantes.

Son pocos los estudios que han abordado el tema del efecto de la localización rural o urbana de las escuelas en el rendimiento de los alumnos; estos comenzaron en Estados Unidos a mediados de los ochenta y hasta el momento no parece haber un consenso sobre la significatividad de esta característica; así, Edington y Martellaro (1984) y Ward y Murray (1985) no encuentran diferencias significativas en el rendimiento de los estudiantes entre escuelas urbanas y rurales del estado de Nuevo México; similares resultados obtiene Monk y Haller (1986) para el estado de New York. Por el contrario, Kleinfeld et al. (1985), para Alaska, y Blackwell y McLaughlin (1999), para Estados Unidos, si encuentran que la variable de localización rural/urbana es significativa a la hora de explicar

el rendimiento. El debate sobre el impacto de esta variable gira en torno a la posibilidad de que las diferencias entre los rendimientos de los estudiantes de escuelas rurales y urbanas no se deba a la localización *per se*, sino al hecho de que las características de los estudiantes, sus familias y las escuelas son diferentes en estos dos grupos. Los estudiantes de las zonas rurales suelen formar parte de familias con pocos recursos económicos, sus padres tienen bajo nivel de educación y las escuelas a las que asisten cuentan con peores dotaciones y generalmente son más pequeñas que las escuelas urbanas. Estudios como Hannaway y Talbert (1993) y Young (1998) sostienen que, más que la variable de localización, son estas diferencias entre las características de las zonas urbanas y rurales las que explican la mayor parte de la diferencia entre los rendimientos de los estudiantes de las escuelas rurales y urbanas. La pregunta es pues, en qué casos (regiones o países), la variable de localización continúa siendo significativa cuando se estudia en conjunto con las variables de contexto.

Durante la última década, la atención se dirigió hacia algunos países suramericanos, esto debido a la mayor disponibilidad de datos y a la importancia del sector rural en la región. La Tabla 1 presenta un resumen de los estudios realizados. Con excepción de los resultados obtenidos por Woßmann (2010), quien no encuentra diferencias significativas entre el rendimiento de los estudiantes de escuelas rurales y urbanas en Argentina, y Santos (2007), quien, para el mismo país, no encuentra diferencias entre dichos rendimientos para las pruebas de lectura; los demás estudios realizados confirman la significatividad de la variable de localización.

TABLA 1

Hasta donde llega nuestro conocimiento, para el caso colombiano sólo se ha realizado un estudio que aborde el tema de las diferencias rural-urbano en el desempeño estudiantil (Woßmann, 2010) a partir de los resultados obtenidos en el test del *Progress in International Reading Literacy Study* en 2001. En dicho estudio se encuentra que los estudiantes de poblaciones con más de tres mil habitantes superan en más de 26 puntos los resultados obtenidos por los estudiantes de zonas rurales<sup>1</sup>.

El informe PISA elaborado por la OCDE con la información recopilada en 2009 analiza parcialmente la importancia de la localización de la escuela para explicar las diferencias en los resultados obtenidos en la prueba de lectura después de controlar por las características socioeconómicas de las familias de los estudiantes. Dichos resultados se recogen en la tabla II.2.4 de dicho informe (OCDE, 2010). Tal y

---

<sup>1</sup> Otros estudios sobre rendimiento académico realizados para Colombia indican que los principales factores que explican el rendimiento académico son el nivel socioeconómico y los recursos de la institución (Piñeros y Rodríguez, 1998; Gaviria y Barrientos, 2001a y Rangel y Lleras, 2010). El nivel de educación de los padres también tiene un impacto significativo en el rendimiento de los estudiantes (Gaviria y Barrientos 2001b). Estos estudios utilizan como principal fuente de información los resultados de las pruebas de ICFES presentadas por todos los estudiantes de último año de secundaria.

como se puede observar en dicha tabla, mientras que para el promedio de la OCDE las diferencias entre la puntuación obtenida en las zonas menos pobladas y las más pobladas se sitúa alrededor del 4%, en Colombia esta diferencia supera el 8% (si bien es cierto que en otras economías como Panamá, Perú o Argentina las diferencias son aún más acusadas).

Sin embargo, un aspecto a tener en cuenta en el caso colombiano es que el estudio de las diferencias de rendimiento de los estudiantes en función de la localización rural/urbana de la escuela toma especial relevancia si se considera la enorme brecha existente en estos dos entornos como consecuencia de un conflicto armado que durante más de cuarenta años ha afectado sobre todo a las comunidades rurales. Desplazamientos forzados<sup>2</sup>; reclutamiento de menores por parte de las guerrillas<sup>3</sup>; enfrentamientos entre ejércitos, guerrillas y paramilitares; ataques a las infraestructuras escolares o utilización de estas como centros de operación o de reclutamiento; numerosos homicidios de docentes<sup>4</sup>; entre otros, han obstaculizado el normal desarrollo de la actividad educativa en las zonas rurales de Colombia. Estos antecedentes abren la posibilidad de que, aún descontando los efectos atribuibles a diferencias en los perfiles de los estudiantes, sus familias y las características de las escuelas, la variable localización sea un factor determinante a la hora de explicar las diferencias en los rendimientos de los estudiantes.

Con el propósito de analizar la posible existencia de diferencias en los resultados educativos de los estudiantes que asisten a escuelas ubicadas en zonas rurales y urbanas de Colombia, se utilizan los datos PISA de 2006 y 2009 para analizar los resultados obtenidos en las materias de matemáticas, ciencias y lectura. Para ello, se aplican metodologías de descomposición del diferencial rural-urbano a partir de la estimación de una función de producción en educación que incluye variables explicativas relacionadas tanto con las características de los estudiantes, sus familias y las escuelas a las que asisten. La aplicación de la descomposición propuesta por Oaxaca (1973) y Blinder (1973), que ha sido ampliamente utilizada en el marco de la economía laboral, por ejemplo, para intentar explicar las causas de las diferencias salariales entre hombres y mujeres, nos permitirá identificar cuales son las variables que contribuyen en mayor medida a explicar la existencia de diferencias en el rendimiento educativo entre zonas rurales y urbanas. Además, la extensión de dicha metodología propuesta por Juhn, Murphy y Pierce (1993) hace posible analizar cuales son los factores que explican los cambios en el diferencial entre zonas rurales y urbanas a lo largo del tiempo ofreciendo así el marco idóneo

---

<sup>2</sup> Según Ibáñez y Vélez (2008), un 29.1% de la población rural es víctima de desplazamiento forzado.

<sup>3</sup> De acuerdo con War Child (2007) uno de cada cuatro miembros de los grupos armados ilegales son menores de 15 años; muchos de ellos son reclutados en pueblos y escuelas rurales.

<sup>4</sup> Colombia, junto con Iraq, Nepal y Tailandia, aparece entre los países con las mayores cifras de homicidios de docentes (O'Malley, 2010).

para explotar la dimensión temporal de la información disponible<sup>5</sup>. Los resultados obtenidos muestran que la mayor parte del diferencial rural-urbano está relacionado con las características de la familia y no tanto con las características de la escuela. Desde una perspectiva de política educativa, la evidencia obtenida refuerza la idoneidad de todas aquellas medidas dirigidas a mejorar las condiciones de la familia y no tanto a la discriminación positiva de las escuelas rurales como medio para mejorar el rendimiento educativo.

El resto del trabajo se organiza de la siguiente manera. En el apartado 2 se presenta la base de datos y se definen las variables de interés para el estudio. A continuación, en el apartado 3, se describe la aproximación metodológica utilizada y se describen los resultados obtenidos. Por último, se resumen las principales conclusiones obtenidas y se plantean algunas líneas de mejora del presente trabajo.

## **2 Rendimiento educativo y características de los estudiantes, de su entorno familiar y escolar.**

La fuente de datos que se utiliza en este trabajo corresponde al *Programme for International Student Assessment (PISA)*, coordinado por la OCDE y que tiene como objetivo evaluar la formación de los estudiantes cuando llegan al final de la etapa de enseñanza obligatoria, a los 15 años, en las áreas de matemáticas, ciencias y lectura, facilitando además información relativa al propio alumno, a su entorno familiar y a su entorno educativo. Se trata de una encuesta de carácter trienal que actualmente proporciona datos correspondientes a las cuatro oleadas existentes: 2000, 2003, 2006 y 2009. El conjunto de países analizados en cada uno de los años ha ido creciendo a lo largo del tiempo hasta abarcar en 2009 a 65 países. Colombia es uno de los países incluidos en las olas más recientes. En concreto, existen datos para 2006 y 2009 que serán los utilizados en este estudio.

Tal y como se ha mencionado anteriormente, el principal objetivo de PISA es evaluar la formación de los estudiantes cuando llegan al final de la etapa de enseñanza obligatoria en las áreas de matemáticas, ciencias y lectura. Para ello, proporciona cinco valores plausibles para cada una de las materias. Los valores plausibles no son las puntuaciones reales de las pruebas sino que se trata de valores imputados a partir de técnicas estadísticas y, por tanto, recogen los valores que podrían ser razonablemente asignados a cada individuo aunque sujetos a una cierta aleatoriedad (incertidumbre). Esta metodología

---

<sup>5</sup> Estas técnicas han sido escasamente utilizadas en este contexto. Algunas excepciones son Burger (2011), Zhang y Lee (2011) y Ammermueller (2007b). De estos tres trabajos, el único que analiza los diferenciales rurales-urbanos es Burger (2011) quien utiliza información sobre el rendimiento educativo en Zambia a partir de una encuesta distinta a PISA. Sus resultados apuntan a que tanto las características de los estudiantes como los rendimientos obtenidos son importantes a la hora de explicar el diferencial rural-urbano.

fue desarrollada por Mislevy y Sheehan (1987, 1989) y está basada en la teoría de imputación de los valores ausentes o perdidos de Rubin (1987). La idea es que cada individuo contesta a un número limitado de cuestiones del test, y, por ese motivo, es preciso estimar como hubiera sido su comportamiento en el caso que hubiera contestado todas las preguntas de la evaluación. Para ello se predicen estos resultados utilizando las respuestas de cuestiones que ha contestado y otras variables obtenidas por los cuestionarios de contexto. En vez de predecir una única puntuación, se genera una distribución de valores para cada individuo con sus probabilidades asociadas obteniendo, por tanto, cinco valores plausibles para cada individuo. De esta manera se previene el sesgo derivado de estimar el resultado a partir de un número reducido de cuestiones de la prueba. Los valores plausibles contienen componentes de la varianza del error aleatorios, que no son óptimos como puntuaciones de las pruebas de manera individual. Así pues, no sirven para el diagnóstico de los sujetos sino solamente para la estimación de parámetros poblacionales consistentes. En nuestro análisis, utilizaremos dichos valores para llevar a cabo el análisis empírico propuesto aunque, para facilitar su lectura, en los descriptivos que se muestran a continuación se ha utilizado únicamente el valor promedio de los mismos. También se ha utilizado en todos los casos los factores de elevación proporcionados por la propia encuesta tanto para 2006 como para 2009.

En lo que se refiere al resto de variables de interés, las características individuales facilitadas en la encuesta y que se consideran en nuestro análisis son género, edad y nacionalidad (nativos e inmigrantes de primera y segunda generación) sin que se haya podido controlar por el tipo de estructura familiar (nuclear, monoparental y mixta) dado que esta información no estaba presente en el cuestionario de 2006. Respecto a las variables relativas a la familia, se incluye el nivel educativo de los progenitores, tal y como se realiza en otros estudios (Meunier 2011, Martins y Veiga, 2010, entre otros) y el entorno cultural del alumno que se aproxima a partir del número de libros disponibles en el hogar. Por último, también se incluyen variables relativas a la escuela como su ubicación en zona urbana o rural (la variable clave en nuestro estudio), el tamaño de la escuela, el número de alumnos por profesor y la titularidad del centro. Además se intenta controlar el efecto de los compañeros sobre el rendimiento del estudiante a través del nivel socio-económico medio de los compañeros de cada una de las escuelas analizadas. En concreto, se utiliza la media del índice ESCS de los estudiantes que pertenecen a la misma escuela del estudiante que se evalúa. El ESCS refleja el estatus económico, social y cultural del alumno, y se construye a partir de tres variables referentes a su background familiar: el mayor nivel educativo de los progenitores de acuerdo con la clasificación ISCED, el mayor índice del estatus ocupacional ISEI de los progenitores y un índice de posesiones culturales en el hogar. Posteriormente, los valores obtenidos se han estandarizado de tal manera que dicho índice tiene un promedio igual a 0 y una desviación típica igual a 1 para los países de la OCDE.



La tabla 2 muestra los principales estadísticos descriptivos para las variables descritas anteriormente. Las primeras columnas de dicha tabla recogen la información relativa al conjunto de las muestras analizadas para 2006 y 2009 mientras que en el resto de columnas se facilita la información desagregada para zonas urbanas y rurales para cada uno de los dos años analizados. Los gráficos 1 y 2 muestran también las diferencias entre las zonas rurales y urbanas para las variables de interés (rendimiento educativo en matemáticas, ciencias y lectura) a lo largo de toda la distribución para 2006 y 2009, respectivamente.

## TABLA 2

## GRÁFICOS 1 y 2

A partir de dichos resultados, se puede observar como existe un claro diferencial en el rendimiento educativo entre los estudiantes de las zonas rurales y de las zonas urbanas a favor de estos últimos. Dicho diferencial es aproximadamente de 30 puntos en ambos periodos, aunque si bien es cierto que cuando se compara la evolución entre 2006 y 2009 se observa una clara mejora en ambas zonas. Existe también una mayor variabilidad en las puntuaciones de la prueba de lectura en relación a las de ciencias y matemáticas. Precisamente, el diferencial en lectura es de los más elevados tanto en 2006 como en 2009 y además, aumenta claramente en este último año en relación al anterior, tendencia no observada para ciencias (donde el diferencial se mantiene prácticamente constante) ni en matemáticas (donde el diferencial se reduce considerablemente). Estos resultados también se confirman a través del análisis de los gráficos 1 y 2.

Si nos centramos en el resto de variables a las que se hace referencia en la tabla 2, se puede observar como el porcentaje de mujeres es ligeramente superior al de los hombres con la excepción de las zonas rurales en 2009. También se observa que la educación de los padres es claramente superior en las zonas urbanas en relación a las rurales sin que existan grandes cambios entre 2006 y 2009. Por último, y en lo que se refiere a la localización de las escuelas y sus características, alrededor del 70% de las escuelas están situadas en las zonas urbanas y que la mayor parte de las escuelas son públicas aunque el porcentaje de las mismas es significativamente superior en las zonas rurales (por encima del 90%) en relación a las zonas urbanas (donde no supera el 80%). También se puede comprobar que las escuelas urbanas tienen, en promedio, un mayor número de alumnos (aunque se ha reducido claramente entre 2006 y 2009) y una mayor ratio de alumnos por profesor que las escuelas rurales, si bien es cierto que en el período analizado se ha producido un claro aumento en esta variable. En lo que se refiere al nivel socioeconómico medio de los compañeros de cada una de las escuelas analizadas, el valor promedio para el conjunto de la muestra para Colombia presenta valores negativos tanto para 2006 como para 2009, lo que indica que la situación socio-económica de las familias de este país están

claramente por debajo del promedio de la OCDE. Destaca, sin embargo, que el valor de dicho índice en las zonas rurales se sitúa muy por debajo del observado para las zonas urbanas, siendo prácticamente el doble tanto para la ola de 2006 como para la ola correspondiente a 2009.

En el siguiente apartado, se aplican técnicas estadísticas y econométricas para analizar la influencia de las distintas variables consideradas sobre el diferencial en el rendimiento educativo de los estudiantes de escuelas rurales en relación a los de las escuelas urbanas.

### 3 Metodología y resultados

Así pues, como un primer paso para analizar si las diferencias observadas en el rendimiento educativo de los estudiantes de zonas rurales y urbanas están relacionadas con factores individuales, del entorno o de las escuelas, a continuación, se especifica y se estima una función de producción educativa donde se incluyen distintos controles tanto a nivel individual como familiar como de las escuelas. En concreto, la función de producción en educación para cada una de las materias que se utilizará en este trabajo toma como punto de partida la siguiente expresión:

$$RTest_i = \alpha + \beta \cdot Z_i + e_i \quad (1)$$

donde  $RTest_i$  hace referencia a los cinco valores plausibles de los resultados de la prueba para cada una de las materias del estudiante  $i$ ,  $Z_i$  es un vector de variables de control referidas a características de los propios individuos, antecedentes familiares y de la escuela, mientras que  $e_i$  es un término de error aleatorio.

Teniendo en cuenta la naturaleza de la variable endógena que ya se ha detallado anteriormente, para poder estimar dicho modelo se ha utilizado un procedimiento que permite realizar múltiples estimaciones de la variable dependiente<sup>6</sup>, que se refiere a los cinco valores plausibles del resultado educativo para cada una de las materias. Además, y debido al complejo diseño muestral utilizado en PISA, es necesario aplicar un procedimiento basado en replicaciones para calcular la varianza de los estimadores. Para este tipo de datos, en OCDE (2009) se recomienda el método de replicación repetida balanceada (BRR) con una particular modificación de Fay (1989), que no modifica los coeficientes

---

<sup>6</sup> Para ello se ha utilizado el módulo de STATA, PV: Stata module to perform estimation with plausible values. <http://ideas.repec.org/c/boc/bocode/s456951.html>

pero mejora la precisión del estimador de la varianza. Este ha sido el procedimiento utilizado en este estudio.

En la tabla 3 se muestran los resultados de estimar el modelo (1) para las principales variables de interés del estudio con los microdatos de 2009 y para las distintas materias analizadas (matemáticas, ciencias y lectura). Los resultados para 2006 no se presentan por cuestiones de espacio pero son similares a los recogidos en la tabla 3 y se encuentran disponibles previa petición a los autores. Los resultados obtenidos son relativamente similares a los obtenidos en otros estudios que han utilizado los microdatos de PISA. En concreto, y como era de esperar, el género del estudiante tiene un efecto estadísticamente significativo sobre el resultado académico del alumno, aunque el signo es diferente dependiendo de la materia que se analice. Las chicas obtienen un resultado académico inferior en matemáticas y ciencias que los chicos, pero superior en lectura. La edad de los alumnos, alrededor de los 15 años y 9 meses con pequeñas oscilaciones de 3 meses, tiene un impacto positivo a medida que va en aumento en todas las materias analizadas. En lo que se refiere al conjunto de variables relativas al entorno familiar del alumno, se observa que las variables ficticias referentes al número de libros en el hogar, incluidas como indicador del entorno cultural, tienen un efecto positivo sobre el rendimiento educativo del alumno, que aumenta a medida que el número de libros en el hogar es mayor. Por otro lado, el nivel educativo de la madre tiene un efecto positivo sobre el rendimiento educativo de los hijos, que toma más importancia para los alumnos con peores resultados educativos. No obstante, no sucede lo mismo en el caso del nivel educativo de los padres, ya que dicha variable no resulta estadísticamente significativa. Entrando ya a analizar las variables relativas a las características de las escuelas, se puede observar que ninguna de las características habituales (titularidad, tamaño y ratio alumno-profesor) resulta estadísticamente significativa. De hecho, la única variable relevante es el nivel socio-económico medio de los compañeros que presenta un signo positivo y estadísticamente significativo a los niveles habituales.

### TABLA 3

Volviendo de nuevo al aspecto que centra el interés de nuestro estudio, la existencia de diferencias entre zonas rurales y urbanas, en el resto de este apartado se procede a descomponer las diferencias en rendimientos educativos entre estudiantes de zonas rurales y de zonas urbanas a partir de la aplicación de la metodología de Oaxaca-Blinder, en primer lugar, y, a continuación de la metodología de Juhn-Murphy-Pierce.

Tal y como se ha comentado en la introducción, la metodología de descomposición salarial de Oaxaca-Blinder ha sido ampliamente empleada para analizar la discriminación laboral por razón de género,

por raza o por otras características de los trabajadores. Como es sabido, esta técnica permite descomponer la diferencia en el nivel medio observado para una determinada variable entre dos colectivos en una parte explicada por las diferencias en las características observadas de ambos colectivos y en una parte originada por diferencias en los rendimientos asociados a dichas características. La extensión de dicha metodología desarrollada por Juhn-Murphy-Pierce, representa un avance importante en las técnicas de descomposición, en la medida en que permite descomponer los cambios en las diferencias a lo largo del tiempo entre los dos colectivos analizados.

A partir de la función de producción educativa estimada conjuntamente para los estudiantes de zonas rurales y urbanas como estructura de referencia en la descomposición, la diferencia en el rendimiento educativo de ambos colectivos se puede expresar como:

$$\overline{RTest}_R - \overline{RTest}_U = (\bar{Z}_R - \bar{Z}_U) \cdot \beta_R + \bar{Z}_U \cdot (\beta_R - \beta_U) + (\bar{\epsilon}_R - \bar{\epsilon}_U) \quad (2)$$

donde los subíndices  $R$  y  $U$ , corresponden respectivamente a zona rural y urbana. La ecuación (2) permite cuantificar en qué medida el origen de las diferencias entre las zonas rurales y urbanas está relacionado con diferencias observadas en las características individuales, de las familias o de las escuelas y con la influencia de factores inobservables. Más concretamente, el primer término del lado derecho de la ecuación corresponde a la parte del diferencial en el rendimiento educativo atribuible a las diferencias entre ambos colectivos en las características observadas y coincide con el componente “explicado” de la descomposición de Oaxaca-Blinder mientras que el segundo y tercer término se corresponden con la diferencia en las retribuciones y diferencias en habilidades no observables y recogería, básicamente, el componente discriminatorio o “no explicado” de dicha descomposición.

Los resultados obtenidos al aplicar la descomposición de Oaxaca-Blinder<sup>7</sup> para la ola de 2009 utilizando como referencia la estimación de la función de producción educativa para el conjunto de la muestra se presentan en la tabla 4<sup>8</sup>. Los resultados para 2006 no se presentan por cuestiones de espacio pero son similares a los recogidos en la tabla 4 y se encuentran disponibles previa petición a los autores. Tal y como se puede observar en dicha tabla, y de acuerdo con los estadísticos descriptivos presentados en el apartado anterior, en 2009 el rendimiento educativo promedio en matemáticas, ciencias y lectura es inferior para los estudiantes de las zonas rurales en relación a los urbanos. Gran parte de dicha diferencia se explica como consecuencia de peores características de los estudiantes de

<sup>7</sup> Para ello se ha utilizado el módulo de STATA, OAXACA: Stata module to compute the Blinder-Oaxaca decomposition, <http://ideas.repec.org/c/boc/bocode/s456936.html>

<sup>8</sup> Se han realizado pruebas de robustez a distintas regresiones de referencia sin que los resultados obtenidos cambien sustancialmente. La ventaja de trabajar con toda la muestra en vez de utilizar la información relativa a los estudiantes de las zonas urbanas o rurales es que los resultados obtenidos son directamente comparables con los obtenidos a través de la descomposición de Juhn-Murphy-Pierce.

las zonas rurales, aunque no tanto individuales sino familiares. Sin embargo, las características más relevantes son las relacionadas con las escuelas ya que son las que contribuyen en mayor medida a explicar el diferencial rural-urbano, aunque las variables tradicionales como titularidad de la escuela, tamaño de la escuela o ratio alumnos-profesor juegan un papel menor cuando se compara su papel con el nivel socio-económico medio de los alumnos de la escuela. Tanto en matemáticas como en ciencias y lectura, esta variable contribuye en más del 90% de la parte explicada del diferencial. También es importante destacar que la parte no explicada contribuye a reducir el diferencial rural-urbano, es decir, existe una diferente retribución entre las escuelas rurales y urbanas que contribuye a disminuir el papel acentuador del diferencial de las características observadas.

TABLA 4

La descomposición propuesta por Juhn-Murphy-Pierce supone que los rendimientos de las características individuales son iguales para ambos grupos. De este modo, el punto de partida de dicha descomposición sería el siguiente:

$$\overline{RTest}_R - \overline{RTest}_U = (\bar{Z}_R - \bar{Z}_U) \cdot \beta_R - \bar{\theta}_U = (\bar{Z}_R - \bar{Z}_U) \cdot \beta_R - \bar{\theta}_U \cdot \sigma_R \quad (3)$$

donde  $\sigma_R$  es la desviación estándar de los residuos ( $e_R$ ) y  $\theta_U = e_U / \sigma_R$ . La interpretación de ambos términos es similar a la descrita anteriormente para la descomposición de Oaxaca-Blinder. Si a partir de dicha ecuación, comparamos los cambios en el diferencial en los rendimientos educativos entre dos momentos distintos del tiempo (por ejemplo,  $t$  y  $t'$ ), obtenemos la siguiente expresión:

$$D_{t'} - D_t = (\Delta Z_{t'} - \Delta Z_t) \cdot \beta_{Rt} + \Delta Z_{t'} (\beta_{Rt'} - \beta_{Rt}) + \\ + (\Delta \theta_{t'} - \Delta \theta_t) \cdot \sigma_{Rt} + \Delta \theta_{t'} (\sigma_{Rt'} - \sigma_{Rt}) \quad (4)$$

donde  $D_{t'}$  representa el diferencial entre el rendimiento educativo medio de los estudiantes de las zonas rurales y los de las zonas urbanas en el instante  $t'$ ,  $D_t$  ese mismo diferencial pero para el instante  $t$  y el símbolo  $\Delta$  denota la variación entre zonas rurales y urbanas para cada una de las variables o parámetros asociados. El resto de símbolos asociados sigue exactamente la misma notación que en (3). El primer término en (4) se corresponde con el cambio observado en las características (efecto cantidades), el segundo término tiene que ver con los cambios en los coeficientes y, por tanto, con variaciones en las remuneraciones (efecto precios), el tercero con la interacción entre ambos mientras que el último término recoge la variación no explicada.

La tabla 5 recoge el resultado de aplicar dicha metodología<sup>9</sup> para explicar las variaciones en el diferencial en el rendimiento educativo entre zonas rurales y urbanas entre 2006 y 2009 para matemáticas, ciencias y lectura. Tal y como se puede observar en dicha tabla, el diferencial entre 2006 y 2009 se ha reducido en matemáticas, se ha mantenido prácticamente constante en ciencias y ha aumentado en lectura. Podemos ver como en todos los casos la variación en la parte explicada ha contribuido a aumentar el diferencial, tanto por cambios en las características como por cambios en los precios aunque este segundo componente ha tenido un mayor impacto en el período analizado. Sin embargo, ambos efectos se han visto parcial o totalmente compensados (como en el caso de las matemáticas) por cambios en la parte no explicada. Así pues, dichos resultados corroboran los resultados obtenidos con la descomposición de Oaxaca dado que refuerzan la idea de que los cambios en las características no son los responsables de la evolución positiva en las escuelas de las zonas rurales sino que probablemente se deban a otro tipo de cambios relacionados con variables no observables como puede ser la mejora en el marco institucional del país, el cese/reducción de la intensidad del conflicto armado y otros aspectos no observables y, por tanto, no incluidos en el modelo.

#### TABLA 5

En síntesis, los resultados obtenidos ponen de manifiesto el escaso impacto de las políticas de discriminación positiva en las escuelas de las zonas rurales (al menos a través de las dimensiones recogidas en este estudio: titularidad, tamaño y ratio profesor/alumno) pero en cambio, sí que se obtiene evidencia sobre el impacto favorable de las condiciones socio-económicas en dicho ámbito así como de otros factores no observables que pueden haber incidido en una mejora del contexto educativo en Colombia.

## 4 Conclusiones

El presente trabajo ha analizado la posible existencia de diferencias en el rendimiento educativo entre los estudiantes de zonas rurales y urbanas de Colombia en las materias de matemáticas, ciencias y lectura. Para ello se han utilizado los datos PISA de 2006 y 2009 y se ha especificado y estimado una función de producción en educación que incluía variables relativas a la ubicación de la escuela y los controles habituales a nivel individual y familiar. Además, y con el fin de identificar los factores que explican este diferencial, se ha utilizado la descomposición de Oaxaca-Blinder y la propuesta por

---

<sup>9</sup> Para ello se ha utilizado el módulo de STATA, JMPIERCE2: Stata module to compute trend decomposition of outcome differentials, <http://ideas.repec.org/c/boc/bocode/s448804.html>

Juhn-Murphy-Pierce para analizar la variación temporal de dichas diferencias. Los resultados obtenidos a partir de la aplicación de ambas metodologías muestran que la mayor parte del diferencial entre escuelas urbanas y rurales está relacionado con las características de la familia y no tanto con las características de la escuela, aunque la explotación de la dimensión temporal también ha permitido destacar el papel de otros factores no observables en la reducción del diferencial rural-urbano.

Desde una perspectiva de política educativa, la evidencia obtenida refuerza la idoneidad de todas aquellas medidas dirigidas a mejorar el contexto educativo general y las condiciones familiares y, quizás, no tanto a la discriminación positiva de las escuelas rurales como medio para mejorar el rendimiento educativo (al menos en lo que se refiere a los indicadores considerados en este estudio: titularidad, tamaño y ratio profesor/alumno).

A partir de estos resultados, se abren, por tanto, además varias vías de continuación. En primer lugar, y dado que las familias parecen ejercer un papel crucial sobre el rendimiento educativo de los estudiantes, sería interesante analizar qué características de las familias resultan más relevantes para explicar el diferencial rural-urbano una vez controlado el resto de los factores. En este sentido, la información sobre las familias que ofrece PISA es insuficiente y sería necesario utilizar una base de datos con mayor riqueza en esta dimensión. En segundo lugar, también sería necesario disponer de una base de datos más rica en lo que se refiere a la información relativa a las características del territorio de residencia de los estudiantes lo que permitiría poder analizar los mecanismos a través de los cuales el entorno geográfico acaba teniendo efectos sobre el rendimiento educativo de los estudiantes y hasta que punto dicho efecto no está capturando la importancia de otras variables que no hemos podido controlar adecuadamente en este estudio (omisión de variables relevantes) que pueden estar relacionados con mejoras institucionales que se han producido en Colombia en el período analizado y que pueden haber tenido mayores efectos en las zonas rurales en relación a las zonas urbanas.

## 5 Referencias bibliográficas

Abdul-Hamid, H. (2004): Assessing Argentina's preparedness for the knowledge economy: Measuring student knowledge and skills in reading, mathematical and scientific literacy. Evidence from PISA 2000. (disponible en: [http://siteresources.worldbank.org/EDUCATION/Resources/278200-1121703274255/1439264-1171379341729/SessionIIa\\_Husein\\_Abdul-Hamid4.pdf](http://siteresources.worldbank.org/EDUCATION/Resources/278200-1121703274255/1439264-1171379341729/SessionIIa_Husein_Abdul-Hamid4.pdf))

Ammermueller, A. (2007a): "Poor background or low returns? Why immigrant students in Germany perform so poorly in the programme for international student assessment", *Education Economics*, 15(2), 215–230.

Ammermueller, A. (2007b): "PISA: What makes the difference?: Explaining the gap in test scores between Finland and Germany", *Empirical Economics*, 33 (2), pp-287.

Arum, R. (2000): "Schools and communities: Ecological and institutional dimensions", *Annual Review of Sociology*, 26:395–418.

Barnett, R. R.; Glass, J. C.; Snowdon, R. I. y Stringer, K. S. (2002): "Size, performance and effectiveness: cost-constrained measures of best-practice performance and secondary-school size", *Education Economics*, 10 (3), 291-311.

Behrman, J.R. (2010): "Investment in education: Inputs and incentives", en Rodrik, D. y Rosenzweig, M. R. (editores), *Handbook of development economics: The economics of development policy*. vol. 5, Amsterdam: North-Holland Publishing. 4883–4975.

Blackwell, D.L. y McLaughlin, D.K. (1999): "Do Rural Youth Attain Their Educational Goals?" *Rural Development Perspectives* 13(3):37–44

Blinder, A. S. (1973): "Wage Discrimination: Reduced Form and Structural Variables", *Journal of Human Resources*, 8, 436-455.

Burger, R. (2011): "School effectiveness in Zambia: The origins of differences between rural and urban outcomes", *Development Southern Africa*, 28 (2), 157-176.

Caldas, S.J. y Bankston, C. (1997): "Effect of school population socioeconomic status on individual academic achievement", *The Journal of Educational Research*, 90(5): 269–277.

Chiswick, B. R. y DebBurman, N. (2004): "Educational attainment: analysis by immigrant generation", *Economics of Education Review*, 23: 361-379.

Coleman, J. S. y Hoffer, T. (1987): *Public and private high schools. The impact of communities*. New York: Basic Books.

Coleman, J. S.; Campbell, J. S.; Campbell, E. Q.; Hobson, C. J.; McPartland, J.; Mood, A. M.; Weinfeld, F. D. y York, R. L. (1966): *Equality of Educational Opportunity*. Washington, D.C.: US Department of Health, Education, and Welfare, US Government Printing Office.

Cueto, S., Guerrero, G., Leon, J, De Silva, M.; Huttly, S. Penny, M. E. Lanata, C. F., Villar, E. (2005). *Social capital and education outcomes in urban and rural Peru*. Working paper # 28, Young Lives, London.



Edington, E.D. y Martellaro, H.C. (1984). Variables affecting academic achievement in New Mexico schools. Las Cruces: New Mexico Center for Rural Education. (Eric Document Reproduction Service No. 271 267)

Entorf, H. y Lauk, M. (2008): “Peer effects, social multipliers and migrants at school: An international comparison”, *Journal of Ethnic and Migration Studies*, 34(4), 633–654.

Entorf, H. y Minoiu, N. (2005): “What a difference immigration policy makes: A comparison of PISA scores in Europe and traditional countries of immigration”, *German Economic Review*, (3), 355–376.

Fay, R. E. (1989): “Theoretical application of weighting for variance calculation”, en *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, 212-217.

Fuller, B. y Clarke, P. (1994): “Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules and pedagogy”, *Review of Educational Research* 64(1):119–157

Feinstein, L. y Symons, J. (1999): “Attainment in secondary education”, *Oxford Economic Papers*, 51, 300-321.

Fertig, M. (2003): “Who’s to Blame? The Determinants of German Students’ Achievement in the PISA 2000 Study”, *IZA Discussion Paper Series*, 739.

Gamoran, A. y Long, D. A. (2007): “Equality of educational opportunity: A 40-year retrospective. International Studies in Educational Inequality”, en Teese, R.; Lamb, S. y Duru-Bellat, M. (Editores). *International studies in educational inequality: Theory and policy*, Dordrecht: Springer, 23–47.

Gaviria, A. y Barrientos, J. H. (2001a): “Determinantes de la calidad de la educación en Colombia”, *Archivos de Economía*, 159.

Gaviria, A. y Barrientos, J. H. (2001b): “Calidad de la educación y rendimiento académico en Bogotá”, *Coyuntura Social*, 24:111–127.

Häkkinen, I, Kirjavainen, T. y Uusitalo, R. (2003): “School resources and student achievement revisited: new evidence from panel data”, *Economics of Education Review*, 22, 329-335.

Hannaway, J. y Talbert, J.E. (1993). Bringing context into effective school research: Urban-suburban differences. *Educational Administration Quarterly*, 29, 164–186.

Hanushek, E. A. (1986): “The economics of schooling”, *Journal of Economic Literature*, 24, 1141-1177.

Hanushek, E. A. (2003): “The failure of input based schooling policies”, *The Economic Journal*, 113, 64-98.

Hanushek, E. A. y Luque, J. A. (2003): “Efficiency and equity in schools around the world”, *Economics of Education Review*, 22, 481-502.

Hanushek, E. A.; Kain, J. F.; Markman, J. M. y Rivkin, S. G. (2003): “Does peer ability affect student achievement?”, *Journal of Applied Econometrics*, 18, 527-544.

Harbison, R. W., Hanushek, E. A. y World Bank. (1992). Educational performance of the poor: Lessons from rural northeast Brazil. Oxford: Published for the World Bank, Oxford University Press.

Haveman, R. y Wolfe, B. (1995): “The determinants of children’s attainment: A review of methods and findings”, *Journal of Economics Literature*, 33(4), 1829–1878

- Henderson, V.; Mieszkowski, P. y Sauvageau, Y. (1978): “Peer group effects in educational production functions”, *Journal of Public Economics*, 10 (1), 97-106.
- Heyneman S. P. y Loxley, W. A. (1983): “The effect of primary school quality on academic achievement across 29 high and low income countries”, *American Journal of Sociology*, 88(6): 1162–1194.
- Ibáñez, A.M. y Vélez, C.E. (2008): “Civil Conflict and Forced Migration: The Micro Determinants and Welfare Losses of Displacement in Colombia”, *World Development*, 36(4): 659–676.
- Juhn, C., Murphy, K.; Pierce, B. (1993): “Wage inequality and the rise in returns to skill”, *Journal of Political Economy*, 101 (31), 410-442.
- Kleinfeld, J.S., McDiarmid, G.W. y Hagstrom, D. (1985). Alaska’s small rural high schools: Are they working? ISER Report Series No. 57, Alaska: University of Alaska.
- Krueger, A. B. (2003): “Economics considerations and class size”, *Economic Journal*, 113, 34-63.
- Lee, V. E., Smith, J. B. y Croninger, R. G. (1997) “How high school organization influences the equitable distribution of learning in mathematics and science”, *Sociology of Education*, 70(2):128–150.
- Martins, L. y Veiga, V. (2010): “Do inequalities in parents’ education play an important role in PISA students’ mathematics achievement test score disparities?”, *Economics of Education Review*, 29 (6), 1016-1033.
- Meunier, M. (2011): “Immigration and student achievement: Evidence from Switzerland”, *Economics of Education Review*, 30 (1), 16-38.
- Mislevy, R. J. y Sheehan, K. M. (1987): “Marginal estimation procedures”, in Beaton, A.E., Editor, 1987. *The NAEP 1983-84 technical report*, National Assessment of Educational Progress, Educational Testing Service, Princeton, 293-360.
- Mislevy, R. J. y Sheehan, K. M. (1989): “Information matrices in latent-variable models”, *Journal of Educational Statistics*, 14, 335-350.
- Mizala, A. y Romaguera, P. (2000): “School performance and Choice: The Chilean Experience.” *Journal of Human Resources* 35(2), 392-417.
- Monk, D. H. y Haller, E. J. (1986): Organizational alternatives for small/rural schools: final report to the new york state legislature. New York, NY: Cornell University, 1986. ED 281 694.
- Noell, J. (1982): “Public and Catholic schools: A re-analysis of public and private schools”, *Sociology of Education*, 55, 123-132.
- Oaxaca, R. (1973): “Male-Female Wage Differentials in Urban labor Markets.”, *International Economic Review*, 14 (3), 139-148.
- OCDE (2009): *Technical report 2006*. OECD, Paris.
- OECD (2010): *PISA 2009 Results: Overcoming Social Background: Equity in Learning Opportunities and Outcomes (Volume II)*, OECD, Paris.

O'Malley, B. (2010), *Education under attack*, UNESCO (disponible en: <http://unesdoc.unesco.org/images/0018/001868/186809e.pdf>)

Opdenakker, M. C. y Van Damme, J. (2006): "Differences between secondary schools: A study about school context, group composition, school practice, and school effects with special attention to public and Catholic schools and types of schools", *School Effectiveness and School Improvement*, 17(1), 87-117.

Piñeros, L. J. y Rodríguez, A. (1999): "Los insumos escolares en la educación secundaria y su efecto sobre el rendimiento académico de los estudiantes: Un estudio en Colombia" LCSHD Paper Series 36. Washington, DC: World Bank.

Rangel, C. y Lleras, C. (2010): "Educational inequality in Colombia: family background, school quality and student achievement in Cartagena", *International Studies in Sociology of Education*, 20(4), 291-317

Rivkin, S. G.; Hanushek, E. A. y Kain, J. F. (2005): "Teachers, Schools and Academic Achievement". *Econometrica*, 73(2), 417-458.

Rubin, D. B. (1987): *Multiple imputation for nonresponse in surveys*, New York, Wiley.

Sander, W. (1996): "Catholic grade schools and academic achievement", *The Journal of Human Resources*, 31 (3), 540-548.

Santos, M.E. (2007): "Calidad de la educación en Argentina: Determinantes y distribución utilizando los resultados de PISA 2000", *Bienestar y política social*, 3(1):79-109.

Somers, M. A.; McEwan, P. J. y Willms, J. D. (2004): "How effective are private schools in Latin America?", *Comparative Education Review*, 48(1): 48-69.

Smith, J. y R. A. Naylor (2005): "Schooling effects on subsequent university performance: evidence for the UK university population", *Economics of Education Review*, 24: 549-562.

Stevans, L. K. y Sessions, D. N. (2000): "Private/public school choice and student performance revisited", *Education Economics*, 8 (2), 169-184.

Vandenberghe, V. y Robin, S. (2004): "Evaluating the effectiveness of private education across countries: A comparison of methods", *Labor Economics* 11(4):487-506.

Ward, A.P. y Murray, L.W. (1985). Factors affecting performance of New Mexico high school students. Paper presented at the Meeting of the Rocky Mountain Educational Research Association, Las Cruces. (Eric Document Reproduction Service No. 271 266)

War Child (2007) Child Soldiers: The Shadow of Their Existence. (disponible en: [http://www.warchild.org/news/News\\_archive/2007/projects/ChildSoldierReport\\_/childsoldierreport\\_.html](http://www.warchild.org/news/News_archive/2007/projects/ChildSoldierReport_/childsoldierreport_.html))

Woßmann, L. (2003): "Schooling resources, educational institutions and student performance: the international evidence", *Oxford Bulletin of Economics and Statistics*, 65 (2), 117-170.

Woßmann, L. (2010): "Families, schools and primary-school learning: evidence for Argentina and Colombia in an international perspective", *Applied Economics* 42, 2645-2665.

Young, D.J. (1998). Rural and urban differences in student achievement in science and mathematics: A multilevel analysis, school effectiveness and school improvement, 9(4), 386-418.

Zhang, L., Lee, K.A.(2011), “Decomposing achievement gaps among OECD countries”, *Asia Pacific Education Review*, 12 (3), 463-474.

## 6. Tablas y gráficos

**Tabla 1. Estudios realizados en países suramericanos en los que se compara el rendimiento de los estudiantes de escuelas rurales y urbanas.**

Estudio	País	Datos	áreas	Método	Características
Harbison y Hanushek (1992)*	Brasil	EDURURAL data collection 1981, 1983, 1985. Segundo y cuarto grado	portugués y matemáticas	Regresión con datos de corte transversal y corrección del sesgo de selección	individuo, familia, profesores, compañeros, infraestructuras, materiales de estudio
Mizala y Romaguera (2000)	Chile	SIMCE Educational Quality Measurement System (average) 1996 cuarto y octavo grado	matemáticas, español	MCO	familiares, personales, profesores, SIMCe 1994
Abdul_Hamid (2004)	Argentina	PISA 2000	matemáticas, lectura, ciencias	Mínimos Cuadrados Generalizados (GLS) y regresiones cuantílicas	familiares, individuales, escuela
Cueto et al. (2005)	Perú	Proyecto "Young Lives" 2002 (niños entre 7.5 y 8.5 años)	lectura, escritura y matemáticas	MCO	familiares, individuales, capital social del hogar y la comunidad
Woßmann (2010)	Argentina, Colombia,	Progress in International Reading Literacy Study (PIRLS). Estudiantes de cuarto grado en 2001	lectura	Mínimos Cuadrados Ponderados y regresiones lineales Clustering-robustas (CRLR)	familiares, individuales, escuela, puntaje del test en el período anterior
Santos (2007)***	Argentina	PISA 2000	lectura y matemática	regresión lineal para datos de encuestas y regresión por cuantiles	familiares, individuales, escuela

\* Estudio centrado en zona la rural (no realiza comparación urbano-rural)

\*\* El valor agregado toma como variable endógena la variación de la puntuación de un alumno entre dos períodos. El resto de modelos toman como variable endógena la puntuación obtenida por el alumno en un único test.

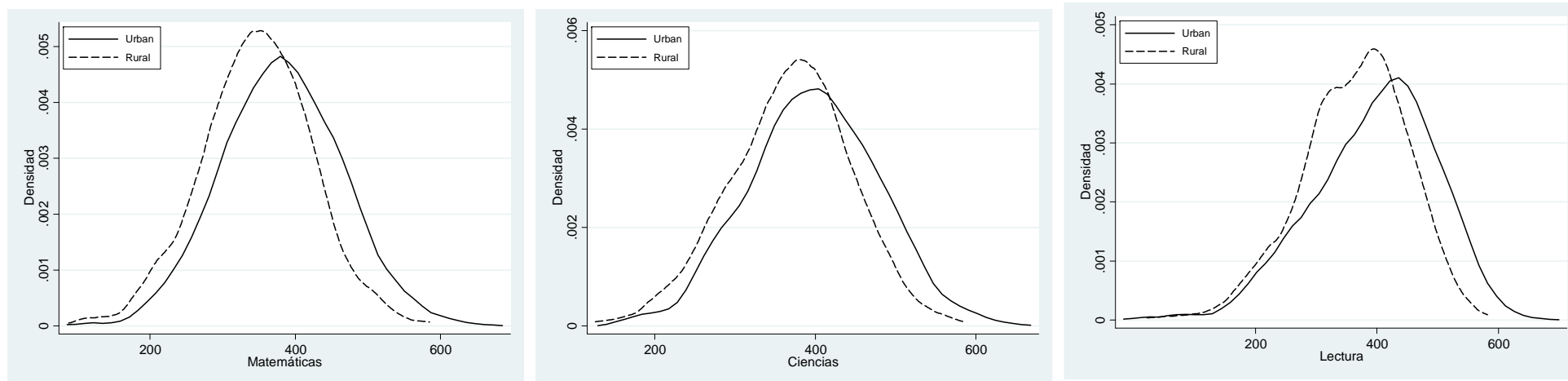
\*\*\* Diferencia rural-urbano significativa para matemáticas pero no para lectura.

**Tabla 2. Análisis descriptivo de las variables utilizadas en el análisis**

	TOTAL 2006		TOTAL 2009		URBANA 2006		URBANA 2009		RURAL 2006		RURAL 2009	
	Media	Desv. est.	Media	Desv. est.	Media	Desv. est.	Media	Desv. est.	Media	Desv. est.	Media	Desv. est.
<b><i>Puntuaciones:</i></b>												
Matemáticas	375.13	80.86	384.99	71.07	385.53	82.88	392.47	72.01	351.96	70.90	369.72	66.55
Ciencias	392.09	79.02	406.63	75.78	400.44	80.57	415.78	75.18	373.50	72.06	387.93	73.53
Lectura	390.89	96.43	416.86	82.92	400.91	99.88	429.33	81.26	368.59	84.05	391.41	80.43
<b><i>Caract. individuales:</i></b>												
Mujer	0.54	0.50	0.51	0.50	0.54	0.50	0.53	0.50	0.53	0.50	0.49	0.50
Edad	15.84	0.29	15.85	0.28	15.84	0.29	15.84	0.28	15.86	0.28	15.87	0.28
Inmigrante 1º generación	0.01	0.04	0.00	0.02	0.00	0.04	0.00	0.03	0.00	0.00	0.00	0.00
Inmigrante 2º generación	0.01	0.05	0.00	0.05	0.00	0.06	0.00	0.06	0.00	0.03	0.00	0.04
Idioma en el hogar	0.01	0.06	0.00	0.06	0.00	0.04	0.00	0.07	0.01	0.08	0.00	0.05
<b><i>Caract. familiares:</i></b>												
Educación madre	9.14	4.56	9.48	4.67	9.78	4.55	10.38	4.52	7.72	4.28	7.65	4.42
Educación padre	9.32	4.84	9.47	4.80	10.06	4.79	10.36	4.69	7.70	4.53	7.66	4.51
<b><i>Entorno cultural:</i></b>												
De 0 a 10 libros en el hogar	0.34	0.47	0.33	0.47	0.26	0.44	0.27	0.44	0.50	0.50	0.46	0.50
de 11 a 25 libros	0.29	0.46	0.31	0.46	0.30	0.46	0.31	0.46	0.28	0.45	0.32	0.46
de 26 a 100 libros	0.26	0.44	0.26	0.44	0.31	0.46	0.30	0.46	0.16	0.36	0.17	0.38
de 101 a 200 libros	0.07	0.26	0.07	0.26	0.08	0.28	0.08	0.28	0.05	0.21	0.04	0.20
de 201 a 500 libros	0.03	0.17	0.02	0.15	0.04	0.19	0.03	0.17	0.01	0.11	0.01	0.08
Más de 500 libros	0.01	0.09	0.01	0.08	0.01	0.10	0.01	0.09	0.00	0.07	0.00	0.03
<b><i>Caract. escuela:</i></b>												
Localización zona urbana	0.69	0.46	0.67	0.47								
Escuela pública	0.83	0.38	0.79	0.41	0.77	0.42	0.72	0.45	0.95	0.21	0.93	0.25
Tamaño escuela	1690.77	1330.57	1356.24	1003.85	2019.98	1362.72	1514.30	1058.02	957.32	893.40	1033.51	790.24
Ratio alumno-profesor	23.90	11.13	27.18	10.00	26.37	9.32	28.49	11.13	18.39	12.75	24.48	6.37
ESCS	-0.96	0.73	-1.08	.85	-0.71	0.69	-0.78	.79	-1.52	.47	-1.69	.60
<b><i>Observaciones</i></b>	409208		425758		282435		285787		126773		139971	

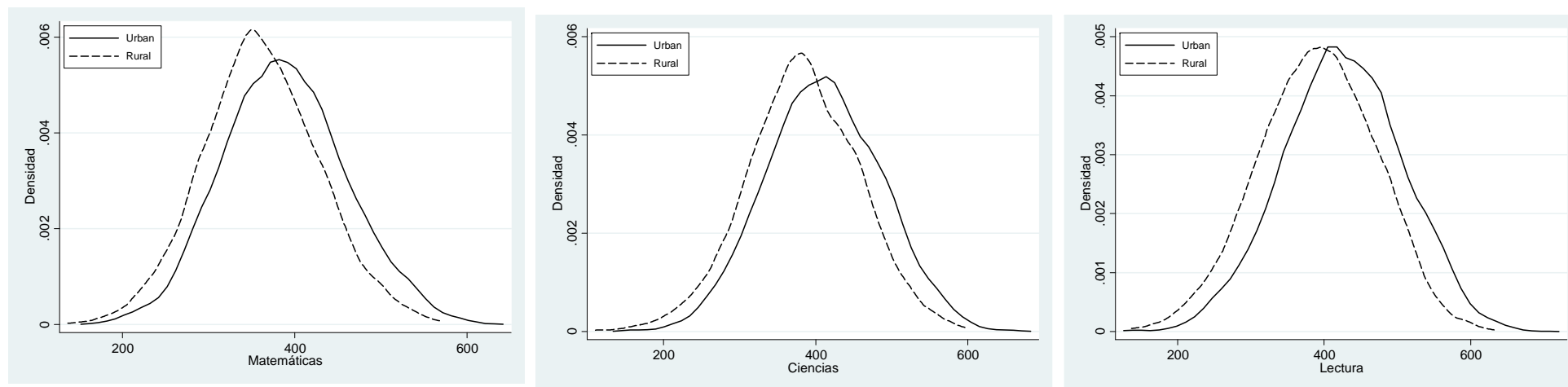
Fuente: Elaboración propia a partir de PISA 2006 y 2009.

**Gráfico 1: Distribución del rendimiento educativo de los alumnos en 2006 según la zona de localización de la escuela (urbana/rural).**



Fuente: Elaboración propia a partir de PISA 2006.

**Gráfico 2: Distribución del rendimiento educativo de los alumnos en 2009 según la zona de localización de la escuela (urbana/rural).**



Fuente: Elaboración propia a partir de PISA 2009.

**Tabla 3. Estimación de una función de producción educativa para Colombia con los datos de PISA 2009**

VARIABLES	Matemáticas	Ciencias	Lectura
Mujer	-30.98*** [2.788]	-20.38*** [2.568]	10.65*** [3.059]
Edad	17.60*** [5.084]	1.557 [5.229]	8.873* [5.221]
Educación madre	1.307*** [0.366]	1.034*** [0.398]	1.030*** [0.372]
Educación padre	0.538 [0.344]	-0.0453 [0.440]	-0.432 [0.396]
Inmigrante 1ª generación	-69.59** [28.29]	-46.09* [23.71]	-84.54** [40.08]
Inmigrante 2ª generación	-67.22*** [23.23]	-63.75** [29.76]	-98.32*** [30.52]
Idioma en el hogar	-31.05* [16.66]	-23.45 [20.66]	-50.46** [22.77]
De 11 a 25 libros en el hogar	15.29*** [3.891]	22.33*** [4.337]	17.42*** [3.791]
De 26 a 100 libros	34.97*** [3.723]	34.16*** [4.782]	30.28*** [4.882]
De 101 a 200 libros	31.11*** [4.782]	33.61*** [5.903]	35.11*** [6.009]
De 201 a 500 libros	38.23*** [8.347]	44.78*** [8.966]	25.95*** [10.03]
Más de 500 libros	20.39 [16.79]	20.07 [16.85]	18.71 [17.66]
Escuela pública	-5.362 [8.302]	-6.907 [6.629]	5.983 [7.194]
Tamaño escuela	0.00298 [0.00207]	-0.000147 [0.00259]	0.000843 [0.00228]
Ratio alumno-profesor	-0.410 [0.258]	-0.652* [0.353]	-0.349 [0.244]
<i>Peer effects</i>	27.52*** [4.028]	31.46*** [4.538]	44.93*** [3.496]
Constante	128.9* [77.97]	421.4*** [85.32]	301.3*** [82.08]
Observaciones	425757	425757	425757

Fuente: Elaboración propia a partir de PISA 2009.



**Tabla 4. Descomposición de Oaxaca del diferencial rural-urbano para Colombia con los datos de PISA 2009**

	Matemáticas	Ciencias	Lectura
Rural (R)	369.7***	387.9***	391.4***
Urbano (U)	392.5***	415.8***	429.3***
Diferencia total (R-U)	-22.75***	-27.86***	-37.92***
Parte explicada (Q)	-34.82***	-37.39***	-48.15***
Parte no explicada (N)	12.07*	9.532	10.23
Desglose de la parte explicada (Q)			
Características individuales (QI)	1.71	0.6046	-0.031
Características familiares (QF)	-10.15***	-8.918***	-8.007***
Características de la escuela (QE)	-26.38***	-29.073***	-40.116***
Desglose de las características de la escuela (QE)			
Escuela pública	0.0205	-2.244	3.175
Tamaño escuela	-1.073	1.457	0.44
Ratio alumno/professor	5.435	5.664	4.619
Peer effect	-30.76***	-33.95***	-48.35***

Fuente: Elaboración propia a partir de PISA 2009.

**Tabla 5. Descomposición de Juhn-Murphy-Pierce del diferencial rural-urbano para Colombia con los datos de PISA 2006 y 2009**

	Matemáticas	Ciencias	Lectura
Diferencial Rural-Urbano 2006 (RU <sub>2006</sub> )	-33.57***	-26.94***	-32.33***
Diferencial Rural-Urbano 2009 (RU <sub>2009</sub> )	-22.75***	-27.86***	-37.92***
Diferencia total (RU <sub>2009</sub> -RU <sub>2006</sub> )	10.82***	-0.91	-5.59***
Parte explicada (QP)	-13.15***	-22.66***	-10.24***
Parte no explicada (N)	23.97***	21.75***	4.64***
Desglose de la parte explicada (QP)			
Efecto cantidad (Q)	-2.59	-3.78	-2.75
Efecto precio (P)	-13.63***	-18.86***	-10.58***
Interacción (QxP)	3.08	-0.013	3.09

Fuente: Elaboración propia a partir de PISA 2006 y 2009.