

# RETORNOS DE LA EDUCACIÓN PARA LOS TRABAJADORES FORMALS E INFORMALES EN CALI: UNA APROXIMACIÓN CON REGRESIONES CUANTÍLICAS Y SPLINES LINEALES<sup>1</sup>

RETURNS OF EDUCATION FOR FORMAL AND INFORMAL WORKERS IN CALI: AN APPROXIMATION WITH QUANTILE REGRESSION AND LINEAR SPLINES

LUCROS DA EDUCAÇÃO PARA OS TRABALHADORES FORMALS E INFORMAIS EM CALI: UMA APROXIMAÇÃO COM CUANTÍLICAS DE REGRESSÕES E SPLINES LINEAR

Iván Andrés Ordóñez-Castaño

Maestrante en Economía Aplicada de la Universidad del Valle. Especialista en Finanzas de la Universidad Eafit. Miembro del grupo de investigación OIKOS. Profesor investigador de la Universidad Cooperativa de Colombia, Cali - Colombia.  
ivan.ordonezc@campusucc.edu.co

Johana Andrea Sanabria-Domínguez

Estudiante de Maestría en Economía Aplicada y Economista de la Universidad del Valle, Cali - Colombia. Profesional Centro Regional de Estudios Económicos Cali, Banco de la República.  
johana.sanabria@correounivalle.edu.co

• **Clasificación JEL:** C5, J15, J31, J45, R10

## RESUMEN

to work in the formal sector. Our findings confirm that returns of education vary across schooling and income levels.

En este artículo se analizan los retornos de la educación para empleados formales e informales con diferentes niveles de ingreso en el municipio de Cali. Se utilizan los datos de la Encuesta de Empleo y Calidad de Vida (EECV) de 2013 para estimar una ecuación de salario con splines lineales, usando la técnica semiparamétrica de regresión cuantílica con corrección del sesgo de selección. Los resultados sugieren que la tasa de retorno de la educación secundaria y técnica en el sector informal supera a la del sector formal para todos los niveles de ingreso. En contraste, con un nivel de educación superior es más rentable emplearse en el sector formal. En conclusión, se corrobora que el rendimiento marginal de la inversión en capital humano no es constante entre niveles educativos y tampoco entre niveles de ingresos.

## KEYWORDS

Returns of education, sample selection, quantile regression, informality, selection bias.

## PALABRAS CLAVE

Retornos de la educación, selección muestral, regresión cuantílica, informalidad, sesgo de selección.

## RESUMO

São analisados os lucros da educação para empregados formais e informais com níveis de entrada diferentes na municipalidade de Cali neste artigo. São usados os dados da Pesquisa de Emprego e Qualidade de Vida (EECV) de 2013 para calcular uma equação de salário com splines linear, enquanto usando o semiparamétrica técnico de quantílica de regressão com correção do preconceito de seleção. Os resultados sugerem que a taxa de retorno da educação secundária e técnica no setor informal superado isso do setor formal para todos os níveis de entrada. Em contraste, com um nível de educação superior é mais lucrativo para ser usado no setor formal. Em conclusão, é confirmado que o rendimento marginal do investimento em capital humano não é constante entre níveis educacionais e nem entre níveis de rendas.

## ABSTRACT

In this article we analyze returns of education for formal and informal workers with different levels of income in Cali. We use data from 2013 "Encuesta de Empleo y Calidad de Vida" (EECV) in order to estimate a semi-parametric quantile regression for wages with linear splines and correcting for selection bias. Our results suggest that the marginal rate of return of high-school and technical education is higher for informal workers than for formal ones, across the whole income distribution. In contrast, for workers with higher education levels, it is more profitable

## PALAVRAS-CHAVE

Lucros da educação, muestral de seleção, cuantílica de regressão, informalidade, preconceito de seleção.

## Introducción

La acción de elegir se fundamenta en la intención y los factores que hacen que la elección sea viable o no para quien la toma. Laboralmente ésta depende de varios factores, como lo expresa Recio (1997), entre otras para cubrir las necesidades propias y familiares y por los valores agregados que puede aportar cada individuo en un contexto laboral, con referencia a la educación y al sector en el que se decide participar.

Hay dos factores importantes que inciden en la elección de intervenir en el mercado laboral, uno es la educación como fundamento del capital humano, el cual debe ser remunerado para dos agentes, según Arvanitis y Loukis (2009): para quien lo provee y por quien lo contrata. Quien lo provee pone al servicio de la unidad productiva, con el objetivo de tener un ingreso como retorno a su capital humano. El segundo, quien contrata o requiere de ese capital humano pretende un mayor impacto en la productividad de su actividad económica.

El otro factor es el sector en el que se desea participar. Portes (1995) denota con claridad dos sectores, el formal e informal, haciendo una análisis en dos posturas, una refiriéndose a la remuneración del trabajador, dependiendo del lugar donde se encuentra como fundamento entre la relación del factor trabajo y capital, este a diferencia del factor analizado en el párrafo anterior hace referencia al capital físico. Es decir, la remuneración al trabajo está asociada al sector al cual se ha vinculado, aunque no es apropiado afirmar que el nivel de ésta se asocia al sector, Uribe, Ortiz y García (2006) dicen que hay empleos precarios formales y algunos trabajadores informales pueden tener ingresos superiores.

Bajo la condición del sector que genera el puesto de trabajo y el retorno de la educación como factor del capital humano, la propuesta en el presente artículo se fundamenta en saber cuál es el retorno a la educación para un individuo que desarrolla su labor en el sector formal o informal: Para responder a la pregunta en el marco teórico se fundamenta la educación como capital humano y la economía informal. Metodológicamente se sigue la técnica semiparamétrica de regresión cuantílica con corrección paramétrica del sesgo de selección a la ecuación de salarios con splines lineales.

Las secciones que se desarrollan en el presente artículo y dan sustento a la respuesta del interrogante planteado

en la investigación son: primera, la revisión de literatura bajo los planteamientos teóricos y trabajos que son pertinentes como referencia para el desarrollo de la investigación. La siguiente es la metodología, donde se presentan las regresiones cuantílicas y la forma como se corrige el sesgo de selección. Posteriormente una descripción de los datos presentes en la Encuesta de Empleo y Calidad de Vida para Cali entre noviembre de 2012 y enero de 2013, que permite la construcción de las variables: logaritmo del salario mensual, nivel de educación, experiencia y su cuadrado, género e informalidad desde el enfoque mixto; y resultados del modelo econométrico. Por último, las conclusiones producto de la repuesta a la pregunta de investigación.

## 1. Revisión de literatura

### 1.1. MARCO TEÓRICO

Según el enfoque económico neoclásico, los factores productivos son remunerados por su productividad marginal; a mayor productividad mayor remuneración. Esta es la base sobre la cual los individuos toman la decisión de educarse. Esto significa que las personas, bien sea por su nivel educativo o producto de experiencias propias o cercanas, tienen algún grado de conocimiento de los beneficios que sobre su nivel de productividad, y por consiguiente sobre su salario, tiene un mayor nivel de educación.

### 1.2. RETORNO DE LA EDUCACIÓN

La educación constituye un componente del capital humano de las personas y como lo plantea Mincer (1958), hay una relación positiva entre el capital humano y los ingresos laborales. Cuando los individuos dedican tiempo a la formación escolar incurren en un costo debido a que dejan de percibir ingresos salariales actuales, con la idea de que las ganancias serán recibidas, en una mayor proporción, a una edad más avanzada. Es decir, la educación genera retornos salariales en relación con el nivel de inversión en capacitación, en un periodo posterior, una vez el individuo participa como ocupado en el mercado laboral.

No obstante, el retorno de la educación no es homogéneo para todos los individuos porque existen heterogeneidades en las productividades (asociadas al nivel educativo y la experiencia), en la estructura del mercado laboral y en las características propias

del individuo. En cuanto a la estructura del mercado laboral, al distinguir entre formal e informal, es posible señalar que la tasa de retorno de la educación es más baja para los trabajadores informales porque desarrollan actividades de baja productividad, tienen empleos inestables y laboran en empresas pequeñas, intensivas en mano de obra y con limitaciones para pagar salarios altos (Uribe y Ortiz, 2006).

En este sentido, el retorno a la educación se ve afectado por el sector económico (formal o informal) y la percepción que se tiene sobre el empleo. lo cual genera entornos propicios para el desarrollo laboral de una persona, reflejado en un ingreso más valorado. Este planteamiento es el fundamento de Gallart (2001) para determinar que hay aspectos del entorno social que afectan los retornos de la formación escolar.

### 1.3. SECTOR ECONÓMICO FORMAL E INFORMAL

La teoría dualista de la economía plantea la existencia de dos sectores: uno moderno y otro tradicional. El sector moderno o formal agrupa grandes empresas, intensivas en capital, con altos niveles de productividad, oferta de salarios altos y empleos de alta calidad. Por su parte, el sector tradicional o informal agrupa pequeñas empresas, con escasa productividad, intensivas en mano de obra, oferta de salarios bajos y empleos de baja calidad (Jiménez, 2012).

Sobre lo anterior se fundamentan dos enfoques para estudiar la informalidad laboral: el estructuralista y el institucionalista, referenciados por Jiménez (2012). El primero sostiene que, en los países en vía de desarrollo, existe una estructura económica y productiva que impide que los demandantes de trabajo en el mercado laboral moderno absorban toda la mano de obra que es ofrecida. Por lo tanto, la informalidad es una opción de supervivencia para los oferentes de trabajo que no alcanzan un empleo en dicho sector.

El segundo enfoque, el institucionalista, afirma que hay empleados y empresas que evaden el cumplimiento de las obligaciones legales y por lo tanto los clasifica como informales. Además, sostiene que quienes son informales lo son por decisión propia, no por subsistencia, pues los costos laborales en el sector formal son muy altos y la opción más racional es el incumplimiento de las disposiciones legales.

No obstante, algunos autores como Freije (2002) y Gasparini y Tornarolli (2009), han hecho trabajos tratando la informalidad desde un enfoque mixto, definido como un híbrido entre las dos visiones anteriores. La informalidad es toda actividad desarrollada fuera de los estándares establecidos por la sociedad y sus leyes del contexto de la empresa, pero a su vez tiene la capacidad de generar ingresos, que para quienes desarrollan esta actividad son su medio para satisfacer las necesidades básicas, y por lo tanto, imprimen un dinamismo económico en el entorno (Portes & Haller, 2004).

Al involucrar el concepto de sector no estructurado para la actividad informal, en la economía informal, según lo cita Cárdenas y Mejía (2007), el Programa de Recuperación de Empleo en América Latina y el Caribe (PRELAC) en 1978, determina que esta actividad se caracteriza desde los trabajos informales. Schneider (2006) hace una clasificación de la informalidad en relación con el tipo de transacción involucrada (monetaria o no monetaria) y si es lícita, desde la evasión y elusión de impuesto, dando una taxonomía de la economía informal.

Por otra parte, en 1972 la Organización Internacional del Trabajo (OIT), define la actividad informal o sector informal como sector no estructurado, para expresar las duras actividades de los trabajadores pobres que no eran reconocidas, registradas, protegidas o reguladas por las autoridades públicas (OIT, 2002). En el mismo informe se define la informalidad en función de las relaciones laborales, así: 1) trabajo independiente en microempresas individuales; 2) trabajo familiar auxiliar; 3) miembros de cooperativas de productores informales; 4) trabajo a domicilio en régimen de tercerización; y, 5) trabajo por cuenta propia en actividades de subsistencia.

### 1.4. MARCO DE REFERENCIA

Méndez (2002), en una publicación sobre empleo informal, sustenta su estudio en las regresiones de ecuaciones mincerianas para estimar el retorno de la educación, denotando que este varía entre hombre y mujeres, al igual que entre el sector económico sobre el cual participan, formal o informal. También tiene en cuenta que la probabilidad de participar en un sector genera el denominado sesgo de selección. Establece además que, según el sector, se pagan salarios diferentes y evidencia que la tasa de retorno de la educación en

el sector formal es inferior al informal, no obstante en el primero se goza de mejores ingresos; la justificación está en que la experiencia es una de las principales características generadoras de ingreso.

Para determinar la rentabilidad de la educación en cada sector, el autor usa un modelo spline para demostrar que individuos que alcanzan una educación superior a quince años tienen cambios significativos en la pendiente, en relación con los que tienen menor educación acumulada.

Por su parte, Zárate (2003) estima una ecuación de salario usando regresiones cuantílicas, técnica semiparamétrica que permite estimar la distribución condicional del salario dado un conjunto de atributos de la variable independiente. La relevancia de la regresión cuantílica es que permite detectar cambios en la respuesta a causa de variables observadas en puntos distintos de la distribución, caracterizando completamente la distribución salarial. En este estudio, la variable dependiente es el logaritmo natural del salario con covariables explicativas que son: la educación, la experiencia y su cuadrado, al igual que una dicotómica de género.

El autor encuentra que los retornos de la educación son mayores en los cuantiles superiores, pero los de la experiencia lo son en los cuantiles inferiores. Además, anota que el incremento de las diferencias salariales está explicado por variaciones en la retribución de las habilidades observadas y modificación en la desigualdad dentro de los grupos.

Prada (2006) intenta responder si es rentable estudiar en Colombia para hombres y mujeres por separado. Para ello estima ecuaciones de salario con splines lineales y regresiones por percentiles. Para controlar el sesgo de selección, se fundamenta en el modelo análogo de Heckman, referenciando en su artículo a Buchinsky (2001), el cual consiste en un modelo probit para la ecuación de participación en una primera etapa y una ecuación de salarios mediante una regresión por percentiles para la segunda.

Las conclusiones a las que llega Prada indican que los retornos de una educación superior son más riesgosos y de allí su volatilidad en diferentes percentiles. Además, comprobó que los empleados con mayores niveles educativos logran una mejor retribución a su educación, con mayor homogeneidad en hombres que en mujeres, así como que los trabajadores con mayor

nivel educativo enfrentan un rendimiento decreciente en el retorno de su educación, cuya justificación está cimentada en la saturación del mercado laboral para algunas profesiones.

Contreras, Melo y Ojeda (2005) examinan el sesgo en su dirección y magnitud en las variables no observables para el retorno de la educación en Chile. Sus planteamientos se fundamentan en que la educación únicamente explicaría una parte de la variación de los ingresos individuales. Los autores plantean una regresión en diferencias sobre datos panel en una ecuación minceriana. Se utilizan este tipo de datos para evitar la correlación positiva entre la experiencia y el nivel de educación. Los resultados de la investigación señalan que el retorno a la educación está entre el 10% y 14% ubicándola como una inversión de alta rentabilidad privada, además que un alto porcentaje está fundamentado en variables no observables.

Correa, Viáfara y Zuluaga (2010) llevan a cabo un estudio sobre la distribución condicional del ingreso según la condición étnico-racial. En la metodología proponen como alternativa de la descomposición Blinder-Oaxaca, el método de Melly que, al tiempo que evalúa las brechas entre diferentes puntos de la distribución del ingreso, descompone el diferencial salarial en tres factores. Los autores concluyen que la endogeneidad entre la educación del individuo y la educación de los padres subestima los verdaderos retornos a la educación y que los afrocolombianos tienen menores retornos a la educación, lo que sugiere la presencia de discriminación.

Galvis (2012) caracteriza la informalidad en Colombia por áreas urbanas, determinando que los empleados informales tienen menor nivel de educación, menor nivel de ingreso y están vinculados a establecimientos de menor tamaño. El autor usa un modelo ANOVA simple por cuantiles, estimando la diferencia salarial entre trabajadores formales e informales para cada percentil de la distribución de salarios. Muestra que la brecha entre cada grupo de la distribución está más pronunciada para niveles bajos y disminuye en la medida que se acerca a la mediana, en adelante se aumenta la brecha hasta el percentil noventa. Entre sus conclusiones señala que mejorar el nivel educativo puede aportar a la reducción en la informalidad.

En resumen, el género y, en menor medida, la condición étnica son variables tradicionales en el análisis de retornos de la educación y en general en



las investigaciones de mercado laboral. La presente investigación propone un análisis de rendimiento de la educación enfocado en el sector económico (formal e informal) donde el sector al cual pertenezca la persona que se emplea será lo relevante, mediante la regresión cuantílica, una técnica semiparamétrica que permite caracterizar completamente la variable dependiente, en función del sector al cual pertenece quien hace parte del mercado laboral.

## 2. Metodología

El propósito del presente trabajo es calcular la tasa de retorno de la educación para los individuos ocupados, considerando el tipo de empleo (formal o informal), bajo el enfoque mixto de informalidad, para lo cual se fundamentará el análisis en el trabajo realizado por Zárate (2003) y la metodología planteada por Prada (2006), que especifica un modelo binario para la ecuación de participación en una primera etapa y una ecuación de salarios mediante regresión por cuantiles en la segunda.

### 2.1. DATOS Y VARIABLES

La información utilizada es proporcionada por la Encuesta de Empleo y Calidad de Vida (EECV) desarrollada por el Ministerio del Trabajo y Seguridad Social para el periodo noviembre de 2012 a enero de 2013. Los datos de la encuesta fueron tomados en la población urbana y rural de Cali. Por el objeto de estudio, se utilizó la Población en Edad de Trabajar (PET) para estimar la ecuación de participación laboral y el criterio para restringir los datos fue la edad de las personas encuestadas, en un rango de 12 a 60 años; el límite inferior se explica basado en la definición expresada por el DANE donde en las zonas urbanas, la población en edad para trabajar es censada desde los 12 años; y el límite superior, por la edad de jubilación, que según el Ministerio del Trabajo es de 60 años promedio por género. Por otra parte, para la estimación de la ecuación de salario se utiliza la Población Económicamente Activa (PEA) que corresponde a ocupados y desempleados.

En este artículo, la informalidad laboral -bajo el enfoque mixto- es una variable dicotómica que toma el valor de 1 para ocupados informales y 0 para ocupados formales, y proviene a su vez de otras dos variables

binarias. Una, de enfoque estructuralista, toma el valor del 1 para los informales e incluye a los empleados y patrones de empresas particulares de hasta 10 trabajadores, los empleados del servicio doméstico, los trabajadores por cuenta propia (sin incluir aquellos con educación superior) y aquellos familiares sin remuneración. La otra, de enfoque institucionalista, toma el valor de 1 para aquellos ocupados que incumplan con al menos uno de los aportes al sistema de seguridad social (afiliación a salud, pensión, ARL y caja de compensación). De esta manera, son informales quienes simultáneamente se ubican dentro de la definición de informalidad estructuralista e institucionalista.

### 2.2. REGRESIONES CUANTÍLICAS

El método de regresión cuantílica permite estimar los efectos marginales de las variables explicativas en distintos puntos de la distribución de probabilidad de la variable de estudio. Este tipo de regresiones fue planteado en 1978 por Koenker y Bassett logrando determinar que este análisis permite caracterizar un punto particular de la distribución condicional; su utilidad radica en lograr analizar distribuciones que no presentan la forma estándar, es decir para asimetría, colas gruesas y truncamientos.

La regresión cuantílica estima el cuantil  $\tau$  de la variable dependiente, en este caso el logaritmo del salario mensual, condicional en diversas variables explicativas que recogen características observables de los individuos. El cuantil  $\tau$  de una variable aleatoria condicional en  $x$  es el valor  $y_\tau$  tal que  $\Pr[y_i < y_\tau | x]$  para  $\tau (0,1)$ . Si se asume que  $y_\tau$  es lineal en  $x$ , es decir si  $y_\tau = x_i' \beta(\tau)$ , entonces el vector  $\beta$  es estimado como la solución a

$$\min_{\beta(\tau)} \sum_{i: y_i \geq x_i' \beta(\tau)} \tau |y_i - x_i' \beta(\tau)| + \sum_{i: y_i < x_i' \beta(\tau)} (1 - \tau) |y_i - x_i' \beta(\tau)| \tag{1}$$

En las regresiones cuantílicas del salario, los coeficientes estimados  $\beta(\tau)$ , se interpretan en forma análoga a los modelos de regresión paramétricos (para la media condicional) pero en referencia al retorno de las distintas características al cuantil  $\tau$  de la distribución de los salarios condicionales.

### 2.3. CORRECCIÓN POR SESGO DE SELECCIÓN

Estimar una ecuación de salario implica que se debe observar el salario, aunque para recibirlo es necesario tener un empleo. Así, el hecho de trabajar con empleados asalariados genera el problema de sesgo de selección muestral porque la muestra no es aleatoria. Para corregir el sesgo de selección se utiliza un proceso en dos etapas. En la primera, se estima una ecuación de selección o ecuación de participación en la cual la variable dependiente toma el valor de 1 si el individuo participa en el mercado laboral y de 0 en caso contrario.

$$\begin{aligned} \text{participa} = & \beta_1 + \beta_2 \text{anoeduca} + \beta_3 \text{edad} + \beta_4 \text{edad}^2 + \beta_5 \text{bsex} \\ & + \beta_6 \text{bjefe} \beta_6 \text{bjefe} + \beta_7 \text{braza} + \beta_8 \text{Ingnolab} + \varepsilon \end{aligned} \quad (2)$$

La ecuación (2) recoge la probabilidad de que una persona participe en el mercado laboral. Las regresoras incluidas en esta etapa son los años de educación aprobados, los ingresos no laborales, la edad y su cuadrado, el género. También las condiciones de jefe de hogar y etnia son incluidas porque múltiples investigaciones las definen como explicativas para el ingreso fruto de la actividad laboral, aunque en este artículo su análisis será menos relevante.

\* *bjefe* es una variable binaria que toma el valor de 1 si el individuo es jefe de hogar, 0 en otro caso. Se espera que una persona que es jefe de hogar tenga mayor probabilidad de participar en el mercado laboral porque sobre ella recae la responsabilidad del sustento económico familiar

\* *Ingnolab* son los ingresos no laborales que recibe un individuo, constituyen una proxy de la riqueza personal. Se espera que el coeficiente de esta variable sea negativo porque si el individuo tiene cierto nivel de ingresos asegurado -que aumenta su salario de reserva- preferirá el ocio y por tanto su probabilidad de participación en el mercado laboral es menor.

Posteriormente, se obtiene un estimador de la variable  $\lambda$ , la razón inversa de Mills, variable que es incluida como regresora en la estimación de la ecuación de salario de los individuos que hacen parte de los sectores formal e informal.

En la segunda etapa, se estima la ecuación de salario de los trabajadores formales e informales en los distintos cuantiles. Las variables en  $x$  son las características observables de los individuos y la variable  $\lambda$  es el término de corrección por sesgo de selección. Se realizan las estimaciones de las regresiones cuantílicas para los percentiles (10, 25, 50, 75 y 90) y se estiman dos modelos separados para formales e informales, considerando dos grupos de covariables. El primero incluye los años de educación, mientras segundo se incluye splines lineales para la educación.

$$\begin{aligned} Q_{\tau}(\log(y_i|x)) = & \log(y_0) + \beta_{1\tau} \text{Edu} + \beta_{2\tau} \text{Exper} + \beta_{3\tau} \text{Exper}^2 + \\ & \beta_{4\tau} \text{Sexo} + \beta_{5\tau} \text{Etnia} + \beta_{6\tau} \lambda + u_{\tau} \end{aligned} \quad (3)$$

$$\begin{aligned} Q_{\tau}(\log(y_i|x)) = & \log(y_0) + \beta_{1\tau} \text{Exper} + \beta_{2\tau} \text{Exper}^2 + \beta_{3\tau} \text{Sexo} + \\ & \beta_{4\tau} \text{Etnia} + \beta_{5\tau} \text{Sec} + \beta_{6\tau} \text{Tec} + \beta_{7\tau} \text{Uni} + \beta_{8\tau} \lambda + \varphi_{\tau} \end{aligned} \quad (4)$$

Donde  $\log(y_i|x)$  representa el logaritmo natural del salario mensual del individuo,  $y_0$  es el salario de una persona sin educación ni experiencia.

\* En (3) *Edu*, representa los años de educación aprobados. Como proxy del capital humano de la persona, se estima que su coeficiente tenga signo positivo lo cual indica que el individuo espera recibir un salario (mayor) acorde con su formación escolar.

\* En (4), primaria, secundaria, técnica y superior son los distintos niveles de educación alcanzados por los individuos. Siendo proxy del capital humano, se espera que su coeficiente tenga signo positivo y que el retorno sea mayor en los niveles educativos más altos.

\* *Exper* representa la experiencia potencial (edad menos años de educación menos 7 años). Como proxy de la formación en la práctica, se espera una relación positiva con el salario.

\* *Exper*<sup>2</sup> es el cuadrado de la experiencia potencial. Se espera que su coeficiente tenga signo negativo lo cual capturaría los rendimientos decrecientes (concavidad) en la experiencia.

\* *Sexo* es una variable binaria que toma el valor de 1 para las mujeres y 0 para los hombres. Se espera un coeficiente negativo en esta variable, debido

a la segmentación del mercado laboral en donde históricamente los hombres han percibido mayores salarios que las mujeres.

\* *Etnia* es una variable dicotómica que toma el valor de 1 si el individuo es blanco y 0 en otro caso (negro, mulato, indígena clasificaciones según la base de datos) y se espera que su coeficiente sea positivo.

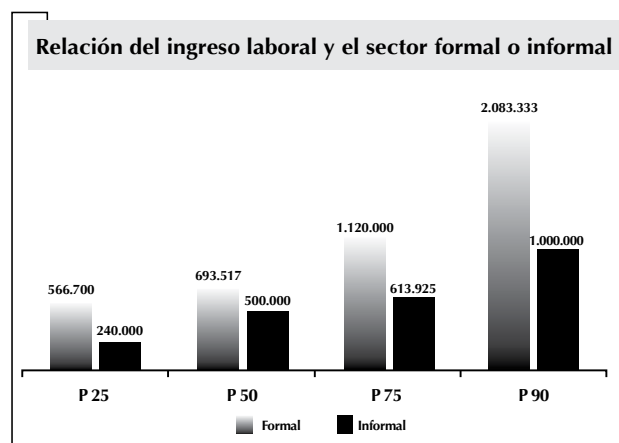
### 3. Resultados

#### 3.1. ANÁLISIS DESCRIPTIVO

En el Gráfico 1 se muestra que las personas del sector informal reciben en promedio 45,49% menos ingreso que los vinculados al sector formal. Por percentiles, una persona en el sector informal tiene un ingreso 3,67 veces mayor si está en el percentil 90 frente a otra que esté ubicada en el 25; en contraparte, si está en el sector formal la diferencia es de 4,16 veces.

En la Tabla 1 se indican las variables regresoras asociadas a la ecuación de selección (ingreso no laboral) y al capital humano (años de educación y experiencia potencial) según la condición de formal o informal. El ingreso no laboral es menor para informales en 43,64% promedio. Comparando el percentil 25 frente al 90, logran 36,4 veces más ingresos los formales, mientras que para los informales la diferencia es de 32 veces.

Gráfico 1. Relación del ingreso laboral y el sector formal o informal



Fuente: EECV Cali - Ministerio de Trabajo. Cálculos de los autores usando Stata 12. Elaboración propia.

La educación acumulada, en promedio, para una persona en el sector formal es de 9,21 años, superior en 9,75% al de una persona del sector informal. En ambas condiciones, en el percentil 25, los años acumulados de educación son 5, en el percentil 50 quienes están en el sector formal ya han logrado el grado de bachillerato y los de informal alcanzan el grado asimilable a media básica. En el percentil 75, hay igualdad de educación acumulada en 11 años aunque el ingreso laboral es 82,43% mayor para la condición formal que informal y en relación al ingreso no laboral es 40,00%. En el percentil 90, los que ostentan la condición de informal siguen con 11 años de educación, mientras que para

Tabla 1. Variables a ser utilizadas en la regresión cuantílica

Variables	Condición	Percentiles				
		Media	P25	P50	P75	P90
Ingreso no laboral	Formal	614.511	41.667	300.000	700.000	1.516.667
	Informal	346.320	25.000	150.000	500.000	800.000
	Total	573.020	41.667	254.896	650.000	1.400.000
Años de educación	Formal	9,21	5,00	11,00	11,00	16,00
	Informal	8,39	5,00	9,00	11,00	11,00
	Total	9,02	5,00	10,00	11,00	15,00
Experiencia potencial	Formal	25,10	5,00	20,00	42,00	59,00
	Informal	27,88	15,00	28,00	39,00	49,00
	Total	25,73	6,00	23,00	41,00	56,00

Fuente: EECV Cali - Ministerio de Trabajo. Cálculos de los autores usando Stata 12. Elaboración propia.

el sector formal son 16 años, es decir, una formación de pregrado que genera un ingreso laboral 2,08 veces superior en la condición formal.

En cuanto a la experiencia potencial, hay una tendencia creciente entre los percentiles 25 y 90 para los dos sectores. Aunque la informal para el primero es 15 años, este dato permite deducir que hay personas vinculadas a este sector con ingresos laborales inferiores al salario mínimo legal vigente, con cinco años de educación acumulada, es decir, son personas dedicadas a oficios mal remunerados y que han permanecido en ellos por bastante tiempo.

### 3.2. REGRESIÓN CUANTÍLICA

En la Tabla 2, ubicada como anexo, se presentan los resultados obtenidos a partir de las estimaciones que incorporan corrección por sesgo de selección muestral. En todas las estimaciones y para todos los puntos de la distribución, los retornos de la educación son significativos y positivos, tomando valores superiores en la cola derecha de la distribución. La comparación por sector indica que los retornos suelen ser mayores para los empleados formales que para los informales.

A su vez, el signo de los coeficientes asociados a la experiencia potencial y su cuadrado son los esperados en todas las estimaciones y para las distintas posiciones analizadas: el primero positivo y el segundo negativo, lo cual significa que el salario aumenta con la experiencia a tasas decrecientes. Además, los coeficientes de ambas variables son significativos, excepto para los formales en el percentil 25, explicado porque en este percentil la experiencia potencial promedio para un trabajador formal es de cinco años, cuatro veces menos que los del siguiente percentil, y para los informales en los percentiles 25 y 50.

El coeficiente asociado con la variable sexo también presenta el signo esperado, negativo, y es significativo en ambas estimaciones y para todos los puntos de la distribución, lo cual señala la existencia de discriminación por género.

Respecto a la variable etnia el coeficiente asociado también presenta el signo esperado (positivo) en todas las estimaciones. Para los empleados formales, su coeficiente es significativo en los tramos superiores de la distribución, en cambio, para los informales sólo es significativo en el percentil 75, lo que indica, que esta

variable no tiene mayor relevancia en la determinación del ingreso salarial, es decir, los ingresos se asocian más a otro tipo de características como el acervo de capital humano.

Por último, el coeficiente del término de corrección por autoselección es significativo para todos los puntos de la distribución del salario de los formales, lo cual indica que se ha corregido el sesgo existente para este grupo. Para los informales, dicho coeficiente no es estadísticamente significativo en ningún punto de la distribución salarial.

El primer resultado observable es el aumento de los retornos de la educación de los formales en los cuantiles superiores de la distribución salarial, es decir, que los retornos son mayores en los cuantiles más altos, coincidente con los resultados de Zárate (2003) y Prada (2006). Un segundo resultado relevante es que para los formales, el retorno de la experiencia es superior, y muy similar, en la parte intermedia de la distribución, mientras para los informales un año adicional de experiencia es mejor remunerado en el extremo derecho de la distribución, percentil 90, siendo máximo el retorno en este tramo. Por género, las mujeres siempre obtienen un salario menor que los hombres, tanto en el sector formal como en el informal y se hace menor en la medida que se avanza hacia un punto superior de la distribución.

Por otro lado, se estiman los retornos de la educación en cada parte de la distribución mediante regresión por cuantiles dividiendo la muestra en 4 cuantiles, incorporando la educación en forma de splines lineales y controlando por autoselección. En la Tabla 3 - como anexo- se presentan los resultados obtenidos al respecto.

Los signos de los coeficientes asociados a la experiencia potencial y su cuadrado son los esperados, positivo y negativo, respectivamente. Además, los coeficientes de ambas variables son significativos, excepto la experiencia para los formales en el percentil 25, mientras que para los informales sólo son significativos en los percentiles 75 y 90, aunque muy cercanos a cero, lo cual indica que para este grupo de trabajadores el efecto de la experiencia sobre el salario es muy pequeño, casi nulo.

El coeficiente de la variable sexo también muestra el signo esperado (negativo) y es significativo en ambas estimaciones y para casi todos los puntos



de la distribución, excepto en el percentil 25 de los ocupados formales, denotando que hay persistencia de discriminación por género.

En cuanto a las tasas de retorno de la educación, para los formales los resultados son los esperados, ya que los rendimientos son mayores en la cola alta de la distribución; sin embargo, para los informales, los resultados no lo son totalmente, pues en las posiciones intermedias se presentan retornos relativamente menores a los del percentil 25, para los niveles educativos secundaria y técnico.

En particular, bajo una comparación vertical, los retornos educativos de los individuos en la cola superior de la distribución salarial son mayores para los empleados informales, excepto para aquellos con educación superior, lo cual señala que su elección es racional. Así mismo, en los percentiles medios, es mayor el retorno de los informales para los niveles de secundaria y técnico, aunque con excepción del primero en el percentil 50.

#### 4. Conclusiones

Se estudió el retorno de la educación desde las regresiones cuantílicas, considerando el tipo de empleo (formal o informal), desde un enfoque mixto de la informalidad en Cali para un trimestre entre noviembre de 2012 y enero de 2013. Las regresiones cuantílicas con splines lineales muestran en general que los retornos de la educación, en cada nivel educativo, son diferentes y crecientes en la medida que se avanza en la distribución del salario. Esto comprueba que el rendimiento marginal de la inversión en capital humano no es constante ni entre niveles ni entre la distribución condicional del salario. Este rendimiento se ve afectado por variables no relacionadas con el capital humano, también evidencia que ser jefe de hogar, el sexo y la etnia influyen dependiendo del punto en que se encuentra el individuo en la distribución. Cali cuenta con una población plural, fruto de su ubicación geográfica que incide como receptora de migraciones, factores determinantes en el comportamiento del sector formal e informal. La ciudad tiene una cobertura en educación primaria y secundaria del 95%, según la Secretaría de Educación Departamental, aunque ésta para la educación de un nivel superior tiene una disminución drástica que implica menor nivel educativo en la población y menor

nivel de ingreso. Según la investigación, se determinó que en la distribución su cola izquierda indica sobre las personas en la ciudad con menor nivel educativo y menor nivel de ingreso.

En la comparación entre la condición de formales e informales, se encuentra un alto rendimiento de la educación en el nivel educativo superior, con una tasa mayor para los individuos informales que para los formales, en todas las posiciones de la distribución. Igualmente, en el nivel técnico, el retorno de los informales supera al de los formales. Por el contrario, la educación secundaria es una mejor inversión en el sector formal. Estos resultados denotan que la informalidad en Cali desde el enfoque mixto, tratado en este artículo, en las personas con un nivel educativo alto es una elección que influye en que este indicador en la ciudad no ceda; adicionado con personas de bajos niveles de educación que hallan en la informalidad, dada su experiencia una forma de sustento. Así la informalidad es causada en la ciudad por dos grupos, uno con bajo nivel de educación pero con retornos –que generan ingresos– dada su experiencia, que los mantienen en labores de baja calificación; el otro por personas con nivel de educación alto que deciden participar en el mercado laboral pero en el sector informal, porque obtienen un mejor rendimiento. No obstante, este resultado puede estar influido por los ingresos no laborales, que según la teoría, elevan el salario de reserva, que finalmente son una complementariedad de los ingresos totales.

Se encuentra además que la experiencia, en términos relativos, tiene baja incidencia sobre el nivel salarial tanto en formales como en informales, aunque es relevante en los percentiles bajos de la distribución factor preponderante para generar un ingreso. Así mismo, la condición étnica tiene ligero efecto en el ingreso laboral en los dos sectores, excepto en la parte superior de la distribución, caso en el que efectivamente los trabajadores blancos –clasificación según el instrumento de sondeo– del sector formal perciben un salario mayor que los indígenas, mulatos o afrodescendientes. Aunque en Cali, según datos del Plan de Desarrollo de la ciudad para el periodo 2012 – 2015, el 43,9% de la población de afrodescendientes e indígenas están distribuidos en quince comunas, las de mayor densidad poblacional, en estratos 1, 2 y 3, se evidencia mediante los resultados que en la cola derecha de la distribución con mayor nivel de educación e ingreso, hay presencia de discriminación por etnia.

Una posible extensión de esta investigación es cambiar los grupos objetivo, por ejemplo, se puede considerar los análisis para trabajadores asalariados y por cuenta propia – en ocasiones asemejados a informales-. Además, estudios posteriores pueden intentar la estimación de la ecuación de participación laboral mediante alguna de las metodologías semiparamétricas reseñadas por Prada (2006) al citar a Buchinsky (2001). Otro frente de trabajo sería profundizar en el retorno de la educación bajo la metodología y definición de esta investigación, pero evaluado en cada comuna de la ciudad. ≡

## CONFLICTO DE INTERESES

Los autores declaran no tener ningún conflicto de intereses.

## NOTAS

1. El artículo hace parte de la investigación: "Modelo para calcular el beneficio económico de las actividades productivas informales como base de la planeación, mediante un producto de software" que se encuentra en curso. Aprobado en la convocatoria SUI-2013 de la Universidad Cooperativa de Colombia.

## REFERENCIAS BIBLIOGRÁFICAS

1. ALCALDÍA DE CALI. Plan de desarrollo de Santiago de Cali 2012 -2015. Cali: Alcaldía de Cali, 2012.
2. ARVANITIS, Spyros, y LOUKIS, Euripidis N. Information and communication technologies, human capital, workplace organization and labour productivity: A comparative study based on firm-level data for Greece and Switzerland. *In: Information Economics and Policy*. 2009, 21, nº 1, 43 - 61.
3. BUCHINSKY, Moshe. Quantile regression with sample selection: Estimating women's return to education in the U.S. *In: Empirical Economics*. 2001, vol. 26, issue 1, p. 87-113
4. CÁRDENAS, Mauricio y MEJÍA, Carolina. Informalidad en Colombia: Nueva Evidencia. Working Papers Series - Documentos de trabajo (Fedesarrollo), 2007, nº 35.
5. CONTRERAS, Dante; MELO, Emerson y OJEDA, Susana. ¿Estimando el retorno a la educación o a los no observables?: Evidencia de datos panel. *En: Estudios de Economía*. 2005, 32, nº 2, 187 - 199.
6. CORREA, Juan Bayron; VIÁFARA Carlos Augusto y ZULUAGA Víctor Hugo. Desigualdad étnico-racial en la distribución de ingreso en Colombia: Un análisis a partir de regresión cuantílica. *En: Sociedad y Economía*. 2010, nº 19, p. 153 - 178.
7. FREIJE, Samuel. El empleo informal en América Latina y el Caribe: causas, consecuencias y recomendaciones de política. Washington D. C.: Banco Interamericano de Desarrollo, 2002.
8. GALLART, María A. Los desafíos de la integración social de los jóvenes pobres: la respuesta de los programas de formación en América Latina. *En: Los jóvenes y el trabajo: la educación frente a la exclusión social*, de Enrique Pieek, 59 - 94. México: Unicef, 2001.
9. GALVIS, Luis Armando. Informalidad laboral en las áreas urbanas de Colombia. *En: Documentos de trabajo sobre Economía Regional*, 2012, nº 164.
10. GASPARINI, Leonardo y TORNAROLLI Leopoldo. Informalidad laboral en América Latina y el Caribe: patrones y tendencias a partir de microdatos de encuestas de hogares. *En: Desarrollo y Sociedad*. 2009, 14, nº 63. 13 - 80.
11. GOBERNACION DEL VALLE DEL CAUCA. Plan sectorial de educación 2012 - 2015. Cali: Imprenta Departamental, 2012.
12. JIMÉNEZ, Diana Marcela. La informalidad laboral en América Latina: ¿Explicación estructuralista o institucionalista? *En: Cuadernos de Economía*. 2012, 31, nº 58.113 - 143.
13. KOENKER, Roger y BASSETT Gilbert. Regression quantiles. *In: Econometrica: journal of the Econometric Society*. 1978, 46, nº 1, 33-50.
14. MÉNDEZ, Jairo Augusto. Empleo informal y evasión fiscal en Colombia. Bogotá: DNP, 2002.
15. MINCER, Jacob. Investment in Human Capital and Personal Income Distribution. *In: The Journal of Political Economy*. 1958, 66, nº 4, 201 - 302.
16. OIT. Informe VI: El trabajo descente y la economía informal (Sexto punto del orden del día). Ginebra: Organización Internacional del Trabajo, 2002.
17. PORTES, Alejandro y HALLER William. La Economía Informal. CEPAL - SERIE Políticas Sociales (Cepal). 2004, nº 100.
18. PORTES, Alejandro. En torno a la informalidad: ensayos sobre teoría y medición de la economía no regulada. México D.F.: FLACSO, 1995.
19. PRADA, Carlos Felipe. ¿Es rentable la decisión de estudiar en Colombia? *En: Ensayos sobre política económica*. 2006, nº 51, p. 267 - 329.
20. RECIO, Albert. Trabajo, personas, mercados: Manual de economía laboral. Barcelona: Icara, 1997.
21. SCHNEIDER, Friedrich. Shadow Economies of 145 Countries all over the World: What do we really know? Austria: Department of Economics, Johannes Kepler University of Linz, 2006.
22. URIBE, José Ignacio; ORTIZ Carlos Humberto y GARCÍA Gustavo Adolfo. La segmentación del mercado laboral colombiano en la década de los noventa. Bogotá: DNP, 2006.
23. URIBE, José Ignacio y ORTIZ Carlos Humberto. Informalidad laboral en Colombia, 1988-2000: Evolución, teorías y modelos. Cali: Universidad del Valle, 2006.
24. ZÁRATE, Héctor. Cambios en la estructura salarial: una historia desde la regresión cuantílica. *En: Borradores de Economía*. 2003 nº 245. p. 1-45

Anexos

Tabla 2. Resultados de la estimación de regresiones cuantílicas con corrección de sesgo

	Variable	P25	P50	P75	P90
Formales	Añosedu	0,070***	0,118***	0,142***	0,152***
	Exper	0,028	0,050***	0,051***	0,041**
	Exper2	-0,001	-0,001***	-0,001**	-0,001*
	Sexo	-0,284**	-0,417***	-0,432***	-0,363***
	Etnia	0,049	0,080	0,112**	0,222***
	Inv mills	0,464*	0,689***	0,887***	0,821***
	Constante	12,113***	11,624***	11,577***	11,926***
	Informales	Añosedu	0,062**	0,072***	0,090***
Exper		0,014	0,031	0,075***	0,102***
Exper2		0,000	-0,001	-0,001**	-0,002***
Sexo		-0,471**	-0,497***	-0,715***	-0,800***
Etnia		-0,050	-0,111	-0,135*	0,038
Inv mills		-0,413	-0,055	0,402	0,413
Constante		12,320***	12,478***	12,160***	12,098***

Significativo \* al 10%, \*\* al 5% , \*\*\* al 1%.

Fuente: EECV Cali - Ministerio de Trabajo. Cálculos de los autores usando Stata 12.

Tabla 3. Resultados de la estimación del salario con splines lineales usando regresión cuantílica corregida por autoselección

	Variable	P25	P50	P75	P90
Formales	N_secundaria	0,073	0,194*	0,293***	0,364*
	N_técnico	0,027	0,349**	0,608***	0,771***
	N_superior	0,595***	1,087***	1,364***	1,434***
	Exper	0,011	0,044***	0,052***	0,087***
	Exper2	0,000	-0,001***	-0,001***	-0,002***
	Sexo	-0,178	-0,330***	-0,462***	-0,373**
	Inv mills	0,220	0,489***	0,693***	1,095***
	Constante	13,075***	12,838***	12,941***	12,785***
Informales	N_secundaria	0,392***	0,272***	0,288**	0,406**
	N_técnico	1,004**	0,822***	0,790	0,891*
	N_superior	0,516	0,609*	1,184***	1,168***
	Exper	0,032	0,004	0,061**	0,088**
	Exper2	0,000	0,000	-0,001**	-0,002**
	Sexo	-0,491***	-0,469***	-0,628***	-0,706***
	Inv mills	-0,410	-0,249	0,215	0,136
	Constante	12,459***	13,284***	13,057***	13,280***

Significativo \* al 10%, \*\* al 5% , \*\*\* al 1%.

Fuente: EECV Cali - Ministerio de Trabajo. Cálculos de los autores usando Stata 12.