

Subempleo por ingresos y funcionamiento del mercado de trabajo en Colombia

Income-related Underemployment and Labor Market Performance in Colombia

Luis Eduardo Arango
Diana Carolina Escobar
Emma Mercedes Monsalve¹

DOI: 10.13043/DYS.72.4

Resumen

El mercado laboral en Colombia funciona mucho mejor de lo que sugiere la alta tasa de subempleo, la cual es sinónimo de "empleo de baja calidad". A juzgar por los componentes de la brecha de ingresos entre quienes no se quejan de su salario y quienes sí lo hacen, la tasa de subempleo efectiva en Colombia debe ser mucho menor que la reportada en las estadísticas oficiales, según las cuales una cuarta parte de la fuerza laboral es subempleada por ingresos.

1 Las opiniones expresadas en este documento son responsabilidad exclusiva de los autores y no comprometen ni al Banco de la República ni a su Junta Directiva, ni al Ministerio de Hacienda y Crédito Público ni al Banco Interamericano de Desarrollo. Se agradece la asistencia de investigación de Gabriela Bonilla, Guillermo Cangrejo y Nataly Obando, así como los valiosos comentarios y sugerencias de Lina Cardona, Francesca Castellani, Hugo Ñopo, Ximena Peña, Carlos Esteban Posada y Carmiña Vargas. También se agradece a los asistentes a los seminarios de la Universidad Católica de Pereira-Maestría en Gestión del Desarrollo Regional, del Banco de la República y del Grupo de Economía y Empresa de la Universidad EAFIT. Los posibles errores del documento son responsabilidad exclusiva de los autores. Se encuentran vinculados al Banco de la República, el Ministerio de Hacienda y Crédito Público y el Banco Interamericano de Desarrollo. Correos electrónicos: Luis Eduardo Arango: larangth@banrep.gov.co; Diana Carolina Escobar: descobar@minhacienda.gov.co; Emma Mercedes Monsalve: emmamo@contractual.iadb.org.

Este artículo fue recibido el 15 de abril de 2013; revisado el 18 de septiembre de 2013 y, finalmente, aceptado el 18 de septiembre de 2013.

Para llegar a concluir lo anterior, se estimó la brecha salarial entre empleados particulares y servidores públicos catalogados como plenamente ocupados y subempleados por ingresos, mediante la ecuación de Mincer, la descomposición Blinder-Oaxaca y la metodología de Ñopo (2008). Los resultados indican que aunque existe un diferencial de salarios entre ambos grupos, del orden de 40% a 44%, el componente no explicado de este es mucho menor: de 36% a 40%. Aunque algunos subempleados por ingresos podrían tener argumentos para considerarse como tales, aquí se presenta evidencia que muestra que la mayor parte de la brecha es explicada por variables asociadas al capital humano y otros controles. El diferencial se presenta con mayor nitidez en la parte alta de la distribución de los salarios (a mayor edad y mayor educación, mayor es el componente no explicado del diferencial) y en los sectores de minas, transporte y construcción.

Palabras clave: subempleo por ingresos, funcionamiento del mercado laboral, descomposición de la brecha salarial.

Clasificación JEL: J31.

Abstract

The labor market in Colombia performs better than suggested by the high underemployment rate which is tantamount to "low quality employment". Given the wage gap among those that complain about their wages and those that do not, the effective underemployment rate in Colombia should be much lower than that reported by official statistics according to which a fourth part of the labor force is income-related underemployed. To arrive to this conclusion, the wages of private salaried workers and public servants were estimated for both income-related underemployed for and not underemployed at all. The wage gap was estimated by means of the Mincer equation, the Blinder-Oaxaca decomposition and the approach of Ñopo (2008). The results show a wage gap between 40% and 44% but the unexplained component is much smaller (between 36% and 40% of it). Although some underemployed might have reasons to be regarded as such, we provide evidence in the sense that the higher portion of the gap is explained by human capital variables and other controls. The differential is more neatly observed in the higher part of the wage distri-

bution (the higher the age and schooling, the higher the unexplained component of it) and in the sectors of mines, transport and building.

Key words: Income-related underemployment, performance of the labor market, wage gap decomposition.

JEL classification: J31.

Introducción

Indicadores como las tasas de ocupación, desempleo y participación, la estructura de los salarios reales, la composición de la fuerza de trabajo y el tamaño del sector informal, entre otros, dan indicios del funcionamiento del mercado laboral. Dichas variables son el resultado de la interacción de las fuerzas del mercado (oferta y demanda) en un marco de incentivos e instituciones (por ejemplo, normas y aplicación de los mecanismos de *enforcement*, etc.) y otros aspectos sociológicos².

En teoría, el desempeño del mercado laboral refleja decisiones óptimas de oferta y demanda de trabajo; sin embargo, también puede mostrar situaciones subóptimas. En el primer caso se encuentran, por ejemplo, las personas cuyas habilidades están siendo utilizadas de manera eficiente y su salario real es, al menos, igual que su salario de reserva. Por otro lado, las empresas están en equilibrio cuando, por ejemplo, a partir de un proceso de maximización de beneficios determinan la cantidad de mano de obra que pueden contratar en el mercado, dados los precios de los factores y otros determinantes como su eventual poder en el mercado del bien o servicio que producen. Como resultado, las empresas pagan a sus empleados un salario real similar a su productividad.

En el segundo caso, el subempleo es una de las manifestaciones de que en el mercado laboral no todos los hechos corresponden al *first best*, un resultado que permite identificar situaciones en las cuales hay desemparejamiento *expost* (es decir, dado un contrato de trabajo) entre lo que se demanda y lo que se ofrece.

2 Kaufman y Hotchkiss (2006) se refieren a resultados y procesos del mercado laboral.

En este sentido, hay personas que estando ocupadas señalan que su empleo es insuficiente en materia de horas o que están viviendo situaciones de empleo inadecuado. Las primeras trabajan menos de 48 horas a la semana, el máximo permitido por la ley laboral según el artículo 161 del Código Sustantivo del Trabajo, y consideran que podrían laborar más horas. Entre las segundas están quienes se sienten subutilizadas porque sus habilidades no se compadecen con las labores que están llevando a cabo y, por tanto, creen que hay espacio para mejorar la utilización de sus capacidades, formación y experiencia; también quienes desean mejorar sus ingresos ya que, posiblemente, consideran que su remuneración es inferior a su salario de reserva o a su productividad laboral³, o que, aun siendo igual a esta última, es inferior a lo que ganan otras personas con similares habilidades, conocimientos y otras características observables. En cualquier caso, debe tenerse en cuenta que la distribución de salarios del mercado y la productividad de la mano de obra, así como las exigencias y riesgos de algunos trabajos, no son observables y que las personas podrían no tener una noción completamente clara acerca de su salario de reserva.

Sin embargo, hay situaciones de subempleo por ingresos que no necesariamente son casos subóptimos en los cuales el salario es inferior a la productividad laboral o al salario de reserva. Dichas situaciones podrían corresponder a brechas salariales entre personas que no fueron clasificadas como subempleadas por alguna razón y subempleados por ingresos explicadas, entre otros factores, por la existencia de salarios hedónicos o por diferencias en: a) la productividad, vinculada, a su vez, a divergencias de calidad en la educación recibida por los individuos con la misma escolaridad⁴, b) el poder de los empleados para negociar el salario en algunos sectores de la producción y c) los rasgos de la personalidad de los empleados premiados por el mercado⁵.

3 Lo cual puede suceder tanto en un mercado de bienes con características de competencia perfecta como en uno en el que la empresa tenga poder de mercado, de suerte que el *mark-up* haga que los trabajadores perciban que su salario es inferior a la productividad del trabajo. En este último caso, el salario real (w) es inferior a la productividad marginal del trabajo ($PMgL$) en virtud de un eventual *mark-up* (μ); esto es: $PMgL/\mu = w$, debido a que $\mu > 1$.

4 López (2012) presenta evidencia del "efecto colegio" en las pruebas del Estado.

5 Cinco grandes rasgos de la personalidad son: extroversión, apertura de mente y tolerancia, estabilidad emocional, responsabilidad y amabilidad o empatía. Wichert y Pohlmeier (2010) hacen una aplicación reciente al caso de la participación laboral femenina. Véanse también Mueller y Plug (2006).

Brechas aparentes también podrían surgir cuando los subempleados tienen como referencia a personas que reciben un salario de eficiencia para estimular en ellas un mayor esfuerzo o premiar características no observables (mejor disposición al trabajo, mejores ideas, mayor compromiso con la firma) que se reflejan en la productividad individual o grupal. Desde luego, si los salarios de eficiencia fueran el esquema prevaleciente, aunque no exclusivo, en el mercado de trabajo de Colombia, los subempleados por ingresos no tendrían incentivos para cuidar su permanencia en el puesto de trabajo, no se esforzarían lo suficiente y estarían generando pérdidas de eficiencia en la economía.

Este artículo pone en duda que la alta tasa de subempleo por ingresos que se observa hoy día en Colombia, correspondiente a la brecha salarial entre los asalariados que no se quejan de su salario (plenamente ocupados⁶) y quienes sí lo hacen (subempleados por ingresos), constituya un indicador apropiado del (mal) funcionamiento del mercado de trabajo. En particular, los autores del presente artículo creen que el mercado laboral funciona mucho mejor de lo que sugiere la alta tasa de subempleo, la cual se considera sinónimo de "empleo de baja calidad" o de "condiciones de empleo inadecuado". A juzgar por los componentes de la brecha de ingresos entre aquellos plenamente ocupados y los subempleados por ingresos, la tasa de subempleo efectiva en Colombia es mucho menor que la reportada en las estadísticas oficiales, según las cuales la población subempleada por ingresos es una cuarta parte de la fuerza laboral.

En tal sentido, el objetivo de la investigación es presentar evidencia sobre el funcionamiento del mercado de trabajo en cuanto a la asignación de salarios entre 2007 y 2012 y plantear algunas hipótesis y conjeturas sobre el subempleo por ingresos.

El artículo somete a prueba la hipótesis de que quienes son catalogados como subempleados por ingresos ganan realmente un salario inferior que el de los plenamente ocupados, una vez se tengan en cuenta las diferencias en las características observables. Dado lo anterior, se estima la magnitud de la brecha no explicada por la teoría del capital humano.

6 Asalariados del sector privado o del sector público que no se declaran subempleados por ninguna de las razones expuestas.

Para cumplir el objetivo planteado se utilizan dos enfoques. El primero se apoya en la ecuación de Mincer (1974) para estimar la parte no explicada de la descomposición Blinder (1974) – Oaxaca (1973), usualmente asociada con diferencias en la remuneración a las características observables de los subempleados por ingresos.

Sin embargo, tal descomposición parece sobrestimar el componente no explicado, al desconocer las diferencias en los *soportes* de las distribuciones empíricas de las características individuales de los plenamente ocupados y los subempleados. Por tal razón, en el segundo enfoque se utiliza la metodología de Ñopo (2003, 2008), la cual permite explorar la importancia de las combinaciones de características individuales para las que existen, por un lado, personas catalogadas como plenamente ocupadas pero no subempleados por ingresos y, por otro, subempleados por ingresos pero no plenamente ocupados. Para tales combinaciones los salarios no son estrictamente comparables.

Los resultados sugieren que el subempleo por ingresos podría no solo ser una percepción subjetiva del individuo, sino el posible reflejo de una ineficiencia del mercado al retribuir las dotaciones de forma diferente entre individuos cuyas características observables son similares, en promedio. Es decir, se presenta evidencia de que quienes se declaran subempleados por ingresos efectivamente ganan, en promedio, menos que otras personas, incluso después de incluir algunas variables que podrían dar cuenta del tiempo de búsqueda del empleo, la antigüedad en el empleo actual y el tamaño de la empresa en la que trabajan.

Sin embargo, la magnitud de la brecha y, en tal sentido, de la ineficiencia desde el punto de vista del salario es mucho menor (entre el 27% y el 40%) que la que surge al observar las cifras gruesas del subempleo por ingresos. En consecuencia, la principal conclusión de este artículo es que el mercado laboral funciona mucho mejor de lo que sugieren las cifras brutas de subempleo por ingresos y que, por tanto, se debería tener más cuidado al identificar el subempleo por ingresos como un "desempleo inadecuado", como se le denomina en el lenguaje oficial. Sin embargo, esta investigación no reestima (corrige) la tasa de subempleo.

El subempleo por ingresos se presenta en personas de salarios situados en la parte alta de la distribución; esto es, con más educación y experiencia, que

desempeñan cargos de dirección y que viven en estratos altos. Es decir, el subempleo está más asociado con una mala formación de expectativas con respecto al salario de mercado⁷, con la subvaloración del salario de reserva, con que la productividad que perciben los empleadores refleja deficiencias en la calidad de la educación recibida por los subempleados, o con que su productividad es inferior a la de alguien a quien se le paga un salario de eficiencia.

Suponiendo que sea adecuada la forma en que se mide el subempleo por ingresos, algunas explicaciones del aparente mal funcionamiento residual del mercado laboral estarían en la falta de información sobre la distribución de salarios, un conocimiento deficiente de la mano de obra sobre lo que es su salario de reserva, la existencia de salarios de eficiencia y hedonistas, diferencias en la calidad de la educación y el poder de mercado de las firmas.

Este artículo está dividido en tres secciones. En la primera se muestra la forma como se mide el subempleo y su evolución, así como algunas características de los subempleados mediante modelos *logit* binomiales y multinomiales. En la segunda se estiman ecuaciones de Mincer para determinar la brecha salarial entre los individuos plenamente ocupados y subempleados por ingresos y se usa la metodología de Ñopo (2003, 2008) para depurar el componente no explicado de la brecha. Por último, en la tercera sección se concluye.

I. Elementos del subempleo y evolución reciente

Para caracterizar el mercado laboral no solo en función del desempleo y la informalidad, en 1970 la Organización Internacional del Trabajo (OIT) complementó las estadísticas de empleo, desempleo e inactividad añadiendo las de subempleo. Entre 1976 y 1999, el Departamento Administrativo Nacional de Estadística (DANE) midió el subempleo y lo clasificó en dos tipos: visible e invisible. El primero incluía personas cuya jornada laboral era inferior a 32 horas semanales, mientras que el segundo tenía en cuenta a las personas cuyo salario era insuficiente para atender sus gastos normales o que realizaban una actividad que no se ajustaba a sus habilidades (Urrutia, 2002, p. 8). Para medir la prevalencia de estas situaciones se utiliza la tasa de subem-

7 Las personas cuyos ingresos están en la parte baja de la distribución tienen menos incertidumbre sobre lo que podría ser su salario y allí hay menos indicios de subempleo por ingresos.

pleo, la cual relaciona la población subempleada con la población económicamente activa (PEA)⁸.

Con la introducción de la Encuesta Continua de Hogares (ECH), en el año 2000 se definieron el subempleo por insuficiencia de horas y el subempleo por situaciones de empleo inadecuado, mencionados en la introducción. La Gran Encuesta Integrada de Hogares (GEIH), utilizada desde 2006⁹, mantuvo las categorías anteriores de subempleo¹⁰, pero además clasificó a cada uno de ellos como subjetivo u objetivo. El primero, constituido por la población subempleada que manifestó el deseo de mejorar sus ingresos, aumentar el número de horas trabajadas o tener una labor más conforme a sus capacidades. El segundo, integrado por individuos subempleados que no solo manifiestan el deseo de mejorar sus condiciones laborales sino que han hecho alguna gestión para materializar su aspiración y están en disposición de efectuar el cambio¹¹ (DANE, 2006).

Entre enero de 2001 y junio de 2006, período de vigencia de la ECH, las tasas de subempleo subjetivo y objetivo fueron 31,4% y 13,0%, respectivamente. Entre julio de 2007 y junio de 2012, el período muestral de la presente investigación, la tasa de subempleo subjetivo fue de 29,5% y las correspondientes de subempleo por ingresos, competencias y horas fueron de 25,4%, 16,9% y 9,2%, respectivamente. Los promedios de las tasas de subempleo objetivo fueron 12,1% para el total, y 10,1%, 7,1% y 4,3% para los casos de ingresos,

8 La PEA es la fuerza laboral: suma de los ocupados (quienes tienen una ocupación) y los desocupados (quienes no la tienen, pero la están buscando).

9 Para determinar si una persona es subempleada por horas o no, la GEIH del DANE pregunta a los entrevistados: "Además de las horas que trabaja actualmente, ¿quiere trabajar más horas?". Después pregunta: "¿Desea cambiar el trabajo que tiene actualmente?". En caso afirmativo, pregunta: "¿Por qué motivos desea hacerlo?" y propone las siguientes opciones: a) para mejorar la utilización de sus capacidades o formación, b) para mejorar sus ingresos, c) porque desea trabajar menos horas, d) porque el trabajo actual es temporal, e) por problemas en el trabajo, f) porque no le gusta el trabajo actual, g) porque su trabajo actual exige mucho esfuerzo físico o mental, h) por problemas ambientales, i) por otra razón, en cuyo caso pide especificar cuál. Además pregunta: "Durante las últimas 4 semanas, ¿[...] hizo diligencias para cambiar de trabajo?" y "si le resultara un nuevo trabajo o empleo a [...] ¿podría empezar a desempeñarlo antes de un mes?".

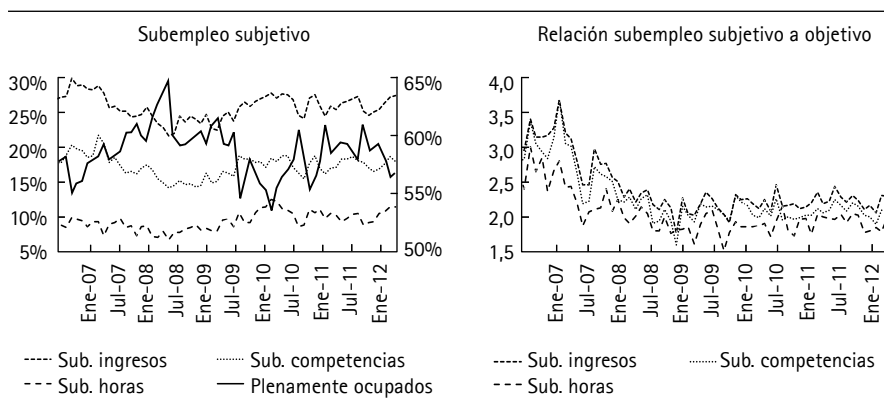
10 Para cada tipo de subempleo, ingresos, capacidades y horas, se calcula la relación entre el número de subempleados y la PEA. Esto es, $TS_i = (\text{población subempleada por ingresos} / \text{PEA}) \times 100$, $TS_c = (\text{población subempleada por capacidades} / \text{PEA}) \times 100$ y $TS_h = (\text{población subempleada por horas} / \text{PEA}) \times 100$.

11 Subempleados objetivos son las personas que buscan un nuevo trabajo estando empleadas (*on-the-job search*).

competencias y horas, en su orden. Según estas cifras, el mercado de trabajo de Colombia estaría alejado del *first best*¹² y, dado que el componente de subempleo por ingresos es el más alto, es este el que se analizará aquí¹³.

El gráfico 1 muestra la evolución reciente de las tasas de subempleo. En el panel izquierdo del gráfico se observa que, efectivamente, el principal componente del subempleo es el asociado con los ingresos y que los tres tipos de subempleo crecieron entre 2008 y 2010, mientras que la tasa de ocupación plena disminuyó en el mismo lapso, aunque posteriormente se recuperó. En el panel de la derecha, se observa que el subempleo objetivo ha sido significativamente menor que el subjetivo, pero la brecha ha disminuido. En efecto, el subempleo subjetivo pasó de ser entre tres y cuatro veces el objetivo a algo menos de 2,4 veces después de 2010.

Gráfico 1. Evolución de las tasas de subempleo y plena ocupación



Nota: el porcentaje de plenamente ocupados se mide en el eje derecho.

Fuente: GEIH (DANE) y cálculos propios.

Por su parte, el gráfico 2 muestra la evolución de los salarios reales por hora de los subempleados y los plenamente ocupados¹⁴. En el período julio de 2007-

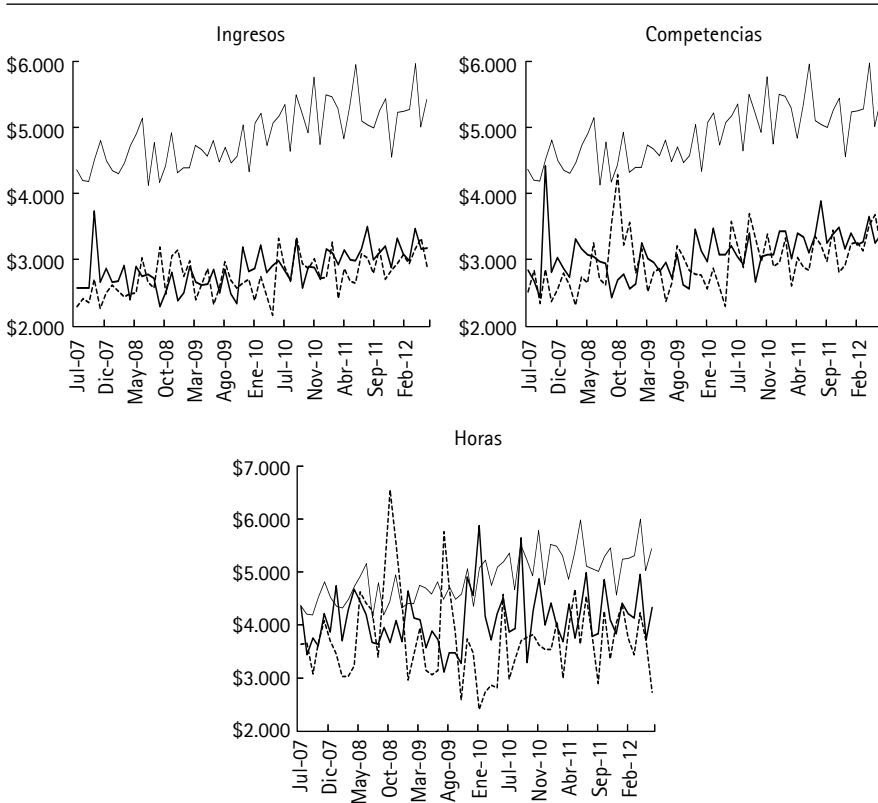
12 Con claras implicaciones en las decisiones de capital humano, consumo y el ahorro representado en pensiones.

13 Puyana, Ramos y Zárate (2013) analizan el subempleo por horas.

14 Para el cálculo del salario real por hora se usaron las siguientes preguntas de la GEIH: a) "¿Cuántas horas a la semana trabaja normalmente [...] en ese trabajo?" (*HTM*) y b) "antes de descuentos, ¿cuánto ganó [...] el mes pasado en este empleo?" (*ILAB*). Las horas trabajadas al mes (*HTM*) se calcularon suponiendo que las personas trabajan cinco días y medio a la semana; además, se controló por días hábiles (*DH*).

junio de 2012, las diferencias promedio entre los salarios de los plenamente ocupados y los subempleados subjetivos y objetivos fueron de \$1.971 y \$2.008, en el caso de los ingresos, \$1.766 y \$1.804, en cuanto a las competencias y \$763 y \$1.086, en lo referente al subempleo por horas. El salario real promedio de los subempleados objetivos es menor que el de los subjetivos y el de estos últimos es inferior en 40% al de los plenamente ocupados.

Gráfico 2. Salarios reales promedio por hora de los subempleados y los plenamente ocupados



Nota: la línea continua delgada identifica los salarios de los plenamente ocupados, la continua gruesa los salarios de los subempleados subjetivos y la punteada los salarios de los subempleados objetivos, (2008:12=100).

Fuente: GEIH (DANE) y cálculos propios.

Por último, el salario mensual (*WH*) fue deflactado utilizando índices de precios por ciudad (*DEC*). Así, el salario real por hora corresponde a: $HTM = (HTN/5,5) \times DH$, $WT = (ILAB/DEC)$, $WH = (WT/HTT)$.

No obstante, pese a su importancia, el subempleo es un fenómeno poco analizado en Colombia. Se destacan las excepciones de Urrutia (2002), López (2000) y Puyana, Ramos y Zárate (2013), quienes estudian los determinantes del subempleo y coinciden en una posible relación de causalidad con el desempleo, pero desde perspectivas diferentes. López (2000) y Urrutia (2002) sugieren que el desempleo puede ser determinante del subempleo agregado, mientras que Puyana *et al.* (2013) plantean un modelo de incertidumbre y riesgo en el que existe un diferencial positivo de salarios para los subempleados, aunque dicha compensación salarial solo existe para los subempleados por horas, por el riesgo de desempleo que enfrentan¹⁵. Por su parte, Uribe, Ortiz y García (2008) encuentran relaciones entre la informalidad y el subempleo; una por el lado de la demanda y otra por el lado de la oferta. Sin embargo, según los autores ambas reflejan deficiencias en la calidad del empleo.

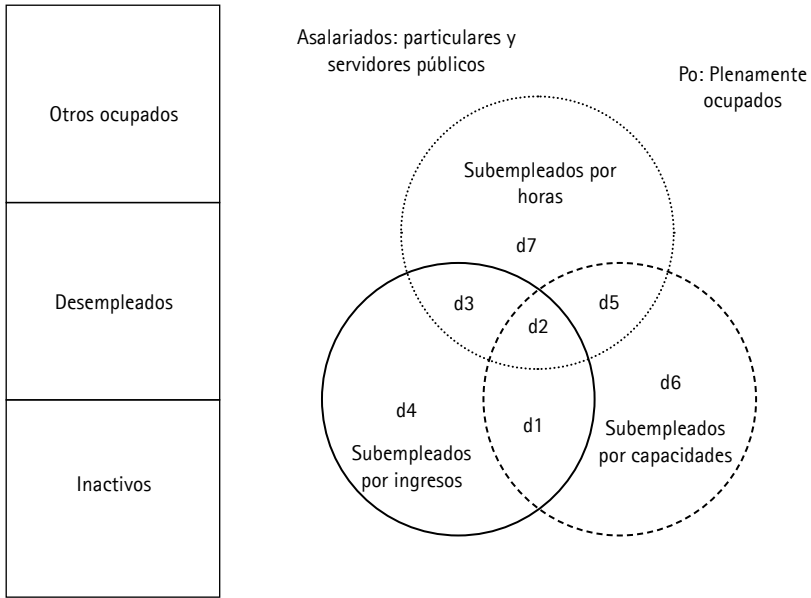
Por otro lado, Adams y Bederman (1974), con información del censo de 1970 identificaron las correlaciones entre el subempleo (trabajadores cuyos ingresos anuales en 1970 fueron inferiores a \$4.000 dólares o que trabajaban tiempo parcial) y algunas características individuales para los empleados de Atlanta. Identificaron que el subempleo afectaba principalmente a las mujeres afrodescendientes, cabezas de hogar y no calificadas, sin importar qué tan cerca se encontraran de las zonas industriales. Osei y Sackey (2006) establecieron cuáles factores demográficos, educativos y de tamaño de la firma son importantes en la incidencia del desempleo y subempleo en Ghana.

A. Algunas características de los clasificados como subempleados por ingresos

Para efectos de esta investigación, la población en edad de trabajar (*PET*) se ha dividido en: a) asalariados particulares y servidores públicos, b) ocupados en posiciones diferentes a los asalariados particulares y empleados del Gobierno, esto es, patrones o empleadores, trabajadores por cuenta propia, empleados domésticos, trabajadores familiares sin remuneración, trabajadores sin remuneración en otras empresas y peones o jornaleros, c) desempleados y d) inactivos (véase gráfico 3).

15 Véanse también las ponencias recopiladas por el DANE (2000), con ocasión del seminario "Subempleo y bienestar social", en las que se hace una caracterización de este fenómeno en Colombia, para el período 1981-1999. El documento incluye ponencias de Alejandro Vivas, Álvaro Suárez, Adriana Mata, Juan C. Guataquí, Liria A. Rodríguez, Hugo López, Marta Luz Henao y Alberto Hernández.

Gráfico 3. Clasificación de la población en edad de trabajar



Nota: el tamaño de las secciones y áreas no representa la proporción de cada grupo en la población en edad de trabajar.

Fuente: elaboración de los autores.

A su vez, el primer grupo (el grupo de interés) se ha dividido en *po*: plenamente ocupados; *d1*: subempleados por ingresos y capacidades; *d2*: subempleados por ingresos, capacidades y horas; *d3*: subempleados por horas e ingresos; *d4*: subempleados por ingresos únicamente; *d5*: subempleados por capacidades y horas; *d6*: subempleados por capacidades únicamente; *d7*: subempleados por horas únicamente.

Para estudiar el subempleo por ingresos, retenemos del grupo de interés a los plenamente ocupados (*po*) y a los subempleados por ingresos (*subi*). En el primer subgrupo se incluyeron aquellos asalariados del sector privado o del sector público que no se consideraron subempleados por ninguno de los motivos (*po*), mientras que en el segundo se incluyeron todos los subempleados por ingresos (*d1*, *d2*, *d3* y *d4*). Dado que las posibilidades de ser subempleado ya sea por horas o por situaciones de empleo inadecuado (ingresos y competencias) no son mutuamente excluyentes, se pueden establecer diferentes combinaciones de subempleo, como se mostró en el gráfico 3.

El cuadro 1 muestra algunos datos extraídos de la GEIH para asalariados particulares y empleados del Gobierno que se catalogan como plenamente ocupados o como subempleados por ingresos, bien sea exclusivos o en combinación con los demás tipos de subempleo (subjetivo), siguiendo la clasificación del gráfico 3. Los valores que aparecen allí deben interpretarse como la proporción de los subempleados en relación con la proporción de los plenamente ocupados; un valor igual a uno significa que las proporciones son iguales.

Como se observa, en la fila correspondiente al género, los hombres representan más o menos la misma proporción entre los subempleados que entre los plenamente ocupados y por ello los valores calculados son cercanos a uno. Las personas entre 12 y 17 años de edad tienen mayor presencia en los subgrupos *d2*, *d3* y en los subempleados totales que en el de los plenamente ocupados. Igual sucede con las personas entre 18 y 24 años, que tienen mayor representación en todos los subgrupos de subempleados que en el de los plenamente ocupados.

Mayor representación que en los plenamente ocupados tienen las personas con educación primaria de los subgrupos *d3* y *d4* y las personas que tienen educación secundaria en todos los subgrupos de subempleados. Ciudades como Cali, Pasto, Ibagué y, en alguna medida, Pereira y sectores como los de la construcción, comercio y transporte tienen mayor prevalencia de subempleo. En contraste, hay muy poca presencia de subempleados en los sectores minero, Gobierno, electricidad e industria.

B. Probabilidad de ser clasificado como subempleado por ingresos

Una manera complementaria de caracterizar a los subempleados es mediante la estimación de la probabilidad de ser subempleado por ingresos, con el propósito de sugerir algunos determinantes, no necesariamente estructurales, a partir de características observables. En este sentido, el cuadro 2 presenta los coeficientes de un modelo *logit* de la probabilidad de que un ocupado sea subempleado por ingresos, en el cual la variable dependiente toma el valor de uno si la persona ocupada es catalogada como tal y de cero en otro caso.

Entre las variables independientes están: edad, educación (escolaridad), género, duración del desempleo¹⁶, tiempo de vinculación a la empresa o entidad actual, tamaño de la empresa, posición en el hogar, estado marital (comprometido o no comprometido), estrato, posición ocupacional, ciudad y sector económico.

Cuadro 1. Algunas características de los subempleados en relación con los plenamente ocupados (GEIH)

Característica	Subempleados por ingresos				Total	Subempleados totales
	Ingresos y capacidades	Ingresos, capacidades y horas	Ingresos y horas	Solo por ingresos		
Subgrupo según notación del gráfico 3	d1	d2	d3	d4		
Género						
Hombre	1,0	0,9	0,9	1,0	1,0	1,0
Edad						
12-17 años	1,0	2,2	2,8	1,1	1,2	1,2
18-24 años	1,6	1,8	1,3	1,2	1,5	1,4
25-34 años	1,2	1,0	0,9	1,0	1,1	1,1
35-44 años	0,8	0,8	1,0	1,0	0,9	0,9
45+ años	0,5	0,6	0,9	0,8	0,6	0,6
Educación						
Primaria (0-5 años)	0,9	1,0	2,2	2,0	1,2	0,7
Secundaria (6-11 años)	1,1	1,2	1,2	1,2	1,2	1,1
Técnico (12-14 años)	1,1	1,0	0,5	0,6	0,9	0,9
Universitario (15 años en adelante)	0,7	0,7	0,4	0,4	0,6	0,7
Parentesco						
Jefe de hogar	0,9	0,7	0,9	1,0	0,9	0,9
Esposo(a)	0,8	0,9	1,0	0,9	0,9	0,9
Hijo	1,3	1,5	1,2	1,0	1,2	1,2
Otro pariente	1,1	1,2	1,1	1,1	1,1	1,1
Cargo desempeñado						
Profesional-técnico	0,6	0,8	0,6	0,4	0,6	0,7
Directores-funcionarios	0,6	0,4	0,2	0,4	0,5	0,5
Personal administrativo	1,0	1,0	0,8	0,7	0,9	0,9
Comerciante	1,1	1,0	0,9	1,0	1,1	1,0
Trabajadores de servicios	1,1	1,4	1,6	1,2	1,2	1,1
Agrícola-forestales	1,0	1,0	1,1	1,5	1,1	1,1

16 Esta variable corresponde a la pregunta de la GEIH: "¿Cuántos meses estuvo sin empleo o trabajo entre el trabajo actual y el anterior?". Dado el signo del coeficiente estimado (cuadro 2), se podría conjeturar que las personas que duran más tiempo desempleadas terminan aceptando ciertos trabajos, tal vez por razones de impaciencia, y después se declaran subempleados por ingresos.

Cuadro 1. Algunas características de los subempleados en relación con los plenamente ocupados (GEIH) (continuación)

Característica	Subempleados por ingresos				Total	Subempleados totales
	Ingresos y capacidades	Ingresos, capacidades y horas	Ingresos y horas	Solo por ingresos		
Subgrupo según notación del gráfico 3	d1	d2	d3	d4		
Ciudad						
Barranquilla	0,4	0,3	0,3	0,5	0,4	0,4
Bucaramanga	1,2	1,1	0,9	1,1	1,2	1,1
Bogotá	1,2	1,1	0,8	0,8	1,1	1,1
Manizales	0,9	1,2	0,7	0,6	0,8	0,9
Medellín	0,9	1,0	1,3	1,1	1,0	1,0
Cali	1,1	1,5	1,6	1,4	1,3	1,3
Pasto	1,1	1,7	2,4	2,4	1,5	1,5
Cartagena	0,3	0,1	0,3	0,9	0,4	0,4
Montería	0,5	0,4	1,3	2,0	0,9	0,9
Villavicencio	1,1	0,6	0,7	1,2	1,1	1,0
Cúcuta	1,1	1,0	0,7	1,2	1,1	1,1
Pereira	0,5	0,4	1,9	1,9	0,9	0,9
Ibagué	1,2	1,6	2,0	1,5	1,4	1,3
Sector económico						
Agrícola	1,0	0,9	0,9	1,3	1,0	1,0
Minas	0,4	0,3	0,8	0,3	0,4	0,4
Industria	1,0	0,7	0,7	1,0	1,0	0,9
Electricidad	0,6	0,3	0,2	0,6	0,6	0,5
Construcción	1,6	1,4	2,2	2,5	1,9	1,7
Comercio	1,3	1,5	1,6	1,3	1,3	1,3
Transporte	1,1	1,1	0,8	1,1	1,1	1,1
Comunicación	1,0	0,9	0,7	0,7	0,9	0,9
Financiero	0,9	1,2	1,0	0,7	0,9	1,0
Gobierno	0,2	0,3	0,2	0,2	0,2	0,3
Salario sector privado	2.317,0	2.408,9	2.234,1	2.234,1	2.262,5	2.323,5
Sal. real mediano/sal. real mediano Gobierno	81,4%	84,6%	78,4%	78,4%	79,4%	81,6%

Nota: un valor igual a uno significa que la proporción correspondiente al subgrupo de subempleados es igual a la proporción entre los plenamente ocupados. Si el valor es mayor que uno, la proporción entre los subempleados es mayor que en los plenamente ocupados y viceversa cuando es menor que uno.

Fuente: GEIH (DANE) y cálculos propios.

En la primera columna del cuadro 2 se observa que la probabilidad de ser subempleado por ingresos se reduce con la antigüedad del individuo en la empresa, para los cónyuges y otros parientes en el hogar y las personas que tienen un cargo de dirección en la empresa donde trabajan. De igual manera,

dicha probabilidad se reduce para quienes residen en vecindarios de los estratos 2 al 6 (en relación con el estrato 1), viven en las ciudades de Barranquilla o Cartagena (en relación con Bogotá), o están vinculados a los sectores de minas, energía o el sector público (en relación con el sector manufacturero). Por el contrario, la probabilidad de ser subempleado por ingresos aumenta con la edad, la educación, la duración del desempleo (el tiempo de búsqueda), la condición de ser mujer y para las personas que viven en Pasto, Cali e Ibagué o que están vinculadas a los sectores de la construcción, comercio y transporte.

Cuadro 2. Probabilidades de ser (clasificado como) subempleado por ingresos

Variable	Modelo <i>logit</i> : probabilidad de ser subempleado por ingresos	Modelos <i>multilogit</i> : probabilidad de ser subempleado por ingresos en relación con la probabilidad de ser plenamente ocupado					
		Subempleado por ingresos exclusivo o en combinación con otras forma de subempleo				Subempleado por ingresos objetivo y subjetivo	
		Ingresos y capacidades: d1	Ingresos, y horas: d2	Ingresos y horas: d3	Solo por ingresos: d4	Subjetivo	Objetivo
Edad	0,025***	0,031***	-0,027***	-0,007	0,043***	0,026***	0,034***
Edad ²	-0,000***	-0,001***	0,000***	0,000	-0,001***	-0,000***	-0,001***
Educación	0,010***	0,062***	0,072***	-0,071***	-0,068***	-0,012***	0,048***
Mujer	0,094***	0,059***	0,149***	0,293***	0,095***	0,103***	0,075***
Duración del desempleo	0,008***	0,007***	0,016***	0,014***	0,004***	0,008***	0,006***
Tiempo de vinculación	-0,007***	-0,006***	-0,011***	-0,012***	-0,005***	-0,004***	-0,013***
Tamaño empresa 1-5	0,722***	0,614***	1,562***	1,338***	0,629***	0,573***	0,937***
Tamaño empresa 6-10	0,379***	0,345***	0,737***	0,627***	0,351***	0,326***	0,458***
Espos(a)	-0,142***	-0,149***	-0,048	0,001	-0,150***	-0,108***	-0,197***
Hijo	-0,008	-0,017	0,138***	0,104**	-0,057***	-0,045***	0,034
Otro pariente	-0,078***	-0,065***	-0,017	-0,052	-0,110***	-0,077***	-0,082***
Comprometido	-0,082***	-0,090***	-0,246***	-0,218***	-0,021	-0,055***	-0,119***
Estrato 2	-0,111***	-0,063***	-0,070**	-0,110**	-0,141***	-0,106***	-0,122***
Estrato 3	-0,172***	-0,106***	-0,143***	-0,355***	-0,251***	-0,170***	-0,176***
Estrato 4	-0,238***	-0,219***	-0,294***	-0,478***	-0,259***	-0,259***	-0,208***
Estrato 5	-0,431***	-0,418***	-0,478***	-1,071***	-0,376***	-0,413***	-0,455***
Estrato 6	-0,632***	-0,654***	-0,704***	-1,346***	-0,621***	-0,558***	-0,736***
Director	-0,299***	-0,142***	-0,992***	-1,401***	-0,273***	-0,298***	-0,287***
Administración	0,086***	0,291***	-0,220***	-0,617***	-0,010	0,066***	0,123***
Comerciante	-0,012	0,195***	-0,501***	-0,980***	0,003	0,004	-0,022
Servicio	0,281***	0,454***	0,082	-0,275***	0,241***	0,277***	0,292***
Agrícola	0,269***	0,516***	0,150	-0,615**	0,140	0,234***	0,336***

Cuadro 2. Probabilidades de ser (clasificado como) subempleado por ingresos (continuación)

Variable	Modelo <i>logit</i> : probabilidad de ser subempleado por ingresos	Modelos <i>multilogit</i> : probabilidad de ser subempleado por ingresos en relación con la probabilidad de ser plenamente ocupado					
		Subempleado por ingresos exclusivo o en combinación con otras formas de subempleo				Subempleado por ingresos objetivo y subjetivo	
		Ingresos y capacidades: d1	Ingresos, capacidades y horas: d2	Ingresos y horas: d3	Solo por ingresos: d4	Subjetivo	Objetivo
Barranquilla	-0,944***	-1,141***	-1,355***	-1,013***	-0,368***	-1,271***	-0,563***
Bucaramanga	0,046**	0,016	-0,020	-0,020	0,212***	0,123***	-0,095***
Manizales	-0,223***	-0,234***	0,183***	-0,155*	-0,356***	-0,256***	-0,170***
Medellín	-0,065***	-0,276***	-0,007	0,504***	0,353***	-0,100***	-0,012
Cali	0,102**	-0,095***	0,399***	0,583***	0,450***	0,122***	0,062**
Pasto	0,274***	-0,196***	0,290***	0,826***	1,010***	0,257***	0,303***
Cartagena	-0,700***	-1,282***	-1,876***	-0,771***	0,345***	-0,853***	-0,478***
Montería	-0,278***	-1,015***	-1,150***	0,225*	0,786***	-0,464***	-0,034
Villavicencio	-0,230***	-0,255***	-0,780***	-0,606***	0,117***	-0,352***	-0,061*
Cúcuta	-0,210***	-0,237***	-0,446***	-0,583***	0,061	-0,087***	-0,448***
Pereira	-0,221***	-0,939***	-0,958***	0,636***	0,756***	-0,308***	-0,088***
Ibagué	0,133***	-0,047*	0,280***	0,697***	0,451***	-0,083***	0,419***
Agropecuario	0,125***	0,079	0,176	0,289	0,140	0,146***	0,080
Minas	-0,552***	-0,641***	-0,364	0,277	-0,625***	-0,560***	-0,554***
Electricidad	-0,176***	-0,165**	-0,599**	-0,413	-0,169*	-0,139*	-0,220***
Construcción	0,321***	0,329***	0,134**	0,119	0,319***	0,269***	0,391***
Comercio	0,095***	0,076***	0,166***	0,342***	0,071***	0,076***	0,128***
Transporte	0,131***	0,123***	0,248***	0,105	0,108***	0,060***	0,232***
Financiero	-0,079***	-0,060***	-0,102	-0,024	-0,155***	-0,085***	-0,067**
Servicios	-0,017	-0,061***	0,344***	0,431***	-0,113***	-0,042**	0,022
Gobierno	-0,683***	-0,739***	0,089	-0,077	-0,924***	-0,622***	-0,793***
Constante	-1,225***	-2,347***	-3,680***	-3,403***	-2,319***	-1,540***	-2,630***
No. observaciones	278.473				278.473		278.473
R ²	0,0940				0,0947		0,0828
Wald	X ² (63)=23.073,73			X ² (240)=36.250,10		X ² (120)=25.712,67	

Nota: coeficientes estimados sin expandir la muestra a la población (sin factores de expansión). Una estrella denota significancia al 10%, dos estrellas al 5% y tres estrellas al 1%.

Fuente: GEIH (DANE) y cálculos propios.

El cuadro 2 contiene también, en el panel central y en las dos columnas de la derecha, dos modelos *logit* multinomiales. El primero estima la probabilidad de ser subempleado por ingresos tanto de manera exclusiva como en combinación con las otras dos formas de subempleo (por competencias y horas), en

relación con la probabilidad de ser plenamente ocupado. El segundo estima la probabilidad de ser subempleado por ingresos subjetivo y objetivo, en relación con la probabilidad de ser plenamente ocupado.

Las probabilidades de ser subempleado por distintos conceptos en relación con la probabilidad de ser plenamente ocupado que se observan en las cuatro columnas del panel central del cuadro 2 son consistentes con las del promedio (primera columna de la izquierda correspondiente al modelo *logit*), en el sentido de que el tiempo de vinculación a la empresa, la condición de estar comprometido (casado o en unión libre), el estrato, desempeñar un oficio de dirección, trabajar en los sectores de minas, eléctrico o en el Gobierno o vivir en la ciudad de Barranquilla reducen, sin ambigüedades, la probabilidad de ser subempleado por ingresos.

Por otro lado, la condición de ser mujer, la duración del desempleo o trabajar en los sectores de la construcción o del comercio aumentan la probabilidad de ser subempleado por ingresos, de manera exclusiva o en combinación con otra modalidad. Con respecto a las demás variables, los signos de los coeficientes varían de un tipo de subempleo a otro. Por ejemplo, la edad no siempre aumenta la probabilidad de ser subempleado, lo cual ocurre solo para $d1$ y $d4$, pero la reduce para $d2$ y $d3$, aunque para estos últimos no es significativa. En el caso de la educación, la probabilidad de ser subempleado para los $d3$ y $d4$ se reduce, pero aumenta para los $d1$ y $d2$ ¹⁷.

En lo que respecta al modelo *logit* multinomial para los subempleados subjetivos y objetivos por ingresos, se observa que la educación reduce la probabilidad de ser subempleado subjetivo en relación con la probabilidad de ser plenamente ocupado, pero aumenta la probabilidad de ser subempleado objetivo por ingresos en relación con la probabilidad de ser plenamente ocupado. Es decir, los subempleados objetivos (aquellos que buscan un nuevo trabajo estando empleados) parecen tener características educativas diferentes a los subempleados subjetivos. La antigüedad en una empresa y la condición de ser mujer actúan, cualitativamente, de la misma manera en las probabilidades de ser subempleado objetivo y subjetivo.

17 Recordemos que $d1$ son subempleados por ingresos y capacidades; $d2$, subempleados por ingresos, capacidades y horas; $d3$, subempleados por horas e ingresos; y $d4$, subempleados por ingresos únicamente.

Las probabilidades de ser catalogado como subempleado objetivo o subjetivo en relación con la probabilidad de ser plenamente ocupado es menor en las ciudades de Barranquilla o Cartagena (frente a Bogotá). Por su parte, vivir en Cali tiene un efecto diferencial en la mayor probabilidad de ser subempleado subjetivo que objetivo.

La probabilidad de ser subempleado por ingresos en relación con la probabilidad de ser plenamente ocupado es menor para quienes residen en vecindarios de estratos dos a seis. Sin embargo, la razón de probabilidades es mayor (en valor absoluto) en el caso de los subempleados objetivos; es decir, la probabilidad de ser subempleado objetivo por ingresos en relación con la probabilidad de ser plenamente ocupado cae más que la de ser subempleado subjetivo. Finalmente, pasar del sector industrial a los sectores de la construcción, comercio o transporte aumenta la probabilidad de ser subempleado en relación con la probabilidad de ser plenamente ocupado.

En suma, según los cuadros 1 y 2 y de acuerdo con la clasificación de subempleados por ingresos que hace el DANE de los asalariados, los más educados, las mujeres, los individuos que habitan en las ciudades de Bogotá, Cali, Pasto e Ibagué o aquellos vinculados a los sectores de la construcción, comercio y transporte caracterizan en mayor medida el subempleo por ingresos. Sin embargo, como veremos en la siguiente sección, pensar que todas estas personas (25,4% de la PEA) son realmente subempleadas por ingresos puede ser una conclusión inexacta que transmite la sensación de un muy mal funcionamiento del mercado de trabajo en Colombia.

II. Brecha salarial entre plenamente ocupados y subempleados por ingresos: verificación y descomposición

La metodología para verificar si los subempleados por ingresos tienen argumentos para ser catalogados como tales consta de dos enfoques. El primero se basa en la ecuación de Mincer (1974) y la descomposición Blinder-Oaxaca, de donde se obtiene el componente no explicado de la brecha salarial entre plenamente ocupados y subempleados por ingresos. El segundo enfoque descompone la brecha de salarios para determinar si esta se sustenta en diferencias en características observables de los individuos. Como predice la teoría de capital humano (véanse Becker, 1993; Mincer, 1958), es resultado de dife-

rencias en los soportes entre plenamente ocupados y subempleados por ingresos, o es producto de la compensación a características no observables que el mercado premia discrecionalmente. Este segundo enfoque sigue la metodología de Ñopo (2003, 2008).

A. Retornos al capital humano y brecha de salarios

Para verificar la validez de la clasificación de una persona como subempleada, se estima una ecuación de Mincer para la población objetivo (asalariados particulares y empleados del Gobierno), la cual toma la forma dada por:

$$w_i = \beta x_i + \varepsilon_i \quad (1)$$

siendo w_i el logaritmo del salario por hora del individuo i , x el vector de características observables, β el vector de retribución de las x y ε el término de error. El conjunto de determinantes incluye variables derivadas de la teoría de capital humano, como escolaridad (años de educación) y experiencia (aproximada por la edad), además de un conjunto de controles como estado marital, género, parentesco con el jefe del hogar, posición ocupacional, ciudad y sector económico. La estrategia adoptada incluye, además, una variable dicotómica ($subi$) que toma el valor de uno si la persona se cataloga como subempleada por ingresos ($d1$, $d2$, $d3$ o $d4$ según el gráfico 3) y de cero si es plenamente ocupada.

En estas condiciones la ecuación de Mincer se especifica como:

$$w_i = \beta x_i + \gamma sub_i + \varepsilon_i \quad (2)$$

El cuadro 3 contiene las estimaciones de la ecuación de Mincer¹⁸, que además de las variables y controles usuales incluye otros como: "subempleo por ingresos", "tiempo de vinculación continua a la empresa actual (antigüedad)", "tamaño de la empresa por número de empleados" (de 1 a 5 empleados, de 6 a 10 y empresas con más de diez empleados) y "duración del desempleo" (el lapso entre el empleo anterior y el actual).

18 Estos resultados se basan en la muestra efectiva, es decir, sin utilizar los factores de expansión. En el anexo 1 se presentan algunas estimaciones utilizando los factores de expansión.

Cuadro 3. Ecuaciones de ingreso

Variable	(1)		(2)		(3)	
	Mincer	Heckman	Mincer	Heckman	Mincer	Heckman
Edad	0,035***	0,048***	0,030***	0,046***	0,026***	0,043***
Edad ²	-0,000***	-0,000***	-0,000***	-0,000***	-0,000***	-0,000***
Educación	0,072***	0,076***	0,065***	0,069***	0,064***	0,067***
Sexo: mujer	-0,076***	-0,112***	-0,071***	-0,113***	-0,066***	-0,115***
Esposa(o)	-0,052***	-0,049***	-0,039***	-0,036***	-0,023***	-0,021***
Hijo	-0,086***	-0,086***	-0,083***	-0,082***	-0,077***	-0,077***
Otro pariente	-0,091***	-0,091***	-0,084***	-0,084***	-0,082***	-0,081***
Comprometido	0,041***	0,039***	0,030***	0,028***	0,025***	0,023***
Director	0,140***	0,139***	0,158***	0,157***	0,165***	0,165***
Administrador	-0,309***	-0,307***	-0,329***	-0,327***	-0,344***	-0,343***
Comerciante	-0,331***	-0,330***	-0,337***	-0,336***	-0,345***	-0,344***
Servicio	-0,430***	-0,429***	-0,442***	-0,441***	-0,462***	-0,461***
Agrícola	-0,324***	-0,322***	-0,315***	-0,312***	-0,338***	-0,335***
Subempleado por ingresos	-0,208***	-0,208***	-0,179***	-0,178***	-0,180***	-0,180***
Duración del desempleo					-0,003***	-0,003***
Tiempo continuo de vinculación en trabajo actual	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***
Tamaño empresa 1-5			-0,239***	-0,240***	-0,227***	-0,227***
Tamaño empresa 6-10			-0,094***	-0,094***	-0,091***	-0,091***
Constante	6,841***	6,543***	7,066***	6,716***	7,199***	6,809***
Número de observaciones	348.636	892.595	348.636	892.595	266.572	892.595
R ²	0,5622		0,5789		0,5666	
λ		0,076***		0,089***		0,089***
Diferencia en descomposición Blinder-Oaxaca	0,4355	0,4355	0,4355	0,4355	0,4167	0,4167
Componente explicado	0,2272	0,2272	0,2568	0,2568	0,2368	0,2368
Componente no explicado	0,2083	0,2083	0,1787	0,1787	0,1798	0,1798

Nota: todos los modelos incluyen efectos fijos por ciudad, sector económico y por trimestre entre 2007 y 2012. Las ecuaciones de Mincer fueron estimadas sin expandir la muestra a la población (sin factores de expansión). Una estrella denota significancia al 10%, dos estrellas al 5% y tres estrellas al 1%.

Fuente: GEIH (DANE) y cálculos de los autores.

Las estimaciones muestran que las personas con mayor escolaridad y experiencia tienen un salario real por hora más alto, en línea con la teoría de capital humano. El retorno de un año de experiencia está entre 2,6% y 3,5%, mientras que el de un año de escolaridad está entre 6,4% y 7,2%. Por otro lado, la condición de ser mujer, esposa(o)¹⁹ e hijo lo reducen, lo cual contrasta con

19 Se refiere a la posición en el hogar, cuyas opciones son: jefe(a), esposo(a), hijo, nieto, otro pariente, otro no pariente.

el efecto en el salario de ser una persona comprometida (casada o en unión libre). Los directivos de las empresas tienen un mayor salario que los profesionales y estos, a su vez, mayor que los comerciantes o trabajadores de servicios o agrícolas.

La antigüedad en la empresa aumenta el salario, así como el trabajar en una firma con más de diez empleados. Se observa que la disminución promedio en el salario real por hora es mayor para quienes trabajan en firmas de menos de cinco empleados (columnas 2 y 3). De igual forma, según el coeficiente correspondiente a la variable "duración del desempleo", cuanto mayor es el número de meses sin ocupación, menor es el salario actual, lo que sugiere que para estas personas el tiempo que están cesantes no les ayuda a encontrar un empleo de mejor remuneración²⁰.

El coeficiente de la variable *subi*, que identifica a los catalogados como subempleados por ingresos, es negativo en todos los casos (se sitúa entre -0,208 y -0,179), lo cual sugiere que ellos ganan un salario inferior a otras personas similares en las dimensiones que muestra el cuadro 3. Este resultado podría ser indicio de que el mercado de trabajo no está funcionando de manera adecuada, ya que: a) una fracción de los salarios en Colombia están mal fijados y su nivel no corresponde a la productividad de los trabajadores o a las exigencias y riesgos inherentes al trabajo; o b) una fracción de las personas están trabajando por un salario inferior a su salario de reserva²¹.

Debido a que los integrantes de la muestra pueden autoseleccionarse, existe la posibilidad de tener estimaciones inconsistentes, lo cual suele ser corregido siguiendo a Heckman (1979). Tal es el caso de quienes reciben un salario, ya que este hecho viene precedido de la decisión de participar en el mercado laboral, la cual toman los individuos comparando el salario de mercado con el salario de reserva. Esto mismo aplica al subempleo, ya que una persona considerada como tal no es seleccionada aleatoriamente y para ser catalogada como subempleada es necesaria la decisión anterior de participar en el mer-

20 Este es un resultado que vale la pena tener en cuenta en las discusiones sobre el establecimiento de un seguro de desempleo en Colombia, para no sobrestimar sus bondades.

21 Es posible que las razones estén en la falta de conocimiento previo sobre las exigencias del trabajo o, inclusive, del propio salario de reserva. Aunque también es posible que haya personas que acepten un trabajo a sabiendas de que el salario que les van a pagar es inferior al precio sombra del ocio, pero que se emplean y continúan buscando (*on-the-job search*).

cado laboral, pese a optar por un salario que, podría ser el caso, no supere su salario de reserva.

Al introducir la corrección de Heckman, se estima un modelo *probit* de participación a partir de una muestra completa de participantes (ocupados y desocupados) y no participantes (inactivos), para así obtener el inverso de la razón de Mills²² (λ) que se incorpora a la ecuación de salarios a estimar, la que ahora está dada por:

$$w_i = \tilde{\beta}x_i + \tilde{\gamma}subi_i + \rho\lambda_i + \epsilon_i \quad (2')$$

Las estimaciones de la ecuación de Mincer con la corrección de Heckman del cuadro 3 sugieren que existe sesgo de selección. Aunque nuestros valores son un poco más similares a los de Guataquí, García y Rodríguez (2009), difieren de manera importante de los estimados por Forero y Gamboa (2007), tanto para los retornos a la educación como para el inverso de la razón de Mills (λ). De todas maneras, el coeficiente asociado al subempleo por ingresos no cambia de manera importante entre las ecuaciones con y sin corrección y sigue estando entre -0,178 y -0,208²³.

El coeficiente de la variable *dummy* "subempleado por ingresos" corresponde al componente no explicado de la brecha salarial de la descomposición Blinder-Oaxaca. En efecto, según este enfoque la brecha del ingreso medio de dos grupos de individuos (subempleados por ingresos, *subi*, y plenamente ocupados, *po*) se descompone en dos términos aditivos: el primero explicado por dife-

22 El inverso de la razón de Mills es la relación de la probabilidad de ser seleccionado en la muestra a la probabilidad de no ser seleccionado. Esto es, $\lambda = \phi(z) / 1 - \Phi(z)$, donde ϕ y Φ son, respectivamente, la función de densidad y la función de distribución acumulada de una variable normal; λ se construye a partir del vector de coeficientes y errores estándar que maximizan una función de máxima verosimilitud.

23 La corrección de Heckman supone que la ecuación de selección de todos los individuos (desocupados e inactivos) diferentes a la población objetivo es la misma y en esta ocasión ese podría no ser el caso ya que el individuo puede ser inactivo, asalariado (plenamente ocupado o subempleado), otro tipo de ocupado (patrón o empleador, cuenta propia, servicio doméstico, trabajador familiar sin remuneración, etc.) y desempleado. Las distintas escalas de la corrección se basan en el gráfico 3. Para dar cuenta de esas múltiples posibilidades, Dubin y McFadden (1984) proponen una corrección cuya selección es multinomial, como es nuestro caso. Nosotros, en particular, consideramos tres posibles estados: inactivo, desocupado y ocupado. Los resultados del modelo con la corrección de Dubin y McFadden (1984) aparecen en el anexo 1, donde no se observan grandes diferencias con respecto al ejercicio anterior en el que no se hacía la corrección multinomial. Bourguignon *et al.* (2007) proveen el código para efectuar dicha corrección.

rencias en las características observables y el segundo, no explicado, debido a diferencias en la retribución a esas características.

De acuerdo con la descomposición Blinder-Oaxaca, el estimador de la brecha de ingresos entre los dos grupos, basado en la ecuación (1), está dado por:

$$\bar{w}^{po} - \bar{w}^{subi} \quad (3)$$

Las ecuaciones de ingresos de dichos grupos están dadas por:

$$w_i^{po} = \beta^{po} x_i^{po} + \varepsilon_i^{po}$$

$$w_i^{subi} = \beta^{subi} x_i^{subi} + \varepsilon_i^{subi}$$

Por tanto, el ingreso esperado de cada grupo es:

$$\bar{w}^{po} = \hat{\beta}^{po} \bar{x}^{po} \quad (4)$$

$$\bar{w}^{subi} = \hat{\beta}^{subi} \bar{x}^{subi} \quad (5)$$

Reemplazando (4) y (5) en (3), tenemos:

$$\bar{w}^{po} - \bar{w}^{subi} = \hat{\beta}^{po} \bar{x}^{po} - \hat{\beta}^{subi} \bar{x}^{subi}$$

El componente no explicado de la brecha se obtiene respondiendo a la pregunta: ¿cuál sería el ingreso de un *subi* si el esquema de compensación de sus características fuera igual al de un *po*? Con tal propósito, a la ecuación anterior se suma y se resta el término $\hat{\beta}^{po} \bar{x}^{subi}$, esto es:

$$\bar{w}^{po} - \bar{w}^{subi} = \hat{\beta}^{po} \bar{x}^{po} - \hat{\beta}^{subi} \bar{x}^{subi} + \hat{\beta}^{po} \bar{x}^{subi} - \hat{\beta}^{po} \bar{x}^{subi}$$

Reordenando se tiene:

$$\Delta = \bar{w}^{po} - \bar{w}^{subi} = \hat{\beta}^{po} (\bar{x}^{po} - \bar{x}^{subi}) + (\hat{\beta}^{po} - \hat{\beta}^{subi}) \bar{x}^{subi} \quad (6)$$

El primer término del componente de la brecha corresponde a la parte explicada ya que el diferencial salarial responde a las diferencias en las características de los dos grupos de individuos. El segundo componente de la brecha es

la parte no explicada debido a que es resultado del diferencial salarial cuando ambos grupos poseen las mismas características²⁴. El coeficiente -0,208 en la columna (1) ó -0,179 en la columna (2) del Cuadro 3, correspondientes a la variable *dummy* "subempleo por ingresos", identifican la parte no explicada de la brecha salarial entre los *po* y los *subi*.

La asociación entre la variable *dummy* de la especificación (2) y el segundo elemento del término de la derecha de (6) es sencilla. Ahora suponga que la ecuación de salarios es:

$$w_i^{subi} = \beta^{po} x_i^{subi} - \gamma sub_i + \varepsilon_i^{subi} \quad (7)$$

En este caso, la contraparte de (5) es:

$$\bar{w}^{subi} = \hat{\beta}^{subi} \bar{x}^{subi} = \hat{\beta}^{po} \bar{x}^{subi} - \hat{\gamma} psub_i$$

donde *psubi* es la proporción de subempleados por ingresos en el total de asalariados. De esta forma, la brecha está dada por:

$$\Delta = \bar{w}^{po} - \bar{w}^{subi} = \hat{\beta}^{po} (\bar{x}^{po} - \bar{x}^{subi}) + \gamma psub_i \quad (8)$$

La comparación de (6) y (8) permite concluir que la *dummy* de la especificación (1) recoge la parte no explicada de la brecha salarial entre plenamente ocupados y subempleados por ingresos. Según la descomposición Blinder-Oaxaca ubicada en la parte baja del Cuadro 3, la brecha es del orden de 42%-44%²⁵, siendo un poco mayor la parte explicada por diferencias de capital humano y otros controles que la parte no explicada (de 23 a 26 puntos porcentuales). De momento, el componente restante de la brecha queda sin explicar.

Lo anterior es evidencia de que, en promedio, una persona catalogada como subempleada gana un salario inferior que un plenamente ocupado, aun con sus mismas características observables. Esto puede ser indicativo de que

24 Algunas veces, dependiendo de los dos grupos que se estén analizando, la parte no explicada de la brecha se asocia con "discriminación". En la notación de Ñopo (2003, 2008), utilizada más adelante, este componente se identifica por Δ_o , mientras que el componente explicado de la brecha se denota por Δ_x .

25 En la sección I se señaló que la diferencia en los salarios de los subempleados subjetivos y los plenamente ocupados era de 40%.

tiene un salario menor que su productividad o que su salario de reserva²⁶. Sin embargo, también podría no ser un indicio suficiente de que, efectivamente, sea subempleada por ingresos sino que simplemente esté ganando un salario menor debido a que: a) la productividad que perciben los empleadores del subempleo refleja deficiencias en la calidad de la educación recibida; b) su productividad es inferior que la de alguien a quien se le paga por su mejor disposición en la empresa, sus mejores ideas u otras características no observables²⁷; c) aunque su productividad sea igual a la de alguien de las mismas características observables, está comparando su salario con el de un individuo vinculado a una empresa o sector en el que se acostumbra el pago de incentivos por resultados o donde resulta más difícil (costoso) verificar el esfuerzo de los trabajadores; d) está comparando su salario con el de alguien que tiene un salario superior en vista del riesgo que enfrenta en el desarrollo de este (teoría hedonista de los salarios) u otra heterogeneidad no observable de los trabajos; y e) el subempleo está comparando su salario con el de alguien que trabaja en un sector en el que los sindicatos tienen poder de negociación en momentos en que el sector está viviendo un período de auge.

Pese a que estas hipótesis son difíciles de verificar con los datos disponibles del mercado de trabajo en Colombia y que no son mutuamente excluyentes, en lo que sigue se tratará de acotar un poco más el componente no explicado de la brecha salarial introduciendo interacciones del subempleo por ingresos con la escolaridad, la posición ocupacional y el sector económico²⁸. Las dos primeras interacciones buscan identificar en qué sector de la distribución de salarios se presenta el subempleo por ingresos y acercarse a las hipótesis a), b) y posiblemente c), mientras que la interacción del subempleo por ingresos con el sector económico busca acercarse a las posibilidades restantes.

26 Porque formó mal sus expectativas sobre el salario de mercado o subvaloró equivocadamente su salario de reserva.

27 De nuevo se puede tratar de alguien que está dotado de los cinco grandes rasgos de la personalidad: extroversión, mente abierta y tolerancia, estabilidad emocional, responsabilidad y amabilidad o empatía.

28 En relación con el cuadro 3, se dejó de utilizar la duración del desempleo debido a que el valor, si bien es significativo, es numéricamente muy pequeño, y además reduce mucho el tamaño de la muestra. Se incluyen dos *dummies* de trimestre para estimar la caída en los salarios reales ocurrida durante la desaceleración que se presentó entre 2008 y 2009 en la economía colombiana.

Según el modelo (1) del cuadro 4, el coeficiente de la *dummy* de subempleo por ingresos cae a -0,014 debido a la incorporación de la interacción de la subempleo por ingresos y educación²⁹. Se observa que el coeficiente de esta última es negativo (-0,015), lo cual afecta de paso el retorno de la educación; sugiere que a medida que aumenta la educación, cae el salario promedio de un subempleado. A mayor educación, más alto el salario de un plenamente ocupado y mayor que el de un *subi*.

Cuadro 4. Ecuaciones de ingreso con interacciones de la variable "subempleo por ingresos"

Variable	(1)		(2)		(3)		(4)	
	Mincer	Heckman	Mincer	Heckman	Mincer	Heckman	Mincer	Heckman
Edad	0,030***	0,045***	0,030***	0,045***	0,030***	0,045***	0,030***	0,045***
Edad ²	-0,000***	-0,000***	-0,000***	-0,000***	-0,000***	-0,000***	-0,000***	-0,000***
Educación	0,070***	0,074***	0,070***	0,074***	0,069***	0,073***	0,070***	0,073***
Sexo: mujer	-0,072***	-0,114***	-0,072***	-0,114***	-0,072***	-0,113***	-0,072***	-0,113***
Espos(a)	-0,039***	-0,036***	-0,039***	-0,036***	-0,039***	-0,036***	-0,039***	-0,039***
Hijo	-0,081***	-0,081***	-0,081***	-0,080***	-0,081***	-0,080***	-0,081***	-0,080***
Otro pariente	-0,083***	-0,083***	-0,083***	-0,083***	-0,083***	-0,082***	-0,083***	-0,082***
Comprometido	0,030***	0,028***	0,030***	0,028***	0,030***	0,028***	0,030***	0,028***
Director	0,157***	0,156***	0,157***	0,156***	0,191***	0,190***	0,190***	0,189***
Administrador	-0,322***	-0,320***	-0,322***	-0,320***	-0,327***	-0,325***	-0,327***	-0,325***
Comerciante	-0,330***	-0,329***	-0,330***	-0,329***	-0,319***	-0,318***	-0,317***	-0,316***
Servicio	-0,434***	-0,433***	-0,434***	-0,433***	-0,445***	-0,444***	-0,441***	-0,441***
Agrícola	-0,307***	-0,305***	-0,307***	-0,305***	-0,300***	-0,297***	-0,298***	-0,296***
Subempleado por ingresos	-0,014**	-0,014**	0,001	0,001	-0,015**	-0,014**	0,002	0,001
Educación x subempleado por ingresos	-0,015***	-0,015***	-0,015***	-0,015***	-0,014***	-0,014***	-0,015***	-0,015***
Subempleado objetivo por ingresos			-0,054***	-0,054***	-0,054***	-0,054***	-0,053***	-0,053***
Tiempo continuo de vinculación en trabajo actual	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***
Tamaño empresa 1-5	-0,242***	-0,242***	-0,240***	-0,240***	-0,240***	-0,240***	-0,240***	-0,240***
Tamaño empresa 6-10	-0,094***	-0,093***	-0,093***	-0,093***	-0,093***	-0,093***	-0,093***	-0,093***

29 Se incluyen *dummies* de trimestre para estimar la caída en los salarios reales ocurrida durante las fases de expansión y desaceleración, como, por ejemplo, la que se presentó entre 2008 y 2009 en la economía colombiana.

Cuadro 4. Ecuaciones de ingreso con interacciones de la variable "subempleo por ingresos" (continuación)

Variable	(1)		(2)		(3)		(4)	
	Mincer	Heckman	Mincer	Heckman	Mincer	Heckman	Mincer	Heckman
Director x subempleo por ingresos					-0,213***	-0,213***	-0,210***	-0,209***
Administrador x subempleo por ingresos					0,023***	0,022***	0,023***	0,023***
Comerciante x subempleo por ingresos					-0,034***	-0,035***	-0,044***	-0,045***
Servicio x subempleo por ingresos					0,039***	0,038***	0,025***	0,025***
Agrícola x subempleo por ingresos					-0,027	-0,026	-0,031	-0,030
Agropecuario	0,020**	0,021**	0,020**	0,021**	0,020**	0,021**	0,022**	0,023**
Minas	0,426***	0,425***	0,426***	0,426***	0,427***	0,427***	0,454***	0,453***
Electricidad	0,109***	0,109***	0,109***	0,109***	0,110***	0,110***	0,110***	0,110***
Construcción	0,138***	0,140***	0,139***	0,141***	0,140***	0,142***	0,149***	0,151***
Comercio	-0,023***	-0,022***	-0,023***	-0,022***	-0,024***	-0,023***	-0,027***	-0,026***
Transporte	-0,074***	-0,073***	-0,074***	-0,073***	-0,073***	-0,072***	-0,058***	-0,057***
Financiero	0,107***	0,108***	0,107***	0,109***	0,107***	0,109***	0,108***	0,109***
Servicios	0,035***	0,036***	0,035***	0,036***	0,037***	0,038***	0,032***	0,033***
Gobierno	0,262***	0,263***	0,262***	0,264***	0,266***	0,268***	0,263***	0,265***
Agropecuario x subempleo por ingresos							-0,009	-0,008
Minas x subempleo por ingresos							-0,184***	-0,182***
Electricidad x subempleo por ingresos							-0,001	-0,001
Construcción x subempleo por ingresos							-0,026***	-0,025***
Comercio x subempleo por ingresos							0,011	0,011*
Transporte x subempleo por ingresos							-0,054***	-0,054***
Financiero x subempleo por ingresos							-0,002	-0,002

Cuadro 4. Ecuaciones de ingreso con interacciones de la variable "subempleo por ingresos" (continuación)

Variable	(1)		(2)		(3)		(4)	
	Mincer	Heckman	Mincer	Heckman	Mincer	Heckman	Mincer	Heckman
Servicios x subempleo por ingresos							0,019***	0,019***
Gobierno x subempleo por ingresos							0,034***	0,031***
Tercer trimestre 2008	-0,046***	-0,046***	-0,046***	-0,045***	-0,046***	-0,0465***	-0,045***	-0,045***
Cuarto trimestre 2008	-0,011**	-0,010**	-0,010**	-0,010**	-0,010**	-0,010**	-0,010**	-0,010**
Constante	7,022***	6,676***	7,021***	6,675**	7,026***	6,683***	7,022***	6,684***
λ		0,088***		0,088***		0,087***		0,086***
Número de observaciones	348.636	892.595	348.636	892.595	348.636	892.595	348.636	892.595
R^2	0,5803		0,5806		0,5811		0,5813	
Diferencia en descomposición Blinder-Oaxaca	0,4355	0,4355	0,4355	0,4355	0,4355	0,4355	0,4355	0,4355
Componente explicado	0,4218	0,4218	0,4370	0,4370	0,4208	0,4208	0,4373	0,4373
Componente no explicado	0,0136	0,0136	-0,0015	-0,0015	0,0146	0,0146	-0,0018	-0,0018

Nota: todos los modelos incluyen efectos fijos por ciudad y trimestre entre 2007 y 2012. Las ecuaciones de Mincer fueron estimadas sin expandir la muestra a la población (sin factores de expansión). Una estrella denota significancia al 10%, dos estrellas al 5% y tres estrellas al 1%.

Fuente: GEIH (DANE) y cálculos de los autores.

En cualquier caso, la parte no explicada de la brecha salarial sigue siendo $-0,179$ (véase la columna 2 del cuadro 3), pero ahora se ha encontrado que el componente no explicado covaría con la escolaridad; por tanto, puede señalarse que el subempleo por ingresos es una característica de la parte derecha de la distribución de salarios³⁰. En la columna (2), se adiciona la *dummy* "subempleado objetivo por ingresos" para diferenciar la caída en sus ingresos frente a los plenamente ocupados y los subempleados subjetivos, y se observa que esta condición deja de ser significativa, pero los subempleados objetivos se muestran ganando 5,4% menos que un plenamente ocupado, mientras que los subempleados subjetivos presentan un menor salario únicamente asociado con una mayor escolaridad.

30 Con el uso de esta interacción pierde significancia la interacción de "mujer-subempleo por ingresos", por lo que dejó de incluirse en las especificaciones del cuadro 4.

El modelo de la columna (3) incluye, adicionalmente, la interacción de "subempleados subjetivos por ingresos" con "posición ocupacional", y se observan algunos resultados importantes, además de cambios en algunos coeficientes. En primer lugar, varían los coeficientes estimados del "subempleo por ingresos", el cual vuelve a ser significativo (y negativo), y el retorno a la educación, que aumenta un poco para los subempleados (el coeficiente se ubica ahora en $-0,015$).

Sin embargo, se observa que el salario de los "directores-subempleados por ingresos" es ahora menor que el de los "profesionales" y demás posiciones ocupacionales. Esto ratifica que el subempleo por ingresos se presenta básicamente en los salarios ubicados en la parte derecha de la distribución, en la que, además, se presenta la mayor dispersión³¹. Las dos interacciones incluidas hasta ahora permiten pensar que la brecha no explicada de los salarios entre los plenamente ocupados y subempleados por ingresos puede estar conectada con las hipótesis a), b) y c) señaladas antes, pero también con el desconocimiento de los salarios en la parte alta de la distribución.

Finalmente, en la última columna del cuadro 4 se incluye una interacción entre "subempleo por ingresos y sector económico", para tratar de discernir algo sobre las hipótesis d) y e), que permitirían dilucidar el componente no explicado de la brecha. Allí, la variable "pivote" es el subempleo por ingresos en el sector industrial. De acuerdo con ello, los subempleados del sector minero ganan sustancialmente menos que los del sector industrial; igual sucede, pero en menor medida, con los subempleados por ingresos de los sectores transporte y comercio.

Es posible que en el sector minero, en el que el poder de negociación de los empleados no es despreciable, haya una alta dispersión salarial debido a que algunos salarios (de los plenamente ocupados) reflejan primas asociadas a riesgos o incentivos que al final hacen que muchas personas sean catalogadas como subempleadas.

31 El anexo 2 presenta el *kernel* de la distribución de los salarios por nivel educativo. Se observa allí que a mayor escolaridad, mayor la dispersión de salarios y, por ende, aumenta la probabilidad de encontrar más personas subempleadas por ingresos.

Sin embargo, debemos recordar que se trata simplemente de hipótesis que se deben probar formalmente ya que, de momento, el resultado es que el mercado de trabajo podría estar operando en condiciones de ineficiencia; no tanto como lo indica el resultado grueso de una diferencia salarial del 40%-42% entre plenamente ocupados y subempleados por ingresos, pero sí alrededor del 18%. De todas maneras, este resultado podría corresponder a una sobrestimación del componente no explicado de la brecha, que se corrige mediante el enfoque de Ñopo (2003, 2008) y que se presenta a continuación.

B. Depuración del componente no explicado de la brecha salarial

Apoyándose en el enfoque Blinder-Oaxaca, Hugo Ñopo (2003, 2008) desarrolló una metodología alternativa, cuyas ventajas radican en no hacer supuestos *a priori* sobre la distribución de los datos y no requerir estimaciones de ecuaciones de ingreso. Esta técnica no paramétrica basada en el emparejamiento³² de individuos con características análogas (soporte común) reconoce la existencia de diferencias en los soportes de la distribución y compara individuos cuyas características son, en efecto, comparables, ya que comparten un soporte común.

Según esta metodología, los ocupados (plenamente ocupados o subempleados) se consideran como tratamiento y se utiliza el emparejamiento para seleccionar la submuestra de unos y otros que no presenta diferencia en las características observables³³. En consecuencia, la descomposición de la brecha salarial basada en las diferencias en la distribución individual de características, que da cuenta de las discrepancias en los soportes, se calcula a partir de la diferencia en el ingreso de un plenamente ocupado y un subempleado por ingresos, así:

$$\Delta = E[w|plenamente ocupado] - E[w|subempleado por ingresos] \quad (9)$$

32 Las técnicas de *matching comparison* buscan emparejar muestras con características observables similares, excepto para la variable de tratamiento que es usada para agrupar observaciones en dos conjuntos: tratamiento y control. Una vez se han controlado las características observables, la técnica de comparación permite medir el impacto del tratamiento en estos grupos bajo diferentes supuestos.

33 En el anexo 3 se presenta más formalmente la metodología de Ñopo (2008) aplicada en este trabajo.

La metodología descompone la brecha en cuatro componentes aditivos:

$$\Delta = \Delta_x + \Delta_{po} + \Delta_{subi} + \Delta_0 \quad (10)$$

siendo Δ_x la fracción de la brecha asociada con diferencias sobre el mismo soporte en la distribución de características individuales de subempleados y plenamente ocupados; Δ_{po} el componente sustentado en combinaciones de características que poseen los plenamente ocupados, pero que no son comparables con las de los subempleados por ingresos; Δ_{subi} la parte de la brecha salarial debida a combinaciones de características que tienen los subempleados pero no los plenamente ocupados; finalmente, Δ_0 representa la porción de la brecha salarial que se explica por características no observables que pueden afectar la productividad.

El componente Δ_x corresponde, así como en Blinder-Oaxaca, a la parte explicada de la brecha salarial mientras que Δ_0 , Δ_{subi} y Δ_{po} suman lo que en esta descomposición se denomina componente no explicado, pero que en el enfoque de Ñopo se depura al reconocer la existencia de individuos que no son comparables.

El procedimiento de emparejamiento consiste en: a) seleccionar de la muestra un individuo del conjunto de subempleados por ingresos; b) tomar todas las personas plenamente ocupadas que poseen las mismas características que el individuo previamente seleccionado; c) construir, a partir de estos, un individuo "sintético", cuyo salario corresponda al promedio de los plenamente ocupados que fueron emparejados; d) llevar al plenamente ocupado sintético y al subempleado por ingresos a una nueva muestra de individuos emparejados; e) finalmente, repetir los pasos anteriores hasta agotar los subempleados por ingresos en la muestra.

Una vez se realiza este ejercicio para todos los subempleados por ingresos, se obtiene una base de datos que contiene personas subempleadas para las cuales no fue posible encontrar como mínimo un plenamente ocupado que tuviera las mismas características, plenamente ocupados que nunca fueron seleccionados dado que no tenían un subempleado semejante, y plenamente ocupados y subempleados emparejados. Como lo señalan Hoyos, Ñopo y Peña (2010), por construcción, el conjunto compuesto por los subempleados por ingresos y plenamente ocupados emparejados no tiene diferencias en la dis-

tribución de las características, es decir que cualquier brecha en los salarios obedece a diferencias en la forma en que el mercado premia características no observables. De ser así, se estaría verificando que las personas que se consideran subempleadas por ingresos ganan un salario inferior que un plenamente ocupado y que esto puede deberse a que su salario es menor que su productividad o su salario de reserva, y no a comparaciones equivocadas en algunas dimensiones con otras personas.

La descomposición de la brecha mediante el emparejamiento de Ñopo fue realizada para ocho especificaciones diferentes, las cuales adicionan una característica observable a la anterior. La primera especificación para seleccionar la muestra del paso a) anterior, la más elemental, considera dos variables: edad y educación, que representan las características que podrían explicar la diferencia salarial en mayor medida³⁴. En la segunda se incluye el género; la tercera agrega el estrato; la cuarta adiciona la posición en el hogar; la quinta incorpora el estado marital; la sexta, el cargo que la persona desempeña en la empresa y, posteriormente, se agregan la ciudad en la que vive el individuo y el sector económico.

El cuadro 5 muestra el porcentaje de subempleados y plenamente ocupados que forman parte del soporte común, es decir, el porcentaje de individuos para los cuales fue posible encontrar una contraparte con las mismas características en el otro grupo, según cada especificación. El resultado indica que al controlar por la mayor cantidad de variables, para 73,3% (=100%-26,7%) de los subempleados y 86,8% (=100%-13,2%) de los plenamente ocupados no fue posible encontrar al menos una persona del grupo opuesto que tuviera las mismas características.

El cuadro 6 muestra la descomposición en función de los cuatro componentes de la brecha de salario, medida en términos relativos (con respecto al salario de los plenamente ocupados). De esta manera, para el período muestral se estima que, en promedio, los subempleados por ingresos ganan 40,1% menos que los plenamente ocupados.

34 En cumplimiento del paso b), se construye un sintético con los *po* que tienen la misma edad y educación y se continúa con el procedimiento descrito antes.

Cuadro 5. Porcentaje de subempleados por ingresos y plenamente ocupados emparejados para diferentes soportes

Asalariados	Controlando por:							
	Edad y educación	(i) + género	(ii) + estrato	(iii) + posición en el hogar	(iv) + estado marital	(v) + ocupación	(vi) + ciudad	(vii) + sector económico
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Subempleados	100,0%	99,9%	99,0%	96,4%	94,0%	83,5%	48,5%	26,7%
Plenamente ocupados	99,3%	98,7%	94,5%	88,4%	84,6%	67,8%	29,0%	13,2%

Fuente: GEIH (DANE) y cálculos propios.

Cuadro 6. Descomposición de la brecha de salario según cada especificación (porcentaje)

	Controlando por:							
	Edad y educación	(i) + género	(ii) + estrato	(iii) + posición en el hogar	(iv) + estado marital	(v) + ocupación	(vi) + ciudad	(vii) + sector económico
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Δ	-40,1	-40,1	-40,1	-40,1	-40,1	-40,1	-40,1	-40,1
Δ_0	-26,8	-26,7	-24,3	-22,6	-21,1	-19,5	-17,4	-14,4
Δ_{subi}	0,0	0,0	0,1	0,3	0,2	0,5	2,1	5,1
Δ_{po}	-1,1	-2,2	-7,1	-9,8	-11,1	-15,8	-25,8	-33,2
Δ_x	-12,1	-11,1	-8,8	-7,9	-7,1	-5,3	1,0	2,4

Fuente: DANE-GEIH cálculos propios.

Es interesante observar que a medida que aumentan los argumentos del soporte, disminuyen los componentes no explicados de la brecha salarial (Δ_0) y explicados por diferencias en la distribución de las características observables (Δ_x). Al mismo tiempo, aumenta la proporción de la brecha explicada por diferencias en los soportes; en especial, por características que tienen los plenamente ocupados pero que no tienen los subempleados (Δ_{po}). Allí está la mayor parte de la brecha salarial³⁵. Este es el principal aporte de la metodología de Ñopo: hacer una estimación más precisa del componente no explicado de la brecha salarial que el que se obtiene con la descomposición Blinder-Oaxaca, y vemos según el soporte, cómo varía entre -14,4 y -26,8.

35 Cuando en los soportes se tienen todos los argumentos de la columna *viii*) del cuadro 5, 31,8 puntos porcentuales de la brecha de 40,2% se deben a características que tienen los *po* pero que no tienen los *subi*.

En el gráfico 4 (panel superior izquierdo) se observa que la brecha es menor para niveles educativos bajos (primaria); pero a medida que el individuo adquiere más educación, el diferencial salarial aumenta y supera el 38,1%, en promedio, para el nivel educativo universitario (personas con más de quince años de educación) y el 31,7% para personas que tienen posgrado.

Por género no se encuentran diferencias importantes en la brecha de salarios entre hombres plenamente ocupados y subempleados por ingresos (38,7%) ni entre mujeres de ambos grupos (41,7%). A diferencia de la descomposición de la brecha por nivel educativo, donde la mayor parte de las brechas salariales corresponde al componente no explicado (Δ_o), en el caso del género el principal componente corresponde al de diferencias en el soporte por características que tienen los plenamente ocupados pero que no poseen los subempleados por ingresos (panel superior derecho).

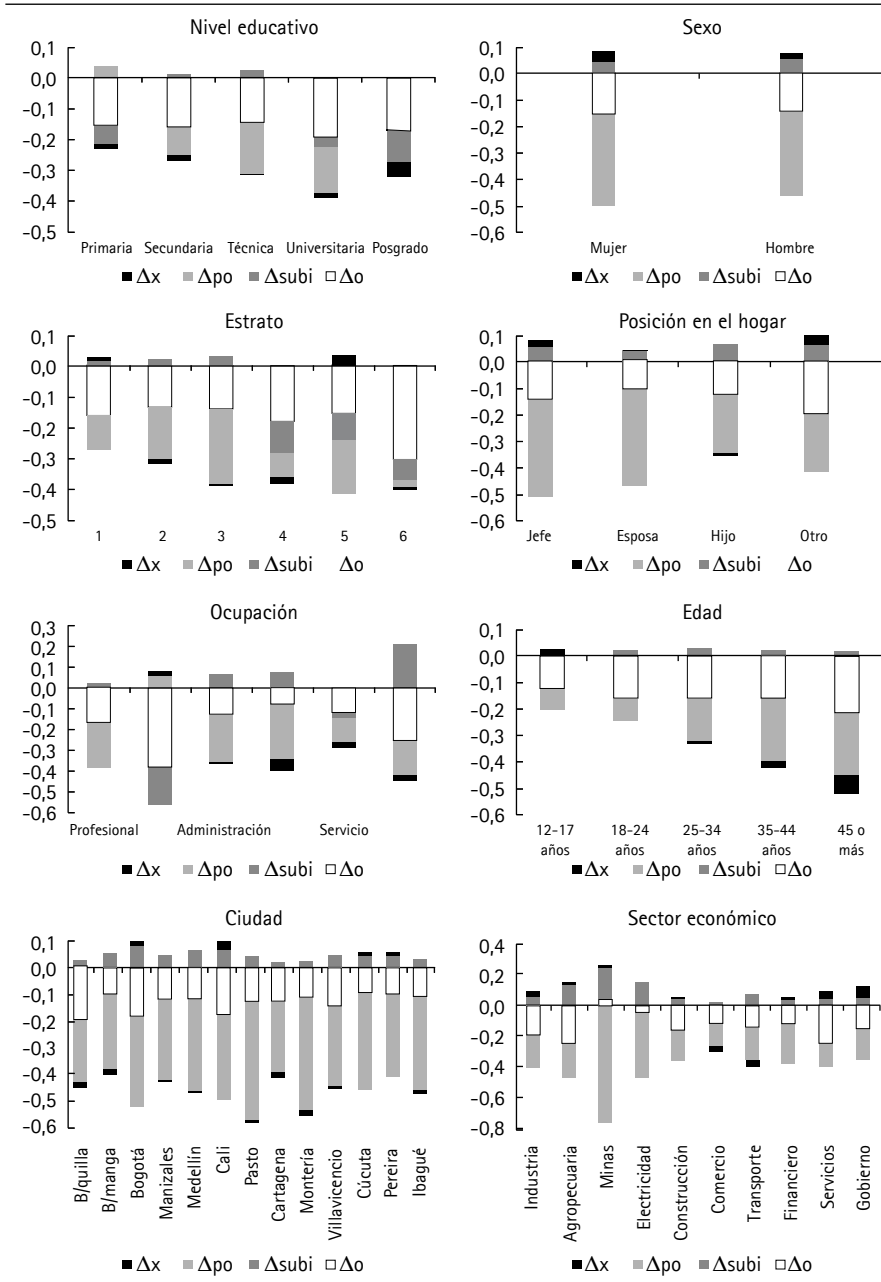
III. Conclusiones

Este artículo pone en duda que la alta tasa de subempleo por ingresos que se observa hoy en día en Colombia, correspondiente a la brecha salarial entre los asalariados que no se quejan de su salario y quienes sí lo hacen (subempleados por ingresos), constituya un buen indicador de algunas de las deficiencias del mercado de trabajo en el país, en cuanto a la asignación de salarios se refiere.

Así, los autores someten a prueba la hipótesis de que quienes son catalogados como subempleados por ingresos ganan realmente un salario inferior que el de los plenamente ocupados, una vez tomadas en cuenta las diferencias en las características observables. Dado lo anterior, se estima la magnitud de la brecha no explicada por la teoría del capital humano.

Para verificar si la brecha salarial se debe a salarios mal asignados, se utiliza la ecuación de Mincer, la descomposición Blinder-Oaxaca y la descomposición de Ñopo, y se verifica en todos los casos la existencia de la brecha, la cual se sitúa entre 40% y 44% del salario real por hora, que tiene un componente explicado y otro no explicado.

Gráfico 4. Descomposición de la brecha salarial: especificación (viii) del cuadro 6



Nota: el soporte correspondiente a la especificación (viii) incluye: edad, educación, género, estrato, posición en el hogar, estado civil, ocupación, ciudad y sector económico.

Fuente: GEIH (DANE) y cálculos propios.

Mediante el primer enfoque (Mincer y la descomposición Blinder-Oaxaca), se estima que, en efecto, existe un componente no explicado de la brecha que está alrededor de 18%, mientras que con el enfoque de Ñopo se estima en cerca de 14%. Es decir, la evidencia apunta a que en promedio la parte no explicada de la brecha, que daría justificación a que una persona se califique como subempleada por ingresos, está entre 36% y 40%. Por tanto, en lo que se refiere a esta dimensión, el mercado de trabajo funciona mejor de lo que, a primera vista, sugieren las cifras de subempleo.

La pregunta que surge es: ¿por qué hay subempleo por ingresos? Las respuestas (hipótesis) que surgen son varias. La primera es que, en efecto, los salarios en Colombia están mal determinados y, en tal sentido, el mercado estaría funcionando de manera menos que eficiente, pero en mucha menor medida de lo que sugieren las cifras brutas. Sin embargo, quedan por verificar las demás hipótesis mencionadas antes.

Pese a que esto es difícil con la información disponible, con la ayuda de algunas interacciones se logró establecer que el subempleo por ingresos se presenta en personas cuyo salario se ubica en la parte derecha de la distribución; esto es, que tienen más educación y experiencia, que tienen cargos de dirección y que viven en vecindarios de estratos altos. Es decir, el subempleo está más asociado con una mala formación de expectativas con respecto al salario de mercado³⁶, o a la subvaloración (equivocada) del salario de reserva, o a que la productividad que perciben los empleadores refleja deficiencias en la calidad de la educación recibida por los subempleados, pese a tener la misma escolaridad que un plenamente ocupado, o a que su productividad es inferior a la de alguien a quien se le paga por su mejor disposición en la empresa, sus mejores ideas u otras características no observables.

Además, los subempleados están vinculados a los sectores de minas y transporte. Si bien Ibagué, Pasto, Montería y Barranquilla tienen las mayores brechas salariales, los componentes no explicados de importancia se encuentran en Pasto, Villavicencio, Barranquilla, Cali y Bogotá. Esto permitiría focalizar un poco más la búsqueda de sectores con márgenes (*mark-up*) importantes o en los que hay que mejorar la información sobre salarios.

36 Las personas cuyos ingresos están en la parte baja de la distribución tienen menos incertidumbre de lo que podría ser su salario y allí hay menos indicios de la existencia de subempleo por ingresos.

Pero de esta investigación también quedan recomendaciones. La primera es que, dado que buena parte de la brecha salarial se explica por condiciones observables de las personas, no es adecuado utilizar todo el subempleo por ingresos como indicador de "condiciones de empleo inadecuado" o "empleo de baja calidad", como se hace hoy en día. Tanto la tasa de subempleo por ingresos como estos calificativos diezman de forma imprecisa e innecesaria el desempeño del mercado laboral. Entre 2007 y 2012, la tasa de subempleo subjetivo por ingresos fue de 25,4%, pero muy seguramente esta cifra sobrestima la población subempleada. La segunda es que un mejor conocimiento de los salarios en el mercado es indispensable para que los individuos tomen decisiones más acertadas. En tal sentido, se debe mejorar la información tanto de oferentes de mano de obra (incluyendo el salario de reserva) como de vacantes y de los puestos ocupados (incluyendo los salarios que pagan).

Referencias

1. ADAMS, J. y BEDERMAN, S. (1974). "Job accessibility and underemployment", *Annals of the Association of American Geographers*, 64(3):378-386.
2. BECKER, G. (1993). *Human capital: A theoretical and empirical analysis, with special reference to education*, NBER, Chicago, Tercera edición.
3. BLINDER, A. (1974). *Toward an economic theory of income distribution*. Cambridge, MA: MIT Press.
4. DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA (DANE) (2000). Ponencias del seminario "Subempleo y bienestar social", en DANE, Pontificia Universidad Javeriana, Instituto de Políticas de Desarrollo, Bogotá.
5. DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA (DANE) (2006). Ficha metodológica Gran Encuesta Integrada de Hogares.
6. FARNÉ, S. (2013). "Determinantes del subempleo en Colombia: un enfoque a través de la compensación salarial, comentarios", en Arango

y Hamann (eds.), *Mercado de trabajo en Colombia: hechos tendencias e instituciones* (pp. 707-709). Bogotá: Banco de la República.

7. FORERO, N. y GAMBOA, L. (2007). "Cambios en los retornos de la educación en Bogotá entre 1997 y 2003", *Lecturas de Economía*, 66:225-250.
8. GUATAQUÍ, J. C., GARCÍA, A. F. y RODRÍGUEZ, M. (2009). Estimaciones de los determinantes de los ingresos laborales en Colombia con consideraciones diferenciales para asalariados y cuenta propia (Documentos de Trabajo 70). Universidad del Rosario.
9. HECKMAN, J. (1979). "Sample selection bias as a specification error", *Econometrica*, 47:153-162.
10. HOYOS, A., ÑOPO, H. y PEÑA, X. (2010). The persistent gender earnings gap in Colombia, 1994-2006 (Documento CEDE 16). Universidad de los Andes, Facultad de Economía.
11. KAUFMAN, B. E. y HOTCHKISS, J. L. (2006). *The economics of labor markets* (7ª ed.). Thompson, Mason, OH.
12. LÓPEZ, H. (2000). "Desempleo y subempleo: desafíos estratégicos para la economía colombiana", en DANE, Subempleo y bienestar social, 131-154.
13. LÓPEZ, S. (2012). "Estimación del efecto colegio en Colombia: 1980-2009", *Estudios Gerenciales*, 28(122):49-68.
14. MINCER, J. (1958). "Investment in human capital and income distribution", *JPE*, 66(4):281-302.
15. MINCER, J. (1974). *Schooling, experience and earnings*. NBER, Chicago.
16. MUELLER, G. y PLUG, E. (2006). "Estimating the effect of personality on male and female earnings", *Industrial and Labor Relations Review*, 60(1):3-22.

17. ÑOPO, H. (2003). "Matching as a tool to decompose wage gaps", *The Review of Economics and Statistics*, Middlebury College Economics Discussion Paper 04-06.
18. ÑOPO, H. (2007). An extension of the Blinder-Oaxaca decomposition to a continuum of comparison groups (Discussion Paper 2921, julio). IZA.
19. ÑOPO, H. (2008). "Matching as a tool to decompose wage gaps", *The Review of Economics and Statistics*, 90(2):290-299.
20. OAXACA, R., (1973). "Male-female wage differentials in urban labor markets", *International Economic Review*, 14:693-709.
21. OSEI, B. y SACKKEY, H. (2006). "Human resource underutilization in an era of poverty reduction: An analysis of unemployment and underemployment in Ghana", *African Development Review*, 18(2):163-182.
22. PUYANA, R., RAMOS, M. y ZÁRATE, H. (2013). "Determinantes del subempleo en Colombia: un enfoque a través de la compensación salarial", en Arango y Hamann (eds.), *Mercado de trabajo en Colombia: hechos, tendencias e instituciones* (p. 707). Bogotá: Banco de la República.
23. URIBE, J., ORTIZ, C. H. y GARCÍA, G. (2008). "Informalidad y subempleo en Colombia: dos caras de la misma moneda", *Cuadernos de Administración*, 21(37):211-241.
24. URRUTIA, M. (2002). "El subempleo en Colombia", *Revista del Banco de la República*, nota editorial.
25. WICHERT, L. y POHLMEIER, W. (2010). Female labor force participation and the big five (Discussion Paper 10-003). IZA, Center for European Economic Research.

Anexo 1. Ecuaciones de ingreso con factores de expansión y corrección del sesgo de selección

Cuadro A1.1. Ecuaciones de ingreso con factores de expansión y corrección de Dubin y McFadden (D-MF)

Variable	(1)		(2)		(3)		(4)	
	Mincer	D-MF	Mincer	D-MF	Mincer	D-MF	Mincer	D-MF
Edad	0,030***	0,045***	0,030***	0,045***	0,030***	0,045***	0,030***	0,045***
Edad ²	-0,000***	-0,000***	-0,000***	-0,000***	-0,000***	-0,000***	-0,000***	-0,000***
Educación	0,075***	0,074***	0,075***	0,074***	0,074***	0,074***	0,075***	0,074***
Sexo: mujer	-0,076***	-0,117***	-0,076***	-0,118***	-0,076***	-0,117***	-0,076***	-0,117***
Esposo(a)	-0,036***	-0,037***	-0,036***	-0,037***	-0,036***	-0,037***	-0,036***	-0,037***
Hijo	-0,098***	-0,081***	-0,097***	-0,081***	-0,097***	-0,080***	-0,097***	-0,080***
Otro pariente	-0,094***	-0,083***	-0,094***	-0,083***	-0,093***	-0,083***	-0,092***	-0,083***
Comprometido	0,020***	0,028***	0,020***	0,028***	0,020***	0,028***	0,020***	0,028***
Director	0,217***	0,156***	0,217***	0,156***	0,253***	0,190***	0,252***	0,189***
Administrador	-0,333***	-0,320***	-0,333***	-0,320***	-0,340***	-0,325***	-0,340***	-0,325***
Comerciante	-0,335***	-0,329***	-0,336***	-0,329***	-0,326***	-0,319***	-0,324***	-0,316***
Servicio	-0,441***	-0,434***	-0,441***	-0,434***	-0,455***	-0,444***	-0,452***	-0,441***
Agrícola	-0,356***	-0,305***	-0,356***	-0,305***	-0,334***	-0,298***	-0,332***	-0,297***
Subempleado por ingresos	0,027***	-0,014***	0,039***	0,001	0,016	-0,015**	0,038***	0,001
Educación x subempleado por ingresos	-0,018***	-0,015***	-0,018***	-0,015***	-0,016***	-0,014***	-0,017***	-0,015***
Subempleado objetivo por ingresos			-0,050***	-0,054***	-0,050***	-0,054***	-0,048***	-0,053***
Tiempo continuo de vinculación en trabajo actual	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***	0,001***
Tamaño empresa 1-5	-0,223***	-0,242***	-0,222***	-0,240***	-0,222***	-0,240***	-0,221***	-0,240***
Tamaño empresa 6-10	-0,083***	-0,093***	-0,083***	-0,093***	-0,083***	-0,093***	-0,083***	-0,093***
Director x subempleado por ingresos					-0,225***	-0,213***	-0,221***	-0,209***
Administrador x subempleado por ingresos					0,029***	0,022***	0,031***	0,023***
Comerciante x subempleado por ingresos					-0,029***	-0,035***	-0,035***	-0,045***
Servicio x subempleado por ingresos					0,048***	0,038***	0,038***	0,025***
Agrícola x subempleado por ingresos					-0,075**	-0,026	-0,082	-0,030

Cuadro A1.1. Ecuaciones de ingreso con factores de expansión y corrección de Dubin y McFadden (D-MF) (continuación)

Variable	(1)		(2)		(3)		(4)	
	Mincer	D-MF	Mincer	D-MF	Mincer	D-MF	Mincer	D-MF
Agropecuario	0,018	0,021**	0,018	0,020**	0,018	0,020**	0,021	0,023**
Minas	0,493***	0,425***	0,493***	0,426***	0,495***	0,427***	0,526***	0,453***
Electricidad	0,113***	0,109***	0,113***	0,109***	0,113***	0,110***	0,119***	0,110***
Construcción	0,131***	0,140***	0,132***	0,141***	0,133***	0,141***	0,145***	0,150***
Comercio	-0,017***	-0,023***	-0,017***	-0,022***	-0,018***	-0,023***	-0,019***	-0,026***
Transporte	-0,069***	-0,073***	-0,068***	-0,073***	-0,068***	-0,073***	-0,054***	-0,057***
Financiero	0,106***	0,108***	0,106***	0,108***	0,106***	0,108***	0,108***	0,109***
Servicios	0,020***	0,036***	0,020***	0,036***	0,021***	0,037***	0,020***	0,033***
Gobierno	0,180***	0,263***	0,180***	0,263***	0,185***	0,268***	0,178***	0,265***
Agropecuario x subempleado por ingresos							-0,008	-0,009
Minas x subempleado por ingresos							-0,227***	-0,182***
Electricidad x subempleado por ingresos							-0,028	-0,001
Construcción x subempleado por ingresos							-0,034***	-0,025***
Comercio x subempleado por ingresos							0,003	0,011
Transporte x subempleado por ingresos							-0,051***	-0,054***
Financiero x subempleado por ingresos							-0,006	-0,002
Servicios x subempleado por ingresos							0,005	0,019***
Gobierno x subempleado por ingresos							0,083***	0,031***
Tercer trimestre 2008	-0,059***	-0,046***	-0,059***	-0,045***	-0,059***	-0,046***	-0,059***	-0,045***
Cuarto trimestre 2008	-0,026***	-0,011**	-0,026***	-0,010**	-0,026***	-0,010**	-0,026***	-0,010**
m1		0,000		0,000		0,000		0,000
m2		-0,160***		-0,163***		-0,161***		-0,160***
m3		0,107***		0,107***		0,106***		0,105***
Constante	6,984***	6,630***	6,984***	6,625***	6,990***	6,635***	6,985***	6,636***
Número de obs.	348.636	348.636	348.636	348.636	348.636	348.636	348.636	348.636
R ²	0,5561	0,5804	0,5563	0,5807	0,5570	0,5812	0,5572	0,5814
σ ²		0,230		0,230		0,230		0,230

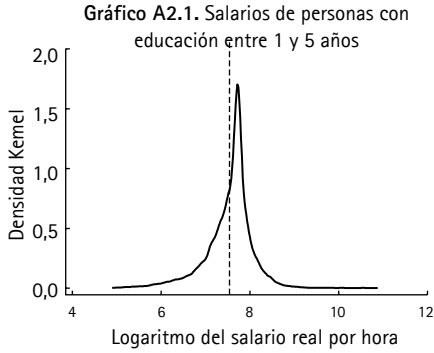
Cuadro A1.1. Ecuaciones de ingreso con factores de expansión y corrección de Dubin y McFadden (D-MF) (continuación)

Variable	(1)		(2)		(3)		(4)	
	Mincer	D-MF	Mincer	D-MF	Mincer	D-MF	Mincer	D-MF
Diferencia en descomposición de Blinder y Oaxaca	0,4355	0,4355	0,4355	0,4355	0,4355	0,4355	0,4355	0,4355
Componente explicado	0,4218	0,4218	0,4370	0,4370	0,4208	0,4208	0,4373	0,4373
Componente no explicado	0,0136	0,0136	-0,0015	-0,0015	0,0146	0,0146	-0,0018	-0,0018

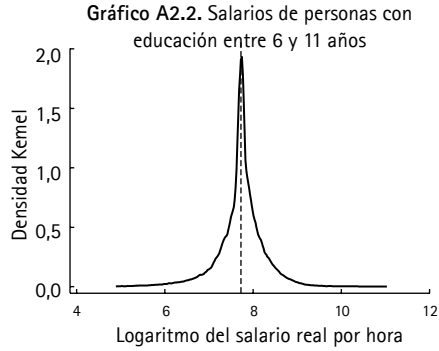
Nota: todos los modelos incluyen efectos fijos por ciudad y trimestre entre 2007 y 2012. Una estrella denota significancia al 10%, dos estrellas al 5% y tres estrellas al 1%. Las variables m_1 , m_2 y m_3 corresponden al valor esperado condicional del residuo de los modelos *logit* multinomiales. Son indicativos de la covarianza entre los residuales del modelo de selección y la ecuación de ingresos.

Fuente: GEIH (DANE) y cálculos de los autores.

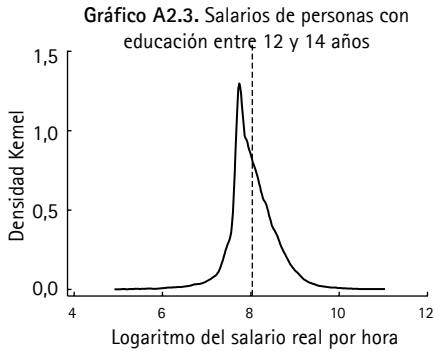
Anexo 2. Distribución de salarios por nivel educativo



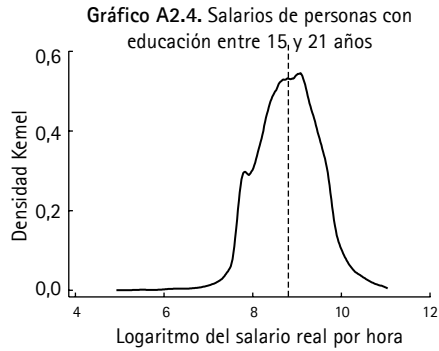
La línea representa la media del salario real por hora: 2,144
 Varianza: 0,2783755. Apuntalamiento: -0,879294
 Curtosis: 6,387618



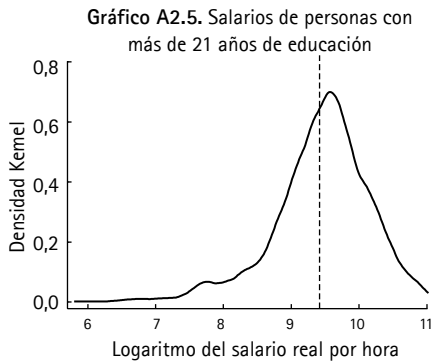
La línea representa la media del salario real por hora: 2,583
 Varianza: 0,2673028. Apuntalamiento: -0,5440438
 Curtosis: 6,83145



La línea representa la media del salario real por hora: 3,608
 Varianza: 0,2809897. Apuntalamiento: 0,2320693
 Curtosis: 6,079942



La línea representa la media del salario real por hora: 8,495
 Varianza: 0,4825766. Apuntalamiento: -0,0582251
 Curtosis: 3,278433



La línea representa la media del salario real por hora: 15,066
 Varianza: 0,4420371. Apuntalamiento: -0,5231063
 Curtosis: 4,0185

Anexo 3. Descomposición de Ñopo

El enfoque de Ñopo (2003, 2007, 2008) parte de definir F^{po} y F^{subi} como las funciones de distribución acumuladas de las características individuales x , condicionales a ser plenamente ocupado (po) o subempleado por ingresos ($subi$), posiblemente relacionadas con el ingreso de unos y otros; y $dF^{po}(\cdot)$ y $dF^{subi}(\cdot)$ como sus funciones de probabilidad.

Sean $\mu^{subi}(L)$ la probabilidad de un conjunto L bajo la distribución $dF^{subi}(\cdot)$, es decir: $\mu^{subi}(L) = \int_L dF^{subi}(x)$ y, de manera análoga, $\mu^{po}(L) = \int_L dF^{po}(x)$. El valor esperado de los ingresos condicionales a las características y a su situación ocupacional (po y $subi$) está dado por:

$$E[w|po, x] = g^{po}(x) \Rightarrow E[w|po] = \int_{L^{po}} g^{po}(x) dF^{po}(x) \quad (A3.1)$$

$$E[w|subi, x] = g^{subi}(x) \Rightarrow E[w|subi] = \int_{L^{subi}} g^{subi}(x) dF^{subi}(x) \quad (A3.2)$$

Sean L^{po} y L^{subi} el soporte de la distribución de características de los plenamente ocupados y subempleados por ingresos. Con estas condiciones, la brecha de salario se define como:

$$\Delta \equiv E[w|po] - E[w|subi] \quad (A3.3)$$

Reemplazando (A3.1) y (A3.2) en (A3.3), se obtiene:

$$\Delta = \int_{L^{po}} g^{po}(x) dF^{po}(x) - \int_{L^{subi}} g^{subi}(x) dF^{subi}(x) \quad (A3.4)$$

Considerando que el soporte de distribución de L^{subi} es diferente a L^{po} , cada integral es separada sobre su respectivo dominio en dos partes:

$$\Delta = \left[\int_{L^{subi} \cap L^{po}} g^{po}(x) dF^{po}(x) + \int_{L^{subi} \cap \overline{L^{po}}} g^{po}(x) dF^{po}(x) \right] - \left[\int_{L^{subi} \cap L^{po}} g^{subi}(x) dF^{subi}(x) + \int_{L^{subi} \cap \overline{L^{po}}} g^{subi}(x) dF^{subi}(x) \right] \quad (A3.5)$$

Como $dF^{po}(\cdot)$ y $dF^{subi}(\cdot)$ son cero por fuera de su soporte, el dominio de la primera y cuarta integrales (con soportes no comunes) puede ser $\overline{L^{subi}}$ y $\overline{L^{po}}$, respectivamente, sin afectar los valores correspondientes. Adicionalmente, cada integral puede ser reescalada para obtener expresiones que involucren los

valores esperados de $g^{subi}(x)$ y $g^{po}(x)$, condicional a su respectivo dominio particionado.

$$\Delta = \left[\int_{L^{subi}} g^{po}(x) \frac{dF^{po}(x)}{\mu^{po}(L^{subi})} \right] \mu^{po}(L^{subi}) - \left[\int_{L^{po} \cap L^{subi}} g^{po}(x) \frac{dF^{po}(x)}{\mu^{po}(L^{subi})} \right] \mu^{po}(L^{subi}) - \left[\int_{L^{po} \cap L^{subi}} g^{subi}(x) \frac{dF^{subi}(x)}{\mu^{subi}(L)} \right] \mu^{subi}(L) - \left[\int_{L^{po}} g^{subi}(x) \frac{dF^{subi}(x)}{\mu^{subi}(L^{po})} \right] \mu^{subi}(L^{po}) \tag{A3.6}$$

Ahora, reemplazando $\mu^{subi}(L^{po})$ por $1 - \mu^{subi}(L^{po})$ y $\mu^{po}(L^{subi})$ por $1 - \mu^{po}(L^{subi})$, la brecha se puede expresar como:

$$\Delta = \left[\int_{L^{subi}} g^{po}(x) \frac{dF^{po}(x)}{\mu^{po}(L^{subi})} - \int_{L^{subi}} g^{po}(x) \frac{dF^{po}(x)}{\mu^{po}(L^{subi})} \right] \mu^{po}(L^{subi}) + \left[\int_{L^{po} \cap L^{subi}} g^{po}(x) \frac{dF^{po}(x)}{\mu^{po}(L^{subi})} - \int_{L^{po} \cap L^{subi}} g^{subi}(x) \frac{dF^{subi}(x)}{\mu^S(L^{po})} \right] + \left[\int_{L^{po}} g^{subi}(x) \frac{dF^{subi}(x)}{\mu^{subi}(L^{po})} - \int_{L^{po}} g^{subi}(x) \frac{dF^{subi}(x)}{\mu^{subi}(L^{po})} \right] \mu^{subi}(L^{po}) \tag{A3.7}$$

Las integrales que son contabilizadas sobre el soporte común se pueden descomponer en cuatro componentes aditivos adicionando y restando $\int_{L^{po}} g^{po}(x) \frac{dF^{subi}(x)}{\mu^{subi}(L^{po})}$, elemento que permite evaluar el contrafactual.

$$\Delta = \left[\int_{L^{subi}} g^{po}(x) \frac{dF^{po}(x)}{\mu^{po}(L^{subi})} - \int_{L^{subi}} g^{po}(x) \frac{dF^{po}(x)}{\mu^{po}(L^{subi})} \right] \mu^{po}(L^{subi}) + \int_{L^{po} \cap L^{subi}} g^{po}(x) \left[\frac{dF^{po}}{\mu^{po}(L^{subi})} - \frac{dF^{subi}}{\mu^S(L^{po})} \right] (x) + \underbrace{\int_{L^{po} \cap L^{subi}} [g^{po}(x) - g^{subi}(x)] \frac{dF^{subi}(x)}{\mu^{subi}(L^{po})}}_{\text{}} + \left[\int_{L^{po}} g^{subi}(x) \frac{dF^{subi}(x)}{\mu^{subi}(L^{po})} - \int_{L^{po}} g^{subi}(x) \frac{dF^{subi}(x)}{\mu^{subi}(L^{po})} \right] \mu^{subi}(L^{po}) \tag{A3.8}$$

Así, la brecha se puede descomponer de la siguiente forma:

$$\Delta = \Delta_{po} + \Delta_x + \Delta_0 + \Delta_{subi} \quad (\text{A3.9})$$

