

Una aproximación econométrica a la tasa de retorno social de la educación

**Carlos E. Castellar P.
José Ignacio Uribe G.¹**

Resumen

Cuando se mide la tasa de retorno de la educación, agrupando según la localización en viviendas contiguas, la rentabilidad social resulta mucho mayor que la individual. Usando una modelización que reconoce simultáneamente al individuo por un lado, y los determinantes del nivel agregado por otro, se encuentra que la diferencia obedece a la presencia de externalidades. La prima diferencial en los ingresos que genera la pertenencia a un segmento, depende a su vez de la concentración de capital humano en el mismo y explica la mayor rentabilidad social. Este resultado contradice anteriores investigaciones que concluyeron que la tasa social de retorno de la educación era inferior a la respectiva tasa privada. Una implicación inmediata de este hallazgo es la recomendación de una urgente revisión de las políticas económicas que pretenden trasladar recursos de la universidad pública a la escuela primaria.

Abstract

The measurement of rates of return to education for groups of individuals from the National Household Survey of Colombia (ENH) yields higher estimates of social returns than individual returns. By modelling simultaneously the individual within his/her group and the determinants at the aggregate level, it is found that the gap between these returns is explained by externalities. The income premium explained by being part of a group depends on the concentration of human capital in the same group and it explains the higher social return. This finding is in contradiction to previous research that has found the social rate of return to less than the private rate. An immediate implication of our results is a change in economic policies aimed at translating resources from professional education in public institutions to primary education.

Palabras claves: educación, economía de la educación, tasas de retorno de la educación, tasa de retorno privada, tasa de retorno social, externalidades, externalidades pecuniarias, capital humano.

¹ Profesores del Departamento de Economía de la Universidad del Valle. Grupo de investigación:

1. Introducción

El objetivo de este artículo es proponer una forma alternativa de medir la rentabilidad social de la educación. El argumento central que se pone a consideración del lector es que la presencia de las externalidades en la educación obliga al análisis simultáneo de dos dimensiones analíticas, una la del individuo y otra, la de los colectivos. Cuando se intenta medir la rentabilidad de la educación, teniendo en cuenta lo anterior, se encuentra que la tasa de retorno en el nivel agregado es mucho más alta que a nivel individual. En este artículo se sugiere una adecuación de la metodología de Datos de Panel, para dar contenido empírico a esta nueva forma de abordar el problema.

El análisis económico lleva en sí la opción de agregación, en una gama tan variada que son infinitas las posibilidades de ir desde lo “micro”, pensado como el individuo, hasta lo “macro”, visto como el colectivo. Cuando dos niveles de agregación se incorporan teóricamente al análisis, es posible mantener una línea de argumentación en la cual coexistan estos dos niveles, configurando las dos dimensiones analíticas: la del individuo dentro del colectivo (intragrupo) y la de los colectivos (intergrupo). El valor agregado que reporta esta concepción teórica es la coexistencia permanente de los dos niveles de agregación y la posibilidad de moverse en el interior de cada uno de ellos, teniendo en cuenta el efecto que uno genera sobre el otro. Por lo tanto, la formulación de políticas públicas en sectores que generan externalidades debe tener en cuenta los dos niveles y hacerlo con uno solo induciría errores en la asignación de recursos.

En el caso de la educación siempre se ha mencionado la existencia de externalidades que hacen que la cuantificación de los beneficios sociales de la misma estén subestimados y por ende su rentabilidad social. No obstante puede pensarse que las externalidades también pueden operar en la mecánica misma del mercado de trabajo. Siguiendo los desarrollos recientes de la microeconomía laboral, puede aseverarse que entre los argumentos de la función de salario de reserva del individuo, aparece un argumento de carácter colectivo, propio del entorno social o económico al cual pertenece la persona. Igualmente en el salario de mercado que se ofrece al individuo aparecería un argumento de iguales características. Consecuente con el papel de la educación se postula que dicho argumento depende del promedio de educación del mencionado entorno. Si es así, este argumento se hace externo al individuo pues él no lo decide, pero en el nivel de agrupación se hace interno, configurando los dos niveles de agregación antes mencionados. Entonces la externalidad generada por la educación es exógena para el individuo pero endógena para el colectivo al cual pertenece.

Economía laboral y sociología del trabajo. Los autores agradecen los valiosos comentarios que los profesores Harvy Vivas y Omar Bello hicieron a una versión preliminar. Este artículo no hubiese sido posible sin la dedicación y consagración de las asistentes de investigación Angie Hernández, Paola Marcela Roldán, Victoria Eugenia Soto y Mariana Uribe. Cualquier error por acción u omisión, es de absoluta responsabilidad de los autores.

Las implicaciones económicas y sociales de ciertas decisiones de política, como el traslado de recursos entre niveles de educación pública hace necesaria, entre otros cosas, una cuantificación específica de las tasas de retorno privado y social de la educación en Colombia. Igualmente debe discutirse la metodología con la cual se hace usualmente este tipo de ejercicio. En la mayoría de casos se tiene como punto de partida la estimación de las llamadas funciones generatrices de ingreso. En el estudio de las funciones de ingreso se han utilizado predominantemente dos tipos de datos. En primer lugar, datos de corte transversal provenientes de encuestas de hogares. En el uso de este tipo de información se ha señalado la presencia de variables no observables, como habilidades y destrezas innatas, cuya no observabilidad dificulta las estimaciones. En segundo término, en tiempos recientes, se ha difundido el uso de Datos de Panel, consistentes con la aplicación sucesiva de la encuesta a un conjunto de individuos. De esta forma se puede aislar el efecto de las variables no observables. La esencia metodológica de esta opción, consiste en la incorporación de dos dimensiones al análisis, en este caso el tiempo y el espacio. (Hsiao, 1986; Baltagi, 1995). La presente investigación propone una aplicación no convencional de la teoría y los métodos de Datos de Panel para las dos dimensiones que configuran el individuo y el colectivo al cual pertenece.

En la siguiente sección se esbozan los elementos del marco teórico que conceptualizan la investigación. Acto seguido se modelan las tasas de retorno, tanto en la perspectiva tradicional que resume el estado del arte como en la sugerida por el nuevo enfoque aquí propuesto. La cuarta sección contiene la adaptación de la metodología de Datos de Panel a la modelización planteada. La siguiente se ocupa de la presentación e interpretación de la evidencia empírica obtenida para el área metropolitana de Cali-Yumbo, en los diciembre de 1988, 1990, 1992, 1994, 1996 y 1998. Una última sección contiene las principales conclusiones, las cuales permiten afirmar que la hipótesis central de un mayor retorno social de la educación frente al privado no se ve rechazada por los datos y de ahí la urgencia de solicitar al Estado la reformulación de la política educativa, en especial con el nivel de educación superior. La bibliografía consultada completa el cuerpo del artículo.

2. Tasas de retorno, externalidades y papel del Estado.

En esta sección se sintetizan los principales elementos teóricos y de contextualización que sustentan la propuesta analítica. Desde el punto de vista de la política económica el objeto de estudio se enmarca en el papel del Estado cuando existen externalidades. Por la íntima conexión entre teoría y método que el tratamiento propuesto contiene, también se dan unas pautas metodológicas.

Uno de los pilares sobre los cuales descansa la nueva política gubernamental de financiación de la educación pública superior en Colombia, es la presunción de que la tasa de retorno social de la educación universitaria es menor que la de la

primaria y de ahí que sea más eficiente trasladar fondos públicos de la superior hacia la básica. El sustento empírico de esta presunción (Clavijo, 1998) se encuentra en los trabajos de Psacharopoulos (1985, 1993) quien establece unos patrones genéricos para los países, de acuerdo a su grado de desarrollo económico. Este último autor afirma que por regla general las tasas de rentabilidad social de la educación son inferiores a las privadas. Es posible que este resultado provenga de la forma como se calcula la rentabilidad social, pues a los costos privados (ingresos no percibidos durante el estudio más los costos directos de la educación) se le adicionan los costos sociales (gasto público por estudiante) pero a los beneficios privados (ingresos durante la vida activa) no se le agregan los beneficios sociales, ni las externalidades positivas que genera, pues se argumenta su difícil cuantificación (Johnes, 1993). Al respecto el mismo Psacharopoulos señala:

“Debido al subsidio público a la educación en todas partes del mundo, típicamente las tasas privadas se encuentran unos cuantos puntos más altas que las tasas sociales de retorno. Por definición, el costo en la estimación de una tasa de retorno privada se refiere solamente a lo que un individuo desembolsa de su bolsillo, mientras que el costo en la estimación de una tasa de retorno social está relacionada al costo total de los recursos de alguien que acude a la escuela”. (Psacharopoulos, 1991, pag 11).

Dos consecuencias que, aunque lógicas y son difíciles de sustentar, del anterior resultado serían: la primera, que hay una presencia de externalidades negativas en la educación; o, la segunda, que la suma de los costos sociales supera los beneficios y externalidades positivas de la misma.

Es oportuno recordar la acepción clásica de externalidad (Marshall, 1920; Ohlin, 1933) formalizada por (Baumol Y Oates, 1975), según la cual ésta se presenta cuando en los argumentos de una función individual se incluyen variables no sujetas a la decisión del individuo y por las cuales no paga (recibe) por los beneficios (costos) que le ocasionen. Para citar un solo pero bien conocido ejemplo, se ha dicho que las externalidades son exógenas para las empresas pero internas para las industrias. Es claro que desde el punto de vista individual, si por pertenecer a un determinado colectivo se percibe una prima diferencial en el ingreso, se estaría en presencia de una externalidad monetaria. Por externalidad monetaria se entiende el efecto de las acciones privadas de los individuos que imponen costes o beneficios monetarios a terceros.

Ahora bien, desde la dimensión del colectivo si la externalidad en cuestión depende del nivel de educación del agregado se estaría en presencia de una externalidad endógena. Es decir, que al pasar del nivel privado al social, hay un proceso de endogenización de la externalidad, la cual pasa de ser exógena al individuo que pertenece a un colectivo, para convertirse en endógena para dicho colectivo. Aunque decir que la externalidad es endógena al colectivo parezca

redundante, el enunciado busca enfatizar la coexistencia de lo exógeno y lo endógeno, dependiendo del nivel de agregación en el cual se sitúe el investigador. Este punto no es banal cuando se quiere diseñar un marco analítico que de cuenta *simultáneamente* de las dos dimensiones.

Sabido es que en presencia de externalidades hay fallos de mercado que obligan a la intervención del Estado. Como se demostrará más adelante, en presencia de externalidades endógenas la tasa interna de retorno de la educación para un individuo es inferior a la respectiva tasa del colectivo. En este orden de ideas el asunto fundamental a resolver para los formuladores de la política económica es si se debe cobrar al individuo por la externalidad que recibe, o mejor, inducir la acción de las externalidades en el nivel colectivo. La pregunta relevante sería: ¿cuál es el nivel de agregación en el cuál debe moverse la política de educación pública: cobrar en el nivel individual o mejorar las condiciones de aparición de externalidades que actúen en los colectivos?

Aún cuando se trabaje en una dimensión espacial, en el problema del retorno de la inversión en educación, y en general en los problemas asociados al mercado de trabajo, coexisten las dos dimensiones: intragrupos e intergrupos. El individuo pertenece a un grupo, bien sea por razones de funcionamiento del mercado (mercados locales, ramas de actividad) o por la localización espacial de los individuos que queda reflejada en la forma en la cual se toma la información (los segmentos de la Encuesta Nacional de Hogares).

La coexistencia de las mencionadas extenalidades en los dos niveles de análisis permite articularlas tanto al marco teórico como al diseño metodológico. Se trata de incorporar en la función de decisión del individuo j ubicado en el grupo i , un elemento a_i no decidido por él, común al agregado al cual pertenece, es decir una externalidad en sentido estricto. Un desarrollo detallado de la conceptualización teórica que fue aplicada al contexto de la producción agropecuaria, la renta de la tierra y los ingresos por venta de fuerza de trabajo se encuentra en Castellar, (1998). En el Anexo 6 de la anterior referencia se ilustra la metodología para los ingresos de trabajo en el área metropolitana de Cali, para septiembre de 1991.

Formalmente, si Y_{ij} es la variable a explicar para el individuo j perteneciente al agregado i , X_{ij} es el vector de variables a explicar y b el vector de parámetros intrasegmento o intragrupos, entonces se modelaría la dimensión intragrupo como

$$Y_{ij} = a_i + b'X_{ij}$$

Si se define un vector de variables de dimensión intergrupos como Z_i , y g como el vector de parámetros de la dimensión intergrupos entonces a la dimensión intergrupos correspondería

$$a_i = g_0 + g_1'Z_i$$

Ahora bien si dentro del vector Z están los promedios del vector X entonces la externalidad se hace endógena en el sentido econométrico. En otras palabras, a_i es variable explicatoria para el individuo y es la variable explicada para el agregado. Si además depende de las X_{ij} (vía promedio), viola el supuesto de exogeneidad de regresores.

La justificación teórica del elemento a_i puede provenir, en el caso de que la agregación sea la ocupación del individuo, de una externalidad que se origine en la función de producción característica de la ocupación, la cual depende positivamente del nivel educativo y que esta externalidad se traslade al mercado de trabajo. Si la agrupación es el segmento, se estaría en presencia de externalidades de origen territorial, las cuales se explicarían por la concentración física de capital humano (Rauch, 1993).

Siguiendo a (Lucas, 1988) en la parte final de su trabajo seminal, podría afirmarse que las interacciones grupales que están en la esencia del papel de la educación como economía externa positiva, proveen los fundamentos microeconómicos del desarrollo (para Colombia un trabajo prometedor e interesante es el de González, Guzmán Y Pachon, 1998). En cualquiera de las dos posibilidades teóricas, las externalidades son endógenas tanto en el sentido conceptual como se ha venido sugiriendo, tanto como en el sentido econométrico propuesto por Hausman (1978), pues estarían correlacionadas con los regresores del modelo. También puede entenderse esta endogeneidad como el hecho de que la probabilidad de que un individuo sea observado en un segmento u ocupación determinada, no es independiente de su nivel educativo.

Cuando se postula que la prima diferencial por pertenecer a un agregado depende del nivel de educación del colectivo o cuando se calcula la rentabilidad de la educación en el nivel de la agrupación, se está cambiando de nivel de agregación y por ende de dimensión de análisis. Si bien se parte del más bajo nivel de desagregación, lo microeconómico, al contemplar la acción de efectos por pertenecer a un segmento o a una ocupación, se está incorporando un nivel de agregación más alto. Puesto que se trata en principio de variables no observables en el nivel individual y dada la coexistencia de dos dimensiones, la teoría y los métodos que subyacen en los modelos de Datos de Panel, están al orden del día.

Aunque la gran mayoría de las aplicaciones de esta novedosa metodología se han hecho en contexto de tiempo y espacio, su diseño interno no la limita a dicho ámbito. De hecho puede afirmarse que siempre que en un problema económico coexistan dos dimensiones, no necesariamente tiempo y espacio, la metodología de Datos de Panel es susceptible de aportar herramientas útiles para la comprensión de los fenómenos a estudiar. Si bien es cierto que la comunidad científica ha reconocido la posibilidad de usar los Datos de Panel para investigaciones basadas en información de corte transversal (Moulton, 1986, 1987), el énfasis se ha hecho en los aspectos técnicos de la estimación y no en la crucial contrapartida teórica que la apreciación de las dos dimensiones lleva consigo.

3. La modelización de las tasas de retorno

En el análisis de la tasa de retorno de la educación, la estimación de las llamadas ecuaciones mincerianas es una referencia obligada. Una excelente síntesis del estado del arte en este tema se encuentra en Berndt (1991, capítulo 5). En esencia se trata de formular un modelo de regresión lineal múltiple cuya variable dependiente es el logaritmo del ingreso de un individuo (LY), en función de sus años de educación (ED), de la experiencia (EX), del cuadrado de la misma, de las horas trabajadas (LH si se toma el logaritmo) y de la convencional perturbación aleatoria (U). Acerca de U se postula las hipótesis convencionales de partida.

Formalmente, para el individuo j se tiene que:

$$LY_j = b_1 + b_2 ED_j + b_3 EX_j + b_4 EX_j^2 + b_5 LH_j + U_j \quad (1)$$

$$U_j \sim N(0, \sigma_u^2)$$

En este modelo no se tiene en cuenta la operación de externalidades y equivale al modelo total de la literatura de Datos de Panel, es decir, en él se postula la hipótesis de inexistencia de economías externas. En la ecuación (1), b_2 , bajo ciertos supuestos simplificadores, suele interpretarse como la tasa de rendimiento privado de un año de educación. Es usual utilizar como variable dependiente el ingreso por hora. Esta opción es un modelo restringido por la hipótesis de que la elasticidad horas trabajadas del ingreso unitaria. Si se cumple la hipótesis de que $b_5 = 1$, es sensato sustituir la variable dependiente por el logaritmo del ingreso por hora, ganándose eficiencia en la estimación. Por el contrario, si no se cumple la anterior hipótesis, imponer dicha restricción puede producir estimadores sesgados. En consecuencia es necesario verificar esta hipótesis en lugar de aceptarla “a priori”. En caso de ser rechazada habría que evaluar si la ganancia en precisión compensa la eventual pérdida en sesgo, si se pretende continuar con la restricción.

Para diferenciar los rendimientos relativos de los niveles de educación se usan funciones quebradas, denominadas Spline en el argot de los econométricos, (Johnston J, 1987) definiendo dos variables falsas:

SEC = 1 si se han estudiado más de cinco años y cero en otro caso.

UNI = 1 si se ha estudiado más de once años y cero en otro caso.

El modelo adopta la forma:

$$LY_j = b_1 + b_2 ED_j + 1_1 SEC_j (ED_j - 5) + 1_2 UNI_j (ED_j - 11) + b_3 EX_j + b_4 EX_j^2 + b_5 LH_j + U_j \quad (2)$$

$$U_j \sim N(0, \sigma_u^2)$$

En este contexto b_2 es la tasa de rendimiento de un año de primaria, en tanto que el de uno de secundaria es $b_2 + l_1$ y el de uno de universitaria viene a ser $b_2 + l_1 + l_2$.

Obsérvese que la variable falsa asociada a la secundaria no se define de la manera tradicional sino que incluye el uno, no sólo para los de secundaria sino también para los que tienen estudios universitarios. La intuición del procedimiento se debe a que para llegar al nivel universitario se debe haber cursado los dos niveles anteriores y un análisis marginal, debe indagar por los aumentos de rentabilidad de un nivel respecto al anterior. La explicación técnica proviene del hecho que el omitir este detalle no permite que la función sea continua. En Colombia diversos trabajos se han ocupado de los diferenciales en la tasa de retorno según niveles de escolaridad Tenjo (1993), Ribero Y Meza (1997) y OIT (1999). En el trabajo de Tenjo se hace una exposición clara de la metodología Spline. No obstante, cuando define las variables falsas (página 87) a la que correspondiente a secundaria la acota entre 5 y 11, y no entre 5 y más años como recomienda la metodología. Un error más delicado aparece en el trabajo de la OIT, pues además de no definir las variables falsas de manera correcta, no se especifica el tratamiento adecuado que debe tener la variable educación. Más adelante al comentar los resultados se volverá sobre el mismo tema.

El mismo error en la definición de las variables falsas para el cálculo de la rentabilidades diferenciales por niveles de educación se comete en OIT (1999, p. 101).

Hasta aquí lo expuesto se circunscribe a ejercicios ya convencionales en el estudio de ecuaciones de ingreso. En adelante, se propone una forma de modelar que permite la incorporación de externalidades en el individuo y la explicación de las mismas en el colectivo.

Supóngase ahora que el individuo j pertenece al segmento i de la población y que en dicho segmento actúa una externalidad monetaria sobre el ingreso de los individuos del segmento. Se entiende por segmento “un área compacta de aproximadamente diez viviendas contiguas, las cuales se investigan en su totalidad” (DANE 1991). Dicha economía externa puede provenir de una “atmósfera” que incida sobre la productividad de los miembros del segmento o puede entenderse como la operación de efectos discriminadores que actúan desde la demanda de trabajo. En esencia se trata de variables no observables en el nivel individual pues su espacio de acción es el agregado al cual pertenece el individuo. En términos teóricos se puede afirmar que tanto los argumentos de la función de utilidad individual y , por ende, de su función de salario de reserva, como en la función de salario de mercado, hay factores que configuran un elemento a_j , externo al individuo pero común a todos los integrantes del segmento. En consecuencia, la función individual minceriana se convierte en:

$$LY_{ij} = a_i + t_2 ED_{ij} + t_3 EX_{ij} + t_4 EX^2_{ij} + t_5 LH_{ij} + U_{ij} \quad (3)$$

$$U_{ij} \sim N(0, s^2_u)$$

a_i representa el efecto neto de todas las externalidades que afectan a todos los individuos j que pertenecen a un segmento i de población. Esta ecuación se puede denominar Modelo de Externalidades Fijas (MEF) y en él, t_2 significa una tasa de rendimiento intrasegmento. En pocas palabras el ingreso del individuo se descompone en una prima diferencial por pertenecer a un segmento y , en otra parte, inducida por las características individuales.

Cuando no es de interés el cómputo de las externalidades ó efectos fijos y éstos son exógenos en el sentido de Hausman, es posible trabajar las externalidades como aleatorias ($a_i = m + v_i$), en cuyo caso se define el Modelo de Externalidades Aleatorias (MEA) como:

$$LY_{ij} = m + t_2 ED_{ij} + t_3 EX_{ij} + t_4 EX^2_{ij} + t_5 LH_{ij} + v_i + U_{ij} \quad (4)$$

$$v_i \sim N(0, s^2_v) \quad U_{ij} \sim N(0, s^2_u)$$

Si a_i , de la ecuación (3) es independiente de las variables explicatorias (exógena), se puede hablar de externalidades aleatorias, en la ecuación (4).

El presuponer que las externalidades son endógenas implica su dependencia de los regresores del modelo. En principio se puede postular una relación de dependencia como la sugerida por Mundlak (1978), en la cual el efecto fijo a_i depende de las medias de todos los regresores; por facilidad en la notación se agrega una M a las variables definidas.

$$a_i = g_1 + g_2 EDM_i + g_3 EXM_i + g_4 EX^2M + g_5 LHM_i + h_i \quad (5)$$

$$h_i \sim N(0, s^2_h)$$

Una consecuencia de la endogeneidad de las externalidades así modeladas es que inducen un cambio paramétrico, en la dimensión agregada. En efecto al sustituir (5) en (3) se obtiene:

$$LY_{ij} = g_1 + t_2 ED_{ij} + t_3 EX_{ij} + t_4 EX^2_{ij} + t_5 LH_{ij} + g_2 EDM_i + g_3 EXM_i + g_4 EX^2M + g_5 LHM_i + U_{ij} + h_i \quad (6)$$

$$U_{ij} \sim N(0, s^2_u) \quad h_i \sim N(0, s^2_h)$$

La anterior ecuación define un Modelo de Externalidades Endógenas (MEE); tomando promedios en i , es decir trasladándose a la dimensión intersegmento, se obtiene:

$$LYM_i = \sigma_1 + (\tau_2 + \sigma_2) EDM_i + (\tau_3 + \sigma_3) EXM_i + (\tau_4 + \sigma_4) EX^2M + (\tau_5 + \sigma_5) LHM_i + UM_i + h_i \quad (7)$$

En el mundo del intersegmento, en presencia de externalidades endógenas, la tasa de rentabilidad de la educación es $\tau_2 + \sigma_2$, es decir la suma de la tasa intrasegmento y el coeficiente que mide el impacto de la educación media en las externalidades fijas. Este cambio paramétrico es sugerido en una aplicación convencional de Datos de Panel para tiempo y espacio por Raymond (1995) y extendida al problema de externalidades territoriales por Castellar (1998).

Obsérvese que si el promedio se hace a partir de la ecuación (1), el mundo intersegmentos en ausencia de externalidades, indicaría:

$$LYM_i = b_1 + b_2 EDM_i + b_3 EXM_i + b_4 EX^2M_i + b_5 LHM_i + UM_i \quad (8)$$

En este mundo, no hay cambios paramétricos y por ende los coeficientes individuales coinciden con los del agregado.

La intuición de estos resultados es inmediata: en ausencia de externalidades endógenas las tasas de retorno de los individuos y de los colectivos a los que pertenecen, coinciden, en tanto que en presencia de dichas externalidades, la tasa individual es inferior a la del colectivo. En estricto rigor, no se trata de una tasa de rendimiento pues la variable dependiente es la media de los logaritmos y no el logaritmo del ingreso medio. No obstante, el logaritmo medio es igual al logaritmo de la media geométrica y si los agregados son relativamente homogéneos en su interior, ambas medidas estarán muy cerca. Adicionalmente la media geométrica es mucho menos sensible a la presencia de valores extremos y puede tener ventajas como medida de tendencia central frente a la media aritmética.

La pregunta relevante para pasar del cálculo de la rentabilidad privada a la rentabilidad social de la educación, es indagar por el impacto en los ingresos medios del segmento i de la población, cuando se incrementa un año promedio de educación en el segmento. Una opción es estimar el equivalente de la ecuación minceriana a nivel de segmentos, tomando medias aritméticas antes de la transformación logarítmica (LMY es logaritmo del ingreso medio y LMH, es el de las horas trabajadas):

$$LMY_i = d_1 + d_2 EDM_i + d_3 EXM_i + d_4 EXM_i^2 + d_5 LMH_i + UM_j \quad (9)$$

En el modelo (9) d_2 sería una alternativa clara para estimar la tasa de rendimiento social, pues está asociada al impacto en el ingreso agregado de una inversión colectiva en educación. En sentido amplio debieran utilizarse varias aproximaciones para llegar a la tasa de rendimiento intersegmento. El punto a resolver sería cómo obtener una adecuada medida de la tendencia central: la mediana, la media geométrica. Así las cosas el usar la media de los logaritmos, es decir, el logaritmo

de la media geométrica del modelo (8) puede tener ventajas como media de tendencia central frente al logaritmo de la media aritmética del modelo (9) en especial si existen valores extremos. Este debe ser un tema de futuras investigaciones.

La conclusión de todo el desarrollo anterior, indica que si existen externalidades que afecten el ingreso y que a su vez dependen de la educación, las dos tasas deben diferir. En ausencia de externalidades no tendría que haber diferencias.

4. La aproximación econométrica

La econometría intenta dar contenido empírico a las relaciones que la teoría económica sugiere. Para ello estima los α_s y las β , valida los supuestos acerca de la perturbación aleatoria y realiza inferencia estadística condicionada para verificar si la observación de una realidad económica particular contradice o no las hipótesis que subyacen en la teoría.

En un pasado se argumentaba que la violación de un supuesto poblacional acerca de perturbación aleatoria (homocedasticidad, no autocorrelación) debía imprescindiblemente llevar a una transformación de los datos (mínimos cuadrados generalizados, por ejemplo) que garantizase el cumplimiento de las hipótesis auxiliares. Esto llevo a una econometría de recetas y a un uso mecánico de los programas de computador. Hoy día, la econometría más madura arguye que la violación del supuesto puede ser en primer lugar un problema de una incorrecta especificación, en cuyo caso lo primero que hay que hacer es reespecificar el modelo; también se contempla la posibilidad de encontrar tipos de estimaciones robustas al problema econométrico para minimizar la transformación de los datos originales; sólo en tercer término se considera las opciones más antiguas.

Antes de pasar a discutir los tópicos pertinentes a la inferencia relacionada con los modelos que se proponen en esta investigación, es preciso dilucidar un punto en el cual ha existido confusión entre algunos econométricos. Se trata del uso de los factores de expansión cuando se trabaja con muestreo aleatorio estratificado. Es evidente que cuando el objetivo es exclusivamente la estimación de un parámetro poblacional, cada individuo debe ser expandido pues representa un estrato de diferente varianza. No obstante, cuando se tiene por objeto hacer inferencia cuyo fundamento es justamente la existencia de una muestra proveniente de la población, expandir las observaciones para que queden del tamaño de la población, es un contrasentido con la esencia del método estadístico. Lo adecuado es controlar por la presencia de heterocedasticidad, puesto que los factores de expansión son diferentes debido a las distintas varianzas de los estratos de la muestra. En épocas anteriores se sugería la aplicación de mínimos cuadrados ponderados con factor de ponderación asociada al factor de expansión. Hoy día, la heterocedasticidad puede ser vista como el resultado de una eventual falta de constancia en los parámetros como bien sugiere Swam Y Taulas (1995) y Pesaran Y Smith (1995).

Bajo este contexto el aporte en White (1980) constituiría la adecuada solución: estimar por mínimos cuadrados ordinarios, realizar el Test de White y en caso de que la hipótesis nula de homocedasticidad resulte rechazada por los datos, utilizar una matriz de varianzas covarianzas robusta a la heterocedasticidad para efectuar la inferencia estadística.

La estimación e inferencia en los modelos de la sección tercera pueden interpretarse como una adaptación de la metodología de Datos de Panel. En efecto, los modelos (1) y (2), se estiman por MCO como regresiones lineales múltiples convencionales y corresponden al modelo total. La ecuación (3), es en esencia también una regresión múltiple con N variables falsas, una por cada agrupación considerada y equivale a un modelo de Efectos Fijos. Para evitar la inclusión de todas las variables ficticias el modelo se estima en desviaciones con respecto a las medias grupales y luego se estiman las externalidades fijas de igual manera que se procede para calcular el intercepto cuando se trabaja en desviaciones.

La estimación del modelo (4) debe hacerse por Mínimos Cuadrados Generalizados y gran parte de la literatura se ha orientado a la estimación de sus varianzas. Se trata de un modelo con independencia intersegmento y autocorrelación intrasegmento, asociado al Modelo de Efectos Aleatorios.

Un punto crucial en el análisis es la verificación de la hipótesis de exogeneidad de las externalidades. Para ello se aplica un test de Hausman, utilizando los modelos de externalidades fijas y aleatorias. La intuición del test es muy sencilla: se tiene un estimador, el de efectos aleatorios, que bajo la hipótesis nula (v_i del Modelo de Externalidades Aleatorias es independiente de los regresores del modelo) es consistente y eficiente en tanto que bajo la hipótesis alterna es inconsistente; se tiene otro estimador, el de efectos fijos, que bajo ambas hipótesis es consistente. El test sigue una distribución chi-cuadrado con tantos grados de libertad como regresores hay.

El tratamiento del Modelo de Externalidades Endógenas también correspondería al de Efectos Aleatorios y algunos investigadores han orientado sus esfuerzos a la estimación de las varianzas para obtener el factor de corrección para el método de Mínimos Cuadrados Generalizados (MCG). Moulton (1986, 1987) y Rauch (1993) proceden en este sentido. Una vía sencilla, intuitiva y práctica, que coincide con la estimación MCG, consiste en proceder en dos etapas. En la primera se estima el Modelo de Externalidades Fijas, aislándose las externalidades. En la segunda se estima el modelo (5), sustituyendo a_i por la estimación de la primera etapa. La demostración para un Panel Convencional Balanceado está en Mundlak (1978), la extensión a un panel no balanceado se encuentra en Baltagi (1985) y la generalización a una situación en la cual la segunda etapa y el modelo intergrupar se estimen por Mínimos Cuadrados Ponderados (MCP) con factor de corrección asociado al MCG, se encuentra en Castellar (1998).

El procedimiento en dos etapas tiene la ventaja de hacer que cada una de ellas haga corresponder, el trabajo econométrico con una dimensión analítica diferente.

En la primera se está en el mundo intrasegmento y de su observación se pueden *cuantificar* las externalidades. En la segunda se salta al mundo intersegmento y en él se pueden *explicar* las externalidades. Además, una consecuencia de las demostraciones reseñadas en el párrafo anterior es que la igualdad paramétrica planteada en la ecuación (7), se reproduce a nivel de la estimación. Este hecho no ha sido resaltado en la literatura sobre el tema, pues no se ha hecho hincapié en que los cambios se dan a nivel paramétrico. También se demuestra que los dos modelos asociados a la dimensión intersegmento (el 5 y el 7 de la sección anterior) comparten la misma estimación del intercepto y todos los errores estándar de estimación.

En consecuencia el diseño metodológico adecuado para el problema planteado debe en primer lugar, estimar el modelo total, controlando por heterocedasticidad; luego cuantificar las externalidades mediante el modelo que las considera fijas. De resultar las anteriores externalidades significativas se debe verificar si son aleatorias o no, mediante el test de Hausman. En caso de no serlo, es procedente intentar modelizarlas y, de resultar explicadas con las concentraciones de capital humano, continuar con la estimación del modelo intersegmentos. Si tanto la conceptualización teórica sugerida como la modelización correspondiente son relevantes se espera que la tasa interna de retorno de un año promedio de educación (d_2) sea superior tanto a la tasa privada total (b_2) como a la tasa intrasegmento (τ_2).

Es importante resaltar que se intenta utilizar la metodología de datos de panel en un contexto distinto al cual fue originalmente pensada. Se debe ser muy cuidadoso cuando se enfrentan problemas econométricos adicionales. Es lógico pensar que en el Modelo de Externalidades Fijas (MEF) también pueda esperarse perturbaciones heterocedásticas y es factible estimar la matriz robusta de White. No obstante, esto implicaría que el test de Hausman debería adecuar la matriz del Modelo de Externalidades Aleatorias (MEA) para que hubiese coherencia en el proceso. También se puede argumentar que las ecuaciones de ingreso tienen un problema de selección que implican un modelo TOBIT para su estimación. Hacerlo por esta vía implicaría combinar el TOBIT con datos de panel. Dado que la propuesta es una exploración teórica-metodológica en este novedoso campo se consideró, siguiendo el ancestral principio econométrico de la parsimonia, que era mejor iniciar con estimaciones relativamente simplificadas. Adicionalmente se ha encontrado que el sesgo de selección no genera variaciones importantes. En el caso de Colombia, Ribero y Mesa (1997) confirman esta afirmación.

Cuadro No.1
Estadísticas descriptivas de la muestra

		NIVEL INDIVIDUAL					
VARIABLES		1988	1990	1992	1994	1996	1998
INGRESOS	Media	0.5060	0.4857	0.4894	0.5547	0.5050	0.4987
	Desviac	0.7512	0.5899	0.5521	0.7397	0.7961	0.6911
	Max	12.239	13.039	4.671	24.925	28.137	10.892
	Min	0.0051	0.0069	0.0059	0.0156	0.0176	0.0010
HORAS	Media	210.94	209.03	208.86	206.97	212.66	207.94
	Desviac	55.715	58.303	63.641	49.966	63.164	78.387
	Max	424.34	424.34	424.34	424.34	424.34	424.34
	Min	12.990	12.990	8.6600	8.6600	8.6600	8.6600
EDUC	Media	7.7746	7.9096	8.1052	8.3886	8.4344	8.4812
	Desviac	4.0251	4.0239	4.1379	4.0511	4.0455	4.1255
	Max	18.000	18.000	18.000	18.000	18.000	19.000
	Min	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EXPER	Media	19.068	19.192	18.697	19.082	19.807	20.157
	Desviac	13.800	13.927	13.636	13.714	14.145	14.245
	Max	85.000	80.000	79.000	75.000	73.000	85.000
	Min	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
NUMERO OBSERVACIONES		3780	3496	3199	2944	2654	3077
		NIVEL DE SEGMENTOS					
VARIABLES		1988	1990	1992	1994	1996	1998
INGRESOS MEDIOS	Media	0.5300	0.5100	0.5500	0.5900	0.5400	0.5500
	Desviac	0.4000	0.3200	0.3600	0.3900	0.4400	0.4300
	Max	3.0900	2.1100	2.1200	3.0400	3.6400	3.0200
	Min	0.2100	0.1800	0.2100	0.1700	0.2100	0.1800
HORAS MEDIAS	Media	210.97	208.93	219.33	208.37	211.94	206.94
	Desviac	17.310	18.290	18.820	18.400	21.530	24.820
	Max	266.29	268.94	282.89	285.78	311.76	319.12
	Min	168.19	151.55	153.17	159.96	127.74	128.46
EDUCAT MEDIA	Media	7.9600	8.1100	8.5100	8.7100	8.7200	8.8100
	Desviac	2.2700	2.5000	2.6000	2.4700	2.4900	2.5200
	Max	14.860	14.290	15.630	15.290	16.000	16.000
	Min	3.2900	3.4500	3.7100	4.2000	3.3300	3.6400
EXPER MEDIA	Media	19.160	19.290	18.620	18.990	19.780	20.130
	Desviac	3.8800	4.2200	4.1600	4.5000	4.8200	4.7000
	Max	35.170	34.880	31.920	33.670	42.000	37.290
	Min	9.5600	8.2200	8.3300	4.2500	5.5000	6.2000
NUMERO OBSERVACIONES		195	198	199	198	198	199

FUENTE: Cálculos en SPSS a partir de la Base de Datos del Proyecto Duración del Desempleo CIDSE COLCIENCIAS.

5. La evidencia empírica

Los modelos sugeridos fueron verificados con datos de seis etapas de la Encuesta Nacional de Hogares del DANE para el área metropolitana de Cali. La información se tomó de la base de datos que los autores construyen en la actualidad para un proyecto sobre determinantes de la duración del desempleo en Cali-Yumbo,

financiado por Colciencias. Se eligieron los trimestres finales cada dos años entre 1988 y 1998, correspondientes a las etapas 62, 70, 78, 86, 94 y 102.

El manejo informático de los datos de la Encuesta Nacional de hogares (ENH) se hizo mediante la descodificación vía SPSS de los archivos de la encuesta hasta convertirlos en un solo archivo pleno por cada uno de ellos. Luego se exportaron las variables relevantes, tanto a nivel individual, como de promedios por segmentos a hojas electrónicas. Estas hojas han sido procesadas en Econometric Views y en Limdep para las estimaciones de índole econométrica.

En el cuadro 1 se ofrecen las principales estadísticas descriptivas de las cuatro variables que intervienen en el modelo: ingresos, educación, experiencia y horas de trabajo. Las medidas se incluyen para las dos dimensiones de análisis propuestos: individual y segmentos. Para poder efectuar comparaciones válidas en el caso de los ingresos se deflactan con el IPC para Cali, base diciembre de 1988. Los valores están en cientos de miles de pesos del último trimestre de 1988. Los ingresos medios de los individuos en el mercado han fluctuados alrededor de 50 mil pesos constantes de 1988, observándose una notable dispersión en los mismos. Esta dispersión se constata por el alto rango que va desde el ingreso mínimo al máximo y puede estar influido por valores no creíbles. Este hecho es importante para la estimación del modelo intersegmento, ecuaciones (8) y (9) de la sección anterior.

La educación media ha ido incrementándose paulatinamente pasando de 7.8 años de escolaridad formal aprobada a 8.5%. Este aumento también se observaba en los valores medios de cada segmento, pasando a un promedio de 8.0 años a 8.8. De manera simplificada podría decirse que los niveles educativos observados pueden ser generados por una distribución de media 8 y varianza 4.

A partir del cuadro 2 se consignan los resultados de la inferencia estadística relacionada en los modelos sometidos a la evidencia empírica. En todas las pruebas de hipótesis se ha incluido el nivel marginal de significación o nivel de significación crítico (NSC), en lugar de los tradicionales asteriscos que denotan significancia al 5% y 1%. Esto obedece también a un cambio en la interpretación de los tests convencionales en lugar de usar mecánicamente los números mágicos del 5% y del 1%, se debe plantear que la prueba de hipótesis nula (H_0) al obtener el valor de contraste tiene un punto en el cual no será rechazada por lo datos. Ese nivel marginal asociado, el mínimo a partir del cual se rechaza H_0 , es conocido como el nivel crítico de significado. En este mundo sólo existe una regla: se rechaza H_0 al nivel e de significación si $NSC < e$.

Cuadro No. 2
Modelo total Mincer estandar

VARIABLES		AÑOS					
		1988	1990	1992	1994	1996	1998
EDUCAT	Coefic.	0.1361	0.1308	0.1326	0.1135	0.1185	0.1234
	Desviac	0.0031	0.0032	0.0032	0.0034	0.0037	0.0040
	Razon t	44.233	40.777	41.804	33.124	31.666	30.571
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EXPER	Coefic.	0.0472	0.0419	0.0426	0.0343	0.0358	0.0393
	Desviac	0.0025	0.0026	0.0026	0.0027	0.0027	0.0029
	Razon t	18.898	15.856	16.605	12.692	13.194	13.753
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EXPER2	Coefic.	-0.0006	-0.0005	-0.0005	-0.0004	-0.0004	-0.0005
	Desviac	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001
	Razon t	-11.450	-9.177	-10.215	-7.097	-8.103	-10.506
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
LOG HORAS	Coefic.	0.4657	0.5677	0.5149	0.3667	0.5038	0.6600
	Desviac	0.0444	0.0418	0.0374	0.0471	0.0405	0.0318
	Razon t	10.478	13.581	13.785	7.793	12.427	20.765
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
INTERCEPTO	Coefic.	-5.1842	-5.6511	-5.3931	-4.2606	-5.1621	-6.0953
	Desviac	0.2442	0.2266	0.2001	0.2546	0.2224	0.1762
	Razon t	-21.228	-24.935	-26.949	16.738	-23.214	-34.597
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
TEST DE WHITE	F-stats	20.107	19.264	17.369	11.450	13.694	11.964
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
TEST B5 = 1	F. CAL	144.56	106.95	168.66	181.16	149.85	114.45
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R2 AJUSTADO		0.4158	0.4004	0.4265	0.3274	0.3820	0.4105
F. CALCULADO		673.31	584.40	595.68	359.13	411.03	536.54
NSC		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

FUENTE: Cálculos en EVIEWS a partir de la Base de Datos del Proyecto Duración del Desempleo. CIDSE COLCIENCIAS

Cuadro No. 3
Cambios en la tasa de retorno por escolaridad

VARIABLES		1988	1990	1992	1994	1996	1998
EDUCAT	Coefic.	0.1107	0.0917	0.1153	0.0908	0.0560	0.0600
	Desviac	0.0109	0.0122	0.0124	0.0155	0.0147	0.0144
	Razon t	10.147	7.4907	9.3187	5.8644	3.8173	4.1733
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
SECUNDAR	Coefic.	0.0079	0.0229	-0.0070	-0.0090	0.0370	0.0368
	Desviac	0.0132	0.0146	0.0150	0.0177	0.0170	0.0174
	Razon t	0.6000	1.5768	-0.4639	-0.5106	-0.1804	2.1175
	NSC	0.5485	0.1172	0.6427	0.6097	0.0293	0.0343
UNIVERSID	Coefic.	0.0682	0.0679	0.0783	0.0940	0.0950	0.1035
	Desviac	0.0112	0.0110	0.0114	0.0120	0.0122	0.0131
	Razon t	6.0928	6.1751	6.8959	7.8786	7.8088	7.8872
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EXPER	Coefic.	0.0464	0.0414	0.0417	0.0322	0.0347	0.0386
	Desviac	0.0025	0.0027	0.0026	0.0027	0.0028	0.0029
	Razon t	18.383	15.402	15.911	11.733	12.359	13.173
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EXPER2	Coefic.	-0.0006	-0.0005	-0.0005	-0.0004	-0.0005	-0.0006
	Desviac	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001
	Razon t	-11.486	-9.287	-10.976	-6.993	-8.239	-10.806
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
LOG HORAS	Coefic.	0.4715	0.5696	0.5192	0.3805	0.5156	0.6727
	Desviac	0.0445	0.0416	0.0373	0.0469	0.0400	0.0319
	Razon t	10.605	13.701	13.927	8.113	12.884	21.101
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
INTERCEPTO	Coefic.	-5.0595	-5.4452	-5.2771	-4.1299	-4.8548	-5.7960
	Desviac	0.2473	0.2274	0.2002	0.2563	0.2288	0.1815
	Razon t	-20.458	-23.941	-26.356	-16.114	-21.216	-31.934
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
TEST DE WHITE	F-stats	12.354	12.302	11.902	8.007	8.498	7.148
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R2 AJUSTADO		0.4241	0.4106	0.4361	0.3440	0.4065	0.4301
F. CALCULADO		434.83	406.55	413.53	258.22	303.91	387.91

FUENTE: Cálculos en EViews a partir de la Base de Datos del Proyecto Duración del Desempleo. CIDSE COLCIENCIAS.

En el cuadro 2 se condensa lo encontrado para el modelo total (ausencia de externalidades), ecuación tradicional de Mincer (número 1 de la sección 3). Lo primero que llama la atención es la estabilidad de los resultados obtenidos en los seis puntos estudiados. En la totalidad de los casos se rechazó la hipótesis de homocedasticidad y por ende la inferencia se realizó con la matriz robusta de White.

En los seis trimestres los signos obtenidos para los coeficientes concuerda con lo que predice la teoría. La rentabilidad privada de la educación ha fluctuado entre 13 y 11% y la del primer año de experiencia lo ha hecho entre 4.7 y 3.5%.

La hipótesis de elasticidad horas de trabajo del ingreso laboral igual a la unidad se ve rechazada siempre. Si se quiere trabajar con el logaritmo de la tasa de salario como variable dependiente debe evaluarse las consecuencias de la imposición de la restricción. En este artículo se eligió la opción de no imponer la elasticidad unitaria.

Cuando se consideran los cambios en la rentabilidad de los niveles secundario y universitario con respecto a la educación primaria (véase la segunda parte del cuadro 3) se encuentra que en 1988 un año de educación elemental reportaba un 11.1%, mientras que uno de secundaria no reportaba un rendimiento marginal. Sin embargo, el rendimiento de la educación primaria ha venido descendiendo hasta ubicarse alrededor del 6% en 1998. En los períodos analizados no se encuentra tampoco cambios en el rendimiento marginal de la secundaria frente a la primaria. Esto quiere decir que el mercado de trabajo no discrimina entre un trabajador con educación primaria y uno que tenga secundaria. Por contra, la educación superior agregaba un 6.8% al rendimiento de la primaria en 1988, en tanto que diez años después lo hacía en 10.4%. El examen del por qué de estos cambios escapa del interés particular del artículo pero es un hecho preocupante que estén aumentando los diferenciales de la rentabilidad. Es decir, que la rentabilidad privada de la educación superior, de acuerdo con la metodología de funciones quebradas o Spline, ha permanecido alrededor del 17%. La comparación con los resultados de la OIT (1999) refuerza la afirmación hecha en la sección 3 cuando se señaló que habían cometido un error. En el cuadro 15 (página 102 de la anterior referencia), aparecen rentabilidades negativas para la secundaria (no se sabe si significativas o no, pues no se aporta la información pertinente). Habida cuenta que las tasas reportadas para la educación universitarias son cercanas al 20% la alternativa de sumar rendimientos no tiene cabida.

Al introducir la hipótesis de existencia de externalidades fijas en la función individual de ingresos (ecuación 3 de la sección 3) esta no resulta rechazada por los datos, tal cual se aprecia en el cuadro 4 cuando se consideran los convencionales test de la F y de razón de verosimilitud. En este modelo la tasa de retorno intrasegmento se estimó entre 12.4% y 9%, en tanto que los coeficientes asociados a las otras variables de la ecuación minceriana no presentan cambios importantes respecto al modelo que niega la existencia de externalidades.

Cuadro No. 4
Externalidades fijas

VARIABLES		1988	1990	1992	1994	1996	1998
EDUCAT	Coefic.	0.1244	0.1103	0.1100	0.0909	0.0951	0.0974
	Desviac	0.0032	0.0035	0.0037	0.0037	0.0039	0.0042
	Razon t	38.817	31.373	29.866	24.378	24.473	23.049
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EXPER	Coefic.	0.0478	0.0419	0.0415	0.0328	0.0334	0.0391
	Desviac	0.0023	0.0024	0.0026	0.0026	0.0025	0.0027
	Razon t	20.750	17.296	16.194	12.631	13.150	13.207
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EXPER2	Coefic.	-5.8e-04	-5e-04	-5e-04	-4e-04	-4e-04	-5.2e-04
	Desviac	4.2e-05	4e-05	5 e-05	5e-05	5e-05	5e-05
	Razon t	-13.774	-11.36	-10.98	-7.847	-9.256	-10.73
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
LOG HORAS	Coefic.	0.4715	0.5708	0.5205	0.3785	0.5381	0.6620
	Desviac	0.0319	0.0305	0.0285	0.0381	0.0311	0.0253
	Razon t	14.743	18.708	18.273	9.936	17.272	26.206
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R2 AJUSTADO		0.443	0.436	0.462	0.3891	0.4479	0.4544
F. CALCULADO		16.20	14.44	14.61	10.33	11.71	13.68
NSC		0.000	0.000	0.000	0.000	0.000	0.000
HIPOTESIS DE EXTERNALIDADES EN LA GENERACION DE INGRESOS							
TEST DE RAZON VEROSIMILITUD	CHI2	382.25	416.95	410.83	487.60	504.23	443.02
	Gdl	194	197	198	197	197	198
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
TEST DE LA F	Fc	1.964	2.118	2.074	2.507	2.604	2.248
	Gdl	194	197	198	197	197	198
		3582	3295	2997	2743	2453	2845
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

FUENTE: Cálculos en LIMDEP a partir de la Base de Datos del Proyecto Duración del Desempleo. CIDSE COLCIENCIAS

Cuadro No. 5
Externalidades aleatorias

VARIABLES		AÑOS					
		1988	1990	1992	1994	1996	1998
EDUCAT	Coefic.	0.1310	0.1220	0.1239	0.1021	0.1066	0.1111
	Desviac	0.0048	0.0031	0.0033	0.0034	0.0035	0.0038
	Razon t	44.730	38.793	37.974	30.017	30.504	28.933
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EXPER	Coefic.	0.0478	0.0418	0.0420	0.0336	0.0348	0.0377
	Desviac	0.0023	0.0024	0.0025	0.0026	0.0025	0.0027
	Razon t	20.994	17.573	16.697	13.148	14.007	14.013
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EXPER2	Coefic.	-6e-04	-5e-04	-5e-04	-4e-04	-4e-04	-5e-04
	Desviac	4e-05	4e-05	5e-05	5e-05	4e-05	5e-05
	Razon t	-13.80	-11.21	-11.06	-7.937	-9.614	-11.03
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
LOG HORAS	Coefic.	0.4696	0.5682	0.5193	0.3746	0.5207	0.6605
	Desviac	0.0314	0.0299	0.0280	0.0373	0.0305	0.0248
	Razon t	14.975	18.962	18.549	10.046	17.081	26.637
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
INTERCEPT	Coefic.	-5.166	-5.574	-5.324	-4.182	-5.125	-5.948
	Desviac	0.1719	0.1645	0.1543	0.2035	0.1682	0.1397
	Razon t	-30.05	-33.88	-34.49	-20.55	-30.47	-42.57
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R2		0.42	0.40	0.42	0.33	0.38	0.41
HIPOTESIS DE EXOGENEIDAD DE LAS EXTERNALIDADES							
TEST DE HAUSMAN		33.13	55.96	72.22	57.84	56.29	60.20
GDL		4	4	4	4	4	4
NSC		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

FUENTE: Cálculos en LIMDEP a partir de la Base de Datos del Proyecto Duración del Desempleo. CIDSE COLCIENCIAS

El descenso en la tasa intrasegmento respecto a la total puede interpretarse como el hecho de que la probabilidad de pertenecer a un determinado segmento no es independiente del nivel educativo del individuo Oliver, Raymond, Roig Y Roca (1998).

El aspecto relevante es que la inclusión de una prima diferencial en los ingresos laborales por vivir en un determinado segmento no es incompatible con la observación de la realidad. Dicha inclusión no cambia los coeficientes asociados a la experiencia y a las horas de trabajo, pero sí lo hace respecto al de la educación.

Es el momento de considerar la posibilidad de que las externalidades sean tratadas como aleatorias. En caso de serlo podrían ser incorporadas al término aleatorio de error ganándose eficiencia en la estimación de los parámetros. De no

serlo su tratamiento como aleatorias induciría inconsistencias en la misma estimación. El cuadro (5) resume los resultados de la estimación del MEA cuyo principal objetivo es la prueba acerca de la exogeneidad de las externalidades. Con esta perspectiva se ejecuta el test de Hausman encontrándose un claro rechazo a la hipótesis de exogeneidad. En otras palabras se puede decir que las externalidades son endógenas al estar correlacionadas con los regresores del modelo. El hecho de que el coeficiente que más cambia sea la tasa de retorno de la educación nuevamente apunta hacia que las externalidades dependan del nivel educativo.

Es muy importante resaltar que el resultado se mantiene en todas las etapas analizadas, lo cual va incrementando la evidencia de que lo que se ha hallado, es una regularidad empírica.

Es procedente entonces intentar modelar las externalidades cuantificadas cuando se consideran fijas. Como se anotó en la sección tercera una primera vía es utilizar como determinantes de las externalidades a las medias de las variables en cada segmento (ecuación 5). En el cuadro 6 se presentan los resultados de este ejercicio, pudiéndose afirmar que el impacto de un año promedio de educación en un segmento inducía en diciembre de 1988 un aumento de 3.2% en la prima diferencial del ingreso de los individuos pertenecientes a dicho segmento, en tanto que en 1998 lo hacía en 8.1%. Los otros regresores, tomados en media, no resultan significativos, comprobando la dependencia que tiene la externalidad con respecto a la concentración de capital humano en el segmento.

Pareciera que el impacto de un año promedio de educación en la prima diferencial viniese aumentando pero la constancia de este hecho supera los límites de este trabajo pero puede ser un tema para futuras investigaciones. Lo importante es que en las seis veces que se verificó este modelo de determinantes de externalidades, la única variable que resultó siempre significativa fue la educación media del segmento.

En el cuadro 7 se estima el modelo de la ecuación (7) para la dimensión agregada, esto es, la intersegmento. La estimación de la tasa de rendimiento del segmento estaría entre 15.4 y 17.9%, muy superior a la rentabilidad individual.

En la metodología propuesta se señalaba que al construir el modelo intersegmento, a partir de las medias grupales de las variables el coeficiente que acompaña a la educación debe ser igual a la suma de la tasa de retorno intra-segmentos y del impacto de un año medio de educación en las externalidades.

Si se suma para cualquier año, por ejemplo 1998, la tasa intrasegmento del cuadro 4 (9.74%) en el impacto de un año medio de educación en la externalidad del cuadro 6 (8.13%) se obtiene exactamente la tasa intersegmento del cuadro 7 (17.87%). Esto lo puede verificar el lector curioso con la totalidad de los coeficientes de los tres modelos involucrados.

También se comprueba que en la estimación de los dos modelos de la dimensión intersegmento que aparecen en los cuadros 5 y 6 se comparten los mismos errores estándar y los interceptos. Como varias veces se ha sugerido el modelo intersegmentos para los ingresos que aparece en el cuadro 7 toma las medias de los

logaritmos tanto del ingreso como de las horas trabajadas. Las medias de los logaritmos corresponden al logaritmo de la media geométrica la cual puede ser una medida de tendencia central cercana a la media aritmética cuando el interior de cada grupo es relativamente homogéneo.

En los distintos modelos de la dimensión intersegmento la evidencia en contra de la hipótesis de homocedasticidad no fue uniforme. En algunos casos no se rechazó (dejando la matriz de varianzas covarianzas en mínimos cuadrados ordinarios) y en otros si. Esto evidencia que las correcciones mecánicas que presuponen un tipo de heterocedasticidad deben evitarse.

Cuadro No. 6
Modelos de la dimensión intersegmento
determinantes de externalidad

VARIABLES		AÑOS					
		1988	1990	1992	1994	1996	1998
EDUCAT	Coefic.	0.0320	0.0560	0.0601	0.0632	0.0538	0.0813
	Desviac	0.0101	0.0114	0.0080	0.0079	0.0094	0.0094
	Razon t	3.1789	4.9253	7.5310	8.0179	5.7195	8.6346
	NSC	0.0017	0.0000	0.0000	0.0000	0.0000	0.0000
EXPER	Coefic.	-0.0153	-0.0029	-0.0021	0.0013	0.0278	-0.003
	Desviac	0.0126	0.0164	0.0118	0.0136	0.0151	0.0153
	Razon t	-1.2181	-0.1741	-0.1793	0.0952	1.8492	-0.219
	NSC	0.2247	0.8620	0.8579	0.9243	0.0660	0.8271
EXPER MEDIA2	Coefic.	0.0001	0.0001	-8.8e-06	67e-08	-0.0005	0.0002
	Desviac	0.0002	0.0003	0.0002	0.0003	0.0002	0.0003
	Razon t	0.4875	0.4607	-0.0459	0.0022	-2.171	0.6707
	NSC	0.6265	0.6456	0.9635	0.9982	0.0312	0.5032
MEDIA LOG HORAS	Coefic.	-0.0037	0.1604	-0.2462	-0.256	-0.343	-0.024
	Desviac	0.1648	0.1491	0.1563	0.1728	0.1565	0.1160
	Razon t	-0.0223	1.0759	-1.5760	-1.306	-2.194	-0.207
	NSC	0.9821	0.2833	0.1167	0.1932	0.0295	0.8365
INTERC	Coefic.	-5.1282	-6.7938	-4.3523	-3.461	-3.955	-6.417
	Desviac	0.9398	0.8083	0.8418	0.9372	0.8552	0.5951
	Razon t	-5.4567	-8.4054	-5.1705	-3.6929	-4.6251	-10.78
	NSC	0.0000	0.0000	0.0000	0.0003	0.0000	0.0000
TEST DE WHITE	F-stats	6.3214	4.3733	3.9116	1.4683	3.6051	3.3522
	NSC	0.0000	0.0001	0.0003	0.1712	0.0006	0.0013
R2 AJUSTADO		0.1951	0.2836	0.3808	0.2750	0.2702	0.3453
F. CALCULADO		12.760	20.500	31.443	19.679	19.238	27.111
NSC		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

FUENTE: Cálculos en EVIEWS a partir de la Base de Datos del Proyecto Duración del Desempleo. CIDSE COLCIENCIAS

Cuadro No.7
Modelos de la dimensión intersegmento
Intersegmentos

VARIABLES		AÑOS					
		1988	1990	1992	1994	1996	1998
EDUCAT	Coefic.	0.1564	0.1663	0.1701	0.1541	0.1489	0.1787
	Desviac	0.0101	0.0114	0.0080	0.0079	0.0094	0.0094
	Razon t	15.531	14.622	21.311	19.539	15.823	18.985
	NSC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EXPER	Coefic.	0.0325	0.0390	0.0394	0.0341	0.0613	0.0328
	Desviac	0.0126	0.0164	0.0118	0.0136	0.0151	0.0153
	Razon t	2.5889	2.3849	3.3454	2.4975	4.0627	2.1402
	NSC	0.0104	0.0181	0.0010	0.0133	0.0001	0.0336
EXPER MEDIA 2	Coefic.	-0.0005	-0.001	-0.001	-0.0004	-0.001	-0.0004
	Desviac	0.0002	0.0003	0.0002	0.0003	0.0002	0.0003
	Razon t	-1.9700	-1.538	-2.784	-1.4648	-3.821	-1.352
	NSC	0.0503	0.1256	0.0059	0.1446	0.0002	0.1781
MEDIA LOG HORAS	Coefic.	0.4678	0.7312	0.2743	0.1529	0.1947	0.6381
	Desviac	0.1648	0.1491	0.1563	0.1728	0.1565	0.1160
	Razon t	2.8389	4.9043	1.7554	0.8844	1.2440	5.5027
	NSC	0.0050	0.0000	0.0808	0.3776	0.2150	0.0000
INTERC	Coefic.	-5.128	-6.794	-4.352	-3.4610	-3.955	-6.417
	Desviac	0.9398	0.8083	0.8418	0.9372	0.8552	0.5951
	Razon t	-5.457	-8.405	-5.171	-3.6929	-4.625	-10.78
	NSC	0.0000	0.0000	0.0000	0.0003	0.0000	0.0000
TEST DE WHITE	F-stats	6.321	4.373	3.912	1.4683	3.6051	3.3522
	NSC	0.0000	0.0001	0.0003	0.1712	0.0006	0.0013
R2 AJUSTADO		0.7447	0.7506	0.8050	0.6809	0.6852	0.7384
F. CALCULADO		142.49	149.23	205.35	106.10	108.20	140.74
NSC		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

FUENTE: Cálculos en EViews a partir de la Base de Datos del Proyecto Duración del Desempleo. CIDSE COLCIENCIAS

6. Conclusiones

La estabilidad de las pruebas realizadas conforma un resultado robusto al tipo de muestra utilizado; es claro que se trata de regularidades empíricas que abren nuevas líneas de investigación en la microeconomía aplicada al mercado de trabajo.

Se comprueba la existencia de una prima diferencial en el ingreso de los individuos como consecuencia de estar residiendo en un determinado segmento de población. Esta prima diferencial puede ser atribuida a la existencia de una externalidad monetaria, la cual a su vez depende del nivel de concentración de

capital humano del segmento, es decir en la dimensión intersegmento la externalidad es endógena.

Como consecuencia de lo anterior, la medición de la rentabilidad de la educación mediante la ecuación minceriana tradicional, el modelo total, reporta valores entre 11.3 y 13.8%, que esconden la interacción de dos dimensiones del problema. En la dimensión intrasegmento, las tasas estarían entre 9 y 12% después de descontar la prima diferencial. En la dimensión intersegmento, cuya medición se acerca más a una tasa de rentabilidad social, se encuentran alrededor del 17%.

El análisis de las externalidades obliga a tener en cuenta, no sólo la dimensión individual, para medir la tasa de retorno de la educación sino también la dimensión colectiva. Las políticas que solamente tienen en cuenta estimativos sociales que parten de la tasa de retorno privada, subvaloran la verdadera tasa social. Para tomar una decisión como el traslado de recursos entre niveles educativos, es preciso clarificar el papel de los distintos grados educativos en las externalidades, estimaciones que los autores tienen en proceso.

El haber encontrado que la educación secundaria no reporta rendimientos marginales significativos respecto a la primaria obliga a repensar la propuesta del gobierno puesto que la universitaria en este contexto si es muy importante. Esto debe ser visto de nuevo en las dos dimensiones. Relacionado con esto es preciso mencionar el costo de oportunidad de estudiar cuando la rentabilidad de no hacerlo esta sobrestimada por narcotráfico y delincuencia común que son alternativas llamativas en nuestro medio.

Una limitación del análisis realizado es que sólo se cuantifican economías externas que se traducen directamente en ingresos individuales. No se tienen en cuenta otro tipo de externalidades que genera la educación entre las que podríamos mencionar :

- La disminución de las tasas de fecundidad.
- A mayor educación, mayor deseo de mantener orden público lo que implica menor costo de mantenerlo.
- Mejoras en el estado de salud de la población.
- Elección de mejores opciones políticas cuando se toman decisiones participativas.
- Mejores decisiones en tiempo libre y mayor acceso a la cultura.

Bibliografía

- Baltagi, B.H. (1985), "Pooling cross-section with unequal time-series length", *Economics Letters* 18, 133- 136.
- Baltagi, B. H (1995), *Econometric Analysis of Panel Data*, John Wiley & Sons, Chichester
- Baumol And Oates (1975), *The theory of Environmental Policy*, Prentice-Hall, New Jersey.
- Berndt, E. (1991), "The practice of Econometrics Classic and Contemporary". Addison Westey Publishing Company, Inc.
- Castellar, C (1998), *Eficiencia productiva, valoración de fincas y externalidades territoriales en la economía campesina colombiana*. Tesis doctoral. Universidad Autónoma de Barcelona.
- Clavijo, S (1998), *Descentralización de la educación y la salud*. Documentos CEDE 98-11.
- DANE (1991), *Veinte años de la encuesta nacional de hogares de Colombia*.
- DNP (1998), *Cambio para construir la paz*, Bogotá.
- Gonzalez F., Guzman C. Y Pachon (1998), *Productividad y retornos sociales del Capital Humano: Microfundamentos y evidencia para Colombia*, Archivos de Macroeconomía, Departamento Nacional de Planeación, Documento 98, Santa Fe de Bogotá, D.C.
- Hausman (1978), *Specification Test in Econometrics*, *Econometría* Vol. 46 Nº 6, pp. 1251-1271
- Hsiao, Ch (1986), *Analysis of Panel Data*, Cambridge University Press, New York
- Johnes G (1993), *Economía de la educación*, Ministerio del Trabajo y Seguridad Social. Madrid, 252 páginas.
- Jhonston (1987), *Métodos de Econometría*, 3ª edición, Editorial. Vicens Vives, Barcelona.
- Lucas R. (1988), "On the Mechanics of Economics Development", *Journal of Monetary Economics*, Vol. 22 Nº 1, July, pp. 3 - 42.
- Marshall A. (1920), *Principles of Economics*, MacMillan, London.
- Moulton (1986), "Random Group Effects and the precision of regression estimates", *Journal of Econometrics*, 32 pp. 385 - 397.
- Moulton (1987), "Diagnostics for Group Effects in Regression Analysis", *Journal of Business & Economics Statistics*, Vol. 5 Nº 2 pp. 275 - 282
- Mundlak (1978), "On the pooling of Time Series and Cross Section Data", *Econometrica* Vol. 46 Nº 1, pp. 69 - 85.
- Ohlin B. (1933), *Interregional International Trade*, Harvard University Press, Cambridge, Massachusetts.
- OIT (1999), *Empleo Un desafío para Colombia*, Organización Internacional del Trabajo, Lima.
- Oliver J., Raymond J.L., Roig J.L Y Roca A. (1998), *Función de Ingresos y rendimiento de la Educación en España 1990*, FIES, Documento de trabajo Nº 100.
- Pesaran, M. And Smith, R. (1995), "Estimating Logn-Run Relationships from Dynamic Heterogenous Panels", *Journal of Econometrics* 68, pp 79-113.
- Swamy, P And Taulas, G (1995), "Random Coefficients Models: Theory and Applications". *Journal of Economics Survey*, Vol. 9, No 2, pp 165-196.
- White, H (1980), "A Heteroskedasticity - Consistent Matriz Estimatore and a Direct test for Heterskedasticity". *Econometrics* 48, pp 817-38.
- Psacharopulus, G. (1981), "Returns to Education: an updated international comparison", *Comparative Education*, Vol. 17 pp. 321-341.

- Psacharopoulos, George (1991), *El impacto económico de la educación*. Centro Internacional para el Crecimiento Económico. Santo Domingo, República Dominicana.
- Psacharopoulos, George (1993), "Returns to Investment in Education", *World Bank Research Observer*. Vol 10, # 2, August.
- Rauch J (1993), "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities". *Journal of Urban Economics*, Vol 34, pp 380-400.
- Raymond J. L. (1995), *Exportaciones y Crecimiento Económico*, FIES Documentos de Trabajo N° 115, Barcelona.
- Rivero, R Y Mesa, C.(1997), *Ingresos laborales de hombres y mujeres en Colombia: 1976 - 1995*; Archivos de Macroeconomía; Departamento Nacional de Planeación; Documento 62, Santa Fe de Bogotá.
- Tenjo, J (1993), "Evolución de los retornos a la inversión en educación 1976-1989", en *Planeación y Desarrollo*, volumen XXIV, diciembre, Bogotá.