

Retornos salariales a la educación: ¿Una fuente de disparidades regionales?

Rodrigo José Galindo Sánchez¹

Resumen

Esta investigación analiza la relación entre retornos a la educación y brecha salarial entre tres *clusters* de ciudades colombianas, agrupadas según la similitud de sus mercados laborales. Los resultados revelan que las ciudades con salarios bajos aún deben invertir mucho más en capital humano pues están por debajo de los niveles de Bogotá y dicha disparidad acentúa la brecha salarial entre regiones. Sin embargo, un mensaje positivo es que en estas ciudades, intermedias y pequeñas, los retornos salariales a la inversión en educación son mayores que en Bogotá. Esto se deduce al aplicar la descomposición de *Blinder-Oaxaca* ya que, mientras el vector de dotaciones amplía la brecha salarial, el vector de coeficientes la reduce.

Palabras clave: Retornos a la educación, Economía regional, Primer empleo.

Clasificación JEL: J21, J24, J61.

¹Economista y Administrador de Empresas de la Universidad de los Andes. Documento de tesis para optar al título de Maestría en Economía de la misma universidad. (Email: rj.galindo87@uniandes.edu.co). Agradezco mucho la guía y paciencia de Ximena Peña. También tengo agradecimientos para Nathalie González, Christian Gómez, Ricardo Salas y Camila Uribe por sus comentarios, siempre pertinentes y generosos. Por último, muchas gracias a Ramón Rosales por el voto de confianza al aceptar dirigir este proyecto.

Wage Returns to Education: A Source of Regional Disparities?

Rodrigo José Galindo Sánchez²

Abstract

This research analyzes the relation between the returns to education and the wage gap among three clusters of Colombian cities, which are grouped according to the similarity of their labor markets. The results show that cities with lower wages should invest more in human capital because they are under the levels of Bogotá, and that disparity emphasizes the wage gap among the regions. However, a positive message is that in those cities, small and medium sizes, the wage returns to education are higher than those in Bogotá. This follows the *Blinder-Oaxaca* decomposition because, while the endowments' vector exacerbates the gap, the coefficients' vector reduces it.

Key words: Returns to education, Regional economy, First employment.

JEL Classification: J21, J24, J61.

² Economist and Business Manager of Los Andes University. Thesis document required for the Master's degree in Economics in the same university. (Email: rj.galindo87@uniandes.edu.co). I am very grateful to Ximena Peña because of her guidance and patience. I am also grateful to Nathalie González, Christian Gómez, Ricardo Salas and Camila Uribe because of their comments, always relevant and generous. Finally, thanks to Ramon Rosales for the vote of confidence by agreeing to lead this project.

1. Introducción

Colombia es un país con profundas disparidades económicas a nivel regional. Por ejemplo, Bogotá en 2013, con sólo el 16% de la población, aportó el 25% del PIB nacional. Otro indicador que refleja la desigualdad regional en materia económica es la tasa de desempleo pues, mientras que en Marzo de 2014 Bogotá reportó un nivel de desocupación del 8.8%, el valor agregado de las 23 ciudades más grandes de Colombia ascendió a 10.1%. Otras ciudades con desempeños económicos superiores a la media nacional son Medellín, Cali y Barranquilla.

Las disparidades regionales son nefastas para el desarrollo económico de un país en la medida que los recursos económicos se concentran en las áreas más productivas, lo que implica que las regiones con más bajos niveles de desarrollo no reciben la inversión necesaria para superar esta barrera. Asimismo, la falta de inversión también se traduce en bajos niveles de cobertura de servicios básicos, mala calidad en la educación y pocas oportunidades laborales, creando un círculo vicioso difícil de romper. Por otro lado, las políticas formuladas de manera centralizada, como en el caso colombiano en su mayoría, al no analizar las brechas de desarrollo, tienen menores impactos en las regiones menos adelantadas. Según Cepeda & Meisel (2013) y Meisel (2011) la forma más efectiva para reducir estas disparidades es invertir en capital humano en las regiones rezagadas.

El mercado laboral es la institución más importante para incentivar la acumulación de capital humano. En particular, sin unos retornos salariales a la educación lo suficientemente atractivos, los individuos elegirán no invertir en capital humano. Este razonamiento tiene respaldo en la teoría de Becker (1964) quien afirma que el capital humano es un *stock* que se incrementa por flujos de inversión en educación y que a cambio espera unos ingresos, en forma de salarios, que compensen dicha inversión. Otro planteamiento similar es el de Mincer (1958) pues afirma que la decisión de invertir en capital humano, para después ofrecerlo en el mercado de trabajo, corresponde a un proceso de optimización en el que los individuos maximizan el valor presente de los salarios futuros.

Las disparidades regionales en Colombia también se manifiestan a través de los salarios. Mientras que el salario promedio por hora en Bogotá durante Diciembre de 2013

era de \$5,571, este mismo indicador y en el mismo período de tiempo, fue de \$3,670 para el caso de Pereira. La presencia de una brecha salarial entre regiones crea una disparidad de incentivos para acumular capital humano que, a su vez, repercute en el desarrollo económico desigual entre las regiones colombianas. Dada la importancia de los salarios en la decisión de acumular capital humano, entender cuáles son los componentes detrás de la brecha regional es un paso necesario para dar recomendaciones de política pública enfocadas a reducir las brechas en materia de desarrollo que aquejan al país.

El primer empleo no se escapa de las disparidades regionales. En Diciembre de 2013, Pasto tenía un salario promedio por hora, para el caso del primer empleo, de sólo el 72% del nivel de Bogotá. Adicionalmente, el análisis del primer empleo también es relevante teniendo en cuenta que la primera experiencia laboral determina, en gran parte, la trayectoria profesional de los individuos. Si desde el primer empleo las personas se ubican en trabajos mal remunerados, no sólo se desincentiva la acumulación de capital humano, sino que también se restringe el desarrollo económico de la región pues los incentivos a innovar, optimizar procesos y, en general, a ser más productivos, estarán ausentes.

Otra consecuencia que podría tener unos bajos retornos a la educación en determinadas regiones es la migración, generalmente de los trabajadores más capacitados, hacia otras regiones que ofrezcan mejores salarios; lo que contribuiría a ampliar la brecha de desarrollo pues, mientras la ciudad receptora recibe los mejores estudiantes y trabajadores que contribuyen al crecimiento de la economía, la ciudad expulsora se queda con personas de bajos desempeños académicos y laborales que no aportan tanto al desarrollo de la región. Esta llamada “Fuga interregional de cerebros” está documentada por Cepeda (2012). A partir de lo anterior, los retornos salariales a la educación en las regiones rezagadas, también deben ser objeto de análisis para buscar mecanismos que motiven a las personas más capacitadas a permanecer en su región de origen.

El mercado laboral de los jóvenes responde a dinámicas diferentes que merecen ser estudiadas con mayor atención. De acuerdo con Sánchez et al. (2009), el mercado laboral colombiano está sesgado en contra de poblaciones vulnerables como los jóvenes. Al ser el colombiano un mercado laboral que sufre de rigideces e inflexibilidades que causan fallas como segmentación entre el sector formal e informal y desempleo [Véase Peña (2013),

Joumard & Londoño Vélez (2013), Mondragón-Vélez et al. (2011) y (2010)], el panorama para los jóvenes es incierto. Mientras que la tasa de desempleo en Diciembre de 2013 para las 13 áreas metropolitanas fue de 9.65%, este mismo indicador, sólo para personas entre 18 y 25 años, asciende a 17.04%. Si los jóvenes anticipan un mercado laboral en su contra, representado por salarios bajos y dificultad para conseguir empleo, la acumulación de capital humano estaría en riesgo y las desigualdades regionales se amplificarían.

El objetivo de este artículo es estimar si los retornos a la educación amplían o reducen la brecha salarial entre regiones para el caso del primer empleo. Si los retornos a la educación acentúan la brecha en lugar de reducirla, la acumulación de capital humano en las regiones rezagadas no estaría logrando su objetivo de reducir las disparidades económicas y se deberían buscar mecanismos que aumenten la rentabilidad de la educación en estas regiones. Por el contrario, si los retornos a la educación contribuyen a cerrar la brecha, las inversiones en capital humano por parte de los sectores público, privado y de la sociedad civil deberían mantenerse y extenderse.

Para responder el interrogante del título y estimar el rol de los retornos a la educación sobre la disparidad salarial regional, se conforman tres *clusters* de ciudades a partir de la similitud en sus mercados laborales. El siguiente paso consiste en utilizar la descomposición conocida como *Blinder-Oaxaca* (*B-O*, en adelante) para desagregar la brecha salarial de tal manera que se aisle y cuantifique el papel que tienen los retornos salariales sobre ésta.

Para cumplir con los objetivos trazados, el documento se divide en 6 secciones, siendo la primera esta introducción. La revisión bibliográfica resume brevemente los hallazgos más relevantes a nivel nacional e internacional para el tema que aquí se concentra. La tercera sección explica los datos utilizados en esta investigación, así como la selección de la muestra. La cuarta sección contiene una descripción del mercado laboral colombiano que amplía la información sobre las disparidades económicas regionales y el sesgo del mercado laboral en contra de los jóvenes. La quinta sección describe la estrategia empírica. La sexta sección presenta y discute los resultados hallados. Por último, el documento resume las principales conclusiones y ofrece algunas recomendaciones de política.

2. Revisión bibliográfica

Las teorías del crecimiento económico señalan a la acumulación de capital humano como el principal motor de crecimiento así como la fuente de divergencia económica más importante [Véase Mankiw et al. (1992), Romer (1990) y Lucas Jr (1988)]. Igualmente, la teoría microeconómica del desarrollo es otro respaldo teórico para el vínculo entre capital humano y desarrollo económico: Sen (1999) afirma que un requisito para el desarrollo es la educación, puesto que esta crea una oportunidad para que las personas desarrollen todas sus capacidades, prosperen como individuos y logren que sus regiones también lo hagan.

En concordancia con lo anterior, la evidencia internacional revela una estrecha relación entre capital humano y desarrollo económico. Acemoglu & Dell (2010) encuentran que cerca de la mitad de las diferencias en ingreso y productividad entre regiones de un mismo país se debe a los niveles asimétricos de acumulación de capital humano. En esa misma línea se ubica el trabajo de Gennaioli et al. (2013) que usa datos para 1,569 regiones sub-nacionales, provenientes de 110 países, cubriendo el 74% de la superficie terrestre y el 96% del PIB mundial. En dicho artículo se concluye que las diferencias en niveles de capital humano son determinantes en las brechas de desarrollo entre regiones sub-nacionales.

El análisis de la brecha salarial dentro un mismo país ha sido ampliamente documentada en países ibéricos. López-Bazo & Motellón (2009) y Motellón et al. (2009) encuentran evidencia de una brecha salarial en España donde las regiones más desarrolladas tienen retornos salariales más altos que las menos adelantadas. Galego & Pereira (2013) encuentran resultados similares para Portugal. Estos artículos también tienen en común la aproximación empírica que utilizan. El uso de la descomposición *B-O* les permite desagregar la brecha regional entre un vector de dotaciones y un vector de coeficientes. Estos tres artículos concluyen que las brechas salariales no sólo se explican porque en las regiones más desarrolladas hay más personas con estudios superiores o, en general, hay más años de educación entre la población (Vector de dotaciones) sino también porque en estas regiones los retornos a la educación son más altos (Vector de coeficientes).

Las brechas económicas entre regiones colombianas han sido estudiadas desde diferentes perspectivas. Cepeda & Meisel (2013) argumentan que las disparidades

regionales tienen un origen institucional y otro geográfico. Las regiones con instituciones inclusivas, varias de ellas originadas en la época de la colonia, son más prósperas en la actualidad que las zonas donde primaron las instituciones extractivas como la esclavitud. Por otro lado, los lugares con climas y topografías aptos para el desarrollo agrícola también tienen un desempeño económico superior que aquellas regiones con una geografía agreste. Por su parte, Galvis & Meisel (2010) discuten las disparidades regionales a partir de un enfoque espacial; dicho trabajo concluye que en Colombia las regiones más pobres son cercanas geográficamente y se tienden a ubicar en la periferia del territorio. Por su parte, Bonet & Meisel (2009) encuentran que hay un proceso de polarización económica entre Bogotá y el resto de los departamentos. Finalmente, Meisel & Romero (2007) concluyen que las políticas centralizadas no funcionan para resolver las disparidades regionales y, por el contrario, varias de ellas las han exacerbado. Esta última conclusión viene acompañada de un llamado para formular políticas públicas diferenciadas entre regiones.

Los estudios sobre disparidad salarial entre regiones colombianas han sido muy escasos ya que las investigaciones que abordan el mercado laboral, en su mayoría, no discuten las dinámicas particulares que tienen las distintas regiones del país. Lo anterior es un vacío que se debe llenar puesto que, ante la ausencia de análisis regionales, se reduce la posibilidad de formular políticas públicas diferenciadas y, sobre todo, acertadas. Este artículo es una contribución a la bibliografía económica colombiana porque primero, analiza y discute la disparidad de salarios entre regiones colombianas y segundo, a diferencia de la mayoría de estudios que documentan la disparidad económica regional y sugiere la acumulación de capital humano como política eficaz para reducirla, esta investigación analiza, para el caso del primer empleo, si la inversión en educación realmente está contribuyendo a cerrar la brecha regional y en qué magnitud lo está haciendo.

3. Datos y muestra

3.1. Datos

Los datos utilizados en esta investigación provienen de la Gran Encuesta Integrada de Hogares (GEIH), recolectada por el Departamento Administrativo Nacional de Estadística (DANE). Esta encuesta se aplica a una muestra aleatoria de hogares cada mes y recoge información sobre las características físicas de la vivienda, información sociodemográfica

para cada uno de los miembros del hogar e información sobre empleo para las personas en edad de trabajar (PET). Asimismo, la GEIH es representativa para las 13 principales áreas metropolitanas³ de Colombia, las mismas que serán analizadas a lo largo del documento. Finalmente, para el período objeto de análisis, 2007 a 2013⁴, este estudio tiene en cuenta los cuatro trimestres del año. Si bien en el primer y último trimestre de cada año se tienen evidencias de estacionalidad en el empleo, dichos problemas son en cantidades, no en precios. Teniendo en cuenta la estructura de la encuesta y el periodo de tiempo bajo análisis, la base de datos usada se cataloga como corte transversal repetido.

3.2. Muestra

El primer criterio de selección de muestra obedece al enfoque de comparación regional de esta investigación pues sólo serán tenidas en cuenta las personas ubicadas en las 13 áreas metropolitanas que analiza el DANE, a saber, Barranquilla, Bogotá, Bucaramanga, Cali, Cartagena, Cúcuta, Ibagué, Manizales, Medellín, Montería, Pasto, Pereira y Villavicencio.

Debido a que esta investigación busca estimar la participación de los retornos a la educación⁵ sobre las disparidades regionales para el caso del primer empleo, los análisis estadísticos y las estimaciones econométricas⁶ se centrarán únicamente en personas que, para el momento que fueron encuestadas, se encontraban en su primer empleo o en la búsqueda del mismo. Asimismo, las personas que ya se encontraban ocupadas en su primer empleo, serán restringidas a aquellas que lleven menos de un año trabajando. Este refinamiento permite analizar de forma más precisa el efecto de la educación sobre el salario, precisamente porque la población sujeta de análisis sólo puede competir por diferencias en niveles educativos y no hay experiencia laboral que afecte la determinación del salario.

³ Durante el documento se usa varias veces el término de ciudades; no obstante, estas incluyen siempre las áreas metropolitanas asociadas a la ciudad.

⁴ Existen 3 razones para seleccionar este periodo de tiempo. La primera es porque el objetivo de esta investigación es entregar resultados con información actualizada. La segunda es porque para períodos anteriores no se podía distinguir entre educación técnica y tecnológica, distinción que es necesaria y relevante para este trabajo. La última razón se debe a que algunos cambios metodológicos entre la Encuesta Continua de Hogares (ECH) y la GEIH pueden tener implicaciones sobre los resultados del estudio.

⁵ Durante todo el documento se utilizan los salarios en efectivo como retornos salariales. A partir de lo anterior, sólo se tienen en cuenta empleados, se descartan los trabajadores cuenta propia o independientes.

⁶ En todas las estimaciones se usan pesos de frecuencia a nivel trimestral pues es a ese nivel que son representativas las 13 áreas metropolitanas.

Otra condición de la muestra es que sólo se incluyen en los análisis personas entre 18 y 28 años de edad. El límite inferior se debe a que el trabajo infantil responde a dinámicas distintas a las que acá se abordan. La selección del límite superior se debe a que, de acuerdo al Ministerio de Educación Nacional, una persona promedio debe concluir su pregrado a los 22 años y, si se suman 6 años como máximo de educación de posgrado, el resultado es 28 años.

En la muestra bajo análisis también se excluyen a las personas analfabetas o que no hayan concluido la educación primaria. Del mismo modo, también se omiten observaciones que reportan trabajar más de 96 horas a la semana y aquellas que tienen un salario por hora superior a \$250,000, a precios de Diciembre de 2013.

Si bien la representatividad de la muestra puede ser motivo de una extensa discusión, estos refinamientos son necesarios en la medida que aíslan otros factores que afectan la relación entre educación y salario como el caso de la experiencia laboral, una variable difícil de cuantificar y con extensa variabilidad entre individuos. En conclusión, los resultados de esta investigación son válidos únicamente para personas sin experiencia laboral.

4. Caracterización del mercado laboral colombiano para los jóvenes

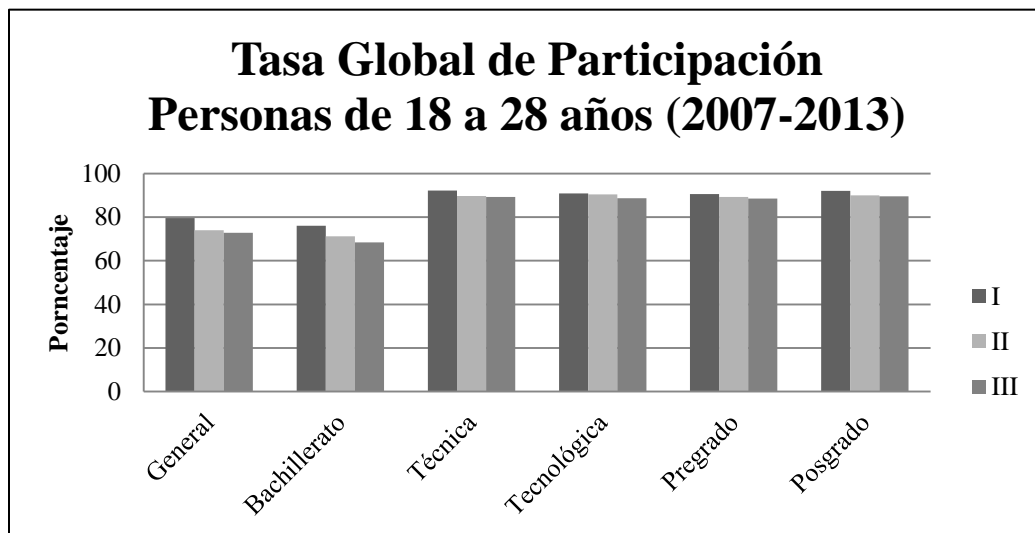
Debido al enfoque regional del documento, el primer paso de esta investigación es agrupar las 13 áreas metropolitanas en tres grupos según la similitud de sus mercados laborales y de esta forma facilitar los análisis reduciendo el número de comparaciones a realizar. Los indicadores que fueron calculados y que sirvieron de inputs para el análisis de *clusters*⁷ son: promedio del salario real por hora, tasa global de participación, tasa de desempleo, coeficiente de Gini de los salarios reales por hora, promedio de los años de educación en el área metropolitana, proporción de los puestos de trabajo en cada uno de los sectores económicos mencionados previamente y población para diciembre de 2013. Cabe resaltar que todos estos cálculos se restringieron a las personas pertenecientes a la muestra objetivo. El resultado de esta asociación es el siguiente (Ver Anexo 1):

⁷ Este análisis agrupó los *clusters* de acuerdo con el promedio de las variables incluidas como *inputs*. Los soportes estadísticos de estos análisis están disponibles por solicitud, en este documento no se incluyen por motivos de espacio.

<i>Cluster</i>	Ciudad
I	Bogotá D.C
	Medellín
II	Cali
	Barranquilla
	Bucaramanga
III	Cartagena
	Cúcuta
	Pereira
	Ibagué
	Villavicencio
	Manizales
	Pasto
	Montería

La tasa de participación laboral es la proporción de personas en edad de trabajar que efectivamente están trabajando o buscando empleo. A partir de la Gráfica 1 se puede afirmar que la proporción de personas en edad de trabajar que participan en el mercado laboral es alrededor de 5 puntos porcentuales más alta en el *cluster* I que en las ciudades medianas y en las pequeñas. Esta misma proporción es muy similar entre el *cluster* II y el III. Asimismo, en los tres *clusters* se observa que entre mayor sea el nivel educativo, la propensión a participar en el mercado laboral aumenta sistemáticamente.

Gráfica 1

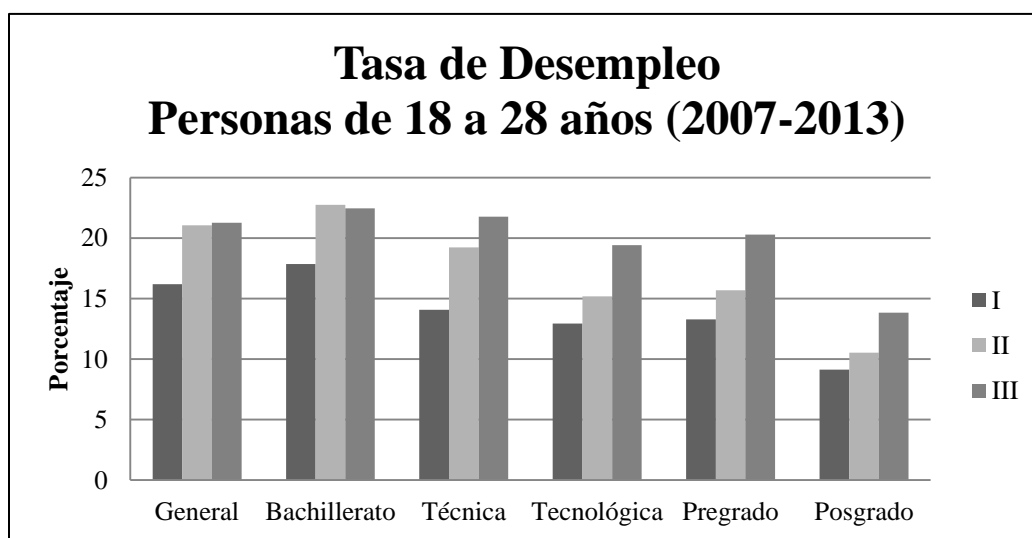


Fuente: GEIH-DANE. Cálculos propios.

La Gráfica 2 contiene la tasa de desempleo según las mismas clasificaciones de la Tabla 1. La tasa de desempleo es la proporción de personas que ofrecen su trabajo pero no

consiguen empleo. Este indicador revela que en Bogotá es más factible conseguir empleo que en las ciudades medianas y pequeñas; dicha diferencia es cercana a los 5 puntos porcentuales. Nuevamente, los resultados para el grupo II y III son similares entre sí. También se puede observar que la tasa de desempleo cae sistemáticamente a medida que aumenta el nivel educativo; este resultado es una señal de los beneficios positivos de invertir en educación.

Gráfica 2



Fuente: GEIH-DANE. Cálculos propios.

Los resultados de la Tabla 1 exhiben la relación directa entre nivel educativo y salario ya que entre más alto es el nivel educativo, mayor es el salario promedio que se percibe. Asimismo, se observa que los salarios más altos, para cada uno de los niveles educativos, los ofrece Bogotá. La proporción de personas en cada nivel educativo revela una tendencia similar, es decir, al comparar Bogotá con las ciudades medianas y pequeñas, es la capital colombiana la que tiene mayor proporción de personas educadas para todos los niveles educativos, a excepción de la educación técnica y tecnológica. Después de Bogotá siguen las ciudades medianas y por último, las pequeñas. En resumen, la Tabla 1 muestra tres resultados relevantes para esta investigación. Primero, ciudades más grandes en población y producción económica ofrecen mejores salarios para todos los niveles educativos. Segundo, estas mismas ciudades tienen mayor proporción de personas educadas para todos los niveles educativos. Tercero, existe una relación directa entre nivel educativo y salario para los tres grupos de ciudades bajo análisis.

Cluster	Bachillerato completo		Técnica completa		Tecnológica completa		Pregrado completo		Posgrado completo	
	%	SR	%	SR	%	SR	%	SR	%	SR
I	83.65	\$ 4,043	12.15	\$ 3,763	5.69	\$ 4,081	12.3	\$ 8,494	4.1	\$ 9,669
II	83.76	\$ 3,539	13.33	\$ 3,599	6.17	\$ 3,830	9.7	\$ 7,309	2.37	\$ 9,502
III	81.01	\$ 3,298	11.95	\$ 3,475	6.5	\$ 3,717	9.1	\$ 6,406	1.82	\$ 8,255
Notación: %: Proporción de la muestra objetivo que tiene el nivel educativo. SR: Salario real por hora, precios de diciembre de 2013.										
Fuente: GEIH-DANE. Cálculos propios.										

5. Estrategia empírica

La estrategia empírica parte de la ecuación de salarios propuesta por Mincer (1974) que consiste en explicar el salario a partir del nivel educativo y la experiencia laboral. Cabe resaltar que la variable dependiente es el logaritmo⁸ del salario real⁹ por hora.

Basado en Tenjo (1993) y Prada (2006), como variable explicativa de la educación se utilizan *splines* lineales por cada nivel educativo relevante para este estudio, a saber, bachillerato, educación técnica, educación tecnológica, pregrado y posgrado. Los *splines* lineales capturan la recompensa por culminar satisfactoriamente el ciclo completo de cada nivel educativo. La presencia de estas variables en el modelo se justifica porque en el ranking salarial de muchos empleadores, entre ellos el gobierno, se especifica la remuneración del trabajador sujeta a los títulos educativos logrados, lo que se puede interpretar como un salto discontinuo que debe ser capturado con los *splines* lineales. En términos econométricos, los *splines* lineales son tantas variables binarias como niveles

⁸ La transformación logarítmica es válida y útil porque el logaritmo suaviza la función de los salarios, pero conserva el orden de la distribución. Asimismo, según Zárate (2003), la transformación logarítmica, primero, aproxima la distribución a la curva normal, lo cual es deseable por los supuestos econométricos necesarios para las estimaciones y, segundo, está catalogada como la mejor transformación en la clase de las Box-Cox. Las transformaciones Box-Cox corrigen parte de la heteroscedasticidad del modelo, mejoran la correlación entre variables y aproximan la distribución de los errores a la curva normal.

⁹ (Base: Diciembre de 2013=100). Sería deseable incorporar índices de paridad de poder adquisitivo para que los salarios no estén afectados por las diferencias en el costo de vida entre las ciudades colombianas; no obstante, estos índices no existen para Colombia. Sin embargo, dos son las razones para argumentar que la falta de estos índices no es un problema crítico para la presente investigación. Primero, Barón (2002) encuentra una convergencia en los niveles de paridad de poder adquisitivo (PPA) de las regiones colombianas; si bien dicho artículo toma un periodo de tiempo distinto al de la presente investigación, teniendo en cuenta la baja inflación en los períodos recientes en todas las ciudades, es de esperarse que esta tendencia a converger no se haya revertido. Segundo, la variable dependiente es un logaritmo, no un nivel, por lo que no afecta tanto la diferencia en PPA.

educativos disponibles que toman el valor de 1 cuando la persona culminó el nivel educativo correspondiente a esa *dummy* y 0 en caso contrario. Adicionalmente, se incluye una variable que acompaña a cada *spline* y contiene los años de educación extra que tiene el individuo del siguiente nivel al máximo alcanzado, para captar el hecho de tener estudios superiores inconclusos, aunque valorados parcialmente en el mercado de trabajo.

Otra modificación al lado derecho de la ecuación es que no se incluye ni la experiencia ni su término cuadrático pues para todas las observaciones de la muestra es cero dado que sólo se toma información del primer empleo. Como variable independiente adicional se incluye el sexo ya que de lo contrario se estaría omitiendo la bien documentada brecha salarial de género en Colombia [Véase Badel & Peña (2010) y Hoyos et al. (2010)]. Otra variable que se incluye al lado derecho de la ecuación es la edad pues, al no haber experiencia laboral, los empleadores pueden valorar los años de vida como un indicador de madurez o experiencia y darle un *premium* al salario por este motivo¹⁰.

Por otro lado, la remuneración salarial está altamente relacionada con el sector económico al que pertenece un trabajador; esto se debe, entre otros, a que los sectores asociados a un *boom* económico o a labores que exigen habilidades muy puntuales ofrecen salarios más altos que otros sectores, para un mismo nivel educativo. La estrategia empírica incorpora este hecho al incluir variables categóricas para el sector minero, agrícola, industrial y de construcción que toman el valor de 1 si el trabajador pertenece a ese sector y 0 de lo contrario. Para evitar multicolinealidad perfecta se excluye la *dummy* que identifica al sector de servicios.

Otra modificación frente a la propuesta original de Mincer (1974) es incluir un índice inverso de Mills como regresor de la ecuación de salarios. En el método propuesto por Heckman (1979), incluir el índice de Mills corrige el sesgo de selección muestral causado por la negativa de algunas personas a participar del mercado laboral cuando el salario de mercado es inferior a su salario de reserva. Teniendo en cuenta la muestra seleccionada para este estudio, la corrección de Heckman cobra aún más relevancia porque si hay un

¹⁰ No se incluye la transformación cuadrática de la edad porque en la muestra bajo análisis el máximo valor de ésta variable es 28 años, por lo que aún no deben aparecer los rendimientos marginales asociados a la edad. Además, esta variable tampoco se incluye en el modelo original de Mincer (1974) ni en las recomendaciones de Lemieux (2003).

segmento que puede permitirse autoexcluirse de participar en el mercado laboral son los jóvenes recién graduados, sin obligaciones ni responsabilidades y respaldados aún por sus acudientes, al considerar que el salario de mercado es inferior a su salario de reserva. Si se omitiera esta variable, el modelo sufriría de endogeneidad por sesgo de selección.

Basado en el planteamiento de Arango & Posada (2002) la ecuación de participación laboral a partir de la que se calcula el inverso del índice de Mills incluye las siguientes variables como regresoras: años de educación, edad, edad al cuadrado, una variable categórica que indica si el individuo está casado o vive en unión libre, otra variable categórica que indica si el individuo estaba estudiando en el período que fue aplicada la encuesta, como *proxy* de riqueza se incluye el valor del arriendo de la vivienda donde vive el individuo o el valor estimado del mismo si la vivienda es propia, las últimas dos variables que se incluyen son la cantidad de menores y la cantidad de desocupados en el hogar. Estas estimaciones se realizan para cada área metropolitana, año y trimestre entre 2007 y 2013, así que también se incluyen efectos fijos de mes. Una vez se calcula la probabilidad predicha de pertenecer a la fuerza laboral, esta se multiplica por el converso de la tasa de desempleo en el trimestre al que pertenece la observación pues ésta ponderación da la probabilidad de que una persona que ofrece su trabajo, efectivamente consiga un empleo y se le pueda observar un salario. Con base en esta nueva probabilidad se calcula el inverso del índice de Mills¹¹. En resumen, la ecuación por estimar es:

$$\begin{aligned} \left[\text{Ln} \left(\frac{w_i}{h_i} \right) \right] = & \beta_0 + \beta_1 D_i^{\text{Bachillerato}} + \beta_2 D_i^{\text{Técnica}} + \beta_3 D_i^{\text{Tecnológica}} + \beta_4 D_i^{\text{Pregrado}} \\ & + \beta_5 D_i^{\text{Posgrado}} + \beta_6 \text{Extra}_i^{\text{Primaria}} + \beta_7 \text{Extra}_i^{\text{Bachillerato}} \\ & + \beta_8 \text{Extra}_i^{\text{Técnica}} + \beta_9 \text{Extra}_i^{\text{Tecnológica}} + \beta_{10} \text{Extra}_i^{\text{Pregrado}} \\ & + \beta_{11} \text{Extra}_i^{\text{Posgrado}} + \beta_{12} \text{Sexo}_i + \beta_{13} \text{Edad}_i + \overrightarrow{\beta_{14} \text{sector}} + \beta_{15} \lambda_i \\ & + \overrightarrow{\beta_{16}} \sum_{t=2}^4 \text{Trim}_t + \overrightarrow{\beta_{17}} \sum_{t=2008}^{2013} \text{Año}_t + \overrightarrow{\beta_{18} \text{area}} + \epsilon_i \end{aligned}$$

¹¹ El índice de Mills se obtiene a partir de la probabilidad estimada (z_i), a través de un Probit, de que una persona en edad de trabajar decida participar en la fuerza laboral. Una vez calculada dicha probabilidad, se multiplica por el converso de la tasa de desempleo y, finalmente, el índice de Mills resulta de reemplazar la ecuación $\lambda_i = \frac{[\phi(z_i)]}{[1-\Phi(-z_i)]}$ siendo ϕ y Φ la función de densidad y de acumulación de una distribución normal, respectivamente.

Siendo w el salario mensual deflactado por IPC (Base Diciembre de 2013=100), h las horas mensuales que reporta trabajar normalmente, $D_i^{Bachillerato}$ una variable binaria que toma el valor de 1 si la persona se graduó de bachillerato y 0 en caso contrario, $D_i^{Técnica}$ una variable binaria que toma el valor de 1 si el individuo posee un grado técnico y 0 en caso contrario, $D_i^{Tecnológica}$ una variable binaria que toma el valor de 1 si el individuo alcanzó un grado tecnológico y 0 en caso contrario, $D_i^{Pregrado}$ una variable binaria que toma el valor de 1 si la persona se graduó del pregrado y 0 en caso contrario, $D_i^{Postgrado}$ una variable binaria que toma el valor de 1 si la persona se graduó de algún posgrado y 0 en caso contrario, $Extra_i^k$ los años de educación que tiene el individuo por encima del último diploma oficial k que obtuvo, $Sexo_i$ una variable binaria que toma el valor de 1 para las mujeres, $Edad_i$ los años cumplidos, \overrightarrow{sector} es un vector de variables categóricas que indican en qué sector de la economía, de los señalados previamente, trabaja el individuo y λ_i el inverso del índice de Mills. Las variables $Trim_t$, $Año_t$ y $area_i$ capturan los efectos fijos de trimestre, año y ciudad, respectivamente.

Con base en la ecuación anterior se aplica la descomposición *B-O* para aislar el vector de coeficientes del vector de dotaciones y observar el rol de los retornos salariales sobre la brecha regional de los salarios. La razón para prestar tanto detalle a la ecuación de salarios y sus variables explicativas es que esta debe cumplir exogeneidad estricta puesto que de lo contrario el vector de coeficientes, que se construye a partir de la porción no observable de la educación de salarios, estaría sesgado. Es por esto que la selección de las variables explicativas de la ecuación de salarios responde tanto a un análisis del mercado laboral colombiano, como a investigaciones teóricas y empíricas anteriores.

La brecha salarial¹² a nivel regional se origina ante una diferencia significativa en los salarios entre dos *clusters* de ciudades:

$$[1] \quad G_{jvs,k} = E(W_j) - E(W_k); \quad \forall j \neq k \text{ y } j, k \in (I, II, III)$$

Siendo $E(W_j)$ el valor esperado del logaritmo del salario por hora de cada *cluster* y $G_{jvs,k}$ la brecha entre cada pareja de grupos de ciudades. Como la descomposición *B-O* sólo

¹² La explicación de la descomposición *B-O* que aquí se presenta está basada, principalmente, en el trabajo de Jann (2008). Para una explicación más clara y detallada se sugiere la lectura de ese documento.

compara dos grupos a la vez, el procedimiento que aquí se expone se realiza tres veces, para comparar los *cluster* I con II, II con III y I con III.

La ecuación de salarios propuesta se puede resumir como $W_j = X_j^T \beta_j + \epsilon_j$. Teniendo en cuenta lo anterior y que $E(\epsilon_j) = 0$, la brecha entre dos *clusters*, aplicando valor esperado a ambos lados de la ecuación de salarios, se puede reescribir como:

$$[2] \quad G_{jvs,k} = E(X_j)^T \beta_j - E(X_k)^T \beta_k$$

A su vez, esta última expresión se puede desarrollar de tal forma que se aislen los componentes de la brecha entre los dos grupos:

$$[3] \quad G_{jvs,k} = [E(X_j) - E(X_k)]^T \beta_k + E(X_k)^T (\beta_j - \beta_k) + [E(X_j) - E(X_k)]^T (\beta_j - \beta_k)$$

Siendo $[E(X_j) - E(X_k)]^T \beta_k = D$ la porción de la brecha explicada por la diferencia en dotaciones, $E(X_k)^T (\beta_j - \beta_k) = C$ el componente de la brecha asociado a la diferencia en coeficientes y $[E(X_j) - E(X_k)]^T (\beta_j - \beta_k) = I$ la interacción entre los dos componentes anteriores ya que estos ocurren de manera simultánea.

Como el objetivo principal de esta investigación es analizar el papel de los retornos a la educación sobre la brecha regional de salarios, el interés sobre el componente D se limita a analizar qué magnitud de la brecha salarial entre los *clusters* j y k se debe a que en el grupo j haya más capital humano¹³ frente al nivel del grupo k . Asimismo, el foco de atención del componente C es el retorno de la educación, es decir, estimar qué tan diferente es la recompensa salarial por cada nivel educativo entre los dos grupos de ciudades bajo análisis. Finalmente, la participación del componente I es más compleja de analizar debido a que combina efectos de ambos vectores. Teniendo en cuenta el objetivo de este trabajo, los componentes relevantes son D y C¹⁴.

¹³ Esta expresión se refiere a que hayan más personas con estudios superiores en el grupo j que en el k o que, en promedio, dos personas de diferentes grupos sean estadísticamente iguales y sólo se diferencien en que la del grupo j tiene más años de educación que la del grupo k .

¹⁴ La descomposición *B-O* se puede interpretar como desagregar la brecha entre componentes observables y no observables. En este enfoque, la parte observable corresponde al vector de dotaciones y una parte de la interacción, mientras que el vector de coeficientes son los no observables y el resto de la interacción.

6. Resultados

Los resultados más relevantes para esta investigación se resumen en los parámetros de los *splines* de las regresiones mincerianas y en los componentes asociados a estos mismos en los vectores de dotaciones y coeficientes de la descomposición *B-O*.

Antes de presentar los resultados, es necesario clarificar la interpretación de los coeficientes asociados a la educación de un individuo¹⁵. β_1 es el retorno salarial de haber completado satisfactoriamente el bachillerato frente a no haberlo hecho, β_2 , β_3 , β_4 y β_5 hacen lo propio para los títulos técnicos, tecnológicos, de pregrado y de posgrado, respectivamente. Así, el retorno salarial de tener un título técnico es $\beta_1 + \beta_2$, el de un título tecnológico es $\beta_1 + \beta_2 + \beta_3$, el de un título de pregrado es $\beta_1 + \beta_4$ y el de un título de posgrado es $\beta_1 + \beta_4 + \beta_5$.

Tabla 2 – Retorno salarial por tener el nivel educativo completo					
<i>Cluster</i>	Bachillerato completo	Técnica completa	Tecnológica completa	Pregrado completo	Posgrado completo
I	0.1574***	0.3662***	0.4547***	0.9804***	1.1027***
II	0.2645***	0.5166***	0.6203***	1.0447***	1.3397***
III	0.3139***	0.5384***	0.6429***	1.0542***	1.3539***
***Significativos al 1%.					
Fuente: GEIH-DANE. Cálculos propios.					

La Tabla 2 resume los parámetros de los *splines* lineales, es decir, los retornos salariales a la educación, obtenidos por el método de Mínimos Cuadrados Ordinarios (MCO) de la regresión de Mincer en cada uno de los *clusters*. En primer lugar, los retornos de todos los niveles educativos son positivos y significativos en los 3 grupos de ciudades. Este es un resultado acorde con la evidencia empírica y la intuición económica.

¹⁵ En la configuración de los *splines* lineales se incorporan varios hechos sobre la secuencia de inversión en educación. En particular, hay que tener presente que para tener un título tecnológico se requiere primero haber completado un nivel técnico y antes el bachillerato. A su vez, para tener un título de posgrado primero se necesita uno de pregrado y previamente el título de bachiller. Otro elemento fundamental para entender esta interpretación es que ni el título de educación técnica ni el de educación tecnológica son prerrequisitos para los estudios de pregrado ni de posgrado.

Del resultado anterior se deduce que los retornos a la educación, sin excepción entre los *clusters*, son crecientes en el nivel educativo de las personas. Otra forma de interpretar este resultado es que entre más alto sea el último nivel educativo del individuo, mayor es la distancia relativa entre su salario y el de una persona que sólo concluyó la básica primaria.

La Tabla 2 anticipa un poco los resultados de la descomposición *B-O*. Las ciudades del grupo III, seguidas de las del grupo II, son las que ofrecen retornos salariales más altos, para cada nivel educativo, con relación al salario de la línea base. De hecho, al practicar pruebas de hipótesis, usando el estadístico t para combinaciones lineales, para verificar si las diferencias en los retornos a la educación son significativas, el resultado es que, en general, sí lo son, como se puede ver en la Tabla 3.

Tabla 3 – Prueba de hipótesis de las diferencias en los retornos a la educación					
<i>Clusters</i>	Bachillerato completo	Técnica completa	Tecnológica completa	Pregrado completo	Posgrado completo
	p-valor				
I vs. II	0.002	0.001	0.001	0.137	0
II vs. III	0.022	0.321	0.358	0.657	0.673
I vs. III	0	0	0	0.034	0

Fuente: GEIH-DANE. Cálculos propios.

La descomposición *B-O* parte de la ecuación de salarios propuesta y divide la brecha entre los tres vectores que se explicaron previamente. Como se había anticipado, teniendo en cuenta el objetivo de esta investigación, los elementos de la descomposición que serán analizados son los asociados con los niveles educativos y los retornos salariales de los mismos. La Tabla 4 presenta los resultados de la descomposición de *B-O*.

La Tabla 4 permite hacer una serie de análisis y conclusiones sobre las disparidades regionales de salarios. En primer lugar, se aprecia que en las tres comparaciones, a saber, entre los *clusters* I con II, II con III y I con III, la diferencia entre los salarios (en logaritmos) es significativa ante pruebas estadísticas. En segundo lugar, se observa que las tres brechas analizadas tienen origen tanto en el vector de dotaciones como en el de coeficientes.

Tabla 4 – Descomposición <i>Blinder-Oaxaca</i>				
		$I_A vs. II_B$	$II_A vs. III_B$	$I_A vs. III_B$
Resumen	$E \left[\ln \left(\frac{w_i}{h_i} \right) \right]_A$	7.9654	7.8617	7.9654
	$E \left[\ln \left(\frac{w_i}{h_i} \right) \right]_B$	7.8617	7.7761	7.7761
	Diferencia A-B	0.1037****	0.0855****	0.1892****
Vector de Dotaciones	Bachillerato	-0.0002****	0.0095****	0.0091****
	Técnica	-0.0029****	0.0033****	0.0004****
	Tecnológica	-0.0004****	-0.0003****	-0.0009****
	Pregrado	0.0196****	0.0044****	0.0236****
	Posgrado	0.0061****	0.0017****	0.0074****
Vector de Coeficientes	Bachillerato	-0.076****	-0.0772****	-0.1507****
	Técnica	-0.0028****	0.0006****	-0.0018****
	Tecnológica	-0.001****	-0.0013****	-0.0024****
	Pregrado	0.0068****	0.0017****	0.0081****
	Posgrado	-0.0051****	0.0005****	-0.0034****
*** Significativos al 1%. Estadístico z para pruebas de hipótesis.				
Fuente: GEIH-DANE. Cálculos propios.				

La diferencia en niveles de los salarios no es un resultado sorprendente sino evidente a partir de las estadísticas descriptivas. No obstante, estas diferencias en niveles no debe ser motivo de confusión con los resultados de la descomposición *B-O*. Como se verá más adelante, las ciudades medianas y pequeñas (*clusters* I y II) tienen retornos a la educación

más altos, es decir, la combinación lineal de parámetros que acompañan a los niveles educativos arroja valores mayores que para el caso de Bogotá. Siendo así, las razones para que Bogotá tenga salarios más altos son otras, distintas a los retornos a la educación.

Al analizar la diferencia salarial entre el grupo I y el II se observa que los componentes del vector de dotaciones alusivos a la educación no tienen una tendencia clara. Si el *cluster* II tuviera la misma cantidad de personas con educación de bachillerato, técnica y tecnológica que el grupo I, los salarios del grupo II serían menores que su nivel actual. Por el contrario, si el *cluster* II tuviera la misma cantidad de ocupados con niveles de educación universitarios y posteriores que el *cluster* I, los salarios en el grupo II serían más altos que su nivel actual. En otras palabras, los componentes del vector de dotaciones asociados a los *splines* de bachillerato, educación técnica y educación tecnológica ayuda a reducir la brecha mientras que los parámetros asociados con los *splines* de educación superior la amplían.

Por el contrario, todos los componentes de los *splines* lineales pertenecientes al vector de coeficientes ayudan a reducir la brecha de salarios entre los grupos I y II, con excepción del caso de la educación de pregrado. Otra forma de interpretar este resultado es que la rentabilidad de la inversión en educación es superior en las ciudades del grupo II que en las del grupo I al comparar los salarios de las personas en los distintos niveles educativos con respecto a los niveles salariales de la línea base.

Una posible explicación detrás de estos dos resultados son las leyes de oferta y demanda. En particular, para el caso de la educación técnica y tecnológica los hallazgos sugieren que este tipo de educación es más valorada en las ciudades intermedias que en Bogotá, un hecho que se puede observar en los diferenciales en las tasas de desempleo para este nivel educativo. El caso de la educación superior no tiene una explicación tan clara porque por el lado del vector de dotaciones se observa que en las ciudades del grupo II no sería del todo deseable tener cantidades de personas con estudios superiores similares a Bogotá, lo cual conlleva a pensar que estas cualificaciones no son tan valoradas en las ciudades intermedias. Sin embargo, el vector de coeficientes revela que es más rentable tener estos niveles educativos en las ciudades intermedias que en Bogotá, aunque sólo para personas con estudios de posgrado.

Otra manera de interpretar los resultados es que el *cluster* II tiene capacidad para aumentar sus salarios, hacia el nivel del *cluster* I, al incrementar la cantidad de personas que tienen educación universitaria. En otras palabras, las ciudades intermedias aún tienen margen para elevar su nivel salarial mediante el aumento en la cobertura de los niveles de educación superior. Con los resultados del vector de coeficientes y la Tabla 2, se puede afirmar que estas inversiones serán rentables y bien remuneradas.

La evidencia que ofrece la descomposición *B-O* ante la brecha entre las ciudades del *cluster* II y III es menos concluyente que en el caso anterior. En cuanto al vector de dotaciones, mientras los parámetros de los *splines* de los niveles de bachillerato y educación técnica contribuyen a reducir la brecha, los demás componentes educativos de este vector la amplían. Del mismo modo, el caso del vector de coeficientes tampoco tiene una tendencia clara puesto que mientras los retornos a la educación de bachillerato y tecnológica reduce la brecha, los premios salariales a la educación técnica, de pregrado y posgrado la amplían. Una posible razón para esta falta de contundencia en los resultados es la similitud en los mercados laborales de estos dos *clusters* de ciudades.

Finalmente, al comparar los resultados del grupo I con el grupo III se observa que los componentes educativos del vector de dotaciones amplían la brecha, mientras que los mismos componentes en el vector de coeficientes la reducen. De nuevo, el canal detrás de este resultado podría ser las leyes del mercado. Bogotá tiene abundancia de personas en todos los niveles educativos por lo que ante este exceso de oferta el precio del trabajo es más bajo que en las ciudades pequeñas donde hay una escasez relativa de personas que han invertido en capital humano.

Asimismo, al comparar Bogotá con las ciudades pequeñas se puede afirmar que estas últimas tienen margen para elevar sus salarios mediante el aumento en la cobertura de todos los niveles educativos, a excepción de la educación tecnológica. Al observar el vector de coeficientes y los resultados de la Tabla 2, se puede deducir que dichas inversiones serán recompensadas con premios salariales más altos que en el caso de Bogotá.

Dos ejercicios adicionales ante estos resultados son, primero, analizar el rol de la segmentación entre el sector público y privado sobre la brecha regional de salarios y,

segundo, dividir la distribución de salarios entre cuantiles para analizar si los resultados se ven afectados según esta división.

El primer ejercicio es intuitivo en la medida que el sector público tiene un tamaño relativo mayor en las ciudades pequeñas que en lugares como Bogotá. Por otro lado, en ciudades pequeñas donde no hay presencia de empresas de gran envergadura, es de esperarse que los mejores salarios los pague el gobierno, de ahí que merezca un análisis particular. No obstante, los resultados de este ejercicio no muestran que segmentar la muestra entre sector público y privado tenga una consecuencia sobre la brecha salarial.

El segundo ejercicio consiste en poner a prueba el supuesto de homogeneidad en la distribución de salarios y analizar si distintas porciones de la distribución arrojan resultados diferentes. Los resultados de este ejercicio, si bien cambian en algunos casos frente a los resultados principales, no sugieren que la brecha regional de salarios sea sistemáticamente distinta según el lugar de la distribución debido a que los cambios que aparecen no siguen una misma tendencia puesto que, según el nivel educativo y el grupo de *clusters* a comparar, algunas veces la brecha se cierra pero en otras se amplía.

Finalmente, vale la pena analizar cómo afecta la calidad en la educación los resultados de esta investigación. En primer lugar, las disparidades en calidad afectan principalmente los niveles de educación universitaria. En segundo lugar, incorporar la calidad de la educación en la estrategia empírica afectaría al vector de coeficientes y no así al de dotaciones. El canal para que suceda esto es que educación de más calidad forma trabajadores más productivos que, a su vez, son recompensados con salarios más altos. Por último, si se incorpora la calidad en la educación, los resultados serían más contundentes de lo que se ha mencionado ya que en Bogotá, donde se concentra la calidad de la educación, los retornos por educación de calidad se aislarían de los retornos salariales a la educación, dejando a estos últimos más bajos de lo que ya muestra esta investigación.

7. Conclusiones

La principal conclusión de este trabajo es que las ciudades intermedias y pequeñas retribuyen con salarios de primer empleo relativos a la línea base, más altos que Bogotá y que dicha tendencia evita que la brecha salarial sea más amplia de lo que es en la

actualidad. En otras palabras, la respuesta al título del artículo es que los retornos a la educación no son una fuente de disparidades regionales sino que, por el contrario, contribuyen a que los niveles de salario converjan entre Bogotá y las ciudades intermedias y pequeñas.

A partir de lo anterior, las recomendaciones de política están encaminadas a extender el acceso a educación técnica, tecnológica y superior en ciudades con rezagos en desarrollo económico pues las personas no sólo serán recompensadas en el mercado laboral, sino que también contribuirán a cerrar las brechas de desarrollo entre estas ciudades y Bogotá. De hecho, como lo demostraron los resultados de la descomposición *B-O*, es el vector de dotaciones el que causa la disparidad de salarios entre regiones.

Una posible estrategia para extender el acceso a los niveles de educación superior es profundizar los beneficios tributarios de las empresas que contraten trabajadores debutantes en el mercado laboral que cuenten con estudios técnicos, tecnológicos y universitarios. Si bien la ley de primer empleo tiene este espíritu, el análisis regional indica que este tipo de iniciativas se deben promulgar de manera diferenciada según las necesidades de cada ciudad. En particular, se observa que en las ciudades del *cluster* II vale la pena extender la cobertura en educación superior, mientras que en las del *cluster* III cualquier inversión encaminada a ampliar la cobertura de todos los niveles educativos será benéfica para los salarios de las ciudades pequeñas.

Bibliografía

- Acemoglu, D., & Dell, M. (2010). Productivity Differences between and within Countries. *American Economic Journal: Macroeconomics*, Vol.2(1), Págs.169-188.
- Arango, L. E., & Posada, C. E. (2002). La participación laboral en Colombia. *Borradores de economía. Banco de la República*.
- Badel, A., & Peña, X. (2010). Decomposing the gender wage gap with sample selection adjustment: evidence from Colombia. *Federal Reserve Bank of St. Louis, Working Papers* 2010-045.
- Barón, J. D. (2002). La inflación en las ciudades de Colombia: Una evaluación de la paridad del poder adquisitivo. *Documentos de trabajo sobre economía regional*, Núm. 31.
- Becker, G. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: National Bureau of Economic Research.

- Bonet, J., & Meisel, A. (2009). Regional economic disparities in Colombia. *Investigaciones regionales. Asociación Española de Ciencia Regional*, 61-80.
- Cepeda, L. (2012). ¿Fuga interregional de cerebros? El caso colombiano. *Documentos de trabajo sobre Economía Regional*, Núm. 167.
- Cepeda, L., & Meisel, A. (2013). ¿Habrá una segunda oportunidad sobre la tierra? Instituciones coloniales y disparidades económicas regionales en Colombia. *Documentos de trabajo sobre economía regional*, Núm.183.
- DANE. (s.f.). Departamento Administrativo Nacional de Estadística. *Gran Encuesta Integrada de Hogares*. Colombia.
- Galego, A., & Pereira, J. (2013). Decomposition of Regional Wage Differences Along the Wage Distribution in Portugal: the Importance of Covariates. *Universidad de Évora - CEFAGE*, Working Paper 2013/16.
- Galvis, L. A., & Meisel, A. (2010). Persistencia de las desigualdades regionales en Colombia: Un análisis espacial. *Documentos de trabajo sobre economía regional*, Vol. 120.
- Gennaioli, N., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2013). Human Capital and Regional Development. *The Quarterly Journal of Economics, Oxford University Press*, Vol.128(1), Págs. 105-164.
- Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, Núm.47, Vol.1, Págs.153-161.
- Hoyos, A., Ñopo, H., & Peña, X. (2010). The Persistent Gender Earnings Gap in Colombia, 1994-2006. *Inter-American Development Bank, Research Department*, Research Department Publications 4673.
- Jann, B. (2008). The Blinder–Oaxaca decomposition for linear regression models. *The Stata Journal*, Vol.8, Núm.4, Págs.453-479.
- Joumard, I., & Londoño Vélez, J. (2013). Income Inequality and Poverty in Colombia - Part 1. The Role of the Labour Market. *OECD Economics Department*, Working Papers Núm. 1036.
- Lemieux, T. (2003). The "Mincer Equation" Thirty Years after Schooling, Experience, and Earnings". *Center for Labor Economics. University of California Berkeley*, Working Paper No.62.
- López-Bazo, E., & Motellón, E. (2009). Human Capital and Regional Wage Gaps. *Research Institute of Applied Economics*, Working Paper 2009/24.
- Lucas Jr, R. E. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, 107(2) Págs.407-437.

- Mankiw, G., Romer, D., & Weil, D. (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics*, 107(2) Págs.407-437.
- Meisel, A. (2011). El sueño de los radicales y las desigualdades regionales en Colombia: La educación de calidad para todos como política de desarrollo de territorio. *Documentos de trabajo sobre economía regional*, Núm. 155.
- Meisel, A., & Romero, J. (2007). Igualdad de oportunidades para todas las regiones. *Documentos de trabajo sobre economía regional*, Núm. 83.
- Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, Vol 66(4), Págs.281-302.
- Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: Columbia University Press.
- Mondagón-Vélez, C., Peña, X., & Wills, D. (2011). Rigideces laborales y salarios en los sectores formal e informal en. *Documento CEDE*.
- Mondragón-Vélez, C., Peña, X., & Wills, D. (2010). Labor Market Rigidities and Informality in Colombia. *Economía, The Journal of the Latin American and Caribbean Economic Association*, Vol11. Núm1.
- Motellón, E., López-Bazo, E., & El-Attar, M. (2009). Regional heterogeneity in wage distributions. Evidence from Spain. *Research Institute of Applied Economics*, Working Paper 2009/03.
- Peña, X. (2013). The formal and informal sectors in Colombia. Country case study on labour market segmentation. *Employment Sector - Employment Analysis and Research Unit - International Labour Organization*, Employment Working Paper Núm.146.
- Prada, C. F. (2006). ¿Es rentable la decisión de estudiar en Colombia? *Ensayos sobre Política Económica*, Vol. 51, Págs. 267-329.
- Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5) Págs. 71-102.
- Sánchez, F., Duque, V., & Ruiz, M. (2009). Costos laborales y no laborales y su impacto sobre el desempleo, la duración del desempleo y la informalidad en Colombia, 1980-2007. *Documento CEDE*, 11.
- Sen, A. (1999). *Development as Freedom*. Oxford: Oxford University Press.
- Tenjo, J. (1993). Evolución de los retornos de la educación 1976-1989. *Desarrollo y Sociedad*, Vol. XXIV, Págs.85-114.
- Zárate, H. M. (2003). Cambios en la estructura salarial: Una historia desde la regresión cuantílica. *Borradores de Economía. Banco de la República*.

Comentarios sobre la construcción de los *Splines* Lineales (SP)

El método que se usa para construir los SP en esta investigación dista un poco del método tradicional. Normalmente, los SP se capturan con variables dicótomas por cada nivel educativo pero la misma persona sólo puede tener un 1 en el último nivel educativo alcanzado y 0 en los demás. En el método que se usa acá, la persona tiene 1 en el último nivel educativo así como en todos los demás hacia atrás. La principal razón para tomar esta decisión fue poder aislar el efecto de cada uno de los niveles educativos, de tal manera de estimar la recompensa exacta de obtener cada diploma frente a no haberlo hecho. Si bien esto mismo se puede estimar con el método tradicional al calcular las diferencias en los parámetros de la ecuación de salarios, en el caso de la Descomposición *B-O* es particularmente útil el método acá usado porque las diferencias que se pueden hacer en la ecuación de salarios no se pueden hacer con los vectores resultantes de la Descomposición. Por último, los resultados de los parámetros no cambian mucho entre un método y otro.

Anexo 1- Análisis de Clusters

Promedio del salario real por hora	Tasa Global de Participación (Población total de 18 a 28 años)	Tasa de Desempleo (Población total de 18 a 28 años)	Coficiente de Gini del salario	Promedio de los años de educación	Población objetivo (Dic-2013)	Cluster	Ciudad	Promedio del salario real por hora	Tasa Global de Participación (Población total de 18 a 28 años)	Tasa de Desempleo (Población total de 18 a 28 años)	Coficiente de Gini del salario	Promedio de los años de educación	Población objetivo (Dic-2013)
\$ 3,798	79.53%	16.19%	0.3380	11.71	466,070	I	Bogotá D.C	\$ 3,798	79.53%	16.19%	0.3380	11.71	466,070
\$ 3,330	74.01%	21.06%	0.3169	11.52	411,491	II	Medellín	\$ 3,578	76.29%	21.03%	0.3187	11.68	228,060
							Cali	\$ 3,085	78.46%	21.92%	0.3052	11.06	135,370
							Barranquilla	\$ 3,008	63.65%	19.73%	0.3167	11.86	48,061
\$ 3,076	72.75%	21.26%	0.3153	11.33	216,459	III	Bucaramanga	\$ 3,445	77.51%	16.23%	0.2879	11.51	44,510
							Pereira	\$ 3,048	77.03%	25.35%	0.2698	11.13	30,051
							Ibagué	\$ 2,933	81.83%	24.96%	0.3192	11.31	25,504
							Manizales	\$ 3,214	66.68%	22.72%	0.3093	11.58	24,901
							Villavicencio	\$ 3,148	75.33%	17.76%	0.2998	10.91	24,636
							Cúcuta	\$ 2,697	75.91%	20.82%	0.2944	10.74	22,712
							Cartagena	\$ 3,358	59.81%	22.44%	0.3333	12.20	17,114
							Montería	\$ 2,611	73.23%	23.24%	0.3865	11.15	13,618
Pasto	\$ 2,383	69.99%	23.95%	0.3946	11.07	13,413							

Proporción de la fuerza laboral en cada uno de los sectores productivos						Cluster	Ciudad	Proporción de la fuerza laboral en cada uno de los sectores productivos					
Agrícola	Minero	Industrial	Construcción	Comercio	Servicios			Agrícola	Minero	Industrial	Construcción	Comercio	Servicios
0.57%	0.43%	18.25%	5.22%	28.09%	47.44%	I	Bogotá D.C	0.57%	0.43%	18.25%	5.22%	28.09%	47.44%
0.61%	0.09%	24.22%	6.89%	29.57%	38.62%	II	Medellín	0.63%	0.14%	25.91%	7.27%	28.88%	37.18%
							Cali	0.70%	0.03%	22.95%	7.93%	30.33%	38.06%
							Barranquilla	0.38%	0.08%	21.34%	3.83%	30.34%	44.03%
1.30%	0.39%	17.75%	8.22%	32.92%	39.41%	III	Bucaramanga	2.03%	0.28%	20.61%	7.95%	31.60%	37.54%
							Pereira	1.67%	0.07%	20.92%	8.41%	33.38%	35.55%
							Ibagué	2.09%	0.15%	17.22%	8.86%	33.05%	38.63%
							Manizales	0.69%	0.09%	19.45%	8.01%	26.99%	44.77%
							Villavicencio	1.49%	0.91%	10.98%	11.97%	38.17%	36.48%
							Cúcuta	0.66%	0.75%	23.12%	9.06%	34.74%	31.68%
							Cartagena	0.50%	0.94%	14.06%	3.20%	32.26%	49.05%
							Montería	0.69%	0.10%	9.20%	10.54%	30.88%	48.59%
Pasto	0.59%	0.07%	12.76%	7.82%	37.93%	40.84%							

Anexo 2 – Estadísticas descriptivas de las variables en la Descomposición de *Blinder-Oaxaza*

Promedio del salario real por hora	Porcentaje de la población objetivo que ha completado el nivel educativo					Porcentaje de mujeres	Promedio de la edad	Cluster	Ciudad	Promedio del salario real por hora	Porcentaje de la población objetivo que ha completado el nivel educativo					Porcentaje de mujeres	Promedio de la edad
	Bachillerato	Técnica	Tecnológica	Pregrado	Posgrado						Bachillerato	Técnica	Tecnológica	Pregrado	Posgrado		
\$ 3,798	83.65%	12.15%	5.69%	12.30%	4.10%	48.31%	23.08	I	Bogotá D.C.	\$ 3,798	83.65%	12.15%	5.69%	12.30%	4.10%	48.31%	23.08
\$ 3,330	83.76%	13.33%	6.17%	9.70%	2.37%	46.21%	23.22	II	Medellín	\$ 3,578	84.60%	15.52%	6.98%	10.06%	3.09%	47.01%	23.27
									Cali	\$ 3,085	80.26%	7.94%	3.88%	7.14%	1.71%	44.82%	23.01
									Barranquilla	\$ 3,008	87.49%	16.29%	7.77%	13.22%	1.32%	46.27%	23.44
\$ 3,076	81.01%	11.95%	6.50%	9.10%	1.82%	47.13%	23.02	III	Bucaramanga	\$ 3,445	82.90%	13.30%	8.12%	8.95%	2.24%	48.62%	22.82
									Pereira	\$ 3,048	80.17%	12.05%	6.70%	7.61%	1.35%	43.60%	23.07
									Ibagué	\$ 2,933	80.92%	11.32%	5.06%	7.88%	1.36%	45.24%	22.73
									Manizales	\$ 3,214	84.46%	14.40%	4.88%	10.42%	2.63%	46.57%	23.14
									Villavicencio	\$ 3,148	78.81%	4.31%	2.65%	6.94%	1.38%	47.27%	22.76
									Cúcuta	\$ 2,697	73.25%	7.20%	3.75%	7.08%	1.69%	44.36%	22.77
									Cartagena	\$ 3,358	88.94%	23.31%	14.72%	13.46%	1.96%	50.69%	23.85
									Montería	\$ 2,611	79.08%	5.93%	3.19%	10.85%	1.42%	49.70%	23.10
Pasto	\$ 2,383	76.82%	9.86%	3.93%	10.36%	2.03%	49.85%	23.19									

Proporción de la fuerza laboral en cada uno de los sectores productivos						IMR promedio	Cluster	Ciudad	Proporción de la fuerza laboral en cada uno de los sectores productivos						IMR promedio
Agrícola	Minero	Industrial	Construcción	Comercio	Servicios				Agrícola	Minero	Industrial	Construcción	Comercio	Servicios	
0.57%	0.43%	18.25%	5.22%	28.09%	47.44%	0.1923	I	Bogotá D.C.	0.57%	0.43%	18.25%	5.22%	28.09%	47.44%	0.1923
0.61%	0.09%	24.22%	6.89%	29.57%	38.62%	0.2813	II	Medellín	0.63%	0.14%	25.91%	7.27%	28.88%	37.18%	0.2538
								Cali	0.70%	0.03%	22.95%	7.93%	30.33%	38.06%	0.2176
								Barranquilla	0.38%	0.08%	21.34%	3.83%	30.34%	44.03%	0.4800
1.30%	0.39%	17.75%	8.22%	32.92%	39.41%	0.2766	III	Bucaramanga	2.03%	0.28%	20.61%	7.95%	31.60%	37.54%	0.2206
								Pereira	1.67%	0.07%	20.92%	8.41%	33.38%	35.55%	0.2444
								Ibagué	2.09%	0.15%	17.22%	8.86%	33.05%	38.63%	0.1940
								Manizales	0.69%	0.09%	19.45%	8.01%	26.99%	44.77%	0.3435
								Villavicencio	1.49%	0.91%	10.98%	11.97%	38.17%	36.48%	0.2455
								Cúcuta	0.66%	0.75%	23.12%	9.06%	34.74%	31.68%	0.2416
								Cartagena	0.50%	0.94%	14.06%	3.20%	32.26%	49.05%	0.4775
								Montería	0.69%	0.10%	9.20%	10.54%	30.88%	48.59%	0.3221
Pasto	0.59%	0.07%	12.76%	7.82%	37.93%	40.84%	0.2842								

Anexo 3

Resultados completos de la regresión de Mincer

ln_salario	Cluster		
	I	II	III
d_bachillerato	0.157*** (0.0337)	0.265*** (0.0231)	0.314*** (0.0181)
d_tecnica	0.209*** (0.0255)	0.252*** (0.0153)	0.225*** (0.0157)
d_tecnologica	0.0886** (0.0410)	0.104*** (0.0213)	0.104*** (0.0203)
d_pregrado	0.823*** (0.0308)	0.780*** (0.0229)	0.740*** (0.0155)
d_posgrado	0.122** (0.0539)	0.295*** (0.0353)	0.300*** (0.0250)
extra_primaria	0.00611 (0.0130)	0.0151* (0.00876)	0.0231*** (0.00692)
extra_bachillerato	0.103*** (0.0107)	0.100*** (0.00646)	0.0935*** (0.00550)
extra_tecnologica	0.107** (0.0471)	0.0674** (0.0304)	0.0357 (0.0576)
sexo	-0.108*** (0.0161)	-0.131*** (0.0100)	-0.164*** (0.00827)
edad	0.0377*** (0.00330)	0.0331*** (0.00210)	0.0331*** (0.00156)
agro	-0.0802 (0.174)	-0.114 (0.0763)	-0.0873** (0.0376)
minero	0.368*** (0.0727)	0.00872 (0.255)	0.310*** (0.100)
industria	0.0175 (0.0200)	0.0108 (0.0119)	-0.0111 (0.0123)
construccion	0.00722 (0.0396)	-0.0741*** (0.0245)	0.0281* (0.0168)
comercio	-0.0733*** (0.0193)	-0.0489*** (0.0117)	-0.0494*** (0.00927)
IMR	0.0964* (0.0565)	0.0358 (0.0295)	0.00623 (0.0212)
Constante	6.694*** (0.0994)	6.780*** (0.0640)	6.531*** (0.0508)
Efectos fijos de tiempo	Sí		
Efectos fijos de ciudad			
*, ** y *** Significativos al 10%, 5% y 1%, respectivamente. Errores estándar entre paréntesis.			
Observaciones (Sin pesos de frecuencia)	14,948	34,714	70,948
R-cuadrado	0.136	0.139	0.140

Anexo 4

Tablas de resultados incluyendo errores estándar

Tabla 2 – Retorno salarial por tener el nivel educativo completo					
<i>Cluster</i>	Bachillerato completo	Técnica completa	Tecnológica completa	Pregrado completo	Posgrado completo
I	0.1574*** (0.0337)	0.3662*** (0.039)	0.4547*** (0.047)	0.9804*** (0.0422)	1.1027*** (0.0562)
II	0.2645*** (0.023)	0.5166*** (0.0263)	0.6203*** (0.0277)	1.0447*** (0.0312)	1.3397*** (0.0359)
III	0.3139*** (0.0181)	0.5384*** (0.0229)	0.6429*** (0.0229)	1.0542*** (0.0227)	1.3539*** (0.0274)
***Significativos al 1%. Errores estándar entre paréntesis.					
Fuente: GEIH-DANE. Cálculos propios.					

Tabla 3 – Prueba de hipótesis de las diferencias en los retornos a la educación					
<i>Clusters</i>	Bachillerato completo	Técnica completa	Tecnológica completa	Pregrado completo	Posgrado completo
I vs. II	-0.1086*** (0.0358)	-0.1473*** (0.0461)	-0.1659*** (0.0506)	-0.0672 (0.0452)	-0.2369 (0.0586)
II vs. III	-0.0533** (0.0233)	-0.0317 (0.0319)	-0.0303 (0.033)	-0.0135 (0.0306)	-0.192 (0.0455)
I vs. III	-0.1581*** (0.0274)	-0.1752*** (0.38)	-0.1913*** (0.0394)	-0.076** (0.0357)	-0.2543*** (0.0516)
*, ** y *** Significativos al 10%, 5% y 1%, respectivamente. Errores estándar entre paréntesis.					
Fuente: GEIH-DANE. Cálculos propios.					

Tabla 4 – Descomposición <i>Blinder-Oaxaca</i>				
		$I_A vs. II_B$	$II_A vs. III_B$	$I_A vs. III_B$
Resumen	$E \left[\ln \left(\frac{w_i}{h_i} \right) \right]_A$	7.9654 (0.0002)	7.8617 (0.0002)	7.9654 (0.0002)
	$E \left[\ln \left(\frac{w_i}{h_i} \right) \right]_B$	7.8617 (0.0002)	7.7761 (0.0004)	7.7761 (0.0004)
	Diferencia A-B	0.1037*** (0.004)	0.0855*** (0.0004)	0.1892*** (0.0004)
Vector de Dotaciones	Bachillerato	-0.0002*** (0)	0.0095*** (0)	0.0091*** (0)
	Técnica	-0.0029*** (0)	0.0033*** (0)	0.0004*** (0)
	Tecnológica	-0.0004*** (0)	-0.0003*** (0)	-0.0009*** (0)
	Pregrado	0.0196*** (0.001)	0.0044*** (0.001)	0.0236*** (0.0001)
	Posgrado	0.0061*** (0)	0.0017*** (0)	0.0074*** (0)
Vector de Coeficientes	Bachillerato	-0.076*** (0.0014)	-0.0772*** (0.0015)	-0.1507*** (0.0015)
	Técnica	-0.0028*** (0.0002)	0.0006*** (0.002)	-0.0018*** (0.0002)
	Tecnológica	-0.001*** (0.0001)	-0.0013*** (0.0001)	-0.0024*** (0.0001)
	Pregrado	0.0068*** (0)	0.0017*** (0.0001)	0.0081*** (0.0001)
	Posgrado	-0.0051*** (0.0002)	0.0005*** (0)	-0.0034*** (0)
*** Significativos al 1%. Errores estándar entre paréntesis. Estadístico z para pruebas de hipótesis.				
Fuente: GEIH-DANE. Cálculos propios.				